FAKULTÄT FÜR
INFORMATIK

Institut für Simulation und Graphik

Visual Analytics of Medical and Biological Data

Habilitationsschrift

zur Erlangung der Venia Legendi für

Informatik

angenommen durch die Fakultät für Informatik
der Otto-von-Guericke-Universität Magdeburg

von: Dr.-Ing. Steffen Oeltze-Jafra

geb. am 22.09.1977 in Magdeburg

Gutachterinnen/Gutachter:

Prof. Dr.-Ing. Bernhard Preim
Prof. Dr.rer.nat. Dr.techn.hc. Dr.-Ing.E.h. Thomas Ertl
Prof. Dr.ir. Jarke J. van Wijk

Magdeburg, den 27. Juni 2016

Oeltze-Jafra, Steffen (Dr.-Ing.):
Visual Analytics of Medical and Biological Data
Habilitation, Otto-von-Guericke-Universität Magdeburg,
27. Juni 2016.

Für Edith, Marlene,
Ursula und Klaus-Dieter

Acknowledgments

I am very grateful for the kind and constant support of my work by many people. My colleagues and collaborators from the field of computer science were a great source of inspiration in solving visualization-related problems. My collaborators from the application domains medicine and biology luckily never stopped creating and working on such problems. My friends helped in taking my mind off these problems whenever needed and my family was the safe haven putting things in perspective when I was at risk of overdoing it.

I am particularly indebted to Bernhard Preim, leader of the visualization group at the Institute of Simulation and Computer Graphics (ISG) at the Otto-von-Guericke University Magdeburg (OVGU), Germany, for paving the way for my career in visualization, the countless advices, and the uninterrupted motivation. I thank my current and former colleagues of the visualization group for the productive and inspiring exchange, the great teamwork also in teaching, as well as for the fun at work and beyond. I would like to give special thanks to my roommates Sylvia Glaßer and Paul Klemm for the friendly, funny, and relaxed work climate in our office. Administrative affairs would have been time-consuming and burdensome without the competent and patient support of Stefanie Quade, Petra Schumann, and Petra Specht from the ISG's secretariat. I deeply appreciate the support of the ISG's technicians Heiko Dorwarth, Volkmar Hinz, and Thomas Rosenburg. One tends to underestimate the significance of smoothly running hardware.

Over the last five years, I have collaborated with many researchers in interdisciplinary projects. Forces were bundled in investigating (1) cerebral aneurysm hemodynamics, (2) protein networks in cells and tissues, (3) socio-demographic, lifestyle-related, and health-related factors of populations, and (4) tissue perfusion.

(1) I thank Gábor Janiga and Philipp Berg from the Institute of Fluid Dynamics and Thermodynamics at the OVGU as well as Juan R. Cebral from the Department of Bioengineering, George Mason University, VA, U.S. for providing simulated hemodynamic data and sharing their knowledge about aneurysm hemodynamics and Computational Fluid Dynamics simulations. I am grateful to Oliver Beuing and Martin Skalej from the University Hospital Magdeburg, Germany for providing most of the underlying medical image data. Their expert knowledge and the valuable comments of Uta Preim from the Community Hospital Magdeburg, Germany led to an increased clinical credibility of my work. I very much appreciate the fruitful discussions about flow topology and visualization techniques with Holger Theisel, Tobias Günther, Dirk J. Lehmann, and Christian Rössl from the visual computing group at the ISG and with Alexander Kuhn from the Zuse Institute, Berlin, Germany. I am indebted to my former colleagues Rocco Gasteiger and Mathias Neugebauer who introduced me to the specifics of visually exploring aneurysm hemodynamics.

(2) I am thankful to Walter Schubert and Reyk Hillert from the section Functional Genomics and Medical Toponomics of the University Hospital Magdeburg, Germany for providing image data of protein networks, sharing their biological knowledge, and many enlightening discussions about cell and tissue phenomena. I am much obliged to Helmut Doleisch, Wolfgang Freiler, and Philipp Muigg from CD-adapco, Vienna, Austria (all were previously with the SimVis GmbH, Vienna, Austria) for their technical support in implementing a visual analytics approach to investigating protein networks within their software framework SimVis.

(3) I would like to thank my colleague Paul Klemm as well as Helwig Hauser and Paolo Angelelli from the visualization group at the Department of Informatics, University of Bergen, Norway for the very inspiring and productive joint work on the investigation of data from the Study of Health in Pomerania (SHiP) and the Norwegian Aging Study, respectively. I gratefully acknowledge the assistance in gaining access to the data of SHiP and the epidemiological counsel by Henry Völzke from the Institute for Community Medicine, University of Greifswald, Germany. I thank

Kathrin Hegenscheid from the Institute for Diagnostic Radiology and Neuroradiology at the same university for her expert guidance in investigating lumbar spine variability and its relation to back pain. I deeply appreciate the granting of access to data from the Norwegian Aging Study by Arvid Lundervold from the Department of Biomedicine, University of Bergen, Norway. I am indebted to Emrah Düzel, Hartmut Schütze, and Anne Maaß from the Institute of Cognitive Neurology and Dementia Research (IKND) at the OVGU for introducing me to dementia research and providing population image data and segmentations of the hippocampus.

(4) I am grateful to my colleague Sylvia Glaßer and to Claire Chalopin from the Innovation Center Computer Assisted Surgery (ICCAS) at the University Leipzig, Germany for the fruitful collaborations within the scope of brain tumor perfusion. Atle Bjørnerud from the Intervention Centre, Oslo University Hospital, Norway and Dirk Lindner from the Clinic and Polyclinic for Neurosurgery at the University Hospital Leipzig, Germany provided the perfusion image data within these collaborations and contributed their medical expert knowledge.

I wish to extend sincere thanks to the bachelor and master students I supervised, in particular, Paul Klemm and Franz Pieper, who spent extra hours investigating protein networks in preparing two publications. I very much appreciate the advices of Timo Ropinski from the Institute of Media Informatics at Ulm University, Germany in preparing this postdoctoral thesis.

Last but not least, I want to express my unwavering love and gratitude to my wonderful wife Edith and my lovely daughter Marlene. You bring joy and happiness to my life contributing to the strength needed for accomplishing this work. Sharing the success of each individual step towards the habilitation with you meant much more than the success itself. Thank you!

Preface

Before the contents of this postdoctoral thesis are presented, the author is briefly introduced and the context of his work and his major contributions are described. A few comments are given on the structure of this thesis and on each publication contained in the cumulative Parts I-V, including where it was published and if the work was accomplished within a specific funded project. Finally, the type-setting and formatting of the cumulated publications are explained and a statement on the author's contributorship is given.

This thesis surveys the research work of Steffen Oeltze-Jafra (né Oeltze) accomplished in the years 2011-2015. After finishing his PhD project on the visual exploration and analysis of medical perfusion data [164], Steffen pursued the idea of integrating visualization, data analysis, and exploration techniques for the investigation of scientific data. He expanded his research by applications from interventional neuroradiology, neurosurgery, molecular biology, and epidemiology.

Early visual data mining and visual analytics research mainly employed techniques from information visualization and focused on non-image data, either with no inherent spatialization or a geo-spatial reference. Not at the least the work of Steffen Oeltze-Jafra conveys the potential of visual analytics and of the related field interactive visual analysis in investigating scientific image and simulation data from medicine and biology. The visual analytics approaches cumulated in Parts I-IV of this thesis help domain scientists to broaden their accustomed data analysis strategies from confirmatory to exploratory analysis and from hypothesis verification to hypothesis generation. The respective case studies demonstrate how insights can be extracted from image data and associated as well as derived attributes by looping through the visual analytics process from raw data to knowledge.

Chapter 1 serves as an introduction to the thesis at hand. It motivates the science of visual analytics in a medical and biological context, gives an application-independent introduction to visual analytics, equips the reader with medical and biological background knowledge, surveys the publications cumulated in this thesis as well as related work, provides a short excursion into further medical and biological applications of visual analytics, which have not been in the focus of this thesis, and closes with a summary, a discussion, and a vision on challenges and future research directions.

Part I comprises two publications dedicated to the visual analytics of cerebrovascular hemodynamic data. An introduction to this medical part is given in Section 1.3. The research presented in both publications was accomplished within an ongoing long-term collaboration with computational fluid dynamics engineers of the University of Magdeburg, Germany. Together, we seek a better understanding of the mechanisms behind cerebral aneurysm rupture and prediction facilities for the therapeutic outcome of aneurysm treatment. *Chapter 2* equals a manuscript that was published in IEEE Transactions on Visualization and Computer Graphics (IEEE TVCG) in 2014 and presented in the same year at the IEEE VIS conference (TVCG track) in Paris, France [168]. It describes the study of different clustering techniques for the visual analysis of blood flow in cerebral aneurysms in the context of virtual stenting. This study was accomplished jointly with visual computing researchers of the University of Magdeburg, Germany (one researcher transfered to the Zuse Insitute, Berlin shortly before publication). *Chapter 3* equals a manuscript that was presented at the IEEE VIS conference (SciVis track) in Chicago, Illinois, USA in 2015 and published in IEEE Transactions on Visualization and Computer Graphics in early 2016 [171]. It presents a clustering-based pipeline for the automatic detection and visualization of vortical blood flow in cerebral aneurysms with a focus on embedded vortices. This pipeline was developed and evaluated in collaboration with a cerebrovascular hemodynamics researcher from George Mason University, Fairfax, Virginia, USA.

Part II comprises three publications dedicated to the visual analytics of toponome data of cells and tissues. An introduction to this biological part is given in Section 1.4. All publications describe results of a long-term collaboration with a molecular biologist and a computer scientist of the medical faculty of the University of Magdeburg, Germany. Together, we pursue an improved understanding of cell structure and function based on protein topology, i.e. the toponome, in fluorescence microscopy images of cells and tissues. *Chapter 4* equals a manuscript that was presented at the VisWeek (SciVis track) in Providence, Rhode Island, USA in 2011 and published in the same year in IEEE Transactions on Visualization and Computer Graphics [166]. It proposes a graph-based interactive visual analysis approach to studying protein topology. This approach was developed in collaboration with two expert researchers in interactive visual analysis of scientific data of the SimVis GmbH, Vienna, Austria (both are now with CD-adapco, Vienna). *Chapter 5* equals a manuscript that was published in the proceedings of the Eurographics Workshop on Visual Computing for Biology and Medicine (VCBM), Norrköping, Sweden, 2012 [167]. It describes visualization and exploration techniques tailored to the analysis of 3D toponome data. The manuscript is based on a master's thesis supervised by Steffen Oeltze-Jafra and the molecular biologist [108]. *Chapter 6* equals a manuscript that has been published in the proceedings of the Eurographics Workshop on Visual Computing for Biology and Medicine (VCBM), Vienna, Austria, 2014 [172]. It presents an approach for the in-place textual and symbolic annotation of protein topology in 2D views. The manuscript is based on a bachelor's thesis supervised by Steffen Oeltze-Jafra and the molecular biologist [177].

Part III comprises five publications dedicated to the visual analytics of population study data. An introduction to this epidemiological part is given in Section 1.5. The first three publications in Chapters 7-9 present results of a long-term collaboration with epidemiologists and a radiologist of the University of Greifswald, Germany and image processing researchers of the University of Magdeburg, Germany. In the scope of the DFG Priority Program 1335 "Scalable Visual Analytics", we jointly developed visual analytics approaches for the concurrent investigation of image and non-image cohort data from the Study of Health in Pomerania (SHiP). *Chapter 7* equals a manuscript that was accepted for but has not been published yet in the proceedings of the Workshop Visualization in Medicine and Life Sciences (VMLS), Leipzig, Germany, 2013 [181]. The manuscript is based on a short paper that was presented at the workshop within the scope of the EG/VGTC Conference on Visualization (EuroVis). It surveys and provides visions on the visual analytics of epidemiological cohort studies incorporating medical image data. The survey was conducted in collaboration with a research expert in interactive visual analysis of cohort study data from the University of Bergen, Norway. *Chapter 8* equals a manuscript that was published in the proceedings of the Workshop on Vision, Modeling and Visualization (VMV), Lugano, Switzerland, 2013 [112]. It proposes a clustering-based approach to the visual analysis of lumbar spine variability in the SHiP cohort. *Chapter 9* equals a manuscript that was presented at the IEEE VIS conference (VAST track) in Paris, France in 2014 and published in the same year in IEEE Transactions on Visualization and Computer Graphics [113]. It describes a web-based visual analytics system for the concurrent investigation of image and non-image cohort study data. *Chapter 10* equals a manuscript that was published in IEEE Computer Graphics and Applications (CG&A) in 2014 and presented in 2015 at the IEEE VIS conference (CG&A track) in Chicago, Illinois, USA [4]. It proposes a data organization model for the seamless integration of heterogeneous cohort study data and their interactive visual analysis. This research was accomplished in a collaborative effort by neuroscientists of the University of Bergen, Norway and researchers in interactive visual analysis of scientific data from the same university and the University of Magdeburg, Germany. *Chapter 11* equals a manuscript that was published in the proceedings of the Workshop Bildverarbeitung für die Medizin (BVM), Aachen, Germany, 2014 [170]. It describes an approach

to measuring the hippocampal substructure Stratum Radiatum/Lacunosum-Moleculare (SRLM) in cohort study image data from ultra-high field 7-Tesla Magnetic Resonance Imaging (MRI). The approach was developed and evaluated in a collaboration with neuroscientists of the medical faculty at the University of Magdeburg, Germany and of the German Centre for Neurodegenerative Diseases (DZNE), site Magdeburg.

Part IV comprises two publications dedicated to the visual analytics of perfusion data. Since these data were not in the research focus of the habilitation, only a brief introduction to this medical part is given in Section 1.6.1. *Chapter 12* equals a manuscript that has been published in the proceedings of the SPIE Conference on Medical Imaging, Lake Buena Vista, Florida, USA, 2013 [63]. It proposes a visual analytics approach to investigating brain tumor perfusion in longitudinal imaging studies. The research was accomplished in the scope of the DFG Priority Program 1335 "Scalable Visual Analytics" and in close collaboration with physicians from the Rikshospitalet-Radiumhospitalet Medical Centre, Oslo, Norway and the Municipal Hospital of Magdeburg, Germany and research experts in interactive visual analysis of perfusion data of the University of Bergen, Norway and the University of Magdeburg, Germany. *Chapter 13* equals a manuscript that was published in the proceedings of the Jahrestagung der Deutschen Gesellschaft für Computer- und Roboterassistierte Chirurgie (CURAC), Innsbruck, Austria, 2013 [28]. It presents a method for the evaluation of intraoperative ultrasound perfusion imaging in brain tumor surgery. The method was developed and evaluated together with neurosurgeons of the University Hospital, Leipzig, Germany and a medical imaging researcher from the Innovation Center Computer Assisted Surgery (ICCAS), Leipzig, Germany.

Part V comprises two surveys which do not build on each other, are self-contained, and hence, require no further introduction within this thesis. *Chapter 14* equals a manuscript that was published in Computers and Graphics in 2011 [191]. It surveys glyph-based visualization techniques for spatial multivariate medical data and provides a taxonomy as well as guidelines for glyph design. The survey has been conducted in a collaboration of research experts in glyph visualization from the University of Münster, Germany (the researcher is now with Ulm University, Germany) and the University of Magdeburg, Germany. *Chapter 15* equals a manuscript that has been published in the proceedings of the Eurographics Workshop on Visual Computing for Biology and Medicine (VCBM), Vienna, Austria, 2014 [172]. It surveys labeling techniques in medical visualizations and provides a taxonomy as well as usage guidelines.

The typesetting of all cumulated publications in Parts I-V has not been modified. The manuscripts were included as they were reviewed and accepted. No changes have been made to the texts and illustrations. Only header, footer, and page numbers were removed and replaced by a chapter header and a running number.

Some final remarks are necessary concerning the authorship and contributorship of all cumulated publications in Parts I-V. None of the papers has been authored solely by Steffen Oeltze-Jafra. The main reason is that Steffen's research is application-driven and inter-disciplinary. He works together with domain scientists in collaborative projects. Furthermore, his research is assisted by undergraduate and graduate students. It can be stated that Steffen Oeltze-Jafra contributed significantly to all publications and often, even took the lead in the related projects. Detailed statements of author contributorship consensus were signed by each co-author of each publication. They are enclosed with the submitted version of this thesis but omitted in the public version to protect the signatures from counterfeiting. The unsigned statements are available at `http://wwwisg.cs.uni-magdeburg.de/~stoeltze/contributorship.zip`.

Steffen Oeltze-Jafra, February 17, 2016.

Contents

Acknowledgments **v**

Preface **vii**

1 Visual Analytics of Medical and Biological Data **1**
1.1 Motivation . . . 1
1.2 Introduction to Visual Analytics . . . 2
1.2.1 The Visual Analytics Process . . . 2
1.2.2 Building Blocks of Visual Analytics Research . . . 4
1.2.3 Interactive Visual Analysis . . . 6
1.3 Simulated Cerebrovascular Hemodynamic Data . . . 8
1.3.1 Cerebral Aneurysms . . . 9
1.3.2 Role of Aneurysm Hemodynamics in Diagnosis and Treatment . . . 10
1.3.3 Hemodynamic Data Generation . . . 11
1.3.4 Hemodynamic Parameters . . . 13
1.3.5 Visual Analytics: Potential and Approaches . . . 14
1.3.6 Relation to Measured Cardiac Hemodynamics . . . 18
1.4 Toponome Data of Cells and Tissues . . . 19
1.4.1 The Toponome and Toponomics . . . 20
1.4.2 Imaging the Toponome and Data Transformation . . . 21
1.4.3 Toponome Analysis . . . 23
1.4.4 Visual Analytics: Potential and Approaches . . . 25
1.5 Epidemiological Population Study Data . . . 28
1.5.1 Role of Population Studies in Epidemiological Research . . . 28
1.5.2 Specifics of Population Study Data and their Acquisition . . . 30
1.5.3 Standard Epidemiological Data Analysis Workflow . . . 31
1.5.4 Visual Analytics: Potential and Approaches . . . 32
1.6 Further Biomedical Data . . . 35
1.6.1 Perfusion Data . . . 35
1.6.2 Electronic Health Records . . . 36
1.6.3 Omics Data . . . 37
1.7 Summary and Discussion . . . 38
1.8 Future Research Directions . . . 45

Part I: Cerebrovascular Hemodynamic Data **49**

2 Oeltze et al., "Blood Flow Clustering and Applications in Virtual Stenting of Intracranial Aneurysms", *IEEE Trans. Vis. Comput. Graphics (IEEE TVCG)*, vol. 20, no. 5, pp. 686-701, 2014. **51**

3 Oeltze-Jafra et al., "Cluster Analysis of Vortical Flow in Simulations of Cerebral Aneurysm Hemodynamics", *IEEE Trans. Vis. Comput. Graphics (IEEE TVCG)*, vol. 22, no. 1, pp. 757-766, 2016. **67**

Part II: Microscopy Data of Cells and Tissues **83**

4 Oeltze et al., "Interactive, Graph-Based Visual Analysis of High-Dimensional, Multi-Parameter Fluorescence Microscopy Data in Toponomics", *IEEE Trans. Vis. Comput. Graphics (IEEE TVCG)*, vol. 17, no. 12, pp. 1882-1891, 2011. **85**

5 Oeltze et al., "Visualization and Exploration of 3D Toponome Data", *Eurographics Workshop on Visual Computing for Biology and Medicine (EG VCBM)*, pp. 115-122, 2012. **95**

6 Oeltze-Jafra et al., "Interactive Labeling of Toponome Data", *Eurographics Workshop on Visual Computing for Biology and Medicine (EG VCBM)*, pp. 79-88, 2014. **103**

Part III: Epidemiological Population Study Data **113**

7 Preim et al., *Visual Analytics of Image-Centric Cohort Studies in Epidemiology*, vol. 3, ch. Visualization in Medicine in Life Sciences, in print. Springer, 2016. **115**

8 Klemm et al., "Visualization and Analysis of Lumbar Spine Canal Variability in Cohort Study Data", *Proc. of Vision, Modeling, and Visualization (VMV)*, pp. 121-128, 2013. **143**

9 Klemm et al.,"Interactive Visual Analysis of Image-Centric Cohort Study Data", *IEEE Trans. Vis. Comput. Graphics (IEEE TVCG)*, vol. 20, no. 12, pp. 1673-1682, 2014. **151**

10 Angelelli et al., "Interactive Visual Analysis of Heterogeneous Cohort Study Data", *IEEE Comput. Graph. Appl. Mag. (IEEE CG&A)*, vol. 34, no. 5, pp. 70-82, 2014. **161**

11 Oeltze et al., "Measurement of the Stratum Radiatum/Lacunosum-Moleculare (SRLM)", *Proc. of Bildverarbeitung für die Medizin*, pp. 264-269, 2014. **171**

Part IV: Perfusion Data **177**

12 Glaßer et al., "Visual Analysis of Longitudinal Brain Tumor Perfusion", *Proc. of the SPIE Medical Imaging*, pp. 86700Z, 2013. **179**

13 Chalopin et al., "Method for the Evaluation of US Perfusion for Brain Tumor Surgery", *Proc. of Jahrestagung der Deutschen Gesellschaft für Computer- und Roboter Assistierte Chirurgie (CURAC)*, pp. 198-202, 2013. 191

Part V: Surveys 197

14 Ropinski et al., "Survey of Glyph-Based Visualization Techniques for Spatial Multivariate Medical Data", *Computers and Graphics (C&G)*, vol. 35, no. 2, pp. 392-401, 2011. 199

15 Oeltze-Jafra and Preim, "Survey of Labeling Techniques in Medical Visualizations", *Eurographics Workshop on Visual Computing for Biology and Medicine (EG VCBM)*, pp. 199-208, 2014. 209

Bibliography 219

Chapter 1

Visual Analytics of Medical and Biological Data

This chapter serves both as an introduction to this postdoctoral thesis and also, a survey of the subsequent cumulative Parts I-IV. At the beginning of the chapter in Section 1.1, the science of *visual analytics* (VA) is motivated in the context of medical and biological data. In Section 1.2, an application-independent introduction to VA is given including a brief description of the scientific disciplines contributing to it and a characterization of the highly-related field of *interactive visual analysis* (IVA). The latter is included since several papers in the cumulative parts are attributable to this particular field. Sections 1.3-1.5 are dedicated to the main types of medical and biological data that have been analyzed in the majority of the cumulated publications: simulated cerebrovascular hemodynamic data (Sec. 1.3), fluorescence microscopy data of cells and tissues (Sec. 1.4), and epidemiological population study data (Sec. 1.5). These sections describe the respective application background, the data characteristics and their acquisition or generation, the standard data processing workflow, the potential of VA in analyzing the data, and existing VA approaches. At the end of this chapter, VA of further types of medical and biological data from other applications is briefly reviewed including a short discussion of the respective seminal and very recent work (Sec. 1.6). Each individual paper of the cumulative Parts I-IV is outlined and related to the respective state-of-the-art within its corresponding Section 1.3-1.5, 1.6.1. The summary and discussion in Section 1.7 recapitulate and comment on the main contributions of the thesis at hand. A visual summary lists the contributions per paper as well as the subset of scientific disciplines co-operating in VA that these contributions can be attributed to (Tab. 1.1). The chapter is concluded in Section 1.8 by a vision on future research directions in visual analytics of medical and biological data.

1.1 Motivation

A data explosion can be observed in many areas of science, industry, business, and public life. At CERN's Large Hadron Collider, 45 terabytes of sensor and machine data are generated per second during experiments and a four-engine Boeing jumbo jet collects 640 terabytes of data from its engines during a transatlantic flight [190]. Monitoring the trading activities at the New York Stock Exchange (NYSE) produces 4-5 terabytes of data per day [68]. A rich source of information for several companies, e.g., in the advertisement industry, is the social network data created and shared by currently 320 million active users per month on Twitter [243] and 1.55 billion monthly active users on Facebook [49].

In medicine and biology, researchers and practitioners are confronted with a steadily growing data complexity paralleling new developments in image and non-image data acquisition. Hospitals need to maintain thousands of electronic medical records protocoling for instance diagnostic procedures, drug prescriptions, and therapeutic measures [174]. Furthermore, medical image data are acquired per patient at several occasions possibly employing multiple modalities, e.g., Ultrasound (US), Computed Tomography (CT), and Magnetic Resonance Imaging (MRI). Body parts can be imaged at a very high resolution and also over time covering dynamic processes. Recent Dual energy CT scanners create full body scans of $\approx$ 20 gigabytes in 2-3 seconds and time-resolved live scans, e.g., of organs in motion, with $\approx$ 5 gigabytes per second [224]. In predictive medicine, simulations of the outcome of pathologic processes or therapeutic interventions are performed besides

imaging and create additional complex data, e.g., simulations of blood flow for planning endovascular interventions [96] and kinetics of implants for surgery planning [204]. In biology, genome sequencers generate tens to hundreds of gigabytes of data to decipher the genome [264]. High-throughput electron microscopy can produce terabytes of streaming data [70]. Multi-variate data with one slice or volume per antibody is acquired in robot-driven fluorescence microscopy [52]. Population studies in epidemiology collect data for thousands of subjects from questionnaires, interviews, laboratory tests, Electrocardiography (ECG) monitoring, DNA sequencing, simple-medical/dental/skin examinations, cardiopulmonary exercise testing, sleep monitoring, and more recently, from medical imaging using non-invasive modalities such as US and whole-body MRI [256]. Often, these studies are carried out over years in several waves generating longitudinal data.

For such complex data, it is impossible to design overview visualizations that convey all the contained interesting patterns. Filtering and analyzing the data are required since Sheiderman's guide to visually exploring data "Overview first, zoom/filter, details on demand" [222] is not applicable here and raw data has very limited value. While fully automatic data analysis techniques perform reliably for well-defined and well-understood problems, they must be combined with the analyst's knowledge and abilities in solving more fuzzy, complex, and opaque issues. This combination is at the heart of the field visual analytics, which has been defined accordingly by Keim et al. [105]:

> "Visual analytics combines automated analysis with interactive visualisations for an effective understanding, reasoning and decision making on the basis of very large and complex datasets."

During the iterative *visual analytics process*, the analyst derives knowledge from the raw data by modifying parameters of the data analysis techniques and steering the analysis process based on the evaluation of intermediate results presented as interactive visualizations. This is reflected by Keim's extension of Sheiderman's guide in the context of visual analytics "Analyse first, show the important, zoom/filter, analyse further, details on demand" [104].

1.2 Introduction to Visual Analytics

At the beginning of this section, the *visual analytics process* from data to knowledge is detailed (Sec. 1.2.1). Then, the building blocks of VA research, i.e. the different scientific disciplines contributing to it, are described (Sec. 1.2.2). Finally, the VA-related field of *interactive visual analysis* spanning a subset of these disciplines is introduced (Sec. 1.2.3).

Sections 1.2.1-1.2.2 are based on the book chapter "Visual Analytics" by Keim et al. [105]. The corresponding VisMaster book provides an excellent comprehensive entry point to VA research. Early visions and principles are described in the pioneering research and development agenda for VA edited by Thomas and Cook [237]. Expanding frontiers and research trends are outlined by leading scientists in a collection of articles edited by Dill et al. [37]. A survey of general visual analytics techniques and applications, which does however attach little importance to the VA of scientific image data, is given by Sun et al. [233]. The proceedings of the annual IEEE conference on Visual Analytics Science and Technology (VAST) are recommend for an overview of very recent VA research. They are accessible from the IEEE digital library [90].

1.2.1 The Visual Analytics Process

The visual analytics process tightly couples automatic data analysis techniques and visualization methods through user interaction for deriving knowledge from data. A schematic overview of the

process is given in Figure 1.1. It is based on the original scheme by Keim et al. [105] and has been modified according to the scheme by Meyer et. al [146]. The latter more explicitly conveys the importance of user interaction at all stages of the process. *Data mining* as the science of data modeling and knowledge discovery is considered fundamental to automated data analysis and hence, both terms are often used interchangeably [105].

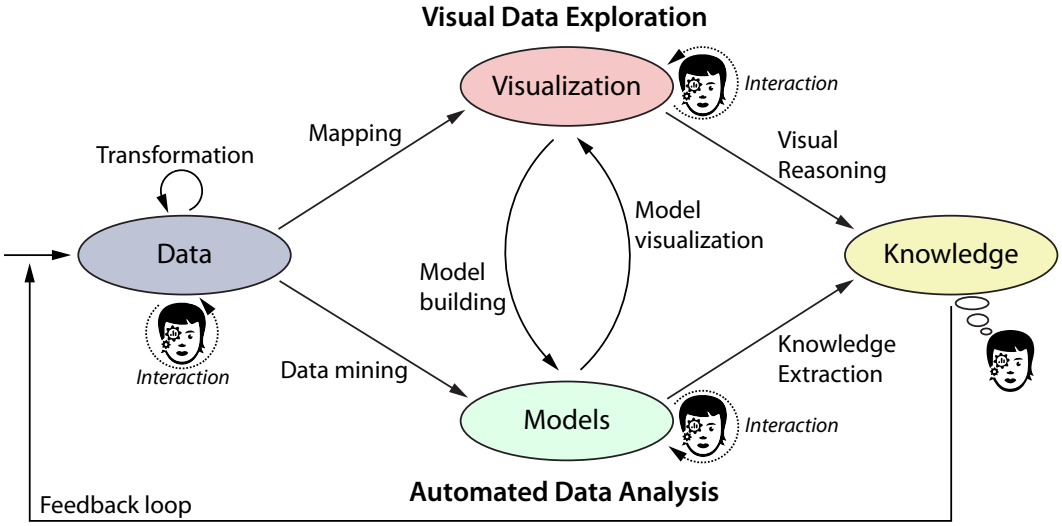


Figure 1.1: The visual analytics process from raw data to knowledge via a tight coupling of automatic data analysis and visualization techniques through user interaction. Adapted from [105; 146].

At first, the raw data may be transformed including the conversion into an organized logical structure, the selection of data subsets, data cleansing, normalization, interpolation, and the merging of multiple heterogeneous sources. The analyst can then start either with (1) an automated data analysis or with (2) visually exploring the data. A population study is used as the working example for explaining both ways in the following. The study collected socio-demographic, lifestyle-related, and health-related attributes for thousands of subjects. (1) The analyst starts by applying a clustering method to three selected attributes including the per cluster computation of an "average" subject and of the subjects' variance around the average. The clusters and the derived information constitute a model of the data from which knowledge can be extracted. In order to evaluate this model, it is visualized for instance by opposing the attributes in a 3D scatter plot where each dot represents a subject and the dot's color indicates the cluster index. Additionally, the average subjects are highlighted. While rotating the plot, the analyst may discover a corrupt cluster separation or average subject computation due to outliers. As a consequence, the parameters of the clustering method may be refined or a method which is less sensitive to outliers is applied. (2) Alternatively, the analyst may start by visually exploring subsets of the data, e.g., by means of a scatter plot matrix opposing multiple attributes in a pair-wise manner. Based on the observation of subject clusters in individual plots, the analyst may reason about the existence of distinct subpopulations. These hypotheses must be verified by building a model of the data using a dedicated clustering method. This model is then visually evaluated and the findings are employed for its iterative refinement by adjusting the parameters of the clustering method.

Knowledge is gained during the visual analytics process, e.g., the existence of subpopulations (clusters) with distinct values for selected attributes. This knowledge may be fed back into the process yielding, e.g., a selection of different input attributes or a subset of subjects. The entire visual analytics process is characterized by alternating between visual and data analysis methods. It facilitates both, the verification of an priori hypothesis as well as hypothesis generation.

1.2.2 Building Blocks of Visual Analytics Research

Visual analytics combines technologies from many different scientific disciplines. An overview of these disciplines is shown in Figure 1.2. It represents a slightly adapted version of the overview by Keim et al. [105]. "Human-Computer Interaction" (HCI) has been added to the set of disciplines to stress the importance of interaction and to acknowledge the increasing interest in exploiting and promoting HCI research work in the realm of visual analytics [6; 266]. The design of interaction techniques and user interfaces are key elements of HCI research finding broad applications in visual analytics frameworks.

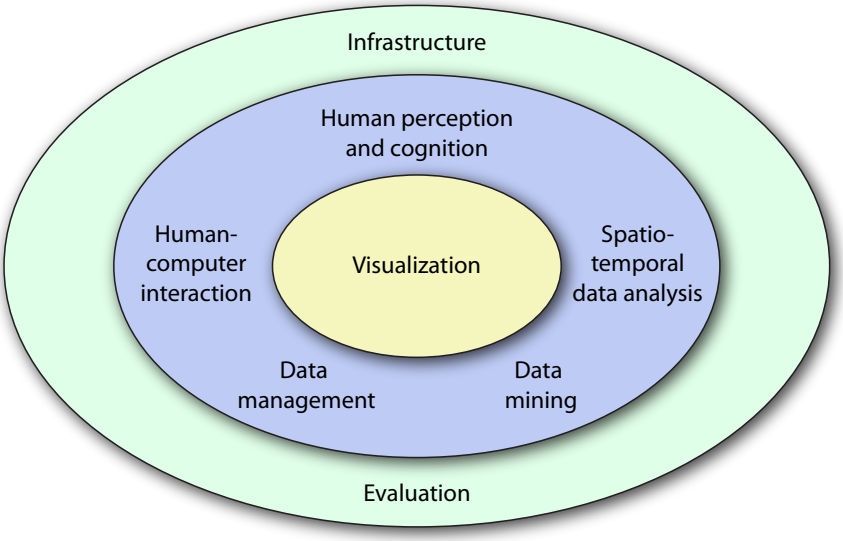


Figure 1.2: Building blocks of visual analytics research. Visualization is at the heart being used by all neighboring disciplines. Visual analytics solutions require an appropriate software infrastructure and their effectiveness, efficiency, and user acceptance must be evaluated. Adapted from [105].

Visualization is at the heart of visual analytics. It serves not only the purpose of presenting final results but also of monitoring and evaluating intermediate steps, for instance, in data generation, management, and mining. Visualization is often classified into scientific (SciVis) and information visualization (InfoVis). While SciVis develops techniques for visualizing data with an inherent spatialization, e.g., measured or simulated data from climate research, medical imaging, and engineering, InfoVis focuses on visualizing data for which a spatialization must be chosen, e.g., business data, social networks, and software structure. Visual analytics solutions employ the best of both worlds and use for instance surface and volume rendering or flow visualization techniques from SciVis [72] and standard plots, parallel coordinates, treemaps or graphs from InfoVis [227].

The user interface and techniques for user interaction with both, the data analysis and the visualizations, are critical components of each visual analytics system. Their development is part of **human-computer interaction** (HCI) research [95]. In a visual analytics framework, often different views on the data are organized as a *coordinated multiple views* system. Coordination is achieved via *linking & brushing* [44], i.e., data items are selected in one view (brushing) and at the same time highlighted in all other views (linking). Coordinated multiple view systems have been surveyed by Roberts [189] with one focus on exploration techniques and another one on user interfaces including window and display management strategies. Guidelines for the design of multiple views on data are given by Baldonado et al. [260]. In particular for data with a

spatial reference, linking & brushing is often coupled with $focus+context$ visualization [75]. The portion of the data within a flexible focus region is emphasized while attenuating but maintaining the surrounding context, e.g., for spatial orientation. Interaction techniques in visual analytics systems may be classified following the categories of interaction with information visualizations suggested by Yi et al. [271]. For instance, interesting data items may be marked by brushing (*select*) and highlighted in multiple views (*connect*), other data may be shown using panning and zooming (*explore*), more or less detail may be displayed (*abstract/elaborate*), and visualization may be restricted to data matching certain conditions (*filter*).

The investigated data in visual analytics is often very heterogeneous, i.e. it stems from multiple sources with different data types (numerical, text, audio, video etc.), formats (RAW, DICOM, HDF, etc.), variable types (nominal, ordinal, quantitative), and dimensions ($1D, 2D, 3D, 3D+time, nD$). The focus of **data management** here is the generation of an integrated and consistent database facilitating queries for automatic analysis and visual exploration. A crucial step is data cleansing dealing with incorrect and missing data. In population studies for instance, participants may refuse to answer specific questions in an interview or they may favor a more socially-accepted answer, e.g., with respect to alcohol consumption, which is however in contrast to further available information, e.g., given by liver function reading.

In **data mining**, methods for an automatic extraction of information from raw data are developed [71]. A multitude of conceptually different approaches exists. *Classification* methods learn a model for classifying unseen data from labeled training samples. Examples are decision trees, Bayesian Networks, Support Vector Machines, and lazy learners. If no a priori knowledge is available, *cluster analysis* methods can be applied grouping data items based on mutual similarity such that intra-group cohesion and inter-group separation are maximized. Examples are partition, hierarchical, density-based, and spectral methods with their representative algorithms k-Means, Agglomerative Hierarchical Clustering, DBSCAN [47], and Normalized Cuts [220]. Some of the existing clustering algorithms are tailored to the identification of outliers, e.g., DBSCAN. Further data mining approaches besides classification and clustering are pattern mining, dimension reduction, and correlation analysis.

Investigating data with references in space and possibly time requires specific **spatio-temporal data analysis** techniques. They aim at the detection of spatial patterns and relations and of temporal patterns, trends, and correlations. For instance, suspicious lesions need to be identified in digitally contrast-enhanced (DCE) MRI mammography data and their tissue heterogeneity must be evaluated based on the spatio-temporal pattern of contrast accumulation [183]. If data of the same phenomenon have been acquired using different devices or parameterizations, the datasets need to be co-registered to a common reference. As an example, pre-operatively acquired MRI data of a brain tumor must be aligned with intra-operative US data for evaluating tumor resection surgery [28]. A crucial aspect of spatio-temporal data analysis is uncertainty originating from the data generation and transformation (Fig. 1.1). Missing or interpolated data, multiple parameterizations of the data generation device or algorithm, imperfectly aligned portions of data showing the same phenomenon, and manual segmentations of a structure in image data are examples for uncertainty that must be considered during analysis and conveyed by visualizations. The user's awareness of the involved uncertainties directly influences the confidence or trust in the analysis results [195].

During the visual analytics process, the user needs to perceive and understand what is displayed on a screen. The sciences of **perception and cognition** research the heavily complex sequence from perceiving to making inferences. *Cognitive psychology* studies "how people perceive, learn, remember, and think about information" [229]. Understanding *visual perception* is essential in this context [225]. It represents the human ability to derive information from visible light reaching

the eye, e.g., emitted by a computer screen. What is known about perception and cognition is exploited in designing visualizations, user interfaces as well as interaction techniques and is hence, also strongly related to HCI [261].

Visual analytics requires an efficient **infrastructure** for combining technology from many different disciplines in a way such that high interactivity is guaranteed during the entire analysis process. Major issues in providing such an infrastructure are incompatibilities between implementations, the missing support for handling large amounts of data, e.g., stored in databases, and accurate but time-consuming computations hampering interactivity. Hence, most visual analytics solutions are custom-built stand-alone applications using in-memory data storage and achieving sufficient computational performance through optimization for a specific type of data and a limited set of tasks. Detailed requirements of visual analytics on software and hardware infrastructure have been discussed by Fekete [50] and classified into requirements on data management, automated data analysis, visualization including hardware, and workflow support for analysts. Furthermore, open obstacles in meeting these requirements have been derived and hints on how to address them in the future were provided.

The **evaluation** of methods and solutions in visual analytics is crucial to assess their effectiveness, efficiency, and usability. A very active research community in this area strives for standardized approaches and solid generalizable results. Both are difficult to achieve due to the broad scope of visual analytics and the complexity of its solutions. An overview of evaluation challenges in visual analytics, current approaches, and recommendations are given by van Wijk [252]. For instance, the user community addressed by visual analytics is very diverse and evaluation results may not carry over, e.g., from laymen to experts. The tasks that need to be solved are often complex and comprise several steps carried out in multiple iterations. For compound visual analytics solutions, it is often difficult to attribute an analysis outcome to one of the components. Furthermore, visual analytics aims at providing insight which is however, ill-defined and therefore hard to measure.

1.2.3 Interactive Visual Analysis

Interactive visual analysis (IVA) is a branch of data visualization whose methodology highly intersects with the one of visual analytics. IVA builds upon visualization, interaction, and human pattern recognition for investigating complex multidimensional and multivariate data. Its focus is on data with an inherent spatialization, i.e. measured and simulated data, and automated data analysis is considered rather optional. IVA is hence tightly related to the visual data exploration part of the visual analytics process (Fig. 1.1). IVA solutions are realized as coordinated multiple view systems comprising two types of views [262]. *Spatio-temporal* or *physical views* show information in the context of the data's spatio-temporal observation space, e.g., by means of volume or surface renderings. *Range* or *attribute views* show relationships between multiple data attributes, e.g., by means of scatter plots or parallel coordinates. Both types of views employ $focus + context$ visualizations and offer brushing facilities for *feature specification*. An interesting pattern spotted by the analyst may hint at a feature, i.e. a distinct characteristic of the data, or alternatively, specific value ranges of selected attributes are suspected to characterize a feature. In both cases, the interesting data portions are brushed (focus) and visualized within their surroundings (context).

One can imagine the information that is hidden in data as a deep ocean [76]. Some information is floating directly beneath the surface and is already accessible through base-level IVA. Other information is buried deeper and their extraction requires more advanced techniques at a higher level. This cascade of information retrieval is reflected by the four levels of IVA [262]:

1. **Show & Brush (level 1)**: This level utilizes at least one physical and one linked attribute view. A potential feature is brushed in one view yielding a $focus + context$ visualization in all linked view(s).

2. **Relational analysis (level 2)**: At this level, multiple brushes from different views are combined using logical operators. This facilitates a more sophisticated feature specification. For instance, places at medium altitude exhibiting low temperature and high precipitation may be extracted from a weather dataset.

3. **Complex analysis (level 3)**: This level is tightly related to visual analytics since it integrates computational data analysis methods for deriving *synthetic attributes* from existing ones. For instance, the temporal variation of an attribute can be computed using derivatives or data items can be clustered and cluster membership can serve as a new attribute. The third level further integrates advanced brushing facilities such as *angular brushing* [77] and *similarity brushing* [154].

4. **Proprietary analysis (level 4)**: At this level, application-specific feature definitions, data analysis methods or visualization techniques are integrated. Examples include local feature detectors from flow field analysis [18] and tailor-made graph views for investigating fluorescence microscopy data [166].

At each IVA level, features may be specified interactively in physical or attribute space for further investigation. The most important IVA feature specification patterns are [165; 262]:

Feature localization refers to the process of brushing a subset of data items in an attribute view and emphasizing the corresponding spatial locations in a physical view. The emphasis reveals whether the selection represents a localized feature such as a specific brain part exhibiting suspicious perfusion characteristics.

Local investigation refers to the selection of interesting spatial locations in a physical view and inspecting the corresponding data items in attribute views. This restricts the analysis to a region of interest such as a suspicious lesion in breast tissue.

Multivariate Analysis is characterized by brushing a subset of data items in one attribute view and observing the selection in views showing different attributes. For instance, tissue with a fast wash-in and wash-out of blood may be selected and its overall blood supply over time may be investigated.

One of the first IVA systems named `WEAVE` has been presented by Gresh et al. [67] for the joint investigation of measured and simulated cardiac data. It comprises a 3D view showing the heart's anatomy as well as histogram, scatter plot, and parallel coordinates [94] views for displaying data attributes. The user may brush an interesting part of the heart's surface in the 3D view (local investigation) or interesting data values in an attribute view (feature localization) and assign a color to the selection. This color is used then, for emphasizing the corresponding data items in all other views. Inspired by WEAVE, Doleisch and Hauser et al. [39; 40] developed the `SimVis` framework for interactive feature specification in computational fluid dynamics data. `SimVis` started as a research prototype and evolved into a mature commercial system highly optimized for IVA of very large data. It comprises a 3D view, histogram, scatter plot, parallel coordinates, and a curve view for exploring time-dependent data attributes [154] (Fig. 1.3). `SimVis` has been applied to and extended for data from a wide range of application areas, e.g., simulation data from

engineering [123] and climate research [103], medical MRI perfusion data [165], and within this thesis, medical US perfusion [28] (Chap. 13) and biological fluorescence microscopy data [166] (Chap. 4). The very flexible IVA system `ComVis` has been developed by Matković et al. [138] for the rapid prototyping and testing of new IVA paradigms and technology. `ComVis` comprises the same types of views as `SimVis` and can easily be extended, e.g., by integrating more advanced views tailored to the IVA of set-typed data [51], families of function graphs [117], or families of surfaces [139].

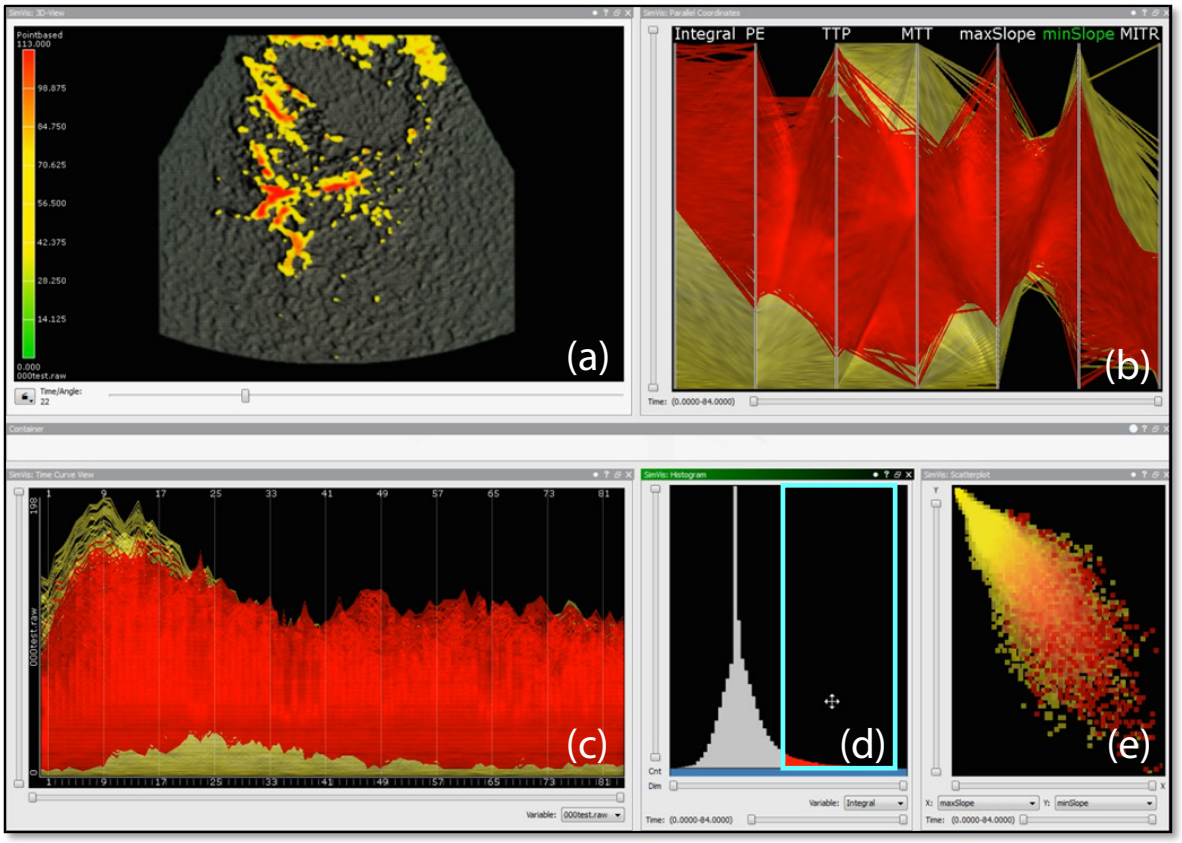


Figure 1.3: Interactive Visual Analysis of Ultrasound perfusion data in the `SimVis` framework [28]. A physical view *(a)* and four linked attribute views *(b-e)* show different aspects of the data. Parallel coordinates provide an overview of all perfusion parameters *(b)*, a curve view conveys changes in tissue perfusion over time *(c)*, a histogram depicts the distribution of a selected perfusion parameter *(d)*, and a scatter plot shows the correlation between two other parameters *(e)*. In a feature localization process, an interesting part of the histogram in *(d)* is brushed (turquoise rectangle). The selection is highlighted in red within all attribute views and colored according to a perfusion parameter in the physical view. The latter shows the selection (focus) within the surrounding tissue (gray context).

1.3 Simulated Cerebrovascular Hemodynamic Data

This section provides context for the cumulated publications of Part I. It introduces cerebral aneurysms including brief descriptions of their pathogenesis, diagnosis, and treatment, as well as related medical research questions (Sec. 1.3.1). Following this, the role of aneurysm hemodynamics in answering these questions, the generation of hemodynamic data, and parameters derived from these data are detailed (Sec. 1.3.2-1.3.4). At the end of this section, the potential of visual

analytics in investigating the hemodynamic parameters as well as existing approaches, including the papers cumulated in this postdoctoral thesis [168; 171] (Chap. 2-3), are surveyed (Sec. 1.3.5). To conclude, the relation to measured cardiac hemodynamics is outlined (Sec. 1.3.6).

The Sections 1.3.2-1.3.4 are based on very comprehensive introductions to the visual exploration of simulated cerebrovascular and more general, of measured and simulated cardiovascular hemodynamics by Neugebauer [157] and Gasteiger [55], respectively. More condensed introductions are provided by Preim et al. [180] and Vilanova et al. [254].

1.3.1 Cerebral Aneurysms

Cerebral aneurysms (also called *intracranial* or *brain aneurysms*) represent a cerebrovascular disorder which is characterized by a localized dilation of the weakened arterial wall. Their pathogenesis is incompletely understood but seems to be related to many factors such as the interplay of hemodynamic stresses and degenerative changes of the arterial wall, inflammatory effects, genetic predispositions, and exogenous factors, e.g., drinking and smoking [121]. The most common subtype *saccular aneurysms* and its morphological features are illustrated in Figure 1(b) on page 52 of Chapter 2. Typical is the balloon-like dilation (the *aneurysm sac*) with a narrow opening (the *aneurysm neck*) [53]. Saccular aneurysms most frequently occur at the base of the brain, either directly at or in the close vicinity of the vascular ring named *Circle of Willis* [54]. While some aneurysms are found along straight vascular segments (*side-wall aneurysm*), most develop at arterial branchings (*bifurcation aneurysm*).

The primary complication of cerebral aneurysms is a progressive weakening of the arterial wall culminating in *aneurysm rupture*. The prevalence of unruptured aneurysms in the general population has been estimated as $\approx 3.2\%$ [255]. Most of them will never rupture and remain clinically silent [248]. However, rupture causes *subarachnoid hemorrhage* (SAH), i.e. bleeding into the *subarachnoid space* filled with *cerebrospinal fluid*, which represents a significant cause of morbidity and mortality world-wide [30].

The diagnoses of most unruptured cerebral aneurysms are incidental findings. In case of SAH, the rupture of a cerebral aneurysm represents a very likely cause and according evidence is explicitly investigated. If a patient enters the hospital with symptoms characteristic for SAH, CT imaging is commonly used to exclude other causes [14]. If SAH has been confirmed, angiography images are acquired to verify the hypothesis of a ruptured aneurysm. Often, CT angiography (CTA) is applied since it can be conducted immediately after the CT imaging [244].

The primary goal of a therapeutic intervention is to stop blood flow into the aneurysm such that rebleeding in ruptured aneurysms and the rupture of clinically silent ones are prevented. *Surgical clipping* has been the standard therapeutic approach for decades [247]. After opening the skull of the patient and cutting through the *dura mater*, the aneurysm is exposed and a self-closing metal clip is attached to its neck. In unruptured or poorly accessible aneurysms, older patients or those who cannot undergo surgery, minimally invasive vascular interventions provide an alternative therapeutic option [184]. They reduce the burden during intervention as well as recovery time and infection risk for the patient. While surgical clipping requires a neurosurgeon, minimally invasive treatment may be performed by an interventional neuroradiologist. In *coiling*, compressed platinum spirals (*detachable coils*) are inserted into the aneurysm via a microcatheter that is guided through a small opening in the *femoral artery*. The coils unfold within the aneurysm sac causing a stagnation of the blood flow ultimately inducing a *thrombosis*, i.e. the formation of a blood clot. The deliberate clogging of the aneurysm is referred to as *embolization*. In patients with complex aneurysms that cannot be filled with coils, deploying a self-expanding, high-profile, flow diverting mesh tube (*Flow-diverter stent*) to the parent arteries in the close vicinity of the aneurysm

provides a promising alternative to coiling (Fig. 1(a), p. 52, Chap. 2). The stent is also guided by a microcatheter and expands with the help of an inner, inflatable balloon until it fits closely with the vessel wall. The blood flow into the aneurysm is strongly diminished and decelerated triggering embolization without inserting artificial material into the aneurysm sac.

Neugebauer collected medical and clinical research questions in an extensive literature review and classified the results into four categories: questions related to the pathogenesis, growth, rupture, and therapy of cerebral aneurysms [157]. The publications in the cumulative Part I of this thesis are dedicated to answering questions in the context of aneurysm rupture [171] (Chap. 3) and stent therapy [168] (Chap. 2). Since aneurysms do not inevitably rupture and surgical as well as endovascular interventions may induce other complications for the patient, e.g., injury of the vessel wall during stent insertion, a bulk of research aims at a reliable assessment of the *risk of rupture*. In considering a stent therapy, it must be clarified whether the individual vascular and aneurysmal morphology is eligible for stenting, which type of stent yields the best outcome, and where for instance in a bifurcation, the stent shall be deployed. The following section illuminates the relation of aneurysm hemodynamics to the risk of rupture and to stent therapy.

1.3.2 Role of Aneurysm Hemodynamics in Diagnosis and Treatment

It is an obvious conjecture that the interplay of a pulsating stream of blood and a potentially destabilized wall is among the factors that effect the risk of aneurysm rupture. The hemodynamics in cerebral arteries cannot directly be measured at a sufficiently high resolution due to the small vessel diameter (1-3 mm at the Circle of Willis) and the limited resolution of today's scanning devices. Hence, *Computational Fluid Dynamics (CFD)* simulations of blood flow are conducted based on patient-specific vascular and aneurysm geometry extracted from medical images [25]. Most studies draw conclusions regarding the role of hemodynamics from comparing simulation results of ruptured and unruptured aneurysms and searching for significant differences. Their underlying image data and aneurysm classification are retrieved from hospital databases and originate in diagnosis and therapy.

Both, abnormally high [27] and low [269] shear stresses of the aneurysm wall have been related to an increased risk of aneurysm rupture. This alleged contradiction is subject of an ongoing debate between researchers [143]. Both results may be valid and depend on different mechanisms of aneurysm growth and rupture. A higher consensus exists regarding the relation of qualitative hemodynamic characteristics to aneurysm rupture. Complex flow patterns, which are unstable over the cardiac cycle [20], as well as a focused influx into the aneurysm with a small region of impact [23] have been related to an increased risk of rupture.

The simulation of aneurysm hemodynamics can contribute to the planning of stent therapy by predicting the hemodynamical outcome of different stent types and locations. In *virtual stenting*, a geometric model of the stent is deployed to the vascular geometry and integrated in the simulation [96; 97]. The simulation results of different stent types and locations are then compared to find the optimal patient-specific configuration, e.g., one that minimizes shear stresses and yields simple and stable flow patterns. In the annual Virtual Intracranial Stenting Challenge (VISC), researchers compete in predicting stenting success and aneurysm risk of rupture based on simulated hemodynamics [86]. They are provided with real clinical cases and stent models and are expected to return well justified treatment decisions and risk estimates, partially, within a time frame that would be acceptable in real-world clinical treatment planning. Assessing a patient's general eligibility for a Flow-diverter stent may also benefit from simulations of aneurysm hemodynamics. For instance, a local increase of pressure yielding an increased risk of rupture was observed in patient-individual simulations of aneurysms that had ruptured after treatment with such a stent [24].

CFD simulations of blood blow provide new insights into the hemodynamics of cerebral aneurysms. They are not (yet) part of the clinical routine. In medical research, they are conducted to better understand the causes of aneurysm initiation, progression, and rupture, to eventually define a rupture risk score for clinical routine use, and to predict the outcome of endovascular interventions in the context of both, virtual stenting but also *virtual coiling* [66; 151]. The following section describes the hemodynamic data generation pipeline from patient-specific image data to CFD simulations of blood flow.

1.3.3 Hemodynamic Data Generation

Several pioneering papers have demonstrated the value of image-based CFD models of hemodynamics for a single aneurysm case [74; 100; 228]. Efficient workflow pipelines were later presented for the reliable, reproducible, and robust processing of data from large-scale studies [5; 25; 55]. While the implementations of individual pipelines differ, they all comprise the following five steps:

1. **Medical imaging** of the patient's cerebral vascular tree.
2. **Segmentation** of the aneurysm and the vasculature in its vicinity.
3. **Surface reconstruction** from the segmentation result, **surface enhancement**, and extraction of anatomical landmarks and geometric features.
4. **Volume grid generation** based on the surface mesh.
5. **CFD simulation** of blood flow on the volume grid considering boundary conditions and an approximate fluid model of blood.

Gasteiger gives a very comprehensive overview of methods that have been proposed in literature for implementing each step [55]. The individual implementations employed throughout Part I of this thesis are described in Section 2.3 on page 52 of Chapter 2 and in Section 2.4 on page 68 of Chapter 3. Based on the overview by Gasteiger [55], important general aspects of each step are summarized in the following. Note that the pipeline steps are usually performed for a priori known aneurysms. In a screening scenario, an aneurysm detection step could be integrated after the medical imaging [81].

Medical imaging is typically conducted for aneurysm diagnosis or treatment in clinical routine. Depicting the patient-specific vascular and aneurysm morphology requires dedicated imaging techniques yielding an enhancement of the vessel lumen. Either a contrast agent is injected, as for instance in 3D-Rotational Angiography (3D-RA), Computed Tomography Angiography (CTA), and Contrast-Enhanced Magnetic Resonance Angiography (CE-MRA), or magnetization properties of the flowing blood are exploited, as in Time-of-Flight Magnetic Resonance Angiography (TOF-MRA). A well-defined boundary of the vessel lumen is required for a reliable subsequent segmentation. However, it is often hampered by artifacts such as a blending of the aneurysm and near-by vessels due to the partial volume effect or an inhomogeneous signal intensity distribution in the lumen due to inconsistent contrast bolus accumulation. Solution strategies often require a tedious manual post-processing of the segmentation or surface reconstruction result.

A myriad of **segmentation** techniques has been proposed for extracting the vessel lumen from 3D image data [126]. *Intensity-based approaches* as well as *deformable models* have been employed in the context of this thesis. Representatives of the former such as *thresholding* and *region growing* are fast and easy to implement but sensitive to the above-mentioned artifacts. They result in a binary voxel mask requiring a subsequent surface reconstruction step. Representatives of the

latter such as *active contours* or *level set segmentations* deform an initial 2D contour or 3D surface with respect to external forces derived from the image data, e.g., gradient information, and internal forces controlling the degree of deformation, e.g., curvature. Deformable models are less sensitive to blending and inhomogeneous signal intensity distributions within the lumen. Furthermore, they obviate the surface reconstruction step and generate a segmentation at sub-voxel level. However, they are computationally more expensive and their parameterization is challenging.

Surface reconstruction of the boundary of the vessel lumen is required if the segmentation resulted in a binary mask. A straightforward approach is to apply the *Marching Cubes (MC)* algorithm to the mask [130]. However, the resulting surface mesh comprises a vast number of triangles among which many exhibit a bad quality, i.e. strongly deviate from being equilateral. Furthermore, the surface suffers from *stair-case artifacts* due to the discrete nature of the image data and the binary nature of the segmentation mask. Alternative approaches generate smoother results by means of implicit surfaces [120; 217; 268]. Furthermore, they optimize triangle quality either in a final remeshing step [216] or by controlling the quality during surface reconstruction [268]. The number of triangles can be optimized by adapting the triangle density to local surface curvature [268].

Different **surface enhancement** techniques may be required depending on the employed segmentation and surface reconstruction approach. Mesh smoothing is necessary for instance to remove stair-case artifacts after an intensity-based segmentation and a MC-based surface reconstruction. Dedicated smoothing approaches preserve the volume of the mesh and instead of treating the entire surface equally, identify and quantify stair-case artifacts to locally adapt the smoothing strength [148]. The correction of surfaces which are erroneously connected due to blending artifacts requires tedious manual adjustments in a mesh processing software. To separate the aneurysm and a blended vessel, the mesh is cut open, holes are filled, and the result is smoothed [149]. A high quality of the surface mesh is necessary to achieve a high quality of the subsequently generated volume grid which in turn is required by the CFD simulation to ensure convergence and minimization of numerical errors [55]. Alliez et al. give an overview of local and global remeshing approaches aiming at close to equilateral triangles, uniform vertex density, and curvature-adaptive triangle sizes [3]. A pipeline for the optimization of vascular surface models in the context of CFD simulations has been proposed by Mönch et al. [149]. It comprises mesh smoothing, the removal of blending, remeshing, and several other steps such as an optimization of the in- and outlets for the simulation.

Anatomical landmarks and geometric features are extracted by decomposing the optimized surface model based on its centerlines [5] and the ostium surface [158] into aneurysm sac, near-, and far-vessel domain [160] (Fig. 1(b), p. 52, Chap. 2). The decomposition results are used for instance to restrict the visual exploration and analysis of blood flow to the aneurysm sac and the near-vessel domain and to derive clinically relevant parameters describing aneurysm size, shape, and spatial relations to the parent vasculature. The ostium surface is often utilized as seeding geometry for integral curves thereby focusing the investigation on flow entering the aneurysm.

The **volume grid generation** for the subsequent CFD simulation is based on the optimized surface mesh. The *advancing layers method* represents the most common approach which fills the vessel lumen with unstructured tetrahedral elements starting from the surface and proceeding into the lumen [199]. In order to capture also the small changes in flow speed at the vessel boundary, a more fine-granular boundary layer of prism elements is recommended [198]. In virtual stenting, a surface mesh of the stent is tightly fitted to the vessel surface and integrated into the volume grid generation [97]. The resulting volume grids typically comprise 100,000 to 3,000,000 elements without and 10,000,000 elements with stent [97]. The resolution of the grid is in the range of 0.02-0.08 mm volume diagonal of a cell [55].

CFD simulations commonly model blood as *incompressible Newtonian fluid*, i.e. a fluid with constant density and viscosity. While this is not true in theory, compression has low impact in practice given the relatively small speed of blood flow and also, non-Newtonian effects seem to be negligible, at least in larger arteries [218]. The corresponding model of blood flow can be described by three-dimensional unsteady incompressible *Navier-Stokes equations*. Numerical solutions of these equations under certain boundary conditions are approximated by means of finite element formulations and a solver. Pre-defined boundary conditions comprise pressure profiles at inlets and outlets, pulsatile flow speed, and no-slip boundary, i.e. zero velocity at the vessel wall. Unsteady simulations run hours or even days and cover the entire cardiac cycle with a temporal resolution in the range of 0.0025-0.005 sec yielding 200-400 time steps [55]. They result in a 4-dimensional (4D) vector field from which further hemodynamic parameters are derived.

1.3.4 Hemodynamic Parameters

Hemodynamic parameters can be classified into **quantitative** and **qualitative**. While the former are computed per grid point, element or region of the volume grid, the latter are derived from a visual inspection of the *blood flow pattern*. The publications in Part I of this thesis are dedicated to the visual analytics of this pattern. The computation of quantitative parameters and their role in characterizing the hemodynamic environment in ruptured and unruptured cerebral aneurysms are detailed in [26]. Qualitative parameters and their associations to aneurysm rupture are discussed in [27]. Here, both classes are briefly summarized starting with the most important **quantitative** hemodynamic parameters [55]:

- *Velocity* encodes flow direction and speed per grid point and time step by a vector $\vec{v}$. All other quantitative and qualitative parameters are derived from $\vec{v}$. SI unit: cm/s or m/s.
- *Fluid Pressure q* constitutes the scalar kinetic energy per unit volume of a fluid particle. It is computed per grid point and time step. SI unit: Pascal (Pa).
- *Wall Shear Stress (WSS)* encodes the force tangential to the vessel wall exerted by the blood flowing past. It is computed for each time step at grid points along the wall as the scalar magnitude of the corresponding WSS vector $\vec{\tau}_{wss}$. SI unit: Pascal (Pa).
- *Oscillatory Shear Index (OSI)* indicates flow disruption by the time average strength of temporal deflection of $\vec{\tau}_{wss}$ from the time-averaged WSS vector. Dimensionless scalar per grid point along the wall.
- *Volumetric flow rate Q* represents the scalar amount of blood volume passing through a predefined surface region per unit time. It is frequently computed at the ostium surface and related to flow rates within the parent arteries. Unit: cm^3/s or m^3/s.
- *Turnover time ToT* encodes the average elapsed time of a blood flow particle from entering to leaving a particular vessel region. A high value of ToT for the aneurysm region indicates flow stasis promoting thrombus formation. Unit: seconds.

Qualitative hemodynamic parameters describe structures and properties of the blood flow pattern. They are often derived in a visual inspection of the pattern mostly based on standard flow visualization techniques such as integral curves and cut planes with color-coding or Line Integral Convolution (LIC) [22]. The most important **qualitative** hemodynamic parameters are [55]:

- **Vortices** represent regions of flow swirling around a straight or curved axis line which is also referred to as *vortex core line*. In a large database of ruptured and unruptured cerebral aneurysms ($n = 210$ [27]), more than 95% of the cases contain at least one vortex [171]. So-called *embedded vortices* consisting of an outer vortical layer flowing in one direction along the vortex core line and enveloping an inner vortical layer flowing in the opposite direction along the core line have been reported by Byrne et al. [19]. As part of this thesis, an automatic detection and comprehensive visualization of this special type of vortex was proposed (Chap. 3 [171]). An illustration and more detailed description of the formation of embedded vortices are given in Figure 2(b) and Section 2.3 on page 68 of Chapter 3.

- **Recirculation** occurs when a forward stream of blood reverses and flows back into a *separation zone*. Its relations to aneurysm growth and rupture have been investigated but remain incompletely understood [223; 234].

- **Inflow jet and impingement zone** refer to the structure of high-speed, parallel inflow into an aneurysm and the associated wall region of first impact. In a large-scale study, ruptured aneurysms showed a more concentrated inflow jet and a rather small impingement zone as compared to unruptured aneurysms [27]. An automatic approach to detecting and quantifying the inflow jet and its impingement zone was proposed by Gasteiger et al. [56].

- **Flow type** can be *laminar* or *turbulent*. In laminar flow, particles move mostly parallel or in a swirling motion along a common axis (vortices) while in turbulent flow, chaotic property changes are observed. The *Reynolds number* indicates the flow type with very high numbers (2300-2400) corresponding to turbulent flow. Since human blood flow exhibits rather small numbers (200-400), it is commonly assumed to be laminar [55].

- **Flow complexity and stability** are related to the number of vortices and separation zones in an aneurysm and to their persistence over the cardiac cycle. A flow is said to be *simple* if only one vortex or separation zone exists. It is said to be *stable* if this vortex or separation zone neither moves nor collapses and reappears over time. In the above-mentioned large-scale study, simple and stable flow patterns were more frequently seen in unruptured aneurysms while complex and unstable patterns were mostly observed in ruptured aneurysms [27]. In a review of this study, a more quantitative assessment based on vortex core line detection was proposed [20]. Flow complexity was expressed by core line length with multiple core lines resulting in a high accumulated length. Flow stability was characterized by a new entropy measure yielding high values in case of significant temporal changes of the flow field. Ruptured aneurysms were associated with higher core line lengths and higher entropies.

The spatio-temporal hemodynamic data comprises a multitude of parameters such as scalar values, vectors, and flow structures, e.g., vortices represented by vortex core lines. Dedicated visualization and analysis approaches are required to derive knowledge from this complex data.

1.3.5 Visual Analytics: Potential and Approaches

Potential An ideal visual analytics solution should not only aim at gaining insights into the final outcome of the hemodynamic data generation pipeline but also at determining the sensitivity of the outcome with respect to adjustments of the many parameterizable pipeline steps (Sec. 1.3.3). Such a sensitivity analysis requires the installation of feedback loops between knowledge generation and each of the steps (Fig. 1.1). However, this is hampered by the partially necessary manual interaction and the long duration of the steps.

A possible solution could provide *simulation ensembles*. They represent a collection of simulation runs employing different parameter settings, not only of the CFD simulation but also of the preceding steps such as segmentation and surface reconstruction. Approaches to the visualization and analysis of simulation ensembles have been surveyed by Kehrer and Hauser [102]. Very recent publications provide guidance in intelligently sampling the very large parameter space to reduce the number of needed simulation runs [140] and allow the analyst to quickly identify local variations in the outcome across individual runs [98].

The potential of interactive visual analysis in investigating spatio-temporal vector fields and associated as well as derived parameters has been demonstrated based on the `SimVis` framework (Fig. 1.3). Applications include automotive engineering [18], climate [103], and medical research [274]. `SimVis` implements a coordinated multiple views system with tailor-made support for very large unstructured grids [153] and for integrating local feature detectors, e.g., vortex extractors [18]. The investigation of CFD simulations of aneurysm hemodynamics would strongly benefit from both features.

Flow partitioning techniques decompose the flow into regions of common structure [201]. This can be exploited for the investigation of aneurysm hemodynamics in multiple ways. Graphical representatives of the regions can be computed and aggregated in a visual summary of the entire flow or a subsequent visualization can be restricted to regions with specific properties, e.g., vortices. The decomposition ("analyse first") and the subsequent visualization of representatives or regions with specific properties ("the important") are in line with the first two steps of Keim's visual analytics mantra "Analyse first, show the important, zoom/filter, analyse further, details on demand" [104]. Many flow partitioning techniques are based on integral curves since in contrast to local vectorial flow information, they represent continuous flow patterns traced over the domain. The papers in Part I also focus on the decomposition of integral curve sets. In the following, potential strategies to achieve such a decomposition are described before published approaches in the context of aneurysm hemodynamics are detailed.

Flow partitioning strategies can be classified into user-guided and automatic (Sec. 3.1-3.2, p. 53, Chap. 2). Examples for user-guided strategies derive attributes from streamlines and pathlines, such as box counting ratio, curvature, and torsion, and employ an interactive visual analysis approach to filter the lines based on interesting values ranges of these attributes [219; 221]. Other strategies filter the lines based on user-defined Boolean combinations of local curve properties, so-called *streamline predicates* [202]. Filtering by pattern matching of a user-defined template streamline with the overall set of lines has recently gained attention [127; 131; 235]. Multi-field pattern matching facilitating the tracking of a template feature over time extends this idea to unsteady simulation data [259]. All user-guided strategies can be employed for *feature localization* representing one of the feature specification patterns of interactive visual analysis (Sec. 1.2.3).

Automatic flow partitioning strategies employ a data-driven approach to decompose the flow. They utilize a clustering algorithm, such as *k-Means*, *agglomerative hierarchical* or *spectral clustering*, to group integral curves based on a measure of similarity [29; 131; 132; 141; 192; 273]. Some approaches include the computation of cluster representatives and their aggregation in an uncluttered visual summary of the flow. Representatives can be streamlines located at cluster boundaries [273], the most dissimilar streamlines of a cluster [141], or the cluster centroid streamline [29]. Partitioning approaches based on hierarchical clustering usually support browsing the generated hierarchy where each level corresponds to a specific number of clusters [131; 141; 273]. As in many other application fields, this is beneficial in aneurysm hemodynamics since the correct number is unknown. A graph-based interface depicting relations among clusters and spatio-temporal regions has been proposed for navigating a cluster hierarchy [132]. Probing vector fields with a seeding rake and clustering the seeded streamlines on-the-fly represents a combined strategy of

automatic and user-guided flow partitioning [141]. It can be employed in a *local investigation* representing one of the feature specification patterns of interactive visual analysis (Sec. 1.2.3).

Approaches The number of proposed visual analytics approaches to investigating cerebral aneurysm hemodynamics is rather small given the great potential of visual analytics. A likely reason is that these data are rare just as the required collaborative efforts between medical doctors, CFD engineers, and visualization experts. A few approaches are described in surveys of the visual exploration of simulated blood flow [180; 254]. They are briefly recapitulated in the following together with the most recent work. The majority of approaches does not incorporate multiple linked views as it is common in visual analytics solutions. However, an automated analysis — often, a flow partitioning approach — is combined with interactive visualizations for investigating complex datasets, which is in line with the definition of visual analytics (Sec. 1.1).

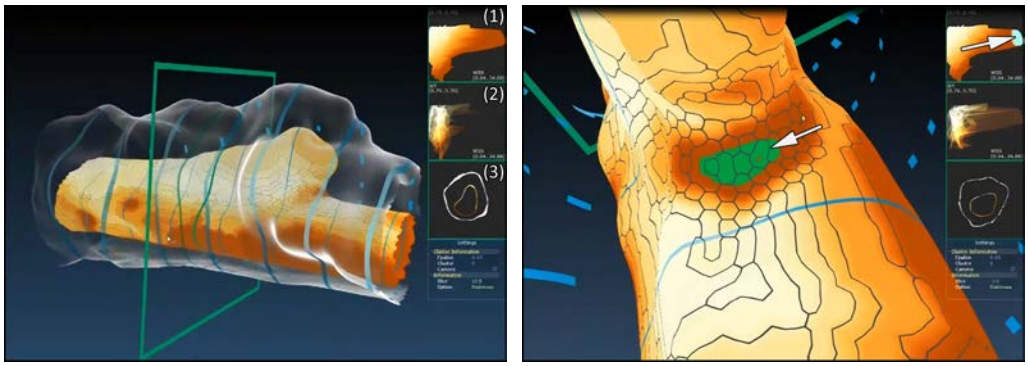
Kuhn et al. [122] decomposed a vector field based on the local bending energy of streamlines and a density function describing the local probability of certain energy values. The function is evaluated over the simulation domain and its minima representing cluster boundaries are computed. Based on an adjustable target number of clusters, cluster neighbors are then merged. Each final cluster is labeled as vortical, laminar or turbulent based on an eigenanalysis of the Jacobian matrix of representative cluster elements. The cluster regions are finally visualized by means of semi-transparent surfaces yielding a quite abstract flow representation.

Gasteiger et al. [56] presented an automatic detection of inflow jet and impingement zone inspired by streamline predicates. The inflow jet hits the aneurysm surface at the impingement zone and is characterized there by a rapid loss of speed and a significant change in flow direction. Hence, streamlines are traced into the aneurysm and multiple parameters are computed along the way such as curvature, acceleration, and minimum distance to the aneurysm surface. Streamlines exhibiting parameter values that are characteristic for the inflow jet are then employed in computing its bounding stream surface which in turn is utilized in a derivation of the impingement zone. Inflow jet and impingement zone are then applied to focus the analysis, e.g., on the Wall Shear Stress inside and in the vicinity of the zone. Later, van Pelt et al. developed comparative blood flow visualizations for virtual stenting assessment incorporating the quantitative definitions of the jet and its impingement [249].

Neugebauer et al. [159] introduced a system for the qualitative exploration of near-wall hemodynamics. They were particularly interested in the relation of high surface curvature and a potentially increased risk of aneurysm rupture as suggested by the investigation of cerebral aneurysms exhibiting *blebs* [142]. In a first step, measures reflecting the surface curvature are computed and automatic thresholding is employed to isolate regions exhibiting high values. The regions are then ranked according to maximum curvature and at highly-ranked regions, streamlines are seeded within the corresponding near-wall boundary layer. They depict the near-wall flow and may form characteristic structures at known or potential rupture sites. The proposed system supports interactive control over thresholding and seeding parameters as well as navigation and flow classification facilities.

Glaßer et al. [62] proposed a combined visualization and analysis of wall thickness and Wall Shear Stress. Since the thickness cannot be measured in vivo yet, they employed a dissected saccular aneurysm phantom whose wall was measured using intravascular ultrasound (IVUS). The hemodynamics were then generated from the IVUS data as described in Section 1.3.3. In a combined visualization of inner and outer vessel wall, Wall Shear Stress is color-coded on the inner wall and wall thickness is encoded by distance ribbons on the semi-transparent outer wall (Fig. 1.4(a)). The visualization is linked to a scatter plot opposing both parameters. Brushing interesting value ranges in the plot causes an emphasis of the corresponding wall parts (Fig. 1.4(b)). A surface clus-

tering approach is integrated to decompose the inner aneurysm wall with respect to a custom risk of rupture score built upon both parameters.



(a) (b)

Figure 1.4: Combined visualization and analysis of wall thickness (WT) and Wall Shear Stress (WSS) in a side-wall aneurysm phantom. *(a)* WSS is color coded on the opaque inner vessel wall. Distance ribbons encode WT on the transparent outer wall. A global (1) and a local (2) scatter plot oppose WSS and WT of the entire wall and the currently visible part, respectively. A contour view (3) illustrates WT and is linked to a widget (green rectangle) that can be dragged along the vessel centerline. *(b)* Brushing a scatter plot (right arrow) results in a colored emphasis of the corresponding wall region (left arrow). Clusters of similar WSS values are indicated along the wall by dark contour lines. Images are courtesy of Sylvia Glaßer, University of Magdeburg.

Lawonn et al. [124] built upon the work of Glaßer et al. [62] and presented the combined visualization of wall thickness and blood flow animations. They encode the thickness on the outer vessel wall using an adjustable discretized color scale (Fig. 1.5). The inner vessel-wall is superimposed using illuminated contours or image-based hatch strokes. Pathline segments depicting the animated flow are drawn as arrow glyphs whose color encodes the distance to the vessel wall based on a discretized scale. A white contour is added to glyphs with a very short distance to hint at near-wall flow. Dynamic cutaway views revealing the flow as it is passing by the vessel and aneurysm wall are automatically generated. The user can interactively adjust parameters of the cutaway generation as well as the discretization of the color scales for wall thickness and distance to the wall.

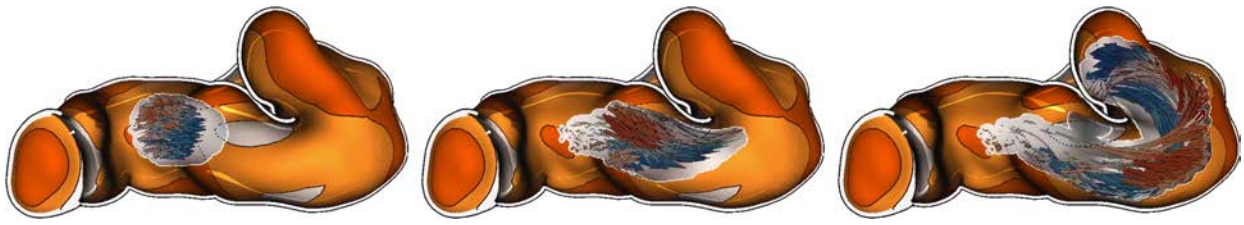


Figure 1.5: Animated pathline segments represented by arrow glyphs illustrate the flow inside a side-wall aneurysm and its parent vasculature at three subsequent time steps. Dynamic cutaways reveal the inner flow as it is passing by. Wall thickness is color-coded on the outer vessel wall using a discretized scale (bright regions indicate a thin wall). The inner vessel wall is superimposed using illuminated contours. Images are courtesy of Kai Lawonn, University of Koblenz.

Oeltze et al. [168; 169] compared multiple streamline clustering approaches in the context of aneurysm hemodynamics and employed the best performing approach to compare different virtual

stenting scenarios (Chap. 2). They conducted a quantitative evaluation of k-Means, agglomerative hierarchical, and spectral clustering based on internal cluster validity measures. They further proposed a visual summary of blood flow, which is composed of one representative streamline per cluster, thereby reducing the visual clutter caused by the original dense streamline sets (Fig. 4(a-b), p. 59, Chap. 2). The clustering was based on streamline geometry as well as application-specific streamline attributes both yielding visual summaries that were beneficial in comparing virtual stenting scenarios (Fig. 5-8, pp. 62-63, Chap. 2). The visual summaries were equipped with various interaction facilities (Sec. 6.3, pp. 59-60, Chap. 2). The user may for instance, adjust the number of clusters and display the original streamlines per cluster. Accompanying videos illustrating the approach are available at `https://www.youtube.com/watch?v=-RVVgqDHzdc` and `https://vimeo.com/102526517`.

In a very recent work, Oeltze-Jafra et al. [171] proposed a clustering-based visual analysis of vortical flow in cerebral aneurysm hemodynamics (Chap. 3). They focused their analysis on *embedded vortices* forming around so-called *saddle-node bifurcations* (Sec. 2.3, p. 68, Chap. 3). A pipeline for the automatic detection and visualization of vortices was presented and tailored to embedded vortices. It incorporates steps for vortex core line extraction and enhancement, a two-step clustering approach generating a coarse grouping of streamlines and an optional group refinement, and the aggregation of custom cluster representatives in a visual summary of vortical blood flow (Fig. 4, p. 70, Chap. 3). In terms of conveying the structure of an embedded vortex and local flow direction, the resulting visual summaries clearly outperform conventional streamline displays (Fig. 1, p. 67, Chap. 3). The summaries were equipped with interaction techniques developed in previous work [168] and a smart visibility strategy for investigating aneurysms with multiple vortices (see the supplemental video at `https://www.youtube.com/watch?v=rAmjHC0zc0c`). The pipeline was successfully demonstrated based on 17 aneurysm cases.

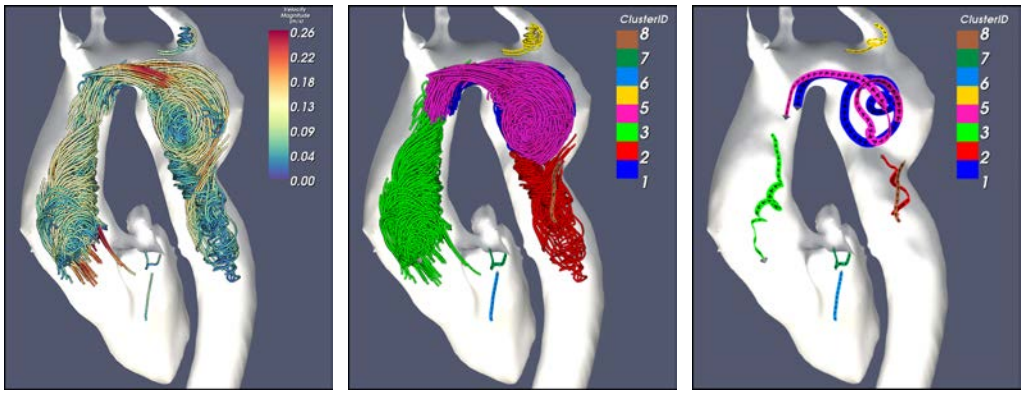
1.3.6 Relation to Measured Cardiac Hemodynamics

Another branch of medical research investigates cardiac hemodynamics with a focus on the aorta distributing oxygenated blood to all body parts trough the circulatory system. Various vascular and cardiac valve pathologies hamper the aortic hemodynamics. Examples are aneurysms and *stenoses* (pathologic narrowing) of the aorta, an *aortic dissection* (separation of aortic wall layers), and a *bicuspid aortic valve* (two of the original three valvular leaflets are fused). Investigating the hemodynamics contributes to an understanding of the initiation and progression of these pathologies as well as to a prediction of their outcome and the patient-specific selection of a suitable therapeutic intervention.

The data generation pipelines of aortic and cerebral aneurysms hemodynamics differ considerably since aortic flow can be directly measured using, e.g., 4D Phase-Contrast (4D-PC) Magnetic Resonance Imaging (MRI). This is feasible due to the much larger diameter of the aorta as compared to the arteries of the Circle of Willis (2.5-3.5 cm vs. 1-3 mm). However, the 4D-PC MRI data suffers from various artifacts such as noise, *phase distortions*, and possibly *phase wraps*. Hence, the pipeline of aortic hemodynamics integrates multiple artifact reduction steps between the imaging and the vessel segmentation step (Sec. 1.3.3). It obviously lacks the volume grid generation and CFD simulation steps. The pipeline results in *phase images* from which a 4-dimensional (3D+time) vector field can be reconstructed. In contrast to simulated cerebral aneurysm hemodynamics, this field is represented on a structured grid (image data). The hemodynamic parameters described in Section 1.3.4 are crucial in both, the cerebral and the cardiac domain. Additional parameters such as the *regurgitant fraction* denoting abnormal reflux of blood, e.g., from the aorta through the aortic valve into the left ventricle, are derived. Comprehensive introductions to cardiac hemodynamic

imaging including a discussion of typical imaging artifacts as well as of image processing steps for artifact reduction and vessel segmentation are given by Gasteiger [55] and Köhler et al. [114].

The potential of visual analytics in investigating cerebral aneurysm hemodynamics also applies to cardiac hemodynamics. The proposed approaches of both domains frequently pursue the same goals, e.g., a flow decomposition and the analysis of vortical flow, and employ the same techniques to achieve these goals. Peculiarities result particularly from the different input grid types and the lower quality of the cardiac hemodynamic data caused by artifacts and the limited resolution of the imaging. For instance, integral curves in measured data are rarely traced over the entire domain due to numerical instabilities yielding much shorter curves than in simulated data. However, meaningful visual summaries of blood flow can still be generated based on these curves, e.g., with the approach by Oeltze et al. [168] (Fig. 1.6). Köhler et al. [114] provide a comprehensive survey of approaches to the visualization of measured cardiac hemodynamics. Exemplary papers in the realm of visual analytics are dedicated to vector pattern matching for vortex detection [79], the clustering-based generation of sparse visual flow summaries [250], the detection of important flow structures such as vortices based on line predicates [16; 115], and the clustering-based classification of vortices [145]. In a discussion of visualization challenges related to understanding cardiac hemodynamics, van Pelt and Vilanova [251] suggested to "supplement the limited spatial and temporal resolution of imaging data with physically based fluid simulations". In joint work with these authors, Hoon et al. [35] proposed an approach to harnessing the mutual benefits of both and showed that a coupled investigation is more accurate and less sensitive to noise.



(a) (b) (c)

Figure 1.6: Clustering of pathlines representing vortical flow in an aorta. *(a)* Pathlines have been seeded everywhere in the aorta and pre-filtered according to a vortex criterion [115]. *(b)* The spectral clustering approach proposed by Oeltze et al. [168] is able to automatically separate the individual vortices as well as outlier pathlines, e.g., the light blue and neighboring dark green cluster. *(c)* Cluster representatives provide a visual summary of the vortical flow inside the aorta. The pathlines in *(a)* are courtesy of Benjamin Köhler.

1.4 Toponome Data of Cells and Tissues

This section provides context for the cumulated publications of Part II. It familiarizes the reader with the *toponome* and with *toponomics* — the associated investigating discipline in systems biology, molecular cell biology, and histology [209] (Sec. 1.4.1). Following this, a dedicated robot-driven fluorescence microscopy technique for imaging the toponome is described including aspects of its clinical role in disease diagnosis and drug design as well as the explanation of frequent toponome image data preprocessing (Sec. 1.4.2). Next, the analysis of the preprocessed toponome

data is detailed including biological tasks, the workflow for accomplishing these tasks, and derived requirements on an improved workflow support (Sec. 1.4.3). At the end of this section, the potential of visual analytics in investigating toponome data as well as related approaches are outlined. Furthermore, the papers cumulated in this postdoctoral thesis [166; 167; 172] (Chap. 4-6) are surveyed and their contributions to an improved workflow support are described. The Sections 1.4.1-1.4.3 are based on various articles of the *Encyclopedia of Systems Biology* [43]. Each article is explicitly referenced at the appropriate position within the sections.

Visual analytics approaches have been developed for the investigation of a broad diversity of biological data including microscopy images [166], genome sequencing data [147], gene expression data [36], and biological pathways and networks [59]. The thesis at hand focuses on microscopy image data but gives a short excursion into visual analytics of other biological data at the end (Sec. 1.6.3). The majority of image data in biology is acquired utilizing microscopy imaging. The increasing complexity of these data is related to high resolution imaging of large samples, streamed imaging, and the acquisition of multiple channels capturing individual properties of the imaged sample. The focus of this section and of the publications in Part II is on the latter aspect, i.e. the visual analytics of multi-channel microscopy data.

1.4.1 The Toponome and Toponomics

This section is based on introductions to the toponome and toponomics by Schubert [209] and Oeltze et. al [166] (Sec. 2.1, pp. 85-86, Chap. 4). Cells represent the basic structural and functional building blocks of all living organisms. Proteins are the basic modules of cells and exist in a huge variety within cell membranes and nuclei. The cell can be considered as an apparatus forming dynamic assemblies of clusters of different proteins (*functional protein patterns*) in order to generate concrete cell functions [205; 208]. The mechanics and rules of this apparatus as well as the functional master plan of cells are incompletely understood.

In the research field *proteomics*, the structure and functions of proteins are investigated. The so-called proteome describes the entirety of proteins in a cell or organism at a given functional state and a specific point in time. It is investigated based on *immunoassays* or using *mass spectronomy*, both resulting in protein profiles from which many insights on the molecular function and structure of proteins could already be inferred. However, the cellular function of a protein cannot simply be derived from its molecular one or from structure since it depends on the spatial context of the protein within a network of proteins inside the cell [205; 207; 208; 210; 211; 212]. This spatial information cannot be extracted from proteome data.

Schubert et al. [205] hence introduced the concept of the *toponome* which also comprises protein topology inside cells and tissues. The toponome is defined as the "[...] spatial network code of proteins and other biomolecules in morphologically intact cells and tissues." [209]. The term has been coined by Walter Schubert and is derived from the ancient Greek nouns "topos" (place or position) and "nomos" (law). It acknowledges that cells adhere to topological rules when forming protein networks. The difference in information encoded by proteome and toponome has been illustrated by Schubert [209] (Fig. 1.7). While protein quantities derived from the proteome of normal and abnormal cell samples may be very similar, the corresponding protein topology may considerably differ having a strong impact in disease-related research. A diseased cell may be characterized by protein rearrangement, e.g., on the cell surface, rather than by up- or downregulation of proteins [209].

In *toponomics*, the inner structure, the biological code, and the semantics of the toponome are investigated [205]. It has been shown that the toponome in cells is hierarchically organized [211]. It comprises protein clusters which are interlocked as a network. One or multiple contained *lead*

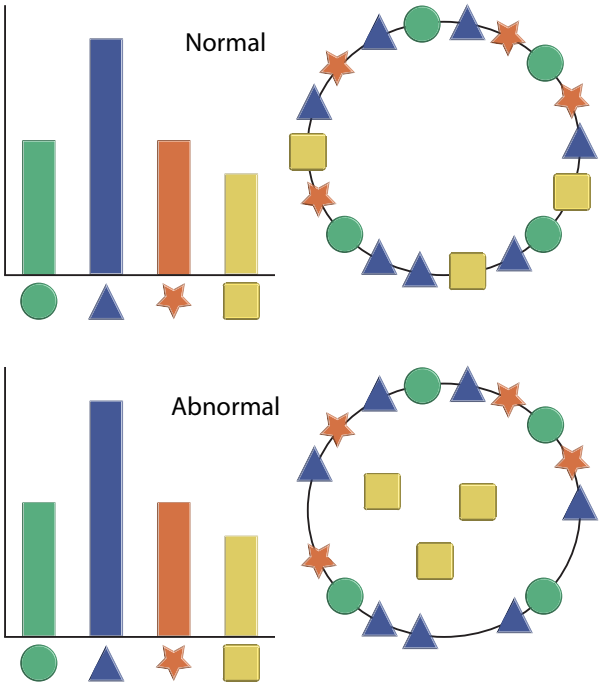


Figure 1.7: Simplified illustration of the difference in information encoded by the *proteome* and the *toponome* for a normal and an abnormal cell of the same type. Extracting proteins and their quantities from protein profiles in *proteomics* shows no difference between the cells (*left*). However, taking protein topology into account in *toponomics*, reveals protein rearrangement, e.g., on the cell surface (*right*). Adapted from [206].

proteins, characterized by an omnipresence in functionally relevant protein clusters, control the topology of the clusters and their function as a network. This has been demonstrated by inhibiting lead proteins using chemical agents [211]. Inhibition caused a disassembly of the corresponding protein clusters which in turn yielded a strongly altered cellular function. Hence, the detection of lead proteins in samples of abnormal cells or tissues is a crucial step in drug design searching for potential *target molecules* in disease treatment. However, neither protein clusters nor lead proteins can be derived or predicted from pure molecular protein profiles. Hence, the toponome must be mapped in cells and tissues employing a dedicated imaging technique.

1.4.2 Imaging the Toponome and Data Transformation

Imaging Imaging the toponome builds upon the general principle of *fluorescence microscopy*: specific molecules, so-called *fluorochromes* or *fluorophores*, absorb light of a particular wavelength, which is then transformed into energy being released again as emitted light with longer wavelength. In a *fluorescence microscope*, light is filtered by an *exciter* such that only radiation with lower wavelengths passes, then hits a *dichroic mirror* and is reflected towards the sample on the *microscope stage* [128]. Fluorochromes in the sample substance now absorb and emit light as described above. The dichroic mirror is designed such that short-wavelength light arriving from the exciter is reflected while light with medium to long wavelengths emitted by the fluorochromes is transmitted. A final *barrier* filters the transmitted radiation such that only long-wavelength light reaches the *ocular*.

Many substances are *autofluorescent*, i.e., no further preparation steps of the sample are necessary. However, labeling of specific components in biological structures often requires dedicated *affinity reagents* or *tags*, e.g., such that recognize individual proteins, which do not exhibit intrinsic fluorescence. In order to induce a fluorescence response, these tags are conjugated to a fluorescent dye and applied to the sample in a preparation step. Microscopic images then convey binding sites of the tag. For instance, locations of a certain protein within a cell are visualized by means of a fluorescence-conjugated tag that is known to bind to this type of protein.

Traditional fluorescence microscopy is not suitable for mapping the toponome since the maximum number of tags recognizing proteins that can be simultaneously imaged in a sample is between five to ten [155]. This is due to the limited spectral separability of the fluorescence response of multiple tags. Already at three tags, emitted wavelengths start to overlap hampering a separation [128]. A possibility to overcome this limitation would be the application and imaging of tags one after another with some sort of neutralization step in between to avoid any energy transfer into the remaining steps. This exactly constitutes the basic idea of the robot-driven *toponome imaging system (TIS)*.

In TIS, the toponome is imaged in a cyclic procedure using multiple fluorescence-conjugated tags recognizing proteins which are organized in a *tag library* [52]. A TIS robot applies the tags one after another to a cell or tissue sample on the microscope stage. In each cycle, the fluorescence response is registered by an *epifluorescence microscope* and recored by a *charge-coupled device (CCD) camera* attached to it. Each cycle is concluded by a neutralization step in which the sample is irradiated for 20 minutes with the excitation light causing a complete bleaching of the fluorescent dye. The imaging yields a resolution in the nanometer range and can be carried out in 2D or in 3D by adjusting the focal plane of the microscope. It results in multi-channel fluorescence data comprising an image or volume per tag/channel. The labeling and mapping of up to 100 proteins in 100 cycles has been demonstrated [211].

Data Transformation Two main strategies in analyzing toponome data can be distinguished: *non-threshold-based* and *threshold-based* [82]. The former employs the raw fluorescence data, e.g., in segmenting the cells of a tissue sample [156], in highlighting local spatial similarities of the multi-channel fluorescence response [214], and in clustering this response [89]. The latter applies a unique threshold to each channel of the raw fluorescence data such that 1 represents protein present and 0 protein absent (Fig. 1(a), p. 104, Chap. 6). Thresholding can be considered a data transformation step in the sense of the visual analytics process (Fig. 1.1). The binarization reduces data complexity and facilitates a clearer interpretation of protein presence but requires the definition of reasonable thresholds, either by an expert [52; 211] or an automatic approach [9]. Within this thesis, only binary toponome data generated by either of the thresholding approaches have been analyzed (Part II).

After binarization, a second data transformation step is carried out. First, the combinatorial binary code (protein pattern) is determined for each pixel or voxel by iterating over all data channels (images or volumes) representing the different proteins. Second, the unique binary codes in the data, which are also referred to as *Combinatorial Molecular Phenotypes (CMPs)*, are computed (Fig. 1(b), p. 104, Chap. 6). Depending on the inspected biology and the number of applied tags recognizing proteins, hundreds to thousands of CMPs may exist. A simple technique for visualizing the toponome is the color-coded representation of the CMPs in a *toponome map* (Fig. 1(c), p. 104, Chap. 6). It reveals the spatial clustering of identical protein patterns with these clusters possibly corresponding to functional cell units, which are of crucial interest.

The clinical role of TIS imaging and investigating the toponome based on CMPs was demonstrated in disease diagnosis and drug design [107]. For instance, distinct protein clusters were found in normal skin and two skin-related diseases *psoriasis* and *atopic dermatitis* [211]. The corresponding toponomes may serve as fingerprints based on which histologists can evaluate skin samples. Specific CMP patterns that may serve as biomarkers in cancer screening and lead proteins representing candidates for target molecules in drug development were determined based on tissue and cell samples of colon cancer [12], prostate cancer [213], and *rhabdomyosarcoma* [211], a malignant tumor of soft tissue in children and adolescents. In rhabdomyosarcoma analysis, it was shown that inhibiting the lead protein stops and reverses as well as prevents malicious cell

transformation [211]. The benefit of TIS in monitoring protein networks for a better understanding of drug actions was demonstrated for the pain reliever and fever reducer *dipyrone* [129]. Very recently, TIS was shown to improve the understanding of impaired immune surveillance by *T lymphocytes* in manifest human skin cancer [83]. Despite these promising results, TIS is so far not used in clinical routine and has been installed at just a few research centers across the world.

1.4.3 Toponome Analysis

Compared to standard Ultrasound, CT or MR medical image data from clinical routine, toponome data are rare and have been acquired for only a few diseases and drugs. In a novel application, often little can be predicted regarding expected protein clusters and lead proteins. While the presence of certain proteins can be presumed based on a priori biological knowledge, which also guides the selection of affinity reagents for the experiment, protein topology may be mostly unknown. Hence, the analysis of toponome data typically starts with a hypothesis-free visual exploration of the extracted CMPs and involves multiple biological tasks. This section starts with an explanation of these tasks and then, elaborates on the biological workflow aiming at their accomplishment. The workflow description is given from a perspective before the contributions of this thesis to clarify the starting point. It discloses various shortcomings from which requirements on an improved workflow support were derived. These requirements motivated the work in Part II of this thesis and are detailed at the end of this section.

Biological Tasks In interviews with toponome experts, observations of their daily work, and reviews of their publications, the following tasks have been identified:

1. detection of selective CMP patterns,
2. comparison of CMP patterns,
3. discovery of lead proteins, and
4. identification and localization of *co-mapping proteins*.

In (1), CMP patterns are searched for that are specific to a certain cell or tissue region of the sample instead of appearing anywhere in the data. This search is part of (2) the pattern comparison between samples of healthy and diseased tissue, cells of different types, individual developmental stages of a cell's life cycle or cells before and after drug administration. The results of (1) and (2) contribute to an understanding of cell composition and function. Furthermore, patterns that are selective for certain pathology, particular cell type or developmental stage of a cell may serve as biomarkers in the screening for diseases. They constitute the basis for (3) lead protein discovery. Lead proteins are omnipresent in functionally relevant units of a cell as indicated by the corresponding CMPs. In rhabdomyosarcoma cells for instance, a specific protein was found to be present in each CMP at cell extensions developing in the course of metastasis formation [211]. In (4), co-mapping proteins, i.e. those being mapped to the same pixel or voxel somewhere in the data, are identified and localized. Co-mapping quantities are determined and the spatial distribution of co-mapping is investigated. The spatial proximity of proteins provides no evidence but suggests binding partners in *protein-protein interactions* [270].

Biological Workflow The biologists accomplish these tasks in their in-house visual analytics framework `MultiCompare` implemented as coordinated multiple views system employing linking & brushing. In early versions, this framework comprised a table view, a toponome map view, and a filter view (Fig. 2, p. 105, Chap. 6; filter view is not shown). In the course of the habilitation, additional views and interaction facilities were added to `MultiCompare` as will be detailed in Section 1.4.4. However, the following workflow description neglects these new features to clarify the starting point.

Since `MultiCompare` does not support the handling of 3D toponome data, individual slices are processed separately and the results are aggregated and matched externally. After 2D toponome data has been loaded into the framework, the table view lists all CMPs as rows, sorted according to frequency of occurrence from top to bottom, and all proteins as columns following their order of acquisition (Fig. 2(a), p. 105, Chap. 6). Table entries equal 1 or 0 and indicate protein present or absent, respectively. Each CMP is assigned a unique color which serves as its visual identifier in all views of the framework. The toponome map view shows a background grayscale phase contrast image which facilitates a rough visual separation of background, cell surface, and cell nucleus and hence, provides a spatial reference for exploration (Fig. 2(b), p. 105, Chap. 6). The filter view allows for the definition of a template CMP, which is then matched with the set of existing CMPs. Each protein of the template can be set to 1, 0 or a wildcard symbol with the latter indicating that both values are accepted.

The biologists familiarize themselves with the spatial domain of the data by inspecting the phase contrast image. Then, they search for selective CMP patterns employing *local investigation* representing one of the feature specification patterns of interactive visual analysis (Sec. 1.2.3). A morphologically interesting focus region is defined in the background image using a rectangular brush (Fig. 2(b), p. 105, Chap. 6). The CMPs inside the brushed region are computed, their corresponding rows are highlighted in the table view, and the respective regions of the toponome map are superimposed on the background image. Highlighting and superimposition both employ the CMPs' unique identifier color. In the toponome map view, the CMPs are superimposed inside but also outside of the brushed region. The latter is required to evaluate if the CMP pattern is selective or appears anywhere in the data. If the pattern is not selective, the exploration proceeds and a new focus region is defined. A new rectangular brush has to be created from scratch since the existing one is neither draggable nor resizable. If a selective pattern is found, it is compared with patterns of other cells or cell parts. Since no multiple brushes are supported, multiple instances of `MultiCompare` must be opened for pattern comparison. The biologists then employ an extra plug-in for computing pattern similarities and differences based on stored copies of two CMP tables, each with labeled rows indicating the respective pattern.

The table view is utilized for further investigating a selective CMP pattern regarding its contributing CMPs and their proteins. Each protein column is checked for omnipresence of entries equal 1 across all highlighted table rows to discover lead proteins. Furthermore, each highlighted row is inspected for multiple entries equal 1 to identify co-mapping proteins. The location and spatial distribution of the co-mapping proteins of a particular CMP can be observed in the toponome map view by searching for the likewise colored regions. If the biologists are interested in specific co-mapping quantities or co-mapping proteins in the entire sample, not just a focus region, they employ *feature localization* representing another one of the feature specification patterns of interactive visual analysis (Sec. 1.2.3). They either select one or multiple CMPs exhibiting interesting co-mapping proteins in the table view or define a template CMP using the filter view both triggering an update of the toponome map view. The co-mapping quantities are displayed in the status bar. In order to protocol the insights gained during an analysis session, the biologists create screenshots of `MultiCompare` and integrate them in their laboratory book.

Requirements on Improved Workflow Support The described workflow suffers from various shortcomings from which the following requirements (R) on an improved workflow support were deduced:

R1 *Exploration and visualization of 3D toponome data*. The slice-wise processing is extremely tedious and error-prone. The comprehension of 3D structures is severely complicated by mental aggregation of 2D slice information. A 3D visualization must offer interaction facilities for focus region definition and for mitigating occlusion problems.

R2 *Dynamic, flexible focus region*. Redefining each new focus region from scratch significantly slows down the detection of selective CMP patterns. It hampers the comprehension of pattern changes between neighboring sample regions. Draggable and resizable regions was well as a paralleling fluent update of all views are required.

R3 *Multiple focus regions*. The missing support for multiple focus regions strongly hampers the comparison of CMP patterns. An elimination of this shortcoming requires techniques for the visual separation of multiple CMP patterns in both, the table and the toponome map view.

R4 *Hints on lead proteins*. Manually searching for lead proteins in a CMP pattern is time-consuming and error-prone since the table view is not restricted to the corresponding rows of the pattern (Fig. 2(a), p. 105, Chap. 6). Tables often comprise hundreds to thousands of CMPs requiring extensive scrolling. The likelihood of being a lead protein could instead easily be computed per protein based on the percentage of CMPs it is registered for.

R5 *Comprehensive overview of protein co-mapping*. Manually searching for co-mapping proteins of a CMP pattern is time-consuming and error-prone since entries of present proteins are not highlighted in the table view (Fig. 2(a), p. 105, Chap. 6). An overview visualization of protein co-mapping is required indicating also co-mapping quantities

R6 *In-place annotation of CMPs and proteins*. Visually matching CMPs in the table view with their corresponding regions in the toponome map view and vice versa, e.g., in a fine-grained investigation of a selective CMP pattern, requires the user to constantly move the focus of attention back and forth between the views. Often, the views are displayed on separate screens to gain maximum space. The investigation would strongly benefit from annotating CMPs and their registered proteins directly in the toponome map or a 3D view (R1).

R7 *Optimized CMP identifier colors*. The matching mentioned in R6 is error-prone due to similar colors assigned to different CMPs. Given the potentially high number of CMPs and the limited ability of humans to reliably discriminate multiple colors, this is inevitable. However, a CMP pattern of interest often comprises just a small subset of all existing CMPs and for its identifier colors, a temporary increase in perceptual difference is required.

1.4.4 Visual Analytics: Potential and Approaches

Potential Toponome data constitute an instance of scientific multivariate data since they represent spatial structures and comprise multiple channels or attributes. Approaches to the visualization and visual analysis of scientific multivariate data have been surveyed by Kehrer and Hauser [102]. None of the reviewed visual analysis approaches offers dedicated support for binary data. However, many oft them build upon the interactive visual analysis methodology and are implemented as coordinated multiple views system comprising a physical and various types of

attribute views (Sec. 1.2.3). The workflow for analyzing toponome data is also realized in such a system (Sec. 1.4.3). The surveyed approaches comprise analysis components, e.g., for computing correlations between attributes or for dimensionality reduction. These components are designed for processing quantitative data and require an adaptation for binary data analysis [33].

Most attribute views in interactive multivariate data analysis employ standard visualization techniques intended for quantitative data, such as scatter plot and parallel coordinates. Binary data require other abstract visualizations as for instance a matrix [80; 197], a graph [166], dedicated glyphs [125], and other custom forms [118]. Matrix and graph are particularly well-suited for representing co-mapping proteins. Entries equal 1 in an adjacency matrix and an edge between two proteins indicate co-mapping pairs. Interesting patterns in matrices can be revealed by matrix reordering [80]. Matrix representations are also applied in comparing multiple binary data subsets [197]. This could be transferred to CMP patterns. Graphs can be coupled with glyphs for encoding co-mapping quantities [166] and for showing protein relations derived from biological pathways [34]. General graph properties can be determined using graph analysis methods and local topology information can be extracted by means of graph signatures [265].

Physical views for visualizing scientific multi-variate data frequently apply glyphs or layering techniques [102]. The latter encode each attribute in a separate layer and employ opacity modulation or different rendering styles in a combined representation of these layers. None of the techniques can cope with the high number of attributes in toponome data (up to 100 proteins [211]). While at no spatial position all imaged proteins co-map, the maximum observed number may still be in the double figures. The high local entropy of toponome data poses further challenges. In contrast to larger homogeneous structures in medical images, phenotypically identical structures in toponome data may cover only a few pixels. This together with the high number of attributes and the binary nature of the data requires dedicated 3D visualization and exploration techniques [167].

As indicated in the paragraph on data transformation in Section 1.4.2, binarization of the raw toponome data is a sensitive task requiring the careful fine-tuning of thresholds. Subtle differences in thresholding may have a considerable impact on the subsequent data analysis (Sec. 1.4.3). Visual analytics methods may help the biologists in investigating this source of uncertainty and in incorporating uncertainty in their toponome analysis workflow [135].

Approaches The approaches presented in Part II of this thesis are motivated by the requirements *R1-R7* on an improved workflow support for analyzing toponome data (Sec. 1.4.3). Their compliance with these requirements is summarized in the following. All approaches have been integrated into the coordinated multiple views framework `MultiCompare` to advance the visual analytics of toponome data (Fig. 2, p. 105, Chap. 6).

Oeltze et al. [166] presented the interactive, graph-based visual analysis of toponome data (Chap. 4). They focused on complying with requirement *R5* and secondary with *R1-R4*, and *R7*. They extended the `SimVis` framework described in Section 1.2.3 by a graph and a table view (Fig. 3(b,e), p. 91, Chap. 4). The graph view was later integrated into `MultiCompare` by Klemm [108]. It shows the proteins of the current CMP pattern as nodes and co-mapping protein pairs as edges (*R5*). The edge width encodes co-mapping quantity (Fig. 3(b,d), p. 91, Chap. 4). Lead proteins are indicated by the filling level of circular glyphs attached to the nodes of the graph (*R4*). A circle is fully filled if the corresponding protein is registered for every CMP of the pattern (Fig. 5, p. 92, Chap. 4). A resizable and draggable focus region can be defined in a scatter plot opposing x-and y-coordinates of the cell sample (*R2*, Fig. 6, p. 92, Chap. 4). Focus region modification triggers an update of all linked views, which is achieved at interactive frame rates by means of parallel programming. `SimVis` allows for the definition of multiple features, e.g., CMP patterns, by means of multiple brushes (*R3*). The graph representation of a feature then visually emphasizes

the differences in comparison with the respective other feature (Fig. 7, p. 93, Chap. 4). `SimVis` copes with both, 2D and 3D toponome data and its $focus + context$ volume rendering view shows a CMP pattern in its spatial context (*R1*, Fig. 3(c), p. 91, Chap. 4). Interaction facilities besides standard navigation are not provided. The color transfer function mapping can be parameterized such that always the full range of available colors is employed in visualizing the current CMP pattern (*R7*). However, this yields multiple colors for the same CMP during an analysis which impairs its recognition. A supplemental video further illustrating the compliance with all mentioned requirements is available at `https://www.youtube.com/watch?v=nU9yLY7lXyM`.

Oeltze et al. [167] introduced dedicated visualization and exploration techniques for 3D toponome data (Chap. 5). They focused on complying with requirements R1 and R7 and secondary with R2 and R6. The developed 3D view, its associated interaction facilities, and the proposed CMP color assignment were developed within `MultiCompare` by Klemm [108]. The 3D view is based on a focus+context ray-casting approach similar to `SimVis` (*R1*). However, instead of mapping CMPs through a color transfer function, it directly employs the unique CMP colors computed by `MultiCompare` (Fig. 3, p. 98, Chap. 5). In a close-up view of the sample, e.g., a zoom in on a specific cell, the visual separability of the currently visible CMPs can be improved on demand (*R7*). They are temporarily assigned a new set of identifier colors with a sufficient perceptual difference (Fig. 5, p. 99, Chap. 5). Once the analysis continues, the original coloring is restored. Similar to delineating a focus region in the 2D toponome map, a focus volume can be defined directly in the 3D view by means of a resizable and draggable brush (*R2*, cf. Fig. 2(b), p. 105, Chap. 6 and Fig. 7, p. 101, Chap. 5). On mouse hover in the 3D view, the visible CMP under the mouse pointer is determined and the names of its co-mapping proteins are displayed in a tooltip (*R6*). In order to mitigate typical occlusion problems in 3D, layering cell structures can be successively peeled off (*R1*, Fig. 8, p. 101, Chap. 5).

More recently, Oeltze-Jafra at al. [172] proposed an interactive labeling method for annotating toponome data (Chap. 6). They focused on complying with requirement *R6* and secondary with *R2-R5*. The labeling was developed within `MultiCompare` by Pieper [177]. The basic idea was to employ a dynamic version of *necklace maps* [226] to provide in-place annotations of CMP patterns within the toponome map (*R6*). A resizable circular focus region can be dragged across the map (*R1*). Simultaneously, circular symbols strung on two surrounding necklaces label the focused CMPs and their proteins (Fig. 3, p. 107, Chap. 6). The color of a protein symbol encodes the percentage of focused CMPs this protein is registered for. A distinct coloring indicates potential lead protein candidates (*R4*, yellow and green in Fig. 3, p. 107, Chap. 6). Clicking on a protein symbol highlights all symbols of co-mapping proteins and all symbols of CMPs this protein is registered for (*R5*). Clicking on a CMP symbol highlights all symbols of proteins registered for this CMP. Multiple focus regions can be defined and their respective necklaces are organized in a separate, storable management view (*R3*, Fig. 6, p. 109, Chap. 6). The non-overlapping display there simplifies a comparison of the necklaces, i.e. the represented CMP patterns. The management view provides a means to structure and log the visual exploration of toponome data. A supplemental video further illustrating the compliance with all mentioned requirements is available at `https://www.youtube.com/watch?v=9sCSPctFRTc`.

No further visual analytics approaches tailored to spatial, binary, multi-variate data are known to the author of this thesis. However, Kölling et al. [116] presented a web-based visual analytics tool for mining protein co-mapping in the raw fluorescence data. They employed *self-organizing maps* for clustering the multi-channel fluorescence response and glyphs for visualizing the clusters. Each glyph encodes the proteins' average fluorescence intensity profile for that cluster. In a coordinated multiple views system, toponome map views, a clustering results view, and a bookmark list logging inspected clusters are synchronized.

A few other approaches related to the visualization and visual analysis of multi-channel 3D microscopy data are briefly described in Section 3.2 on page 97 of Chapter 5. More recent work in visual analytics of biological image data is dedicated to the investigation of kinetic changes in cells deduced from fluorescence microscopy data [99], to the study of neuronal connectivity derived from electron microscopy data [1; 11], and to the discovery of potential relationships in digital histology image collections [38]. However, none of these approaches can readily process neither raw nor binarized toponome data.

1.5 Epidemiological Population Study Data

This section provides context for the cumulated publications of Part III. It briefly introduces *population studies* in epidemiology (Sec. 1.5.1), describes specifics of the study data and their acquisition (Sec. 1.5.2), and elucidates the standard epidemiological workflow for analyzing the data (Sec. 1.5.3) as well as the potential of visual analytics (Sec. 1.5.4). Further details regarding all these aspects are given in the work by Preim et al. [181] which corresponds to Chapter 7. Since we submitted this work in February 2014, a short update on the most recent work in visual analytics of population studies, including the publications cumulated in this postdoctoral thesis [4; 112; 113; 170] (Chap. 8-11), is integrated in Section 1.5.4. The update focuses on visual analytics of population studies in epidemiology, and in particular, on those comprising the acquisition of medical image data. Another scientific field frequently conducting such *image-centric studies*, which however include considerably less individuals (n), is neuroscience. For instance in dementia research, MRI data of the brain is acquired and hippocampal activity ($n = 22$ [133]) and morphology ($n = 9$ [106]) are related to results of memory performance tests. The publication in Chapter 11 [170] was created in this context. *Electronic health records* stored in hospital databases provide information about health-related events in patient populations. Unlike population study data, these records are not compiled for the investigation of specific diseases or health-related effects and risks but mainly for process optimization and quality control in health care. However, they provide an invaluable source of information for epidemiologists, their investigation is a hot topic in visual analytics, and the approaches developed there are transferable to population study data. Hence, a brief introduction to the visual analytics of electronic heath records is given in Section 1.6.2 although these data have not been in the focus of this thesis.

1.5.1 Role of Population Studies in Epidemiological Research

A comprehensive introduction to epidemiology is given by Merrill [144] and a dictionary of epidemiological terms has been compiled by Porta et al. [179]. Epidemiology is a branch of medicine. In contrast to clinical medicine, it does not focus on the diagnosis and treatment of a specific disease, such as a cerebral aneurysm, in a single patient (Sec. 1.3.1). Instead, epidemiology is a scientific discipline investigating the occurrence and distribution of health-related events in *defined populations* [179]. The term "defined population" refers to individuals sharing a common characteristic such as gender, ethnicity, health condition or residential area (close to a nuclear plant or the seashore etc.). Epidemiological investigations typically pursue at least one of the following goals:

1. determination of *prevalence* and *incidence* of diseases, i.e. number of diseased people at a certain point in time and number of people falling ill within a specified time period,

2. identification and characterization of *risk factors* being casually related to changes in relevant health conditions and based on that

3. assessing per individual the risk of falling ill with a specific disease,
4. efficiency evaluation of preventive and therapeutic measures, and
5. determination of differences between healthy aging and pathologies in an early stage.

Exemplary results of such investigations are the linear relationships between cancer risks and ionizing radiation dose [61] or nicotine intake [41], the U-shaped relationship between the risk of dementia and alcohol consumption [193], and the J-shaped relationship between the risk of coronary heart disease and alcohol consumption [31]. Further results prove that epidemiology is not just a pure academic endeavor but has huge consequences, also for clinical medicine. Mildred Vera Peters demonstrated that in treating early-stage breast cancer, breast-conserving surgery followed by radiation therapy is as effective as radical mastectomy while at the same time having a much lower emotional impact [175]. Tukenova et al. showed that radiation dose must be minimized in treating cancer in children and adolescents since higher doses correlate with an increased late mortality from sarcoma, carcinoma, and hematological malignancies [240]. A wide range of public campaigns, screening recommendations, and treatment suggestions as for instance, passive smoking protection, safer sex education, vaccination plans, breast and prostate cancer screening, and suggestions for diabetes treatment, are all based on epidemiological research.

Population studies are the main vehicle of epidemiological research. They collect hundreds of socio-demographic, lifestyle-related, and health-related variables for thousands of individuals in a defined population by means of interviews, questionnaires, and various medical examinations. Studies in modern epidemiological research, such as the Study of Health in Pomerania (SHiP) [256], comprise laboratory tests of blood, urine, and DNA, sleep monitoring, electrocardiogram (ECG) recording, and also medical imaging. In 2008, SHiP was the first study to even include whole-body imaging [78]. The inclusion of medical imaging in a population study bears a great potential since it facilitates a survey of the broad variability in vital organ anatomy and physiology, an improved characterization of health and disease, and a differentiation between physical effects of normal aging and pathologies. Due to ethical reasons, non-invasive imaging techniques such as Magnetic Resonance Imaging (MRI) and Ultrasound (US) are commonly employed for investigating a healthy population.

Epidemiological studies are often *longitudinal*, i.e. carried out in multiple *waves* over years. If a defined population is traced over time, it is referred to as a *cohort* and its study is termed *cohort study* [179]. The visual analysis of cohort study data in chapters 8-9 had to be restricted to one wave due to a lack of medical image data for the remaining waves. Accordingly, the term population study might have been more appropriate in the corresponding publications. Examples for cohort studies which comprise medical imaging, so-called *image-centric cohort studies* [181], are:

SHiP [256] *Initial population*: 4,308 adults of all age groups; *Focus*: explanation of health-related differences between East and West Germany after the German reunion; *Imaging*: US, whole-body MRI, contrast-enhanced MRI [78]

Rotterdam Study [84] *Initial population*: ≈8,000 adults older than 45 years; *Focus*: neurological, cardiovascular, loco motor, and opthalmic diseases; *Imaging*: US, MRI including Diffusion Tensor Imaging (DTI), resting-state functional MRI (rs-fMRI), perfusion MRI [92]

UK Biobank [2] *Initial population*: ≈500,000 adults aged between 40-69 years; *Focus*: diseases with high prevalence in aging society such as cancer, heart diseases, stroke, diabetes, and dementia; *Imaging*: brain, cardiovascular, and abdominal MRI, carotid US, and Dual-energy X-ray absorptiometry (DEXA) of ≈100,000 participants [176]

Norwegian Cognitive Aging Study [272] *Initial population*: 170 adults (120 female) aged between 46-77 years; *Focus*: understanding of relationship between brain anatomy, cognitive function, and genetics; *Imaging*: MRI including DTI and rs-fMRI

German National Cohort [60] *Initial population*: ≈200,000 adults aged between 20-69 years; *Focus*: understanding the causes of widespread diseases such as cancer, dementia, diabetes, and cardiovascular diseases; *Imaging*: US including 3D-echocardiography and brain, cardiovascular, and abdominal MRI of ≈30,000 participants

Honolulu-Asia Aging Study [58] *Initial population*: 3,734 Japanese-American men *Focus*: dementia screening of an originally dementia-free population; *Imaging*: brain MRI [93]

The cumulated publications of Part III are largely based on the SHiP data pool (Chapters 8-9) but also on data from the Norwegian Cognitive Aging Study (Chapter 10).

1.5.2 Specifics of Population Study Data and their Acquisition

All epidemiological instruments need to be applied in a highly standardized manner in order to guaranty comparability of the data across subjects as well as over time. *Standardization* is a crucial aspect already in the planning phase of a study. For instance, multiple physicians and nurses are trained to perform a medical examination and the interpretation of results, e.g., the reading of imaging data, exactly in the same way to minimize inter-observer variability. Vendors of medical scanners are indentured to refrain any software or hardware updates being usually applied over the lifespan of a device. This is to ensure the comparability of image data acquired at different waves of a study. Standardization of the imaging including calibration of the scanner becomes even more difficult if multiple devices at different institutions are employed. For instance in the German National Cohort, imaging is distributed over five cities [60].

Epidemiological data are very complex and heterogeneous. Image data and derived data such as segmentation masks even increase this complexity. The collected variables relate to physical measures, such as blood pressure, heart rate or plasma glucose concentration, aspects of lifestyle, such as drinking and smoking habits, socio-demographic factors, such as education level and occupation, and to visual, hearing, and cognitive function. Some variables have been collected only at later waves due to technical advances or have been removed after a reevaluation of their reliability. Other variables are only available for a subpopulation, e.g., childbirth status and menstrual period of women, or are based on follow-up questions, e.g., the number of cigarettes smoked per day is only recorded for individuals who smoke. The variables differ with respect to their data type: nominal, ordinal, or quantitative. Examples are occupation, income level, and body height. Dichotomous variables represent a subtype of nominal variables and assume only binary values. Examples are gender and questions with the only possible responses being “yes” or “no”. The data type of a variable determines the set of appropriate visualization techniques. Sophisticated approaches are necessary for the joint analysis and visualization of variables with different types.

The analysis of epidemiological studies is hampered by unreliable and missing data. A classical example of the former are self-reports of drinking behavior and sexual practices, which tend to be biased towards socially accepted answers. A good questionnaire design incorporates redundant questions to counteract this effect. Study data may be incomplete due to individuals who drop out of the study since they pass away or move. Further reasons may be the denial of answering particular interview questions or the ineligibility for a certain imaging procedure, e.g., MRI due to claustrophobia. Missing data are a very critical issue and several guidelines such as the Strengthening the Reporting of Observational Studies in Epidemiology (STROBE) statement

exist recommending that cohort studies report on the amount of missing data, the reasons, and the solution strategies [257]. Missing data can be handled up to a certain degree. Donders et al. [42] demonstrate that straightforward techniques such as *complete case analysis* and *overall mean imputation* introduce bias and discuss more sophisticated *imputation* strategies.

The handling of missing data is part of data cleansing and can hence be considered a data transformation step in the visual analytics process (Fig. 1.1). Another common step, the conversion of raw data into a logical structure, is not necessary in case of epidemiological cohort studies due to the highly standardized and formalized data acquisition and subsequent data preparation and quality control by experts. The data are comprehensively described by a *data dictionary* precisely defining all variables and their value ranges. Visual analytics approaches still need to consider outliers, potentially unreliable/uncertain data, and imputed data values.

1.5.3 Standard Epidemiological Data Analysis Workflow

The following description of the standard workflow in epidemiological data analysis is based on discussions with collaborators and work observations by Thew et al. [236]. A schematic overview of the workflow is provided in Figure 1(a) on page 152 of Chapter 9.

The workflow is driven by an a priori hypothesis, which often evolved from observations of physicians in their clinical routine. For instance, the onset of a particular disease may be seen more frequently in people with a specific lifestyle. In the course of verifying the hypothesis, either a new population study is initiated or a large-scale study is analyzed retrospectively given that the prevalence of the disease is high enough in the studied population. A list of variables potentially related to the hypothesis is then extracted from the study data. Next up, *subgroups* of individuals are defined by *categorizing* variables, e.g., age is divided into 20 years bins in order to determine a per-group risk of falling ill with the disease. Larger groups increase the statistical significance of the results as compared to determining risks for specific ages. Categorization yields typical epidemiological statements such as people between 20 and 40 years have a 30% lower risk of falling ill with the disease than people between 40 and 60 years.

From the list of variables, *confounders* must be identified and later considered when interpreting the results. For instance, the effects of alcohol consumption on a certain health condition may be overestimated when the potentially confounding variables smoking and diet are neglected. Without the control of confounders, wrong causal relationships may be derived from the associations of variables. As an example, a causal relationship between gingerbread consumption and leg fractures may be derived from their positive correlation, e.g., gingerbread causes leg fractures, when ignoring the confounding variable season. During winter time, gingerbread is included in the range of products offered by food stores and more people go skiing yielding an increased number of leg fractures.

In the final, most important step, associations between the variables and the investigated disease are determined using statistical methods such as regression analysis. The associations indicate whether variables/factors influence the disease thereby confirming or disproving the given hypothesis. The statistical significance of this influence (*p-value*) and the *relative risk* of falling ill with the disease relative to a factor represent important outcomes of epidemiological research. Visualization plays a minor role in the standard epidemiological workflow and is restricted to conventional statistical plots such as histograms, box plots, and scatter plots. More sophisticated graphical representations such a *Kaplan-Meier curves* and *odds ratio plots* are mainly employed in disseminating epidemiological research results (Fig. 1-2, p. 125, Chap. 7).

Incorporating the analysis of medical image data of a population into the workflow poses additional challenges [238]. The images itself are too complex to be fed directly into the standard

analysis pipeline which is tailored to variables aggregated in data tables. However, epidemiologists are not interested in the images themselves but in deriving numerical variables characterizing the image data content and integrating these into the table for further processing. Example variables are measures related to organ shape and size, angle, location, and neighborhood of structures, and tissue density. The computation of these measures for data of a population requires robust, fully automatic segmentation and quantification algorithms. Even semi-automatic approaches are unsuitable and approximations instead of fine-granular segmentations are accepted. Rak et al. [187] describe a rough detection of the lumbar spine by means of a hierarchical finite element model. While the individual shape of the vertebrae is not captured by the model, their position and orientation as well as the bending of the spinal canal are represented. The model-based detection approach and its applications to cohort study data are detailed in Chapter 7 [181], pp. 129-132 and Chapter 8 [112], respectively.

As part of this thesis, an automatic approach to sampling the thickness along a hippocampal subfield, the Stratum Radiatum/Lacunosum-Moleculare (SRLM), was developed (Chap. 11 [170]). The approach was evaluated based on a small population ($n = 27$) for which the hippocampus and its subfields were manually segmented in ultra-high field 7-Tesla MR images. However, an automatic segmentation of large-scale population study data seems feasible [91] and could be coupled with the automatic thickness sampling in the future.

Reproducibility and statistical soundness are crucial aspects of the epidemiological workflow. Running the workflow again on the same data must yield the same results and running it on data of different study waves must yield contrastable results. All steps of the analysis must be monitored and reported such that other epidemiologists can run the same workflow on a new cohort study and compare their results to previous work. A crucial aspect of statistical soundness is the proper consideration of the underlying data distribution. Several statistical methods are only applicable to normally distributed data, which are less frequent than generally assumed. Different types of statistical tests, such as *normality tests*, exist for verifying assumptions about the underlying distribution.

1.5.4 Visual Analytics: Potential and Approaches

Potential Visual analytics bears a great potential in evaluating cohort study data by providing a methodology for their visual exploration and automated analysis. Klemm et al. [113] propose the integration of an interactive visual analysis step into the standard epidemiological workflow (Fig. 1(b), p. 152, Chap. 9). In their coordinated multiple views system, variables can be added or removed from the analysis via drag and drop and the definition of subgroups can be adjusted interactively, both triggering an update of the statistical analysis (Fig. 3, p. 155, Chap. 9). Very recently, Krause et al. [119] presented an approach to subgroup definition based on temporal patterns of interest and pattern matching.

The data mining component of the visual analytics process facilitates a data-driven definition of subgroups revealing more complex, hidden patterns. For instance, individuals may be clustered according to variables derived from image data or socio-demographic and lifestyle-related variables [109; 113]. The resulting clusters can then be interrelated with other variables. Klemm et al. [113] clustered a SHiP subcohort regarding their spinal canal bending and related each cluster, e.g., to self-reported back pain and the level of physically heavy work (Sec. 6, pp. 157-159, Chap. 9).

While the traditional epidemiological workflow requires an a priori hypothesis, visual analytics supports hypothesis generation as part of an extended workflow. Turkay et al. [242] investigated data of the Norwegian Cognitive Aging Study employing their *dual analysis* coordinated multiple views framework. They demonstrated the generation of new hypotheses in pair analysis sessions

with domain experts. Bernard et al. [10] recently presented a visual-interactive system for the analysis of prostate cancer cohorts and demonstrated its integration in an epidemiological workflow for hypothesis testing and generation. Since reproducibility plays a crucial role in epidemiological research, monitoring and protocoling the workflow is of utter importance. A suitable approach based on a filmstrip metaphor has been proposed in the context of coordinated multiple views systems by van den Elzen et al. [246]. It supports capturing all intermediate steps of an analysis process which can then be revisited and adapted using a rewind and fast-forward navigation mechanism.

The potential of visual analytics in investigating cohort study data was realized by Gotz et al. who hold a patent on the "Iterative Refinement of Cohorts Using Visual Exploration and Data Analytics" [65]. The preamble of the pattern states: "A need exists for an integrated system that combines visual exploration and data analytics to interactively visualize and refine cohorts, request analytics on those cohorts, and make new discoveries." Several claims on such a system are formulated including the integration of methods for reducing a cohort using one or more visual filters, for visualizing a cohort using selected views, for expanding a cohort by one or more analytics, and for determining whether a cohort should be further modified based on, e.g., statistical measures.

Approaches Medical image data of a large-scale population provide a survey of the broad variability in vital organ anatomy and physiology and allow for an improved characterization of health and disease as well as a differentiation between physical effects of normal aging and pathologies. Klemm et al. [112; 113] investigated the SHiP data with a focus on lower back pain (Chap. 8-9). They were interested in studying the shape variability of the lumbar spine, the relation of the variability to socio-demographic and biomedical factors, and the differentiation between pathologic and aging-related spine deformations.

In order to analyze lumbar spine shape, a detection algorithm based on a hierarchical finite element model was applied to MRI data of the spine scanning protocol [187]. The detection resulted in the centerline of the spinal canal as well as position and orientation of the vertebrae (Fig. 1, p. 145, Chap. 8). The subsequent analysis focused on the spinal canal bending, i.e. the curvature of the centerline [112]. It is related to typical deformations such as *lumbar hyperlordosis* (abnormal inward curvature in the sagittal plane), *hyperkyphosis* (abnormal outward curvature in the sagittal plane), and *scoliosis* (abnormal S- or C-shaped curvature in the coronal plane). The bending was approximated by the mean curvature and related to the SHiP variables body height, gender, age, and weight. The reproduced textbook knowledge and a new potentially interesting finding are discussed in Chapter 8, p. 149.

Later, Klemm et al. [110] computed associations between further image-derived variables, such as spinal canal torsion and curvature angle along each projection axis, and non-image variables, e.g., pain indicators [110]. They trained decision tree classifiers to predict pain indicators based on image-derived variables and compared the predictive power of different types of decision trees by means of generalized pairs plots [45]. The results indicate that the tested image-derived variables are not sufficient for characterizing lower back pain in the SHiP data. Better predictive power may be achieved when incorporating information about the vertebrae, such as position, orientation, shape, and distance between neighboring vertebrae. An interactive approach to decision tree construction and analysis could further improve the prediction quality [245].

Klemm et al. [112] also investigated the grouping of SHiP participants according to spinal canal bending (Chap. 8). They applied an agglomerative hierarchical clustering approach coupled with a method for automatically computing the number of clusters to the centerline geometry [200] (Fig. 3, p. 147, Chap. 8). One representative centerline was then computed per cluster and all *representatives* were aggregated in a visual summary (Fig. 4, p. 148, Chap. 8). The summary adheres to radiological viewing conventions by showing a sagittal view of the representatives. While this

facilitates an easy assessment of bending within the sagittal plane (lordosis and kyphosis), special depth cues were integrated to convey also bending orthogonal to it (scoliosis). Inspired by orthopedics, manual measurement tools such as ruler and goniometer were integrated (Fig. 5, p. 149, Chap. 8). Automatic measurements of characteristic distances and angles would be desirable and could be integrated as derived variables in the analysis process.

In a follow-up work, Klemm et al. [113] developed a coordinated multiple views system and integrated the shape analysis and the centerline clustering (Chap. 9). By using the cluster ID as a new variable, the association between a specific bending of the spinal canal and socio-demographic, lifestyle- or health-related variables could be investigated. For instance, does spine shape correlate with age and lower back pain? The results of this hypothesis-driven analysis as well as the generation of hypotheses in pair analytics sessions with a radiologist are described in Chapter 9, pp. 158-159. Klemm et al. [109] extended their system by clustering techniques for mixed numeric and categorical variables. They investigated three clustering algorithms and took special care of missing data.

The coordinated multiple views system comprises several statistical plots for numerical variables, such as histogram, scatter plot, and parallel coordinates, and a mosaic plot for categorical variables [73]. Klemm et al. [113] augmented the histogram and mosaic plot by interactive *small multiples* of average lumbar spine shape. For instance in a histogram of body height, the mean shape per bin is displayed and colored according to its deviation from the mean of the entire population (Fig. 3(b), p. 155, Chap. 9). Each shape can be rotated and zoomed in on causing a synchronized update of the other shapes.

Epidemiologists are very much interested in correlations between variables. Klemm et al. [113] integrated a *contingency view* comprehensively displaying the correlation strengths of all pairs of study variables as a matrix (Fig. 4, p. 156, Chap. 9). Specific measures for correlations between variables of mixed types (numerical and categorical) were employed. The contingency view inspired the 3D regression heat map view, which was very recently added to the system by Klemm et al. [111]. It shows the results of a complex regression analysis as a *regression cube*. The basic idea is that the epidemiologist specifies a dependent variable, usually a pathology. Then, the correlations between all possible combinations of two or three independent variables and the dependent variable are computed. Investigating the regression cube by means of a movable plane assist in understanding how the dependent variable is effected when one independent variable is varied (plane position) while the others are fixed.

Fekete [50] emphasized the importance of a data model fulfilling the specific requirements of visual analytics, in particular very fast responses to queries and seamless integration of heterogeneous data. A model based on *data-cubes* was proposed by Angelelli et al. [4] and integrated in a prototype for the analysis of the Norwegian Cognitive Aging Study (Chap. 10). The model and an aggregation engine allow for a seamless integration of data entities with only partially overlapping dimensions (Fig. 1, p. 162, Chap. 10). For instance, fiber tracts and cortical regions share the IDs of subject and study wave but fractional anisotropy (FA) is only available for fiber tracts and thickness measures have only been computed for cortical regions. Taking a scatter plot as an example, how can FA and cortical thickness be compared across the subjects of a study? The proposed solution employs a flexible and fast aggregation of these dimensions based on statistical estimators. As a simple example, the mean of FA and cortical thickness values can be computed across all fibers and cortical regions, respectively. The resulting values for each subject can then be opposed in the scatter plot. A more fine granular analysis can be achieved by considering the cortical region through which a fiber segment is passing, which is also stored in the data.

The data cube model and the aggregation engine are implemented as part of a prototypical coordinated multiple views system. A 3D view shows the brain regions and fiber tracts of a tem-

plate brain (Fig. 2(g), p. 164, Chap. 10). The fibers have been grouped using a spectral clustering approach [17] and a representative fiber is shown per cluster to reduce visual clutter. Once interesting variable values have been brushed in one of the other views, the associated fibers and brain regions are highlighted and values averaged over the selection are encoded by color or saturation. Two case studies demonstrate the potential of the prototype in verifying as well as in generating hypotheses (Chap. 10, pp. 165-167).

1.6 Further Biomedical Data

Visual analytics approaches for the investigation of many other types of data acquired within further medical and biological applications have been presented in the past. A complete overview is out of the scope of this postdoctoral thesis. In this section, short excursions into the visual analytics of perfusion data (Sec. 1.6.1), electronic health records (Sec. 1.6.2), and *omics data* (Sec. 1.6.3) are given. The former provides context for the cumulated publications of Part IV.

1.6.1 Perfusion Data

Perfusion data represent a specific type of dynamic image data characterizing the regional blood flow in tissue. They facilitate early-stage detection and improved differentiability of diseases and are acquired for the diagnosis of breast tumors, ischemic stroke, renovascular diseases, and early detection and diagnosis of Coronary Heart Disease (CHD). Perfusion data characterize microcirculation through tissue capillaries which is in contrast to simulated cerebral and measured cardiac hemodynamic data conveying the macrocirculation of blood through larger vessels (Sec. 1.3.2 and 1.3.6). Since the average diameter of capillaries (6 µm) is below the resolution of today's medical imaging devices, macroscopic parameters characterizing the microcirculation are derived from the measured perfusion data. In perfusion imaging, either an injected contrast agent is traced over time or magnetically labeled arterial blood water protons are employed as an endogenous tracer in Arterial Spin Labeling (ASL). In time-dependent data from contrast-enhanced perfusion imaging, plotting the recorded signal intensity over time yields a *time-intensity curve* that specifies per voxel wash-in and wash-out of the contrast agent. Perfusion parameters, such as area under the curve, peak enhancement, and time to this peak, characterize perfusion and are derived from theses curves. The original spatio-temporal data and up to seven derived parameter volumes must be integrated for diagnosis.

Oeltze gives a comprehensive introduction to perfusion imaging and data processing in the context of ischemic stroke and breast tumor diagnosis [164]. Preim et al. [182] and Oeltze [164] survey exploration, visualization, and visual analytics approaches for the investigation of perfusion data. Pioneering work in terms of visual analytics has been accomplished by Grzesik et al. [69] and Subramanian et al. [232] who both presented a coordinated multiple views system for the interactive visual analysis of cerebral and breast tumor perfusion, respectively. Later, Oeltze et al. [165] coupled a dimension reduction scheme with the `SimVis` framework described in Section 1.2.3 to reduce the complexity of the perfusion parameter space. Glaßer et al. [64] proposed a clustering-based visual analytics approach to investigating heterogeneity of breast tumor tissue.

In recent work, Glaßer et al. [63] presented a coordinated multiple views system for the investigation of longitudinal glioma perfusion studies (Chap. 12). A glioma is a malignant type of brain tumor that arises from glial cells. Longitudinal studies monitor progression of the glioma in case its removal would be associated with a considerable risk for the patient. Grading the tumor and early identification of a potential transformation from low grade to high grade play an important

role in treatment planning. Glaßer et al. introduced a pipeline for the co-registration of perfusion data from different time steps of the study and dedicated views for monitoring the temporal evolution of perfusion parameters and tumor growth (Fig. 3, p. 184, and Fig. 5, p. 185, Chap. 12).

Chalopin et al. [28] employed the `SimVis` framework for evaluating intraoperative ultrasound (iUS) perfusion imaging in brain tumor surgery (Chap. 13). iUS perfusion imaging is a promising tool in tumor resection control since it enables a more accurate depiction of the tumor border as compared to B-mode ultrasound imaging. For evaluation, iUS perfusion data was co-registered to preoperative Magnet Resonance Imaging (MRI) perfusion data. Then, perfusion parameters were derived from both modalities and concurrently investigated in statistical plots as well as in the spatial context of the tumor tissue (Fig. 2, p. 194, Chap. 13).

Raidou et al. [185] introduced a novel visual representation for comparing different pharmakokinetic models of tumor tissue perfusion. These models are typically fit to time-intensity curves reconstructed from Dynamic Contrast Enhanced (DCE) Magnetic Resonance Imaging (MRI) data. Multiple parameters of the fitted model then characterize the tumor perfusion. To study the variability of these parameters across the different models, a novel plot integrating parallel coordinates with cobweb charts has been designed and linked to 2D and 3D physical views showing the image data.

In very recent work, Raidou et al. [186] presented a visual analytics tool for the exploration of tumor tissue characterization. A 2D embedding view showing the dimensionality reduced parameter space of a fitted pharmakokinetic model constitutes the central component of the tool. It facilitates the visual separation of tissue clusters with similar perfusion characteristics. Clusters can be brushed and further inspected in a linked physical, multiple statistical, and a tailor-made cluster analysis view.

1.6.2 Electronic Health Records

Electronic health records (EHR) stored in hospital databases detail patient-individual histories of diagnosis and treatment. They are mainly compiled for patient care, hospital process optimization, and quality control in health care. Examples of stored information are medical examinations and tests, monitored vital signs, e.g., blood pressure, body temperature, and heart rate, medical image data, diagnosis results, drug prescriptions, and therapeutic measures. Physicians can benefit from this wealth of information in daily decision making. With the advent of digitized health records, they became an attractive resource for clinical researchers searching for interesting patterns, e.g., complications after interventions correlated with specific prior symptoms. Just like population study data discussed in Section 1.5], EHR data bear a great potential for epidemiologists in investigating specific diseases or health-related effects and risks. However, a combination of interaction, exploration, visualization, and analysis techniques is required to gain insights from these complex data thereby supporting physicians, clinical researchers, and epidemiologists. A multitude of visual analytics solutions for EHR data has been presented in the past [188; 263]. The workshop on "Visual Analytics in Healthcare" annually provides a platform for the most recent work in the field [87].

Rind et al. [188] surveyed visualization systems for exploring and querying EHR data. They compared 14 systems in detail regarding the covered types of data, the support for multivariate analysis, the number of EHRs that can be processed (one or many), and user intents. Additional 32 systems were described more compactly. Most systems focus on EHR data with no inherent spatialization and solely apply techniques from information visualization. While medical image data is stored for individual patients in an EHR database, their detailed visualization is of minor interest in analyzing the database. Instead, clinicians may have derived quantitative attributes

such as volumes, lengths, and distances from the image data to assess the severity of a disease or the eligibility of a therapeutic option. These attributes are stored in the database and can be incorporated into the analysis. This situation is identical to the investigation of epidemiological population study data 1.5.3. Rind et al. conclude their survey by recommendations and future research directions for developing EHR systems.

West et al. [263] very recently provided a systematic review of visualization approaches reported for electronic health record data between 1996 and 2013. The focus of early approaches was on the visualization and analysis of the complex EHR data of a single patient [178]. This was later extended to multiple patients and larger numbers of events [258]. The most common interaction techniques implemented in the reviewed systems are filtering, scaling, and zooming and the most frequently used visual attributes are color and density. The most mature system for analyzing EHR data is `LifeLines`[178] with its extensions and applications `LifeLines 2` [258], `LifeFlow` [267], and `EventFlow` [150]. West et al. conclude their review by challenges that will drive future EHR-related visualization research.

The same interaction, exploration, visualization, and analysis techniques have been employed in investigating both, EHR as well as population study data. Most existing systems are realized as coordinated multiple views integrating data mining techniques for clustering, pattern detection, and dimension reduction. However, there are two major differences that must be considered in designing an analysis approach tailored to one or the other type of data. First, in contrast to population study data, EHR data are not acquired in a highly standardized manner within the scope of a specific study (Sec. 1.5.2). They comprise redundant, irrelevant, as well as subjective measures challenging users in synthesizing information [21]. Additional data transformation steps such as data cleansing may be necessary, e.g., to establish comparability between patients or between results of one patient over time. If this is possible at all, at least such steps contribute to uncertainty. Second, events in population studies are inherently synchronized across subjects and correspond to the individual data acquisition waves. Events in EHR data are not necessarily predictable, e.g., a heart attack, and must be synchronized between patients to facilitate a comparison, e.g., of symptoms or drug intake within the immediate period before the heart attack [258].

1.6.3 Omics Data

The term *omics* refers to a collection of disciplines in modern systems biology that study the sum of similar individual elements in a particular biological sample. Example disciplines investigate the sum of all genes (*genome*) in *genomics*, the sum of all messenger ribonucleic acid (mRNA) molecules (*transcriptome*) in *transcriptomics*, the sum of all proteins (*proteome*) in *proteomics*, and the sum of all metabolites (*metabolome*) in *metabolomics* [88]. In *toponomics*, the entirety of spatial protein networks (*toponome*) in a sample is investigated (Sec. 1.4). A broad variety of applications for omics data exists in biomedical, agricultural, and environmental sciences ranging from biomarkers for disease and drug design over crop plant improvement to the assessment of environmental pollutants [8]. Dedicated omics technology is utilized to acquire the data [88]. In genomics and transcriptomics, *microarray* technology is employed to simultaneously measure the expression of thousands of genes. The measurement process results in image data that encodes the expression level of each gene. Expression levels are then analyzed to determine differences in DNA sequence between individuals and to detect abnormalities, e.g., chromosomal insertions and deletions [88]. The most common method for the detection of molecules in both, proteomics and metabolomics, is *mass spectrometry*. It is capable of measuring hundreds to thousands of molecules in a single experiment. The resulting data are applied for identifying and quantifying proteins and metabolites, for characterizing protein structure, determining protein-protein as

well as protein-metabolite interactions, and in mapping metabolites and their associated chemical reactions to metabolic pathways.

Omics data are huge and their analysis requires statistical as well as data mining methods, e.g., for identifying outliers, clusters, and patterns [88]. The likelihood of false positives in the measured data is highly demanding data validation. Mining and validation should both facilitate the incorporation of expert knowledge during the data analysis process. Various omics data have no spatial reference. Information visualization techniques are hence particularly suited to depict these data. A plethora of visualization tools exists [57; 161]. Many visual analytics approaches to investigating omics data have been proposed [231; 241]. A part of the most recent related work is presented at the annual "Symposium on Biological Data Visualization" [85].

Turkey et al. [241] surveyed approaches integrating automated data analysis methods and interactive visualizations for the investigation of biomedical data. They presented a categorization with respect to the level of integration (visualization as presentation, semi-interactive methods, tight integration) and the analytical task (summarizing information/groups, classification/dependence, prediction). Many of the surveyed approaches were dedicated to the visual analytics of omics data. Sturm et al. [231] very recently extended the survey of Turkey et al. [241] by including more such approaches. Furthermore, they introduced two additional categorization dimensions: visualization technique (geometric, table-based, icon/glyph-based, pixel-based, graph) and data type (genomics, proteomics, metabolomics, text, graph, image, multivariate data). The most common visualization technique in genomics and transcriptomics is the heat map encoding up- and down-regulated gene expressions by intensities of red and green. Graph representations are the main vehicle in proteomics and metabolomics for depicting protein interaction networks and metabolic pathways. Sturm at al. also elaborate on data mining approaches for different kinds of omics data. They conclude their survey by a description of open problems and future goals of systems biology such as the seamless integrated analysis from organs to molecules based on linking medical image and all types of omics data.

1.7 Summary and Discussion

In medicine and biology, a steadily growing data complexity often paralleling new developments in image and non-image data acquisition is observed. This complexity poses many challenges on data processing, visualization, and exploration. It renders the design of overview visualizations, which convey all interesting patterns contained in the data, impossible. Instead, automatic data analysis techniques must be tweaked based on expert knowledge and combined with interactive visualizations for the retrieval of such patterns. This approach is at the heart of the field visual analytics. In the visual analytics process, analysts derive knowledge from the raw data in feedback loops. They modify parameters of the automatic data analysis techniques and steer the analysis process based on the evaluation of intermediate results presented as interactive visualizations (Fig. 1.1). Software solutions implementing this process build upon techniques from multiple scientific disciplines cooperating in visual analytics such as visualization, data mining, and human-computer interaction (Fig. 1.2). The field of interactive visual analysis is tightly connected to visual analytics since their methodologies highly intersect (Sec. 1.2.3). It stresses the importance of interaction and human pattern recognition, focuses on data with an inherent spatialization, i.e. measured and simulated data, and considers automated data analysis rather optional.

In the realm of visual analytics, the postdoctoral thesis at hand contributes to the investigation of complex data from clinical medicine, biology, and epidemiology. The visual summary in Table 1.1 lists the main contributions per paper of the cumulative Parts I-IV as well as the subset of

scientific disciplines co-operating in visual analytics that these contributions can be attributed to (columns 2-3). The introductory and survey papers in Chapter 7 and Part V are not included in the visual summary. The thesis at hand focuses on and contributes to:

- the combination of data mining and interactive visualizations for the visual analysis of simulated cerebrovascular hemodynamic data (Part I introduced in Section 1.3 and summarized by rows 1-2 of Table 1.1),
- the interactive visual analysis of toponome data (Part II introduced in Section 1.4 and summarized by rows 3-5 of Table 1.1),
- the joint visual analytics of image and non-image epidemiological population study data (Part III introduced in Section 1.5 and summarized by rows 6-9 of Table 1.1), and
- the interactive visual analysis of perfusion data (Part IV introduced in Section 1.6.1 and summarized by rows 10-11 of Table 1.1).

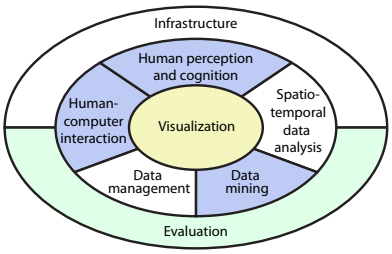
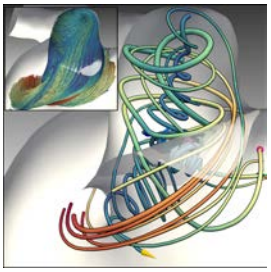
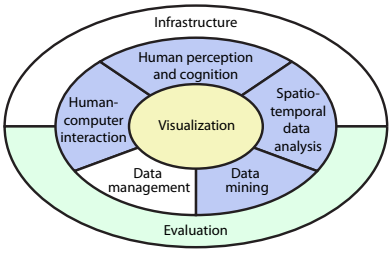
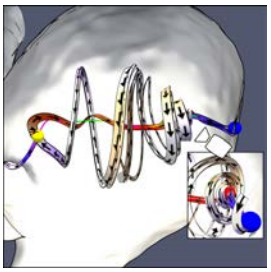
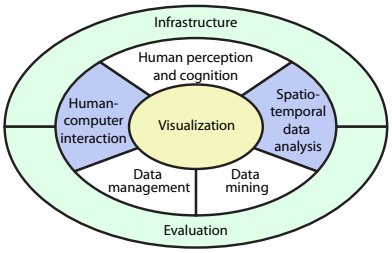
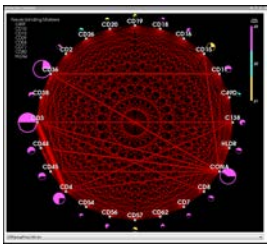
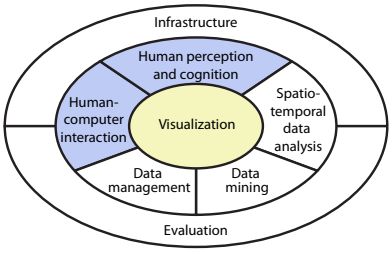
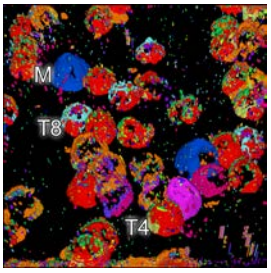
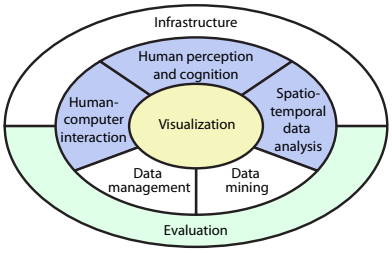
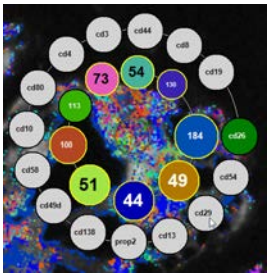
Short excursions on visual analytics of further types of medical and biological data from other applications round out the thesis (Sec. 1.6.2, 1.6.3).

Simulated Cerebrovascular Hemodynamic Data Patient-individual Computational Fluid Dynamics (CFD) simulations of cerebrovascular hemodynamics are conducted to better understand the causes of cerebral aneurysm initiation, progression, and rupture, to eventually define a rupture risk score for clinical routine use, and to predict the outcome of endovascular interventions. The complexity of the blood flow pattern and its stability during the cardiac cycle were both found to be related to the risk of rupture. In treatment planning, they should be compared across different configurations of a virtual intervention in order to determine the optimal configuration. The blood flow pattern is commonly visualized by a dense set of integral curves resulting in so-called spaghetti plots suffering from considerable visual clutter.

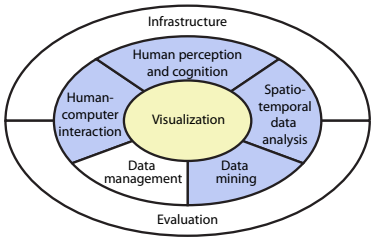
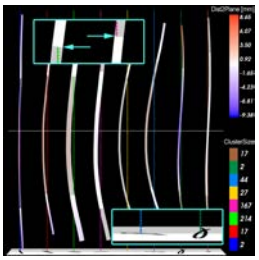
In Chapter 2, the reduction of visual clutter based on clustering streamlines followed by a computation of cluster representatives and their aggregation in a visual summary of blood flow was proposed (Tab. 1.1, row 1). Different clustering algorithms coupled with techniques for estimating the number of clusters were quantitatively evaluated. An expert evaluation of the visual summaries created by the best performing algorithm was conducted and the usefulness of the summaries in comparing different configurations of virtual stenting was assessed. Triggered by one of these comparisons, a hypothesis regarding thrombosis development favored by a specific blood flow pattern was generated. In Chapter 3, the focus was on the uncluttered visualization of vortical flow, in particular, of embedded vortical layers forming around saddle-node bifurcations (Tab. 1.1, row 2). Flow complexity and stability are strongly related to the existence and number of vortices and their persistence over the cardiac cycle. A pipeline for the automatic clustering-based detection and visualization of (embedded) vortices was presented and evaluated in 17 cases studies.

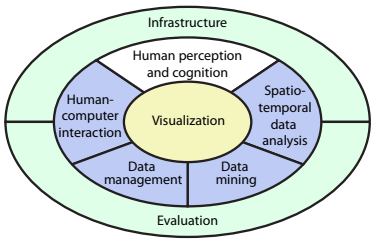
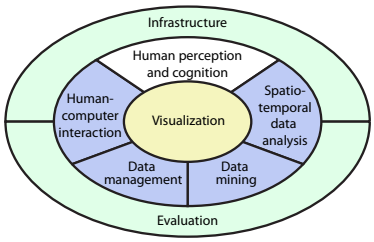
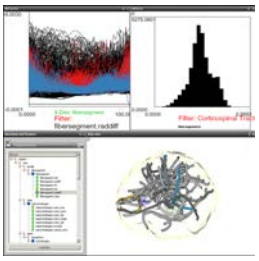
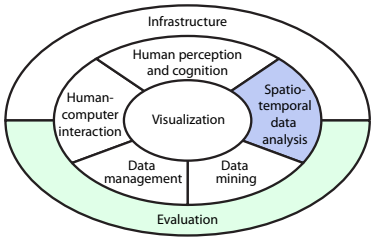
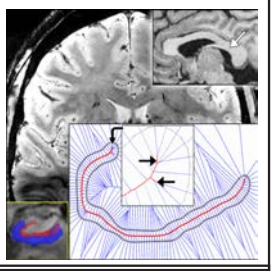
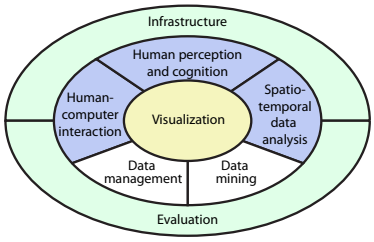
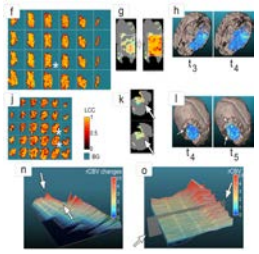
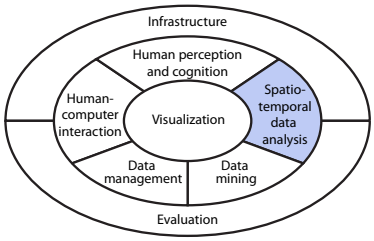
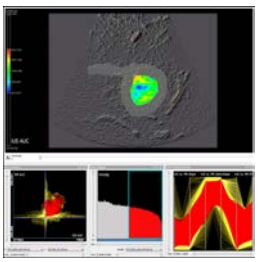
Both chapters restrict the visualization and analysis to a single point in time of the cardiac cycle. While this facilitates an assessment of flow complexity at this particular point, it does not indicate flow stability. In fact, embedded vortical flow is known to collapse and reappear over the cardiac cycle. The investigation of flow stability requires an evaluation of the overall unsteady simulation data and dedicated techniques for tracking flow features over time [203]. The presented cluster representatives very well indicate the overall shape of their cluster but fail to convey its spatial extent. Hence, exploration facilities are offered for displaying the cluster's original streamlines on demand. Depth cues should be added to the representatives, e.g., indicating their distance to the

Table 1.1: Main contributions per paper of the cumulative Parts I-IV. The subsets of scientific disciplines/building blocks co-operating in visual analytics that these contributions can be attributed to are color-coded in *column 2*. Visualization is at the heart of almost each paper and it is being combined with techniques from the surrounding disciplines. Some papers contribute or extend a software infrastructure and most comprise an evaluation of the proposed approach. The individual disciplines and their cooperation in visual analytics are detailed in Section 1.2.2. Brief listings of the main contributions of each paper are given in *column 3*. The visual reminders in *column 4* are meant to strengthen recognition of the papers. This table continues on the next page.

Paper	VA Building Blocks	Main Contributions	Visual Reminder
Oeltze et al. [168] Part I, Chapter 2		• Quantitative evaluation of three streamline clustering techniques • Custom cluster representatives • Visual summaries of blood flow • Expert evaluation of summaries • Application and evaluation in the context of virtual stenting	
Oeltze-Jafra et al. [171] Part I, Chapter 3		• Clustering-based pipeline for visual analysis of (embedded) vortical flow • Vortex core line enhancement • Custom cluster representatives • Visual summaries of vortical flow • Evaluation of summaries • Smart visibility technique	
Oeltze et al. [166] Part II, Chapter 4		• Interactive visual analysis approach for toponome data • Extension of SimVis framework by graph view depicting protein co-mapping • Glyphs encode protein quantities • Brushing facilities for graph view • Two case studies	
Oeltze et al. [167] Part II, Chapter 5		• Volume rendering approach for 3D toponome data • Perceptually-optimized coloring of unique protein patterns • Technique for brushing protein patterns of interest in 3D • Peeling off clusters of protein patterns to resolve occlusions	
Oeltze-Jafra et al. [172] Part II, Chapter 6		• Approach for in-place annotation of protein patterns in 2D views of toponome data • Interactive annotations • Management view for logging and comparing annotations • Two cases studies	

Paper	VA Building Blocks	Main Contributions	Visual Reminder

Klemm et al. [112] Part III, Chapter 8		• Clustering-based approach to studying lumbar spine canal variability in a cohort
		• Visualization of cluster representatives augmented with
		  • visual hints conveying 3D curvature in 2D views and
		  • measurement techniques
			

Klemm et al. [113] Part III, Chapter 9		• Web-based coordinated multiple views system for the integrated analysis of image and non-image cohort study data • Augmentation of InfoVis views with 3D shape renderings • Clustering for subgroup definition • Two case studies	
Angelelli et al. [4] Part III, Chapter 10		• Data model for interactive visual analysis of heterogeneous cohort study data • Model implementation in coordinated multiple views prototype • Clustering-based atlas view of fiber tracts and brain regions • Two case studies and evaluation	
Oeltze et al. [170] Part III, Chapter 11		• Semi-automatic approach to measure thickness of the Stratum Radiatum/Lacunosum-Moleculare in coronal sections from ultra-high field 7-Tesla MRI • Evaluation and statistical analysis based on a small cohort	
Glaßer et al. [63] Part IV, Chapter 12		• Pipeline for co-registration of time steps in longitudinal brain tumor perfusion studies • Coodinated multiple views system for concurrently monitoring temporal evolution of perfusion parameters and tumor growth • Expert evaluation	
Chalopin et al. [28] Part IV, Chapter 13		• Pipeline for co-registration of perfusion data from intraoperative Ultrasound and preoperative MR imaging • Interactive visual analysis approach for comparison of perfusion parameters derived from both modalities	

aneurysm wall, to improve the perception of spatial relationships. The investigation of vortical flow strongly depends on the approach used for vortex extraction. No algorithm or vortex criterion are known to guarantee vortex detection [13]. The existing ones generate different results which should be compared.

The CFD simulation of cerebrovascular hemodynamics is a pure research tool. Most clinicians are neither familiar with these simulations nor with the techniques for visualizing their results. This situation may change if the pipeline from patient-specific image data to simulated hemodynamics is largely automated and simulations were shown to significantly contribute to a reliable risk of rupture prediction. Ideally, the simulation results in a score based on quantitative and qualitative hemodynamic parameters. This score is then used in clinical routine diagnosis for aneurysm grading together with other factors such as aneurysm size, wall thickness, and signs of inflammation. The road to using CFD simulations in clinical treatment planing is a bit more advanced. Virtual stenting was shown to assist decision making within an acceptable time frame before the real intervention. Clinicians can test different stent types and positions for the stent and receive visual feedback potentially including visualizations of the blood flow pattern and its changes after the virtual intervention. In cardiac research, 4D Phase-Contrast (4D-PC) Magnetic Resonance Imaging (MRI) is applied to measure primarily aortic hemodynamics (Sec. 1.3.6). The clustering-based reduction of visual clutter presented in Chapter 2 can be readily transfered to pathline visualizations of these data (Fig. 1.6).

Toponome Data of Cells and Tissues Toponome data are acquired to identify protein networks that are characteristic for a certain disease and to pinpoint lead proteins within these networks, which may represent target molecules in drug design. In toponome imaging, multiple fluorescence-conjugated tags recognizing proteins are applied to a cell sample or tissue probe and the respective fluorescence response is recorded by a CCD camera attached to an epifluorescence microscope. The resulting dataset comprises either an image or a volume per tag depending on whether the acquisition was carried out in 2D or 3D. In a preprocessing step, the data is binarized resulting in a combinatorial binary code (protein pattern) per pixel or voxel. If multiple proteins are registered in a code, i.e. their entries are equal one, they co-map in space indicating protein-protein interactions. The unique binary codes in the data are referred to a Combinatorial Molecular Phenotypes (CMPs).

In Chapter 4, the graph-based interactive visual analysis of protein co-mapping was presented (Tab. 1.1, row 3). A graph represents the proteins as nodes and co-mapping pairs of proteins as edges. Glyphs attached to the nodes encode protein quantities and may hint on lead proteins. The graph is equipped with brushing facilities such that nodes of interest may be selected. The brushes can be combined by logical operators thereby defining a template CMP that can be matched with the data. The graph is integrated in a coordinated multiple views system. Previously, protein co-mapping had to be derived from a tabular view of the CMPs comprising tens of columns and hundreds to thousands of rows (Fig. 2(a), p. 105, Chap. 6).

In Chapter 5, a dedicated ray-casting approach for 3D toponome data was proposed (Tab. 1.1, row 4). Along each ray, it determines the CMP closest to the camera and assigns its precomputed unique identifier color to the pixel. In close-up views of the data, the set of precomputed colors can be temporarily replaced by a perceptually optimized set that is computed only for the visible CMPs. This increases the visual separability of cell or tissue parts with very similar identifier colors. The 3D visualization of toponome data was integrated in a coordinated multiple views system. Previously, 3D toponome data had to be processed slice-by-slice and 3D visualizations were crafted in a separate program. The 3D view was equipped with a brushing facility and an exploration technique for peeling off clusters of protein patterns to mitigate occlusion problems.

In Chapter 6, an interactive labeling approach for the in-place annotation of protein patterns in 2D views was presented (Tab. 1.1, row 5). A resizable and draggable focus regions is surrounded by dynamically updated symbols indicating the currently focused CMPs and their registered proteins. A management view logs annotations of interesting protein patterns and arranges them in a non-overlapping fashion to simplify their comparison. Two cases studies have demonstrated the usefulness of the labeling approach in detecting lead proteins. Previously, the CMPs in a focus region and their registered proteins had to be retrieved in a tedious and error-prone color-based mental matching of the 2D view and a tabular view listing the CMPs (Fig. 2, p. 105, Chap. 6).

The visual scalability of the graph view and the labeling approach is limited by the available screen space and minimum requirements on the visual separability of nodes and edges and the readability of symbol color and text, respectively. The perceptually optimized coloring of toponome data works well for smaller numbers of CMPs. Its view-dependence however, yields multiple colors for the same CMP hampering recognition during analysis. The interactive visual analysis of toponome data is a biomedical research endeavor. Case studies have shown that it can contribute to the disease-related generation of hypotheses regarding characteristic protein networks and contained lead proteins. Larger studies are necessary to verify such hypotheses. In the long run, toponome imaging may enter clinical routine and be used by histologists to compare a patient-individual toponome to a library of disease-specific ones.

Epidemiological Population Study Data Population studies are conducted in epidemiology to investigate the occurrence and distribution of health-related events in a group of individuals sharing a common characteristic. They are often carried out in waves over years (cohort study) and collect hundreds of socio-demographic, lifestyle-related, and health-related variables for thousands of individuals. Recently, such studies also include medical image data posing many new challenges on data analysis and visualization. Variables characterizing the phenomena of interest must be derived from the image data by largely automated and robust algorithms. The joint investigation of these variables and the non-image data reveals interrelations between measurable in vivo phenomena and for instance, age, gender, socio-demographic background, lifestyle, and health consciousness. The traditional epidemiological workflow is driven by an a priori hypothesis and focuses on a subset of the data pool related to this hypothesis. It is mainly based on the statistical analysis of data tables and employs non-interactive visualizations for the display of results. Interactive visual analysis and visual analytics can complement the workflow by offering methodologies for the generation of new hypotheses.

In Chapter 8, a clustering-based approach to studying lumbar spine canal variability in a cohort was proposed (Tab. 1.1, row 6). The variability is assessed based on the bending of the spinal canal. The spine is segmented in the image data of the cohort, the centerline of each spinal canal is determined, and the centerline geometries are clustered. Visualizations of cluster representatives are augmented with specific depth cues and measurement facilities. The clusters were related to non-image variables in order to evaluate whether a distinct lumbar spine bending correlates, e.g., with physically heavy work or self-reports of lower back pain.

In Chapter 9, a web-based coordinated multiple views system for the integrated analysis of image and non-image cohort study data was presented (Tab. 1.1, row 7). The system comprises different views for quantitative and categorical data variables. It includes the previously described clustering approach for a case study on lower back pain. A further approach to the integration of image data into the analysis is realized by augmenting the histogram and mosaic plot view by 3D shape renderings of the lumbar spine. A mean shape per subgroup of individuals is superimposed on its representing element of the plot, i.e. bar or mosaic piece. The system was evaluated based on two case studies demonstrating its potential in both, hypothesis verification as well as generation.

In Chapter 10, a data-cube model was presented for the interactive visual analysis of heterogeneous cohort study data (Tab. 1.1, row 8). The model and an aggregation engine allow for a seamless integration of data with only partially overlapping dimensions. For instance, ECG data acquired at the multiple waves of a cohort study may be stored in one data-cube with the dimensions subject ID, year, and ECG measurement values. The HDL cholesterol level may be stored in a similar manner in another data-cube. The dimensions of both cubes partially overlap (subject ID and year). However, the joint investigation of ECG data and HDL cholesterol level is hampered by the former being time-resolved and the latter being instantaneous. It is not possible for instance, to visually inspect the correlation of both at a selected study wave in a scatter plot. The ECG data is hence aggregated, e.g., through categorization into normal, atrial fibrillation, atrial flutter, and premature ventricular contraction, before it is correlated with the HDL cholesterol level, which may also be categorized into risky, borderline, and protective. The data-cube model was implemented in a coordinated multiple views prototype. A tailor-made, clustering-based atlas view of brain regions and fiber tracts was added to the prototype for investigating a cognitive aging study. An evaluation of the data model and the prototype based on two case studies showed their potential in hypothesis verification and generation.

In Chapter 11, a semi-automatic approach to measuring the thickness of the hippocampal subfield Stratum Radiatum/Lacunosum-Moleculare (SRLM) was described (Tab. 1.1, row 9). While the measurement is fully automatic, it requires a segmentation of the SRLM, which was so far obtained manually. The approach was evaluated based on a smaller cohort ($n = 27$). The results indicate that coupled with an automatic SRLM segmentation, the measurement can be applied in large-scale cohort studies of mild Alzheimer disease or Mild Cognitive Impairment. These studies should correlate thickness, e.g., with performance measures of recognition memory tests, since a reduced thickness was observed previously in individuals with earliest cognitive symptoms of Alzheimer disease.

Visual analytics and interactive visual analysis are relatively new concepts to epidemiologists. However, with initial guidance by a computer scientist, they can exploit their potential in verifying and generating hypotheses. The combination with statistics and statistical graphics turned out to greatly increase the acceptance of these concepts. This is comparable to coupling new, colorful 3D visualizations of medical data with the common gray-scale slice views in radiology. The definition of subgroups or subcohorts is a frequent task in epidemiological workflows. The per subgroup display of a mean shape or any other averaged phenomenon extracted from image data suits this group-wise analysis. The integration of data mining techniques such as clustering facilitates a data-driven definition of subgroups. However, the fine-tuning of clustering parameters represents a hurdle since epidemiologists are in general no data mining experts. Aspects of preprocessing, visualizing, and mining image data are often specific for an organ or structure of interest. General approaches are required to prevent the development of many highly specialized visual analytics solutions. A web-based visual analytics framework simplifies the collaboration with epidemiologists over longer distances and in particular, shortens evaluation and prototyping cycles.

Perfusion Data Perfusion data characterize the regional microcirculation of blood through tissue capillaries which is in contrast to simulated cerebral and measured cardiac hemodynamic data conveying the macrocirculation through larger vessels. They facilitate early-stage detection and improved differentiability of diseases and furthermore, serve disease monitoring and resection control in surgery. In perfusion imaging, often a contrast agent is injected and works as a tracer of the blood. The anatomical part of interest is then rapidly imaged at multiple time steps to capture the wash-in and wash-out of the contrast agent. A time-intensity curve per voxel of the image data describes this temporal behavior. Perfusion parameters, that are substitutes for physiological

parameters such as regional blood flow, regional blood volume and capillary permeability, are derived from these curves.

In Chapter 12, a pipeline for the co-registration of multiple time steps in longitudinal brain perfusion studies and a coordinated multiple views system for monitoring tissue perfusion in gliomas were presented (Tab. 1.1, row 10). Tailor-made views depict the temporal evolution of perfusion parameters and tumor growth. Based on a local correlation coefficient, tumor tissue heterogeneity can be assessed in an additional view. It was hypothesized that this view may help in guiding stereotactic biopsy. The system was evaluated based on four case studies.

In Chapter 13, a pipeline for the co-registration of perfusion data from intraoperative Ultrasound and preoperative Magnetic Resonance (MR) imaging was described (Tab. 1.1, row 11). An interactive visual analysis approach was utilized in comparing the perfusion parameters derived from both modalities.

The approaches presented in both chapters require a series of manual preprocessing steps and the use of multiple tools. This is less critical in the comparison of Ultrasound and MR since this is a pure research endeavor aiming at an evaluation of the former for resection control in brain tumor surgery. However, monitoring gliomas in clinical routine demands a more automated realization within a single tool.

1.8 Future Research Directions

Massive amounts of data with an increasing complexity will be generated in the future, not at the least fostered by *Big Data* initiatives, also in medicine and biology. At the same time, the human cognitive capability will remain constant [266]. Automatic data mining techniques are certainly powerful in classifying, modeling, and summarizing data as well as in detecting anomalies, associations, groups, and structures. However, the process of generating knowledge from raw data in solving fuzzy, complex, and opaque issues will always require human interaction with these techniques. Compounded by an increasing data size, this poses many challenges on human-computer interaction including user interfaces [266]. Furthermore, it demands the development of approaches making data mining more accessible to non-experts and helping them in finding an optimal solutions without knowing details of the underlying algorithm. Related work in image processing provides visual guidance through and abstraction of the parameter space of segmentation and clustering techniques [215; 239]. In order to optimally support an analyst's work, be it a clinician, medical researcher, biologist or epidemiologist, it is vital to better understand and support the human reasoning part of the visual analytics process (Fig. 1.1). Recently, a detailed model describing how the analytical components, i.e. data, models, and visualization, support this knowledge generation part has been proposed [196]. Future research may use this model in designing new visual analytics applications. In order to recognize an analyst's workflow, these applications should adhere to the "human *is* the loop" rather than the "human *in* the loop" philosophy [46]. That is, visual analytics approaches are more likely to get accepted when they are seamlessly fitted to the workflow, e.g. of epidemiologists [113], instead of requiring analysts to adopt their strategies to available tools.

Important research directions in visual analytics, which find applications and open up new perspectives in medicine and biology, are *uncertainty-aware visual analytics*, *predictive visual analytics*, *progressive visual analytics*, and *pair analytics*. Uncertainty-aware visual analytics supports the analyst in making informed decisions taking the inherent uncertainty of data and such that is added by data preprocessing, analysis, and visualization into account [32]. Uncertainty inherent to the data either results from the acquisition process or the variability of the represented

phenomenon [15]. Rather than just visualizing uncertainty, the reasoning under uncertainty must be supported [135]. The analyst's awareness of the involved uncertainties directly influences the confidence or trust in the analysis results [195]. A visual analytics framework should quantify and present the aggregated uncertainty of analysis results and the impact of uncertainty sources to the analyst [32]. Predictive visual analytics has its origins in the business world where *predictive analytics* is utilized to forecast economic developments, e.g., the rise and fall of prices. The predictive capabilities lead to a paradigm shift from solely reactive to proactive visual analytics [136]. Instead of reasoning on events that have already occurred, an analyst can forecast events thereby triggering proactive measures. For instance in heath care, syndromic surveillance data can be exploited to detect regions in time and space with an abnormally high occurrence of events (hotspots) and to predict the growth of these regions as well as new hotspots [136]. Progressive visual analytics enables analysts to investigate partial results of a time-consuming automatic analysis while it is still running [230]. Analysts can act based on intermediate results and adapt parameters of the analysis or prioritize subspaces of interest. In pair analytics, a domain expert and a computer scientist collaboratively analyze the data [7]. The computer scientist knows the specifics of data analysis and visualization and can tweak the corresponding parameters. The domain expert knows how to fine-tune application-specific analytical components, can interpret intermediate analysis results, and poses the relevant questions to proceed. The analysis can be performed at the same (localized) or at separate (distributed) locations. As a consequence, the visual analytics system must run on one or two workplaces each with an input device requiring synchronization and an appropriate GUI-design. Recently, a field experiment methodology has been proposed for pair analytics studies [101].

These important research directions in visual analytics are in the following related to the data which has been investigated within this thesis. In addition, developments in the corresponding scientific fields increasing data complexity and hence, further encouraging the design and utilization of visual analytics approaches are briefly outlined.

Simulated and Measured Hemodynamic Data Sources of uncertainty in simulated and measured hemodynamics are the image acquisition process, yielding data that suffers from noise and partial volume effects, and the individual steps of the respective hemodynamic data generation pipeline, each requiring manual parameter adjustments (Sec. 1.3.3,1.3.6). In simulated hemodynamics, simplifying assumptions of the fluid properties, the boundary conditions, and the simulation model, as well as numerical inaccuracies add to the aggregated uncertainty. The exact impact of these sources could be investigated in the visual analysis of vector fields resulting from ensemble simulations [98]. The analysis and visualization of hemodynamic data have so far widely neglected uncertainty. Analyzing the time-dependent evolution of flow features over all time steps of unsteady CFD simulation data can be time-consuming. Instead of waiting for the final result, a progressive visual analytics approach would facilitate the investigation of intermediate results. This may lead to an adjustment of the analysis strategy or algorithm. Epidemiological population studies may at some point include the simulation and/or measurement of hemodynamics. A visual analytics approach should then support the joint investigation of non-image and simulation/image data as shown in this thesis (Chap. 9,10). Monitoring a very large population over time will facilitate the generation of an atlas of typical blood flow patterns among which some may correlate with a higher risk of aneurysm rupture [23]. Predictive visual analytics should then investigate the role of the blood flow pattern and other cohort study variables in forecasting aneurysm rupture.

In the future, multiple developments will add to the complexity of simulated and measured hemodynamic data. Intravascular imaging provides detailed information about the thickness of the aneurysm wall which should be jointly investigated with the hemodynamics [62; 124]. In the car-

diac domain, the limited spatial and temporal resolution of measured data will be complemented by patient-individual CFD simulations [251]. Their ability to provide metrics which cannot directly be measured such as the wall shear stress is highly appreciated [152]. Whole heart simulations, i.e. including the heart chambers, coronaries, and valves, will contribute to an understanding of intracardiac blood flow phenomena and to the diagnosis of a broader range of pathologies, e.g., ventricular septal defect or mitral valve diseases. Treatment planning will benefit from coupling simulations of cardiac hemodynamics with mechanical simulations of implantations, e.g., transcatheter aortic valve implantations [194].

Biological Multivariate Imaging Data Uncertainty of biological multivariate data results from the measurement inaccuracies and the limited resolution of the respective imaging technique. The uncertainty-aware visual analytics of toponome data has to consider an additional source of uncertainty which is the binarization of the measured fluorescence response. The impact of this binarization on the reasoning of biologists and the final outcome of the analysis must be investigated. In this course, an uncertainty of protein co-mapping should be computed and integrated for instance, in the graph-based visualization presented in Chapter 4. Qualitative user studies with biologists must then be conducted to evaluate the consequences of incorporating uncertainty [253]. Within the thesis at hand, toponome data were investigated in localized pair analytics sessions involving a computer scientist and a molecular biologist (Chap. 4). The results achieved within the investigated case studies demonstrate the potential of this strategy. The pair analytics approach released the biologist from the need of understanding every detail of the analysis and visualization techniques and allowed for focusing on the biological research questions. It offered the computer scientist a deeper understanding of the biologist's mindset, which helped in improving the visual analytics system.

Recent developments in bioimaging will lead to a more frequent acquisition of biological multivariate imaging data. Examples of techniques recording N-dimensional intensity arrays representing the local mapping of molecules, residues or interaction patterns per pixel are Raman imaging and Matrix Assisted Laser Desorption / Ionization (MALDI) imaging [116]. An increasing resolution of fluorescence microscopy may yield larger images per tagged protein in toponome imaging. The resolution of conventional fluorescence microscopy is limited to $\approx$250 nm by the diffraction properties of the light. Super resolution fluorescence microscopy with resolutions down to 10-20 nm produces gigabytes of data in a single run [134]. Time-resolved toponome data of living cells will pose additional challenges on a visual analytics approach [211]. The integration of toponome data with other omics data and with medical image data is required to achieve one of the future goals of systems biology namely the seamless integrated analysis of biological systems from organs to molecules [163].

Epidemiological Cohort Study Data Various sources contribute to the uncertainty of epidemiological cohort study data. Regardless of the applied acquisition technique, the study's medical image data suffer from noise and partial volume effects. Measurement inaccuracies, e.g. of blood pressure, or unreliable self-reports, e.g. on eating and drinking habits, represent uncertainties inherent to the non-image data. The data transformation and also the data analysis itself further contribute to the overall uncertainty. Examples are the padding of missing data with population-derived statistics, statistical tests yielding measures of uncertainty, and fuzzy clustering techniques for subgroup definition. In an uncertainty-aware visual analytics approach, these uncertainties must be modeled, considered in the analysis, and conveyed in the visualizations. Predictive visual analytics can contribute to forecasting the onset of a disease based on the evolution of socio-

demographic, lifestyle-related, and health-related factors. In a retrospective analysis of cohort study data, models can be build that describe how similar populations evolve and predict how they will evolve in the future. For instance, decision trees classifying study participants with respect to a target outcome, i.e. a disease reported in the last wave of a cohort study [162], can be interactively constructed and refined [245]. Cohort study data are very complex and computations involving all the hundreds of variables can be very time-consuming. The benefit of a progressive visual analytics approach has been demonstrated in investigating all correlations of possible variable pairs with all possible target outcomes [111]. Instead of waiting for the whole computation to finish, the analyst can inspect the results of one target outcome the moment they have been computed. Within the thesis at hand, cohort study data were investigated in localized (Chap. 10) and distributed pair analytics sessions (Chap. 9) involving a computer scientist and either a neuroscientist, a radiologist or an epidemiologist. Multiple hypotheses were generated harnessing the pair analytics approach. While quantitative evidence is lacking that neither the computer scientist nor the domain scientist could have achieved the same results by oneself, there is reason to believe that more hypotheses could be generated in shorter time by the collaborative effort.

One of the major joint developments in medicine and epidemiology is *population imaging* shifting the focus from curative to preventive medicine. Forces are bundled to provide a pan-European infrastructure for population imaging since more and more cohort studies include medical and also biomolecular imaging [48]. Analyzing the acquired image and omics data requires dedicated processing methods to extract structures of interest and derive meaningful parameters characterizing them. Since population imaging may cover the entire human body, processing methods should be as modular as possible instead of being tailored to a specific organ or structure. Ideally, domain experts can compose the modules for the task at hand by themselves. The massive amounts of data being acquired in population imaging will make visual analytics approaches indispensable. In order to acquaint non-specialists with the data mining part and its parameters, abstractions guiding parameter fine-tuning must be provided. Epidemiologists are accustomed to dividing a cohort into subcohorts, e.g., by age or gender, prior to investigation and then, compute and compare statistics of these subcohorts. The infrastructure as well as the data analysis and visualization components of a visual analytics framework must support this subdivision strategy [109; 137]. So far, most frameworks developed in the context of epidemiology and public health, facilitate the definition and investigation of only a single subcohort at a time.

Part I

Cerebrovascular Hemodynamic Data

This part of the postdoctoral thesis cumulates the following publications:

Chapter 2 [168] **S. Oeltze**, D.J. Lehmann, A. Kuhn, G. Janiga, H. Theisel, and B. Preim, "Blood Flow Clustering and Applications in Virtual Stenting of Intracranial Aneurysms", *IEEE Trans. Vis. Comput. Graphics (IEEE TVCG)*, vol. 20, no. 5, pp. 686-701, 2014.

Chapter 3 [171] **S. Oeltze-Jafra**, J.R. Cebral, G. Janiga, and B. Preim, "Cluster Analysis of Vortical Flow in Simulations of Cerebral Aneurysm Hemodynamics", *IEEE Trans. Vis. Comput. Graphics (IEEE TVCG)*, vol. 22, no. 1, pp. 757-766, 2016.

Blood Flow Clustering and Applications in Virtual Stenting of Intracranial Aneurysms

Steffen Oeltze, Dirk J. Lehmann, Alexander Kuhn, Gábor Janiga, Holger Theisel, Bernhard Preim

Abstract—Understanding the hemodynamics of blood flow in vascular pathologies such as intracranial aneurysms is essential for both their diagnosis and treatment. Computational Fluid Dynamics (CFD) simulations of blood flow based on patient-individual data are performed to better understand aneurysm initiation and progression and more recently, for predicting treatment success. In *virtual stenting*, a flow-diverting mesh tube (*stent*) is modeled inside the reconstructed vasculature and integrated in the simulation. We focus on steady-state simulation and the resulting complex multiparameter data. The blood flow pattern captured therein is assumed to be related to the success of stenting. It is often visualized by a dense and cluttered set of streamlines.
We present a fully automatic approach for reducing visual clutter and exposing characteristic flow structures by clustering streamlines and computing cluster representatives. While individual clustering techniques have been applied before to streamlines in 3D flow fields, we contribute a general quantitative and a domain-specific qualitative evaluation of three state-of-the-art techniques. We show that clustering based on streamline geometry as well as on domain-specific streamline attributes contributes to comparing and evaluating different virtual stenting strategies. With our work, we aim at supporting CFD engineers and interventional neuroradiologists.

Index Terms—Blood Flow, Aneurysm, Virtual Stenting, Clustering, Evaluation.

1 INTRODUCTION

INTRACRANIAL aneurysms, also referred to as cerebral aneurysms, represent a pathological, balloon like dilation of cerebral vasculature due to a weakness of the arterial wall. They occur with a prevalence of about 2% in Western Europe [1]. Their rupture is associated with a mortality rate of $\approx 50\%$. Among other treatment options, *stenting* plays an increasingly important role. In stenting, the flow is diverted around the aneurysm by an expandable mesh tube (*stent*), thereby reducing and decelerating its inflow (Fig. 1(a)).

The blood flow pattern is among the hemodynamical parameters that are assumed to be related to the success of stenting [2], [3], the development of thrombosis (blood clotting, which is a desirable outcome of stenting) [4], and the risk of aneurysm rupture [5]. A better understanding of these relations may contribute to patient selection for flow diverting stents. While they often lead to thrombosis and reverse remodeling, adverse effects leading to late rupture were also observed [3]. With the increased number of treatment options and available types of stents, the need for decision support is strongly increased.

Computational Fluid Dynamics (CFD) simulations, which generate patient-specific hemodynamic data, are employed to better understand the effect of stents on aneurysmal hemodynamics and for predicting treatment success [2], [6], [7]. In *virtual stenting*, different types of stents are modeled at different locations inside the reconstructed vascular anatomy and integrated in the simulation. We focus on steady-state simulations since major aspects of aneurysmal hemodynamics may be inferred from steady flow [8]. The simulation results in a complex multiparameter dataset comprising several scalar and vectorial attributes. The blood flow pattern captured therein, is often visualized for investigation by a dense and cluttered set of streamlines colored according to one of the scalar attributes.

We present a fully automatic approach for reducing visual clutter and exposing characteristic flow structures by grouping similar streamlines and computing group representatives. We quantitatively evaluate three conceptually different techniques for the grouping: *k-means* clustering, *Agglomerative Hierarchical Clustering* in four variations (single link, complete link, average link, and Ward's method), and *Spectral Clustering*. While each individual technique has been applied to streamlines in 3D flow fields [9], [10], [11], [12], the quality of their results has not been compared before. The gained insight is valuable for all applications employing streamline clustering.

Cluster representatives, which summarize the complex blood flow, are derived from the clustering result. We adapt a type of representative that is employed in clustering fiber tracts of the human brain. In a qualitative expert evaluation of visual blood flow summaries, we compare the quantitatively best performing clustering techniques and the corresponding representatives. Furthermore, we show that clustering streamlines also based on domain-specific attributes supports the evaluation of virtual stenting strategies. For instance, clustering based on the local residence time of blood flow within the aneurysm gives hints on potential locations of thrombosis initiation.

- *S. Oeltze, D.J. Lehmann, H. Theisel, and B. Preim are with the Dept. of Simulation and Graphics, University of Magdeburg, Germany. E-mail: stoeltze@isg.cs.uni-magdeburg.de*
- *A. Kuhn is with the Zuse Insitute, Berlin, Germany.*
- *G. Janiga is with the Institute of Fluid Dynamics and Thermodynamics, University of Magdeburg, Germany.*

In summary, our contributions are:

- Quantitative evaluation of three conceptually different streamline clustering techniques
- Visual summary of flow patterns and design lessons
- Expert evaluation of visual flow summaries
- Application-specific insight from clustering domain-specific streamline attributes
- A tailor-made type of cluster representative

We aim at supporting CFD engineers in investigating simulation results. In a dense sampling of aneurysmal flow by thousands of streamlines, they rely on filtering these lines, a locally restricted streamline seeding or on global hemodynamic parameters. Minor, local changes of the flow pattern yet influencing the success of stenting, may remain unnoticed. We further aim at supporting interventional neuroradiologists in developing a patient-specific treatment strategy. CFD results are not yet part of the clinical routine. Hence, the physicians have little experience in investigating flow data. Our visual flow summary simplifies the access to flow data, it is easy to read, and it contributes to the communication between CFD engineers and physicians. We employ our approach amongst others to data of the Virtual Intracranial Stenting Challenges in 2009 and 2010.

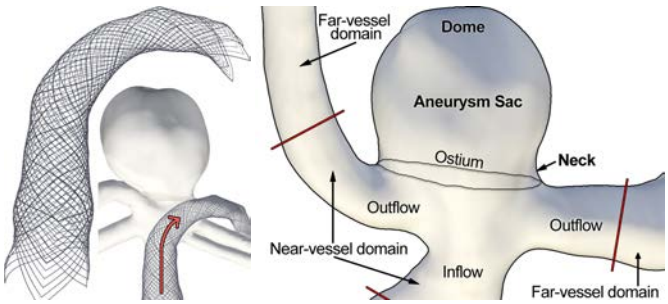
2 Medical and Technical Background

This section briefly overviews the treatment of intracranial aneurysms, introduces the research field virtual stenting, and describes our data generation pipeline.

2.1 Treatment of Intracranial Aneurysms

Intracranial aneurysms usually develop somewhere at the *Circle of Willis*. Their shape may be characterized as saccular, fusiform or dissecting with saccular having by far the highest prevalence [13]. The morphological features of a saccular aneurysm are illustrated by Figure 1(b). Most aneurysms remain undetected until rupture. While surgical clipping has been the gold standard in treatment for decades, the number of endovascular interventions is increasing. They bear less intraoperative risk and may be applied, e.g., by an interventional neuroradiologist, to aneurysms which are difficult or impossible to reach for a surgeon [14]. In *coiling*, the aneurysm is filled with platinum coils to promote thrombosis, which may eventually seal the aneurysm. Self expanding, high-profile, flow-diverting stents provide a promising alternative to coiling in patients with complex aneurysms (Fig. 1(a)). They reduce and decelerate the blood circulation into the aneurysm, thereby causing a prolonged residence time, which in turn promotes thrombosis formation [13].

Despite the progress in interventional techniques, the associated risks persist, e.g., injury of the aneurysmal wall during stent insertion. A detailed risk and benefit estimation and a deeper insight into the hemodynamics of blood flow that cause aneurysm development and rupture are necessary.



(a) (b)

Fig. 1. *(a)* Flow diverting stent and its deployment (arrow indicates flow direction). *(b)* Morphological features of a saccular aneurysm (bold) and subdivision of the surrounding vascular domain (red lines).

2.2 Virtual Stenting

Virtual stenting (VS) is a collaborative effort between CFD engineers, physicians, and computer scientists. Its main objectives are supporting clinical decision making and stent design. In the former, questions such as "Is the vascular and aneurysmal morphology eligible for stenting?" and "Which stent should be used and where should it be placed?" need to be answered. In stent design, different properties, e.g., grade of mesh porosity and strut size, and their impact on the hemodynamics of blood flow are investigated.

One challenge in VS is comparing results of different CFD simulations, e.g., before and after stenting [6]. We support a comparison by visual summaries of blood flow. So far, it is often based on global values such as aneurysmal inflow rate [15]. Sometimes, the aneurysm wall is colored according to a hemodynamic parameter and presented in a side-by-side view [6]. Streamlines are employed for comparing flow patterns. They are often seeded on the ostium and displayed side-by-side [2], [6], [16]. However, either the entire set of lines is displayed leading to visual clutter or representative lines must be selected manually.

2.3 Hemodynamic Data Generation Pipeline

We briefly summarize our hemodynamic data generation pipeline (see [15], [17] for details). First, image data of the aneurysm morphology including the vasculature in the close surrounding are acquired, e.g., by 3D rotational angiography or Computed Tomography (CT) angiography. Next, the aneurysm and the vasculature are segmented via thresholding. Afterwards, a surface mesh of the vessel wall is reconstructed from the segmentation result and optimized [18]. Then, the *ostium* is extracted [19]. It separates the aneurysm from the parent vessel and approximates the original vessel wall (Fig. 1(b)). It is frequently used to explore the flow into the aneurysm, e.g., by seeding streamlines there [20]. Next, the stent geometry is modeled and deployed to the vessel wall. Finally, a volume mesh is constructed based on the surface meshes of the vessel wall and the stent using ANSYS IcemCFD (Ansys Inc., Canonsburg, PA, U.S.). Fluid flow simulations are performed in ANSYS Fluent 12 (Ansys Inc., Canonsburg, PA, U.S.).

3 Related Work on Partition-Based Flow Visualization

Flow visualization techniques have been categorized by Post et al. [21] into direct, texture-based, geometric, and feature-based techniques. Salzbrunn et al. [22] added the class of partition-based techniques, which decompose a flow field based on vector values, integral curve properties or topological features. Blood flow clustering based on vector values has been presented in the context of cardiac blood flow [23]. However, we follow the arguments in [10] and advocate the use of integral curves since they represent continuous flow patterns traced over the domain instead of a very local vectorial flow information. We briefly recapitulate approaches for flow decomposition based on integral curves and classify them into user-guided and automatic partitioning. For State-of-the-Art reports on topology-based decomposition and visualization of flow, see [24], [25].

3.1 User-Guided Flow Partitioning

The approaches in this class decompose a set of integral curves guided by the user. Salzbrunn and Scheuermann [26] propose combined Boolean predicates based on predefined scalar quantities, which determine for each streamline whether it has a desired property. Predicates on pathlines are applied to the visual analysis of measured blood flow in aortic aneurysms [27]. A residence time predicate is used for evaluating blood clotting. In [28], a visual analytics approach is proposed for filtering pathlines based on local and global pathline attributes, e.g., curvature and Lyapunov exponent. Pobitzer et al. [29] demonstrate the application of dimension reduction to the set of attributes in order to detect relevant, independent ones. Two other approaches let the user specify interesting integral curves or curve parts in observation instead of attribute space. Advanced virtual probing of measured cardiovascular flow by seeding integral curves on a flexible probing geometry is presented in [30]. Gasteiger et al. employ a lens metaphor for generating focus-and-context visualizations of streamline parts [17].

The lens metaphor facilitates only a local and view-dependent inspection of the flow pattern. It emphasizes or attenuates all streamline parts inside the lens but it does not reduce visual clutter with respect to the flow pattern. Neither lens nor virtual probing deliver reproducible and quantifiable results. Line predicates and the visual analytics of pathline attributes require the user to define attributes and attribute value ranges of interest in order to compose sets of lines, which are homogenous with respect to a certain attribute or a combination of attributes. Automatic flow partitioning approaches employ a data-driven strategy for creating such sets and are hence self-tuning with respect to differences in the flow across aneurysms.

3.2 Automatic Flow Partitioning

Our work is strongly related to approaches, which automatically partition a set of integral curves by means of *clustering*, i.e. grouping similar curves. These approaches differ in the clustering technique and in the similarity measure. Chen et al. propose a two-stage k-means clustering [9]. The initial rough geometry-based partitioning is refined by taking vector and shape properties into account. Both stages are based on Euclidean distance as the similarity measure. Cluster representatives are the streamlines closest to the cluster centroids. In [12], Agglomerative Hierarchical Clustering (AHC) with average link has been used for partitioning. The authors propose a similarity measure that facilitates an interactive, cluster-based exploration of flow with seeding rakes. A saliency-guided streamline seeding is followed by AHC with single link in [10]. Streamlines at cluster boundaries are displayed as representatives. Gasteiger et al. employ local streamline properties to identify and group lines that constitute the *inflow jet*, which is correlated with aneurysm rupture [31]. Rössl and Theisel discuss a *spectral embedding* of streamlines [11]. They demonstrate Spectral Clustering (SC) in the embedding space and compare various similarity measures. Similar to the clustering of integral curves is the clustering of fiber tracts extracted from Diffusion Tensor Imaging (DTI) data. In [32], fiber tracts are partitioned by means of a specialized SC approach. Three types of cluster representatives are investigated in [33]. Moberts et al. evaluate three variants of AHC and four similarity measures for clustering fiber tracts [34]. A new similarity measure in conjunction with AHC using single link is introduced in [35].

AHC, k-means, and SC are the most widely used techniques for clustering streamlines (and fiber tracts). However, the quality of their results in this context has not been individually assessed and compared. We quantitatively evaluate the three techniques, including four AHC variants, by means of internal cluster validity indices (Sec. 5.4). In a qualitative expert evaluation of the best performing techniques, we identify the most appropriate one for clustering blood flow (Sec. 6.4). While the clustering in related work is mostly restricted to streamline geometry and derived geometrical attributes, we extend it to domain-specific attributes. We adopt the idea of cluster representatives for reducing visual clutter and assess the approaches in [33].

4 Streamline Generation & Similarity

In this section, we describe our generation of streamlines, their properties, and our streamline similarity measures.

4.1 Domain, Tracing, and Line Properties

The input of the streamline generation is the volume mesh from the CFD simulation (Sec. 2.3). It is represented as an unstructured grid composed of tetrahedral cells. A vector is stored at each cell point. Before streamlines are generated, the mesh is manually cropped such that it contains only the aneurysm and the *near-vessel domain* [20] (Fig. 1(b)). This enables us to focus the analysis and strongly improves the expressiveness of the clustering. It is very likely that streamlines follow a similar course in the feeding vessel (inflow) and they may also follow a similar course in a draining vessel (outflow). However, depending on where

they enter the aneurysm, their course may strongly differ inside. If the *far-vessel domain* was also considered in clustering, these differences would have less impact.

To assess the in- and outflow of the aneurysm, streamlines have been seeded on the ostium. The ostium is represented by a triangle mesh whose vertices have been homogeneously distributed such that the under- and over-representation of flow parts are avoided [19]. The number of vertices is adjusted such that the mesh resembles the former vessel wall. Streamlines were traced in ParaView (Kitware, Clifton Park, NY, U.S.). A 5th order Runge-Kutta method has been employed with an integration step size that was constantly adjusted according to an estimated error. The tracing was carried out in backward and forward direction. The resulting two lines were merged such that a linear traversal of the vertices from in- to outflow is possible.

Line Properties: The streamlines differ in their number of vertices and in their length. The former has a strong impact on the computational time of most inter-streamline similarity measures. The similarity itself is strongly influenced by streamline length. Two lines may follow a similar course for a long time but then, one of them is terminated. Most similarity measures assign a much higher weight to the difference in length than to the similarity over a long run. In all our datasets, a few lines follow a course very similar to a large set of neighboring lines but are considerably shorter. They occur close to the vessel wall due to early termination of the integration. We consider them as incomplete rather than incorrect data entities. Hence, the clustering should group them with the streamlines having a similar course. Still, we term them outliers in the following.

4.2 Geometry-Based Streamline Similarity

Geometry-based streamline similarity (or dissimilarity) is often expressed by a distance measure. The choice of a measure depends on the application. General requirements are positive-definiteness and symmetry. A valid example is the Hausdorff distance. However, this distance is very sensitive to streamline length, since it outputs the maximum of point-wise distances [11]. A less sensitive measure is the *Mean of Closest Point Distances* (MCPD) [36]:

$$d_M(s_i,s_j) = mean(d_m(s_i,s_j), d_m(s_j,s_i)) \qquad (1)$$
$$with\ d_m(s_i,s_j) = mean_{p_l \in s_i} \min_{p_k \in s_j} \|p_k - p_l\|$$

Moberts et al. evaluate four similarity measures for clustering fiber tracts and favor MCPD [34]. Yu et al. apply MCPD for clustering streamlines and report that the clusters comprise important flow features [10]. In [11], five similarity measures adopted from the clustering of fiber tracts are evaluated for clustering streamlines. The rather qualitative evaluation includes MCPD and shows no drawbacks compared to the other measures. In [12], a new similarity measure is compared to three other measures including MCPD. The new measure performs one to two orders of magnitude faster but no advantage in terms of cluster quality is reported. However, MCPD is subjectively rated as producing good quality clusterings. We adopted MCPD and applied it to blood flow clustering. Initial tests showed good results but also revealed that MCPD is still too sensitive to streamline length, in particular when being used with clustering techniques being sensitive to outliers (Tab. 1). Very small-sized, outlier-corrupted clusters were generated whose representatives distorted the flow summary. We further reduce MCPD's sensitivity by replacing the outer mean in Equation 1 by a minimum computation:

$$d_M(s_i,s_j) = min(d_m(s_i,s_j), d_m(s_j,s_i)) \qquad (2)$$
$$with\ d_m(s_i,s_j) = mean_{p_l \in s_i} \min_{p_k \in s_j} \|p_k - p_l\|$$

If two lines are very similar but one is shorter, d_m from the shorter to the longer line is chosen. The resulting high similarity increases the chance of being assembled.

4.3 Attribute-Based Streamline Similarity

Besides streamline geometry, we employ streamline attributes for clustering. They describe **(1)** the underlying vector field, **(2)** line bending or **(3)** domain-specific aspects:

1 pressure, velocity magnitude, velocity gradient magnitude, angular velocity, vorticity magnitude
2 curvature, torsion
3 distance to ostium, distance to aneurysm wall, local residence time

In the following, we focus on the domain-specific attributes **(3)** since their clustering revealed the most interesting aspects. The distance to the ostium is computed in order to separate flow structures that occur close to the aneurysm's neck from those that occur close to its dome (Fig. 1(b)). The distance to the aneurysm wall is determined in order to separate flow close to the wall from flow close to the center. Both are inspired by discussions with a neuroradiologist and by clinical research results such as a close correspondence between near-wall flow and wall-shear stress. They have been computed only at streamline vertices located inside the aneurysm as the distance between the vertex and its closest point (not vertex) on the respective surface.

The residence time of flow inside the aneurysm is crucial in thrombosis formation [4]. We compute it by aggregating partial timing results along each streamline. For each line segment inside the aneurysm, the two associated velocity magnitudes are retrieved from the data. Based on their difference and the segment length, the partial residence time is computed. If a line segment intersects the ostium, the velocity is interpolated at the intersection point. While the other streamline attributes are computed per vertex, the residence time is a single scalar per line.

What is left is the definition of a streamline similarity measure on the attributes. For the local residence time, we employ the absolute difference of two scalars. For the remaining attributes, we first compute a simple statistic that approximates the attribute information along a streamline, e.g., minimum, maximum, mean, or median. Since this breaks down the information to a scalar value, we can apply the same similarity measure as for the residence time.

5 Streamline Clustering Techniques: A Quantitative Evaluation

This section is dedicated to the quantitative evaluation of techniques often used for clustering streamlines (Sec. 3.2). It starts with descriptions of Agglomerative Hierarchical Clustering (AHC) and k-means based on [37] and an introduction to Spectral Clustering (SC) based on [38].

5.1 Agglomerative Hierarchical Clustering

AHC starts with each streamline being a cluster and then, repeatedly merges the two closest clusters until a single cluster is formed. The resulting hierarchy is stored and may be visualized by a *dendrogram*. All merge steps rely on a squared, symmetric distance matrix $\mathbf{M}$ and a measure of cluster proximity. In our case, $\mathbf{M}$ contains the pairwise inter-streamline distances (Eq. 2). Various cluster proximity measures have been published among which *single link*, *complete link*, *average link*, and *Ward's method* are most popular. In single link, the proximity of two clusters is defined as the minimum distance between any two points in the different clusters. This approach can handle clusters of arbitrary shape, it tolerates considerable differences in cluster size but it is sensitive to outliers. Furthermore, it is infamous for the *chaining effect* leading to clusters containing very dissimilar elements which are connected by a chain of similar elements via some transitive relationship. In complete link, the proximity of two clusters is computed as the maximum distance between any two points in the different clusters. Complete link is less susceptible to outliers but tends to break large clusters and it favors globular cluster shapes. Average link is an intermediate approach between single and complete link. It also strives for globular compact clusters [39]. The proximity of two clusters is defined as the average proximity between pairs of points in the different clusters. Ward's method aims at minimizing the total within-cluster variance. It defines the proximity of two clusters as the sum of squared distances between any two points in the different clusters (SSE: sum of the squared error). Due to the SSE-based proximity, Ward's method favors globular clusters. It was shown to prefer clusters with similar size and to be robust against outliers in the context of 2D curves [40].

All AHC variants lack a global objective function to be optimized (Tab. 1). They decide locally which clusters are merged. These decisions cannot be undone such that bad decisions, e.g., involving outliers, are propagated throughout the entire clustering process. A strength of AHC is its ability to rapidly generate different numbers of clusters k by cutting the cluster hierarchy at respective levels. Furthermore, it is non-parametric except for k and the proximity measure. Both strengths explain its frequent use when the "correct" number of clusters is unknown. The user then sequentially browses through the levels. Visually comparing consecutive clustering results is simplified by the locally restricted change (split/merge). AHC's bottleneck in terms of time complexity is the computation of $\mathbf{M}$, which often requires a vast number of Euclidean distance tests.

TABLE 1
Comparison of clustering algorithms with respect to the type of objective function (OF) and the capabilities to handle arbitrarily-shaped clusters, clusters of significantly different size, and outliers.

Property	Spectr. Clust.	Agglomerative Hierarchical Clustering				k-means
		Single	Compl.	Avg.	Ward	
OF	global	local	local	local	local	global
Shape	+	+	−	−	−	−
Size	*o*	+	−	*o*	*o*	−
Outlier	+	−	*o*	*o*	+	−

5.2 k-means Clustering

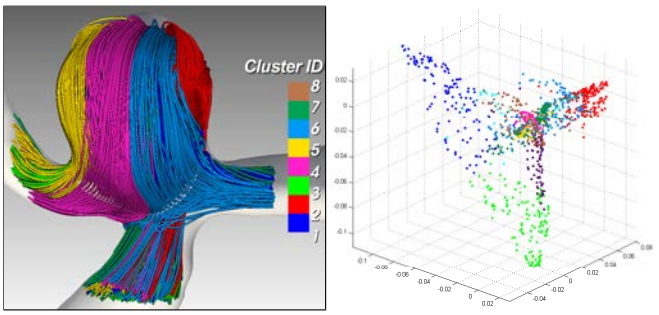
k-means requires an a priori definition of the number of clusters k by the user. Then, k initial cluster centroids are chosen, often by a random selection of k data entities. Each entity is now assigned to the closest centroid, e.g, by comparing squared Euclidean distances. Finally, each centroid is updated to the mean of its assigned data entities (which rarely corresponds to an existing entity). The assignments and updates are repeated until the goal of a global objective function has been achieved. For squared Euclidean distances, the objective function usually aims at minimizing the sum of the squared distances of data entities to their cluster centroid (SSE).

Streamlines cannot be directly plugged into k-means since the computation of their mean is undefined. Feature vectors must be derived representing the lines in a new n-dimensional space. A straightforward approach is to use the 3D coordinates of their vertices. Since the number of vertices varies (Sec. 4.1), each line must be equidistantly resampled to a uniform number. We employ the average number of vertices of all streamlines. A lower-dimensional alternative has been proposed by Chen et al. [9]. Two scalar streamline entropy measures together with the coordinates of start-, middle, and endpoint of the line constitute an 11-dimensional feature vector. Contrary to [9], we employ all dimensions in a single clustering stage since the proposed two stages hamper a user-defined choice of k. However, the latter is required for our quantitative evaluation.

k-means is often computationally faster than AHC since it does not require the computation of pairwise distances between data entities. However, it is sensitive to outliers and fails in handling non-globular clusters and clusters of widely different sizes (Tab. 1). Its results are dependent on the random initialization of the centroids. A "bad" choice causes the algorithm to get stuck in a local minimum of the objective function. We mitigate this problem by running the algorithm ten times and choosing the result with the minimum SSE.

5.3 Spectral Clustering

Spectral Clustering (SC) maps the original streamlines to a *spectral embedding* space where each line is represented by a point (Fig. 2). Key features of the mapping are the preservation of local distance relations between nearby lines and the enhancement of the data's cluster properties, i.e. an



(a) (b)

Fig. 2. *(a)* Spectral Clustering of streamlines in a basilar tip aneurysm. *(b)* Spectral embedding of the lines. The first three largest eigenvectors are shown.

improved cluster separability. In the following, we use the terms distance and difference interchangeably.

SC can be formulated as a graph partitioning problem [41]. Streamlines are represented by a weighted, fully-connected, undirected graph. The nodes are the streamlines and the edge weights are computed according to Equation 2. The weights are then transformed from difference to affinity such that similar streamlines have a high and dissimilar a low pairwise affinity. Next, the graph is partitioned into subgraphs. Shi and Malik [41] propose to use a *normalized cut* which minimizes the sum of weights of the edges that need to be removed (cut) and at the same time balances the sum of edge weights of the partitions. While this problem is NP hard, a relaxed version is solved by spectral graph partitioning using *Graph Laplacians*.

Given a dataset S with n streamlines as a graph and a number of clusters k, **(1)** the $n \times n$ distance matrix $\mathbf{M}$ is computed by a pairwise application of Equation 2 to the lines in S. The same matrix is employed for AHC (Sec. 5.1). **(2)** Based on $\mathbf{M}$, the $n \times n$ weighted adjacency matrix of the graph is constructed by applying a function f to the entries of $\mathbf{M}$ that gives high values in case of small differences and converges to zero for high differences. The resulting matrix $\mathbf{W}$ is referred to as *affinity matrix*. As f, the Gaussian similarity function is used:

$$f(m_{ij}) = f(m_{ji}) = \exp(-(m_{ij})^2/(2\sigma^2)) \qquad (3)$$

The parameter σ controls the width of f thereby steering how rapidly the affinity falls off. **(3)** Next, a $n \times n$ diagonal degree matrix $\mathbf{D}$ is constructed with each diagonal entry d_{ii} being the degree of the node that represents streamline i in the graph. The degree is computed as the sum of weights of the edges incident to the node. **(4)** Now, the *normalized Graph Laplacian* $\mathbf{L}$ is computed [41]: $\mathbf{L} = \mathbf{I} - \mathbf{D}^{-1}\mathbf{W}$ with $\mathbf{I}$ being the identity matrix. **(5)** Then, the eigenvectors and eigenvalues of $\mathbf{L}$ are determined. The eigenvectors corresponding to the smallest k eigenvalues are used for clustering. **(6)** Let $\mathbf{U}$ be the $n \times k$ matrix that contains the k eigenvectors as columns. Each row i of $\mathbf{U}$ then represents the coordinates of a point that corresponds to streamline i in the $\mathbb{R}^k$ spectral embedding space spanned by the eigenvectors. **(7)** In the embedding, clusters can be detected, e.g., by k-means or an eigenvector rotation [42]. We employ the latter since it suggests an optimum number of clusters based on a user-defined range for k. Since it is based on the largest eigenvectors of $\mathbf{L}$, we change the formulation of $\mathbf{L}$ to:

$$\mathbf{L} = \mathbf{D}^{-1}\mathbf{W} \qquad (4)$$

Local scaling: Zelnik-Manor and Perona propose a local determination of σ since a global value (Eq. 3) only works well if all clusters are of the same density [42]. Since we cannot guarantee this for our streamlines, we adopt their *local scaling*. A local σ_i is computed for each line i based on the difference between i and its N'th neighbor. A value of $N = 7$ is reported to give good results [42]. However, our experiments indicated that N must be adjusted to each dataset. In very dense sets of streamlines, SC partially failed to separate clusters. With increasing density, the local neighborhood of a line contains an increasing number of very similar lines. However, the number of neighbors with an affinity $\gg 0$ should not be "too small and not too large" for SC to work properly [38]. Based on ten datasets, we identified $N = 5\%$ of the streamline count as appropriate.

SC strives for a globally optimal partitioning while AHC is bound to locally optimal decisions (Tab. 1). It can handle arbitrary cluster shapes while most AHC variations and k-means favor globular shapes. SC with local scaling considers the local streamline density. This is useful, e.g., if streamlines are seeded with a higher density close to the aneurysm wall. Our implementation of SC is parameter-free except for the range of values for k. Since the eigenvector rotation computes all partitionings within this range, the user can browse also the suboptimal results. SC is biased towards clusters of similar size due to the balancing of edge weights in the graph cutting. On the other hand, this property makes it robust against outliers which was acknowledged in the context of fiber tract length [35]. As for AHC, the bottleneck of SC is the computation of $\mathbf{M}$.

5.4 Quantitative Evaluation

We quantitatively evaluated four variants of Agglomerative Hierarchical Clustering (AHC), k-means, and Spectral Clustering (SC) for clustering streamlines. The evaluation was based on five clinical cases together comprising ten datasets and representing the prevailing types of aneurysms: basilar tip and side-wall aneurysms. Three cases were simulated without virtual stenting (two are shown in Fig. 2(a) and 4(a)). Two cases have been simulated before and after stenting, one of them with two types of stents in two different positions (Sec. 7.1 and 7.2). The streamline count was between 1138 and 2929. The evaluation was restricted to geometry-based clustering (Sec. 4.2). For each combination of clustering algorithm ($n = 6$) and dataset ($n = 10$), streamlines were clustered with the number of clusters being in the range $[2, 20]$ ($n = 19$). This resulted in $6 \times 10 \times 19 = 1140$ partitionings.

Different measures for assessing the quality of a clustering result have been proposed. In the absence of a

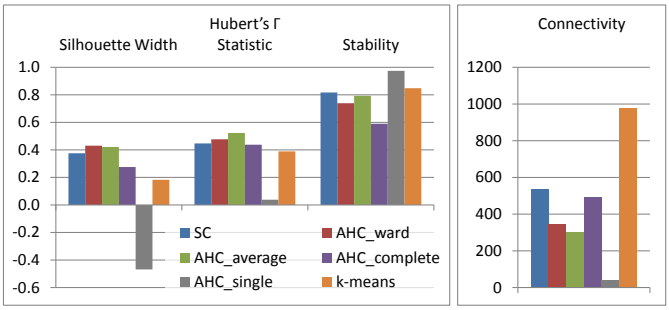


Fig. 3. Average internal cluster validity measures based on ten datasets. Spectral Clustering (SC), four variants of Agglomerative Hierarchical Clustering (AHC), and k-means are compared.

ground truth, e.g., external labels provided by an expert, unsupervised measures of cluster validity are appropriate [37]. They are also called *internal validity measures* since they are purely based on information present in the data. We employ four internal measures which together cover the most important aspects of cluster quality [39]:

- *Silhouette Width*: Non-linear combination measure of cluster cohesion and separation. Values are in the range $[-1,+1]$ and should be maximized.
- *Connectivity*: Local measure reflecting to which degree the L most similar neighbors of a streamline are placed in the same cluster. Values are in the range $[0,+\infty]$ and should be minimized. We define $L = 20$.
- *Hubert's* Γ *Statistic*: Measure of correlation between the distance matrix $\mathbf{M}$ and an idealized distance matrix (distance is 0 for streamlines in the same cluster and 1, otherwise). Values of the normalized statistic are in the range $[-1,+1]$ and should be maximized.
- *Stability*: Measure reflecting the stability and hence, the significance of the clusters. Random overlapping subsamples of the data are repeatedly drawn and clustered using the same algorithm. We draw 20 subsamples. Their clusters are then compared to the partitioning of the original data via the Adjusted Rand Index whose values are in the range $[-1,+1]$ and should be maximized [43].

To ensure comparability of the algorithms, all measures were computed in 3D streamline space although k-means and SC cluster in different spaces, i.e. in feature vector space and in the spectral embedding. The first three measures employ the similarity of two streamlines which is inferred from the distance matrix $\mathbf{M}$. Clustering by k-means has been based on two types of feature vectors (Sec. 5.2). The type based on streamline resampling consistently achieved better internal measures, which is likely due to the very sparse representation of the streamline course by the other type (only three vertices). Hence, we restrict the presentation of evaluation results to the former. For each algorithm, the internal validity measures were averaged over the 19 partitionings and the 10 datasets (Fig. 3).

Silhouette Width: AHC with single link exhibits a very poor silhouette width (-0.47). This is due to the chaining effect, which leads to a single huge heterogeneous cluster containing almost every streamline (Sec. 5.1). Hence, cluster cohesion as well as separation are small. Chaining has been observed for all datasets and most numbers of clusters. K-means performs better but sill exhibits a rather low value (0.18). The reason is that simply resampling all streamlines to a uniform number of vertices amplifies differences in streamline length and position offset for otherwise very similar lines. This counteracts our streamline similarity measure, which has been tailored to tolerate these differences (Eq. 2). As a consequence, similar lines are assigned to different clusters. Complete link also achieves a low silhouette width of 0.28. This is likely due to its tendency to break large clusters leading to a low inter-cluster separation between the resulting parts. This effect could be observed on a sample basis. Average link, Ward's method and SC perform equally well and exhibit the highest silhouette widths: $0.42, 0.43, 0.38$.

The silhouette width is biased towards globular clusters [39]. In case of elongated or concave clusters, algorithms correctly identifying them, e.g., single link and SC, may be assigned a lower silhouette width than failing algorithms. Since the cluster shape in streamline space is not clear, the silhouette width must be employed carefully. For fiber tracts, the non-globular nature of clusters has already been acknowledged [44].

Connectivity: Single link clustering by far achieves the best connectivity value due to its proximity measure which strives for a merge with the nearest neighbor. This bias has already been acknowledged in [39]. The second and third best connectivity values are achieved by average link and Ward's method. Complete link exhibits the worst value of all AHC variants. It more often adds similar neighbors of a streamline to another cluster, which may again be due to the breaking of large clusters. This leads to streamlines at the joint cluster border, which have similar neighbors in both clusters. The connectivity of SC is worse than for all AHC variants. However, this is to a great extent caused by the functioning of the algorithms and the way of computing connectivity. The computation adds the highest penalty value if the most similar neighbor is not in the same cluster. This rarely occurs in AHC since each variant starts by locally aggregating the nearest singleton clusters. SC aims at a global optimization and occasionally adds the most similar line to another cluster. A preliminary investigation revealed this phenomenon at the joint border of closely spaced clusters. Due to the bias of connectivity towards the AHC approaches, its usefulness in assessing SC is questionable. The connectivity of k-means is worst for the same reason as for the silhouette width.

Hubert's Γ **Statistic**: Hubert's Γ Statistic shows a poor result for single link due to the chaining effect (0.04). In the one large cluster, very dissimilar streamlines are grouped together leading to negative correlation values. The performance of k-means is considerably better (0.39) but still worse than for the remaining algorithms since the above-mentioned assignment of similar streamlines to different clusters leads to negative correlation values.

Complete link, SC, and Ward's method reach similar results on average $(0.44, 0.45, 0.48)$. The highest value is measured for average link by a rather narrow margin (0.52).

Stability: Single link's stability (0.97) is not expressive since the entire set of streamlines is always grouped in a single cluster. Complete linked achieves the lowest stability (0.59) due to the maximum computation in the proximity measure (Sec. 5.1). Since random subsamples are drawn from the original data to measure stability, different streamlines are missing each time. While the maximum computation is considerably affected by missing lines, the average and the variance computation in Average link and Ward's method, respectively are less sensitive $(0.79, 0.74)$. SC and k-means achieve the highest meaningful stability values $(0.82, 0.85)$. Both apply a global objective function and are hence, less sensitive to local changes than AHC. However, the stability of k-means is dependent on the number of runs ($= 10$, Sec. 5.2) and decreases to 0.72 for a single run. Even with a high number of runs, k-means may generate different results if started several times due to the random initialization of cluster centroids. The result of all AHC variants is dependent on the order of the input streamlines. If the proximity measure happens to be equal for two pairs of clusters, the first encountered pair is merged. However, we did not observe this problem.

Summary: Single link is not suitable for clustering blood flow due to the chaining effect which requires dedicated post-processing [10]. Complete link generates better clusters but tends to break large clusters. This has a negative impact on inter-cluster separation, which is reflected by lower silhouette widths. Further, the clustering results of complete link show a rather low stability. Stability becomes an important issue if the seeding density is varied, e.g., along the ostium, or in interactively sampling a region-of-interest by overlapping seeding regions, e.g., the aneurysmal near-wall region. In both cases, pairs of similar streamlines that survive the modifications should consistently be assigned to a joint cluster. K-means performed particularly poor with respect to the silhouette width and connectivity. Also, the stability of its clusters is less predictable due to the random initialization. Average link, Ward's method, and SC performed equally well except for the connectivity which is however biased towards AHC. An extended evaluation may investigate the overlap of their clustering results to gain further insight into their principles of operation and the data.

Average link's sensitivity to outliers was significantly reduced by our adapted streamline similarity measure (Sec. 4.2). While the original measure (Eq. 1) lead to small-sized, outlier-corrupted clusters (< 6 streamlines) in each dataset, this effect was only observed in three datasets with the new measure. Ward's method and SC proved to be rather insensitive to outliers. Overall, we recommend Average link, Ward's method, and SC for clustering blood flow. Visual blood flow summaries based on each of them are qualitatively evaluated by domain experts in Section 6.4.

6 Visual Summary of Blood Flow

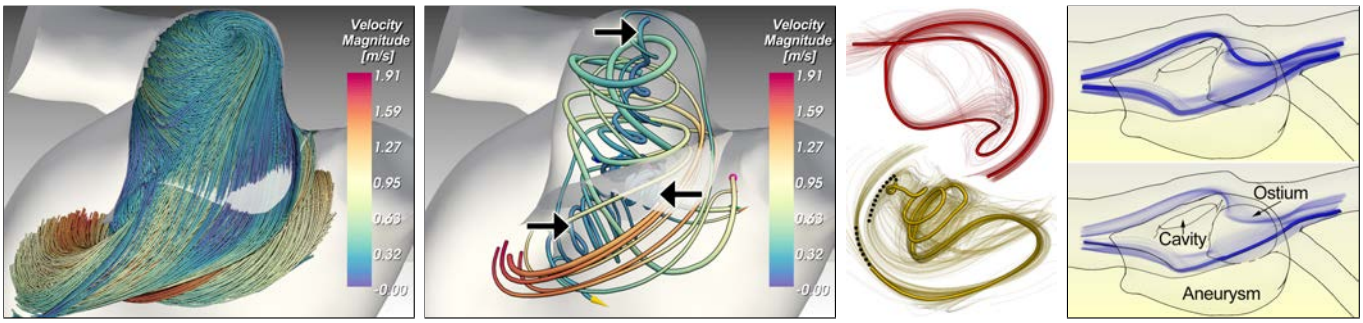
This section is dedicated to the computation of cluster representatives, their aggregation in a visual flow summary, the interaction with the summary, the expert evaluation of the summary, and our development environment.

6.1 Cluster Representatives

Displaying thousands of streamlines leads to a cluttered visualization hampering particularly the interpretation of inner flow structures (Fig. 4(a)). Cluster representatives summarize the flow and show these structures (Fig. 4(b)). In the context of clustering fiber tracts, different types of representatives have been discussed [33]. O'Donnel et al. employ Spectral Clustering and determine an *embedding-based* representative for each fiber bundle in spectral embedding space (Fig. 2(b)). The centroid of the bundle's point cloud is computed and the fiber closest to it is chosen. This is feasible due to the high density and number of embedded fibers (up to $25,000$ per brain). In our case, the streamline count is often < 3000. Furthermore, given a non-globular cluster, e.g., banana-shaped, the streamline closest to the cluster centroid may provide a weak representative.

As an alternative computed in the original 3D space, we chose the streamline with the smallest average distance to all other lines of the cluster. While often well representing the clusters, this *distance-based representative* is prone to outlier streamlines due to the outer minimum in the distance measure (Eq. 2). A short outlier, running very similar to all streamlines in its cluster, is assigned a small distance to all of them. Longer streamlines are more likely to deviate from the other lines in their cluster. Hence, the outlier is a more likely candidate for representative selection.

O'Donnell et al. propose another approach for computing representatives in 3D space [33]. For each cluster of fibers, a local Cartesian grid is aligned with the cluster's axis-aligned bounding box. For each voxel of the grid, the number of fibers that pass through is recorded leading to a density volume. For each fiber, the density is integrated along the line and the result is weighted with the fiber's length. The fiber with the highest value is the *density-based representative*. Several problems occur in transferring this approach to streamlines. The lines in a cluster may follow the same course over a long range but extend beyond either end of this range (Fig. 4(c), bottom). No line may exist that faithfully represents the entire cluster. The lines may also differ significantly in length. Furthermore, a few very long lines may exist in helical flow. Hence, we consider only density and for now neglect the weighting with length. Note that length is still inherently considered, since longer lines may accumulate more densities. The primarily density-based representatives well indicate the densest parts of the clusters which often occur in regions of helical or turbulent flow being of high interest. In an initial flow summary and in the remainder of this paper, we employ density-based representatives. However, the user may change the flow summary by modifying weights $[0, 1]$, which we assigned to density, length, and distance. For instance, setting the



(a) (b) (c) (d)

Fig. 4. *(a)* Full set of streamlines in a side-wall aneurysm. *(b)* Streamlines in (a) clustered according to geometry. One representative is displayed for each cluster ($n = 9$). A prominent swirl in the center of the aneurysm and laminar helical and complex flow below the ostium (transparent surface) are revealed (top, left, and right arrow). *(c)* Examples for good (red) and amendable (yellow) representatives. Dots indicate parts of the cluster which are not represented. *(d)* Flow around a cavity clustered according to local residence time. A selected cluster is visualized by semi-transparent streamlines. Its attribute-based representative indicates only the lower branch (bottom). A representation of cluster shape is obtained by further clustering based on streamline geometry (top).

weight of density to zero and the weight of length to one leads to length-based representatives, which may better illustrate the entire extent of the cluster.

If streamlines were clustered according to a streamline attribute, we employ *attribute-based representatives*. For each cluster, the mean of the attribute or of the statistic that has been employed is computed and the line with an attribute value closest to the mean is chosen (Sec. 4.3). The representative then indicates the clusters attribute range instead of its shape. Since the course of streamlines inside a cluster may be rather heterogeneous, we conduct a further partitioning according to streamline geometry (Fig. 4(d)).

6.2 Number of Clusters

A crucial question in generating the blood flow summary is how many representatives should be displayed, i.e. how many clusters must be computed? For blood flow data, the "correct" number of clusters is not known. Agglomerative Hierarchical Clustering (AHC) is well suited here since the cluster hierarchy may be cut at consecutive levels in order to interactively browse through a range of cluster numbers (Sec. 5.1). Spectral clustering (SC) and k-means require rerunning the algorithm each time. Merging and splitting clusters in AHC occurs locally in space and is hence easier to track visually. However, our practical experience with highly intertwined 3D streamline clusters shows that it is still difficult to grasp the change between consecutive cluster numbers without visual guidance.

We aim at minimizing the workload of physicians by making a "good guess" with respect to the number of clusters. A default number increases the reproducibility of our approach, which is a key requirement for entering clinical routine. Further, it facilitates a more standardized comparison of the flow before and after stenting and it supports a categorization of blood flow patterns. A good guess leads to clusters representing all significantly distinct flow structures – overrepresented structures are tolerable while missing structures are not – and each cluster is homogeneous such that the representative indeed represents all contained streamlines. Translated into clustering language, the inter-cluster separation and the intra-cluster cohesion should be high. We couple the quantitatively best performing streamline clustering techniques, AHC with average link, AHC with Ward's method, and SC (Sec. 5.4), with state-of-the-art techniques computing the number of clusters k that best satisfies both requirements.

Salvador and Chan propose the *L-method* for computing k in hierarchical clustering algorithms [45]. The method is based on detecting the *knee* in a graph that opposes numbers of clusters and a cluster evaluation metric. Since the location of the knee depends on the shape of the graph which again depends on the number of tested cluster numbers, a full evaluation graph, ranging from two clusters to the number of data elements, is recommended. We compute the full graph based on the evaluation metric suggested in [45]. Zelnik-Manor and Perona propose an algorithm for computing k in SC [42]. The algorithm iterates over a user-defined range $[a,b]$ for k and determines the optimal value. The optimization is based on finding the optimal rotation between the set of the first $k_i, i \in [a,b]$ largest eigenvectors of the Graph Laplacian (Eq. 4) and the canonical coordinate system. We empirically determined the range $[4,20]$ for detecting all relevant flow structures in ten datasets.

6.3 Visualization and Interaction

In the initial blood flow summary, cluster representatives corresponding to the optimal partitioning are shown (Fig. 4(b)). The user may inspect the suboptimal partitionings by browsing AHC's hierarchy or SC's range $[a,b]$. A representative can be picked causing the corresponding cluster to be displayed. For browsing all clusters, the user may scroll the mouse wheel. If the clustering was based on a streamline attribute, the set of geometry-based representatives per cluster is displayed after picking and during browsing (Sec. 6.1).

The streamline visualization is embedded in a surface rendering of the vessel wall. The wall is reconstructed from the unstructured grid of the CFD simulation. It is rendered opaque with culled front faces. The opaque back faces prevent a look through the aneurysm on lines in the near-vessel domain. The ostium and the stent surface are integrated. The ostium is rendered highly transparent.

Streamlines are rendered with GPU support as sets of quads and halos are added to improve spatial perception [46] (Fig. 4(a)). The halo color is either set to black or encodes the cluster ID. The latter is useful to distinguish clusters when the line color is modified according to a streamline attribute. However, our collaborators criticized the interference of halo and line color hampering the readability of the attribute. We initially color all halos in black and optionally allow an encoding of the cluster ID.

For visualizing the representatives, we evaluated stream ribbons and tubes. While ribbons additionally show rotation about the flow axis, color-mapped values are easier to read from tubes during a change of the viewing perspective. Our collaborators rated the readability as more important and hence, we employ tubes. In order to illustrate the flow direction, arrowhead glyphs are attached to the end of each tube pointing in outflow direction. The tube radius encodes the cluster size, i.e., the number of grouped streamlines. Halos are added to the representatives and initially colored in black. While this solves the color interference problem, it hampers visually tracking a tube through the set of highly intertwined representatives. Hence, we offer an optional color encoding of the cluster ID. Alternatively, only the halo of the representative under the pointer is colored according to cluster ID during mouse hover and the other representatives are rendered semi-transparently.

An important aspect is the coloring of streamlines and representatives. In geometry-based clustering, streamline color is modified according to a user-defined attribute. In attribute-based clustering, the statistic that has been employed for clustering is displayed per line, e.g., the maximum or mean of the attribute (Sec. 4.3). Two approaches are implemented for coloring the representatives: (1) simply copying the attribute values of the corresponding streamline, and (2) averaging the attribute values over all lines in the cluster. If the clustering has been based on streamline geometry, we apply (1) for attributes being defined as a series of values along each streamline and (2) for single scalar attributes. Note that (1) provides a reasonable approximation of the entire cluster for most flow attributes since their change in value is similar across all streamlines in the cluster due to the common underlying flow pattern. If the clustering has been based on an attribute, we directly apply (2) for single scalar attributes and for a series of values, we average over the statistic that has been employed in computing streamline similarity (Sec. 4.3).

6.4 Qualitative Evaluation

We let domain experts evaluate blood flow summaries generated by means of the quantitatively best performing streamline clustering techniques (Sec. 5.4): Agglomerative Hierarchical Clustering (AHC) with average link, AHC with Ward's method, and Spectral Clustering (SC). The number of clusters in the summary and hence, the number of representatives, has been computed automatically (Sec. 6.4). The evaluation is based on three clinical cases together comprising five datasets and representing the prevailing types of aneurysms. One case has been simulated without virtual stenting (Fig. 4(a)). Two cases have been simulated with and without stenting, one of them with two types of stents in two different positions (Sec. 7.1 and 7.2). For the latter case, we considered only the most beneficial type of stent and position. The blood flow summaries were evaluated by two board certified (BC) senior interventional neuroradiologists, a BC senior radiologist with a strong background in aortic aneurysms, two CFD engineers with a strong background in cerebral blood flow (one being coauthor of the paper), and one computer scientist working on experimental 7-Tesla Magnetic Resonance Imaging (MRI) of cerebral blood flow. The CFD engineers and one of the neuroradiologists participated in the Virtual Intracranial Stenting challenges in 2009 and 2010 (Sec. 7.1 and 7.2). The case without virtual stenting was stented by the neuroradiologist in real life.

Flow Summary: At first, the experts were asked to familiarize with the original data, i.e., the streamlines. All of them had seen streamline visualizations of blood flow before. However, the two neuroradiologists had no and only limited experience, respectively in interacting with such visualizations, e.g., filtering lines and probing by interactive seeding. The streamlines were visualized as in Figure 4(a). The experts could filter lines by thresholding their average distance to the vessel wall. This offered browsing through the lines from the vessel wall to the center in order to grasp the path of the flow through the near-vessel domain (Fig. 1(b)) and to detect characteristic flow structures, such as swirls. The experts were asked to sketch the flow path and annotate all structures that they consider to be relevant in a drawing of the aneurysmal silhouette.

Then, the flow summaries based on the three clustering algorithms were presented in a random, blinded side-by-side arrangement. In addition, a *control summary* was generated and mixed in to eliminate coincidence. This summary was generated based on a random number k of clusters, with k being in the range of the numbers computed for the three algorithms. Cluster size, the assignment of streamlines to clusters, and the selection of cluster representatives were also randomized. The experts were asked to rate each flow summary. Zero points were given if the sketched flow was in no way represented by the summary, one point was given if it was partially represented and two points in case of full representation. Finally, the experts should check whether the summary reveals other important patterns than they had discovered. Additional comments were recorded during the evaluation. The overall time exposure for the experts was ≈ 60 minutes.

The results of the evaluation are summarized in Table 2. SC consistently achieves the best results. Except for one

TABLE 2
Average expert ratings of blood flow summaries ($\in \{0,1,2\}$, 2=best). Comparison of Spectral Clustering (SC), Agglomerative Hierarchical Clustering (AHC) with average link (avg) and Ward's method, and random generation (RAND).

	Datasets (NVS=Not Virtually Stented,V09/V10=Virtual Intracranial Stenting Challenge 2009/10, S=SILK, R=right posterior cerebral artery)					
Algorithm	NVS	$V09$	$V09S$	$V10$	$V10S_R$	$\oslash$
SC	2.0	2.0	1.5	2.0	2.0	1.9
AHC_avg	1.5	1.8	2.0	2.0	1.2	1.7
AHC_ward	1.8	1.8	0.8	1.8	1.2	1.5
RAND	1.0	1.0	0.0	0.7	1.0	0.7

dataset, its flow summaries fully represent the flow sketched by the experts. For this specific dataset, half of the participants considered a swirl as "not really visible" (one point) while the other half considered it to be "slightly indicated" (two points). AHC with average link and with Ward's method show the second and third best results, respectively. However, Ward's method never achieves the full score on average for none of the datasets. The control summary (RAND) performs significantly worse than the rest, which confirms that the other summaries indeed provide non-random, meaningful insight. In 33 flow summaries out of 90 (5 datasets times 6 participants times 3 algorithms, excluding RAND), the experts detected more interesting flow patterns than they had discovered during streamline filtering further indicating the summary's benefit. The 33 summaries were generated in equal shares by the algorithms thus not indicating a unique feature.

Number of Clusters: The CFD engineers and the computer scientist were given an extra task before the assessment of the flow summaries. This time-consuming task did not fit into the tight schedule of the physicians since it extended the evaluation time to $90-120$ minutes. In a sequence, the flow summaries based on the range of possible numbers of clusters $[4,20]$ were presented and the experts were asked to select the number k_{sel} that fully represents their sketched flow, possibly shows more important flow structures, and is still clearly readable. To reduce time exposure, each expert assessed each dataset only based on one alternately chosen algorithm A with the control summary being left out (3 experts times 5 algorithms results in 15 ratings). After k_{sel} had been determined, the experts were asked to rate the flow summaries as explained above. Afterwards, the summary corresponding to A was pointed out and the expert was asked to compare the associated computed number of clusters k_{cmp} to k_{sel}.

For SC, k_{sel} was preferred once over k_{cmp}, namely for the only dataset for which SC's flow summary did not achieve the full score on average (Table 2, $V09S$). For both AHC with average link and AHC with Ward's method, k_{sel} was preferred three times over k_{cmp} since important flow structures were missing based on k_{cmp}. The remaining 8 comparisons assessed k_{cmp} as appropriate for generating an uncluttered summary, which is complete with respect to characteristic flow structures. In 5 (of 8) comparisons, these structures were overrepresented ($k_{cmp} > k_{sel}$) but still clearly visible. In the remaining 3 comparisons, k_{sel} was higher than k_{cmp} because one specific swirl was seen based on both but even more clearly based on k_{sel}.

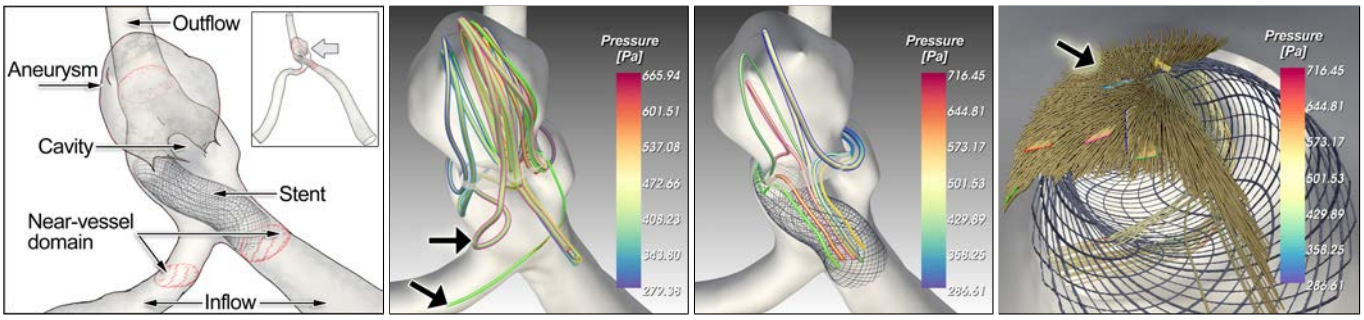
In conclusion, the blood flow summaries based on SC have achieved the best evaluation results by a narrow margin. The applied clustering algorithm, the number of clusters, and the type of representative effect the success of the summary. Hence, we recommend and employ in the remainder SC, its associated technique for computing a reliable number of clusters, and density-based representatives (Sec. 5.3, 6.1, 6.2). Since k_{cmp} was assessed as inappropriate in one case of SC, we offer interactively browsing the range of possible cluster numbers $[4,20]$ starting from k_{cmp}.

Anecdotal Feedback: All experts agreed that the flow summary is much faster to interpret than the entire set of streamlines and reveals flow features which are hidden inside the streamline clutter. They appreciated the workload reduction by avoiding the tedious iterative procedure of selectively seeding and/or filtering streamlines. Displaying streamline clusters on demand was rated as very valuable to get an impression of the spatial region that is represented by a cluster representative. Supporting the visual tracking of individual representatives by coloring the halo of the representative under the mouse pointer was preferred over temporarily modifying the halo color of all representatives according to cluster ID (Sec. 6.1). The physicians agreed that the comparison of flow before and after stenting is greatly simplified by the flow summaries in Figure 5 and 7.

6.5 Design Lessons

We carefully designed the flow summary in a tight feedback loop with our collaborators. The design lessons learned help other visualization practitioners working with similar data.

(1) Restrict the clustering domain to the region-of-interest. We restrict it to the aneurysm and the near-vessel domain. Otherwise, long sections of straight in- and outflow would lead to high streamline similarities while differences inside the aneurysm would have less impact (Sec. 4.1). (2) Choose a similarity measure that is less sensitive to streamline length if the course of streamlines is the primary concern. (3) Provide a good initial guess of the number of clusters since visually tracking the changes while browsing through different numbers of clusters is a tedious task especially for highly intertwined streamlines. (4) Use tubes as cluster representatives instead of ribbons if the readability of attribute values is crucial. (5) Add halos to streamlines and representatives in order to enhance their spatial perception. (6) Use black as halo color to avoid visual interference with color-coded streamline attributes. (7) Support visual tracking of tubes through a set of intertwined representatives by assigning a striking color to the halo of the representative under the mouse pointer. (8) Allow the user to see the original clusters since the representatives well encode the general course of the contained



(a) (b) (c) (d)

Fig. 5. VISC 2009. *(a)* Virtual stent placement, morphological features, subdivision of vascular domain (red circles) and flow conditions. *(b,c)* Streamlines clustered based on geometry (b) before and (c) after stenting. Arrows point at interesting differences, e.g., reflux (upper arrow). Flow from the right artery is not completely diverted (c). *(d)* Inside the artery. Flow bypassing the stent (arrow) reveals a gap between stent and vessel wall.

streamlines but fail in illustrating the cluster extent. (9) Encode the direction of the flow, e.g., by arrowhead glyphs. (10) Attribute-based clustering may require the computation of several representatives per cluster since the streamlines in a cluster may be quite heterogeneous with respect to their geometry (Fig. 8 (b,d)).

6.6 Development Environment

The clustering algorithms, the similarity measures and the computation of cluster representatives are implemented in MATLAB (MathWorks, Natick, MA, U.S.). Source code for local scaling and determining the number of clusters is provided by Zelnik-Manor and Perona [47]. All MATLAB code is exported as a shared library and accessed from custom C++ code. The three categories of streamline attributes are computed using (1) ANSYS Fluent 12 and ParaView, (2) the Vascular Modeling Toolkit (www.vmtk.org), and (3) custom C++ code (Sec. 4.3). The visualization is implemented in C++ and the Visualization Toolkit (Kitware, Inc., Clifton Park, NY, U.S.).

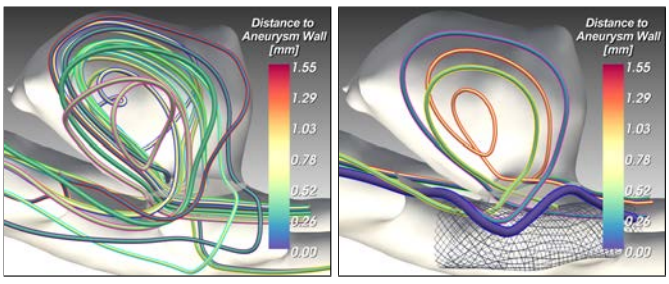
7 Application

We applied our approach to data of the Virtual Intracranial Stenting Challenges (VISC) in 2009 and 2010 [48]. Please consider the following advices when reading the figures of this section. The color scales refer to the representatives, not their halos. The annotated range of values is based on the entire set of streamlines. Halo colors must not be employed for establishing correspondence between clusters in different figures or figure parts. They are assigned independently to each clustering result and simplify the visual tracking of representatives in a non-interactive display.

7.1 Virtual Intracranial Stenting Challenge 2009

For the VISC 2009, teams were invited to compete in predicting stenting success based on simulated hemodynamic data. Two cases and a model description of the flow diverting *SILK* stent (Balt, Montmorency, France) were provided. Due to space restrictions, we only discuss the first case with a saccular side-wall aneurysm located at a bifurcation (Fig. 5(a)). A rare anatomical variant is the cavity (*fenestration*) behind the aneurysm. Our medical collaborators suggested placing the stent in the right artery and circumventing the aneurysm to the left. The stent geometry was modeled in a CAD program and manually fitted to the vessel wall. The hemodynamic data generation resulted in volume meshes with 4.3 and 4.6 (with stent) million tetrahedral elements (Sec. 2.3). The meshes constituted the input for streamline generation (Sec. 4.1).

The resulting lines have been clustered based on geometry (Sec. 4.2). The flow summaries are displayed in Figure 5 (b,c). A higher number of clusters can be observed in the untreated aneurysm indicating a more complex flow pattern (Fig. 5(b)). After stenting, the flow is less complex which decreases the risk of aneurysm rupture [5]. In the stented configuration, flow arriving from the right artery is not completely diverted but still enters the aneurysm (Fig. 5(c)). A closer look from inside the vessel at the location where this flow enters the stent reveals that the stent model does not perfectly adhere to the vessel wall (Fig. 5(d)). A considerable gap exists through which flow with high pressure is bypassing the stent. A neuroradiologist commented that such gaps indeed occur in real stenting due to a sharp bending of the vessel. Their prediction would be of great value. The flow that travels through the virtual stent, exits the stent at its aneurysm-near inflection point



(a) (b)

Fig. 6. Clustering streamlines according to their mean distance to the aneurysm wall before *(a)* and after virtual stenting *(b)*. Stenting reduces near-wall flow.

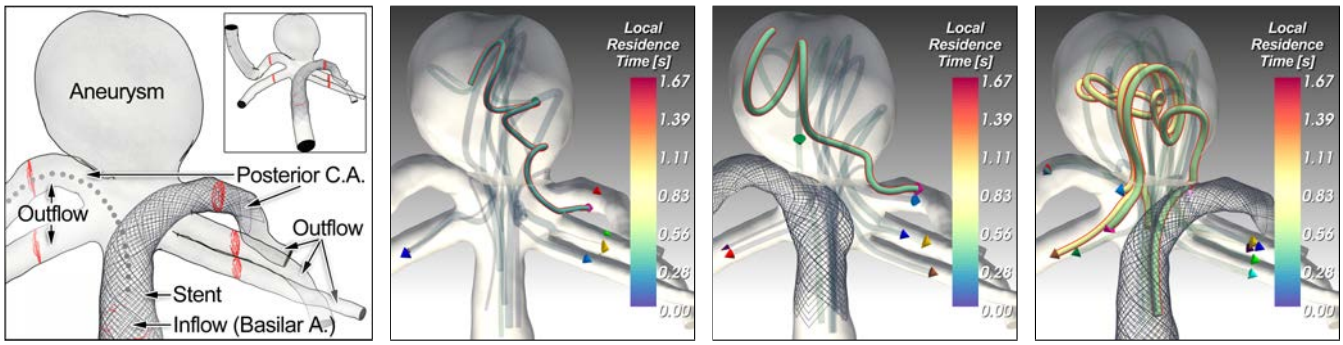


Fig. 7. VISC 2010. *(a)* Virtual stent placement, morphological features, subdivision of the vascular domain (red circles) and flow conditions. *(b-d)* Clustering of streamlines based on geometry before (b) and after stenting (c,d). Representatives indicating a major difference between the flow patterns are rendered opaque. While a "simple" swirl is characteristic for the first two patterns (b,c), a double helical swirl is observed in the third one (d).

and enters the aneurysm (Fig. 5(c)). This may be mitigated by a higher general or local mesh density. Before stenting, reflux is observed below the ostium (Fig. 5(b), top arrow). Furthermore, flow is entering the aneurysm from the left artery with high pressure (Fig. 5(b), bottom arrow). After stenting, this flow is obstructed by the stent and circumvents the aneurysm. This is a convenient side effect of diverting the flow arriving from the right branch.

A comparison of the aneurysmal wall-shear stress (WSS) before and after stenting revealed lower values in the latter which indicates a benefit. We investigate the near-wall flow by clustering the streamlines based on their mean distance to the aneurysm wall (Sec. 4.3). The results before and after stenting are compared in Figure 6. To support a visual comparison, the color mapping and the radius scaling of the representatives after stenting are applied uniformly to both configurations. The comparison shows that more flow hits the wall and is traveling through the near-wall region before stenting. This is in accordance with the higher WSS [49]. After stenting, a considerable amount of the flow barely enters the aneurysm (thick blue tube in Fig. 6(b)). Note that attribute-based representatives have been applied well indicating a cluster's range of attribute values (Sec. 6.1).

7.2 Virtual Intracranial Stenting Challenge 2010

For the VISC in 2010, research teams were invited to find the optimal placement of a stent in treating a basilar tip aneurysm (Fig. 7 (a)). We considered two types of stents and two different positions, both covering the end of the basilar artery and then extending to the beginning of the left and the right posterior cerebral artery (LPCA/RPCA), respectively. We restrict our discussion to the most beneficial type of stent (SILK). The hemodynamic data of the two stented configurations and the untreated case has been generated as described in Section 2.3. The biggest tetrahedral mesh consists of 13.5 million elements (including stent). The stent geometry was modeled in a CAD program. Learning from the issues of a manual stent deployment (Sec. 7.1), we applied an automatic wall-tight deployment using polyharmonic splines for free-form deformation [15].

For the detection of flow structures in the untreated aneurysm and in the two stented configurations, the near-vessel domain is specified (Fig. 7 (a)) and the data is cropped. Then, streamlines are seeded at the ostium and clustered based on geometry. Cluster representatives are displayed and colored according to local residence time (RT, Sec. 4.3). The color scale has been set for all configurations to mapping the range of RT in the untreated configuration (Fig. 7 (b-d)). Thus, regions of prolonged RT after stenting can be easily spotted. Before we focus on RT, we study the detected flow structures.

In Figure 7 (b-d), representatives indicating a major difference between the flow patterns are rendered opaque. Before and after stenting along the LPCA, parts of the flow enter the aneurysm and after a swirling motion inside, exit via the RPCA (Fig. 7 (b,c)). Stenting along the RPCA considerably alters the flow pattern and generates a double helical swirl in the center of the aneurysm. A closer look at the highlighted representative(s) of each configuration revealed that they always represent those clusters with the highest RT values on average. Comparing their coloring indicates that SILK stenting along the RPCA causes the most prolonged RT and hence represents the preferred strategy (Fig. 7(d)). Further evidence is given by plotting the percentage of streamlines over discrete RT values (Fig. 8 (a)) and by Janiga et al. [15] who report the most prolonged *turnover time* for this configuration. The turnover time is a global scalar measure which is proportional to RT and both characterize intra-aneurysmal flow stasis [50]. In the following, we focus on stenting along the RPCA.

In order to investigate RT more locally, the streamlines of the stented configuration have been clustered based on it. The cluster with the highest RT values is shown in Figure 8 (b). Its streamlines are rendered semi-transparent such that the inner swirl is easier to perceive. Flow enters the aneurysm, is attracted by opposing wall parts, converges in a swirl in the center, and leaves the aneurysm (the swirling motion is also indicated in Fig. 7(d)). Since correspondences between a low WSS and thrombosis development as well as between a high RT and thrombosis development

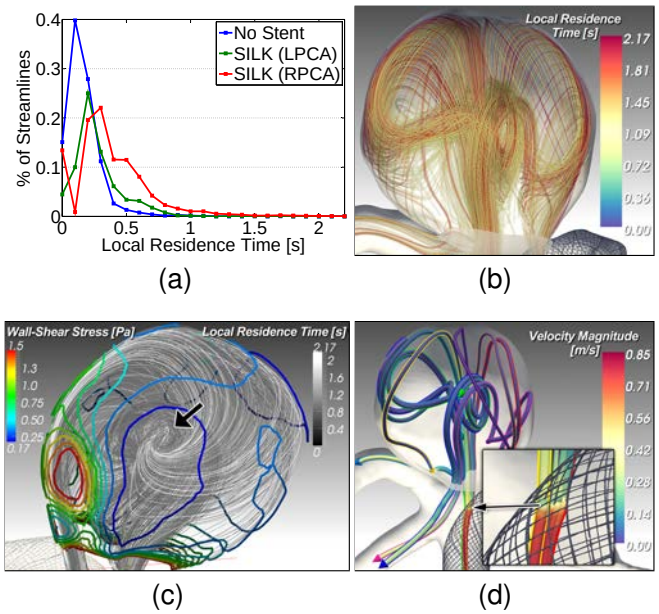


Fig. 8. VISC 2010. *(a)* Comparison of local residence times (RT) before and after stenting along the left and right posterior cerebral artery (LPCA, RPCA). The % of streamlines is plotted over discrete RT values. Stenting causes prolonged RT. *(b)* Cluster with the highest average RT in RPCA stenting. *(c)* Investigating this cluster in the context of wall-shear stress (iso-contours). *(d)* Partitioning the cluster based on streamline geometry. Flow is strongly decelerated at the stent wires (inset).

are known [4], the cluster has been further investigated in the context of WSS (Fig. 8 (c)). WSS is mapped to the surface of the aneurysm and visualized by contour lines. In agreement with [4], a value of 1.5 is chosen as the upper limit for color mapping. Values above are clamped to this limit. As can be observed, low WSS values occur in a large region where the flow streaks the wall (arrow). The same is true for the opposite side of the wall. It should be further investigated whether these regions are potential candidates for thrombosis initiation and whether a double helical swirl particularly encourages flow stasis.

It was shown in [15] that the SILK stent diverts a considerable amount of blood. However, parts of the blood flow exit the wired mesh and enter the aneurysm. In Figure 8 (d), the cluster with the highest values of RT is refined by a clustering based on geometry and the new representatives are colored according to velocity magnitude. The structure of the swirl is easier to perceive as compared to Figure 8 (b). Furthermore, it can be observed that the flow exiting the stent is strongly decelerated at its wires thus leading to a prolonged RT and a slow inflow (strong red to green jump in the inset of Fig. 8 (d)).

7.3 Performance

This section reports on the performance of our approach. The focus is on computation time since memory consumption is not critical. The time is dependent on the number of streamlines and their number of vertices (columns 2-3 in Table 3). While the first varies with the sampling density of the ostium, the latter depends on the streamline length and integration step size (Sec. 4.1). We measured the computation time of Spectral Clustering and of the visualization. In clustering, we differentiated between the computation of the distance matrix and the actual clustering. The latter also comprises the determination of cluster representatives. In distance matrix computation, we focused on geometry-based distances since attribute-based distances are much faster to compute. The timings were taken on a 3.07GHz Intel 8-core PC with 8GB RAM and a 64bit Windows operating system (Table 3).

TABLE 3
Dataset characteristics and timings $[s]$ of geometry-based clustering and visualization. V09/V10=Virtual Intracranial Stenting Challenge 2009/2010, S=SILK stent, N=Neuroform stent, L/R=left/right posterior cerebral artery.

Dataset	#Stream-lines	#Vertices (∅)	Distance Matrix	Clustering	Visualiza-tion
$V09$	2254	505	4547	34.7	9.5
$V09S$	2207	249	732	29.9	4.4
$V10$	2929	265	1567	57.2	6.2
$V10N_R$	2923	275	1581	55.9	6.8
$V10N_L$	1153	283	256	12.8	4.1
$V10S_R$	2891	234	1128	51.8	7.4
$V10S_L$	1138	212	142	14.1	3.7

As expected, the computation of the distance matrix represents the bottleneck. However, the matrix can be reused for different clustering settings. In attribute-based clustering, the time for computing the matrix depends on the applied statistic (Sec. 4.3). For simple statistics such as min/max, the computation is two orders of magnitude faster than in geometry-based clustering. The timings for the clustering itself are in the range of seconds. The most time-consuming part of the visualization is the geometry computation for the GPU-based streamline rendering.

8 Summary and Discussion

We presented an approach for reducing visual clutter in streamline visualizations of simulated blood flow. The approach is based on clustering streamlines and computing cluster representatives, which are compiled into a flow summary. To determine the most appropriate clustering algorithm, we carried out a quantitative evaluation of Spectral Clustering (SC), four variants of Agglomerative Hierarchical Clustering (AHC), and k-means. Based on cluster validity measures, we identified SC and AHC with average link and Ward's method, respectively as superior. In an expert evaluation of blood flow summaries generated by these algorithms, SC achieved the best ratings by a narrow margin. Its summaries are complete with respect to the relevant flow structures in most cases. In a tight feedback loop with our collaborators, we carefully designed the flow summary. The design lessons learned help scientists, e.g., in exploring flow in other vascular pathologies.

The computation of the summary is fully automatic. Only a range for the number of clusters that possibly exist in the dataset must be provided. The optimal number is identified automatically. We empirically determined a range of $[4, 20]$ for detecting all relevant flow structures in ten datasets. The time needed for the clustering and the visualization is within the range of minutes. Compared to the Computational Fluid Dynamics (CFD) simulation, which takes hours, it is of little consequence with respect to a possible therapeutic workflow.

Results from CFD simulations are not yet part of the clinical decision pipeline although they can be generated within a clinically acceptable time frame for planning an intervention. Neuroradiologists have little experience in investigating flow data. Our flow summary simplifies the access to the data, it is easier to read than full streamline visualizations, and it contributes to the communication between CFD engineers and physicians. The latter is of crucial importance in understanding "How stent properties affect flow patterns?", "How the change in flow patterns after stenting is related to treatment success?", and "How flow patterns are related to the risk of aneurysm rupture and the development of thrombosis?". Once these questions can be answered, stenting may not be planned solely based on the coverage of the aneurysm neck by the stent, but also based on CFD results and the flow summary. The concept of the summary can be readily transferred to (virtual) coiling. However, the joint visualization of coils and cluster representatives will cause serious occlusion problems.

The success of virtual stenting is so far evaluated based on global measures, e.g., the turnover time. However, if a certain stented configuration does not indicate a benefit for the patient, global measures fail to explain why. We cluster streamlines also based on locally derived domain-specific attributes, e.g., the distance to the aneurysm wall and the local residence time (RT). The latter was considered a useful extension to the turnover time. Clusters with a high RT may forecast locations of thrombosis initiation.

A limitation of our approach is that a few cluster representatives do not capture the entire structure of their cluster. They faithfully represent its densest part but fail to represent all parts in the in- and outflow regions of the near-vessel domain. Hence, the clusters itself should also be inspected.

ACKNOWLEDGMENTS

We thank Cordula Scherlach, Oliver Beuing, Martin Skalej (University Hospital Magdeburg, Germany), Uta Preim (Community Hospital Magdeburg), Daniel Stucht (Institute of Biomedical Magnetic Resonance, University Magdeburg), and Philipp Berg (Institute of Fluid Mechanics and Thermodynamics, University Magdeburg) for participating in the evaluation and providing valuable feedback. We are grateful to Christian Rössl and Max Zimmermann (University Magdeburg) for discussions on clustering issues.

REFERENCES

[1] R. D. Brown, J. Huston, R. Hornung, T. Foroud, D. F. Kallmes, D. Kleindorfer, I. Meissner, D. Woo, L. Sauerbeck, and J. Broderick, "Screening for Brain Aneurysm in the Familial Intracranial Aneurysm Study: Frequency and Predictors of Lesion Detection," *Neurosurgery: Pediatrics*, vol. 108, no. 6, pp. 1132–1138, 2008.

[2] I. Larrabide, M. L. Aguilar, H. G. Morales, A. J. Geers, Z. Kulcsár, D. Rüfenacht, and A. F. Frangi, "Intra-Aneurysmal Pressure and Flow Changes Induced by Flow Diverters: Relation to Aneurysm Size and Shape," *Am J Neuroradiol*, vol. 34, no. 4, pp. 816–822, 2012.

[3] J. Cebral, F. Mut, M. Raschi, E. Scrivano, R. Ceratto, P. Lylyk, and C. Putman, "Aneurysm rupture following treatment with flow-diverting stents: Computational hemodynamics analysis of treatment," *Am J Neuroradiol*, vol. 32, no. 1, pp. 27–33, 2011.

[4] V. Rayz, L. Boussel, L. Ge, J. Leach, A. Martin, M. Lawton, C. McCulloch, and D. Saloner, "Flow residence time and regions of intraluminal thrombus deposition in intracranial aneurysms." *Ann Biomed Eng*, vol. 38, no. 10, pp. 3058–69, 2010.

[5] J. R. Cebral, F. Mut, J. Weir, and C. M. Putman, "Association of Hemodynamic Characteristics and Cerebral Aneurysm Rupture." *Am J Neuroradiol*, vol. 32, no. 2, pp. 264–270, 2011.

[6] S. Appanaboyina, F. Mut, R. Löhner, C. Putman, and J. Cebral, "Simulation of intracranial aneurysm stenting: Techniques and challenges," *Computer Methods in Applied Mechanics and Engineering*, vol. 198, no. 45-46, pp. 3567 – 3582, 2009.

[7] M. Kim, E. I. Levy, H. Meng, and L. N. Hopkins, "Quantification of hemodynamic changes induced by virtual placement of multiple stents across a wide-necked basilar trunk aneurysm." *Neurosurgery*, vol. 61, no. 6, pp. 1305–1312; discussion 1312–1313, 2007.

[8] A. R. Mantha, G. Benndorf, A. Hernandez, and R. W. Metcalfe, "Stability of pulsatile blood flow at the ostium of cerebral aneurysms." *J Biomech*, vol. 42, no. 8, pp. 1081–1087, May 2009.

[9] C.-K. Chen, S. Yan, H. Yu, N. Max, and K.-L. Ma, "An illustrative visualization framework for 3d vector fields." *Comput. Graph. Forum*, vol. 30, no. 7, pp. 1941–1951, 2011.

[10] H. Yu, C. Wang, C.-K. Shene, and J. H. Chen, "Hierarchical streamline bundles." *IEEE Trans Vis Comput Graph*, vol. 18, no. 8, pp. 1353–67, 2012.

[11] C. Rössl and H. Theisel, "Streamline Embedding for 3D Vector Field Exploration," *IEEE Trans Vis Comput Graph*, vol. 18, no. 3, pp. 407–420, 2012.

[12] T. McLoughlin, M. W. Jones, R. S. Laramee, R. Malki, I. Masters, and C. D. Hansen, "Similarity Measures for Enhancing Interactive Streamline Seeding." *IEEE Trans. Vis. Comput. Graphics*, vol. 19, no. 8, pp. 1342–53, 2013.

[13] M. Forsting and I. Wanke, *Intracranial Vascular Malformations and Aneurysms: From Diagnostic Work-Up to Endovascular Therapy*, ser. Medical Radiology. Springer, 2008.

[14] N. Andaluz and M. Zuccarello, "Recent Trends in the Treatment of Cerebral Aneurysms: Analysis of a Nationwide Inpatient Database," *Neurosurgery: Pediatrics*, vol. 108, no. 6, pp. 1163–1169, 2008.

[15] G. Janiga, C. Rössl, M. Skalej, and D. Thévenin, "Realistic Virtual Intracranial Stenting and Computational Fluid Dynamics for Treatment Analysis," *J Biomech*, vol. 46, no. 1, pp. 7–12, 2012.

[16] S. Seshadhri, G. Janiga, O. Beuing, M. Skalej, and D. Thévenin, "Impact of stents and flow diverters on hemodynamics in idealized aneurysm models." *J Biomech Eng*, vol. 133, no. 7, p. 071005, 2011.

[17] R. Gasteiger, M. Neugebauer, O. Beuing, and B. Preim, "The FLOWLENS: A Focus-and-Context Visualization Approach for Exploration of Blood Flow in Cerebral Aneurysms," *IEEE Trans. Vis. Comput. Graphics*, vol. 17, no. 12, pp. 2183–2192, 2011.

[18] J. Schöberl, "NETGEN An Advancing Front 2D/3D-Mesh Generator Based on Abstract Rules," *Computing and Visualization in Science*, vol. 1, pp. 41–52, 1997.

[19] M. Neugebauer, G. Janiga, O. Beuing, M. Skalej, and B. Preim, "Anatomy-Guided Multi-Level Exploration of Blood Flow in Cerebral Aneurysms," *Comput. Graph. Forum*, vol. 30, no. 3, pp. 1041–1050, 2011.

[20] J. R. Cebral, F. Mut, J. Weir, and C. M. Putman, "Quantitative Characterization of the Hemodynamic Environment in Ruptured and Unruptured Brain Aneurysms," *Am J Neuroradiol*, vol. 32, no. 1, pp. 145–151, 2011.

[21] F. H. Post, B. Vrolijk, H. Hauser, R. S. Laramee, and H. Doleisch, "The State of the Art in Flow Visualisation: Feature Extraction and Tracking," *Comput. Graph. Forum*, vol. 22, no. 4, pp. 775–792, 2003.

[22] T. Salzbrunn, H. Jänicke, T. Wischgoll, and G. Scheuermann, "The state of the art in flow visualization: Partition-based techniques," in *SimVis*, 2008, pp. 75–92.

[23] R. F. P. van Pelt, S. S. A. M. Jacobs, B. M. ter Haar Romeny, and A. Vilanova, "Visualization of 4D Blood-Flow Fields by Spatiotemporal Hierarchical Clustering," *Comp. Graph. Forum*, vol. 31, no. 3pt2, pp. 1065–1074, 2012.
[24] R. S. Laramee, H. Hauser, L. Zhao, and F. H. Post, "Topology-Based Flow Visualization, the State of the Art," in *Topology-based Methods in Visualization*, ser. Mathematics and Visualization, H. Hauser, H. Hagen, and H. Theisel, Eds. Springer, 2007, pp. 1–19.
[25] A. Pobitzer, R. Peikert, R. Fuchs, B. Schindler, A. Kuhn, H. Theisel, K. Matković, and H. Hauser, "On the Way Towards Topology-Based Visualization of Unsteady Flow – the State of the Art," in *State of the Art Reports*, H. Hauser and E. Reinhard, Eds. Eurographics Association, 2010, pp. 137–154.
[26] T. Salzbrunn and G. Scheuermann, "Streamline predicates," *IEEE Trans. Vis. Comput. Graphics*, vol. 12, no. 6, pp. 1601 –1612, 2006.
[27] S. Born, M. Pfeifle, M. Markl, and G. Scheuermann, "Visual 4D MRI Blood Flow Analysis with Line Predicates," in *Pacific Vis*, 2012, pp. 105–112.
[28] K. Shi, H. Theisel, H. Hauser, T. Weinkauf, K. Matković, H.-C. Hege, and H.-P. Seidel, "Path line attributes - an information visualization approach to analyzing the dynamic behavior of 3D time-dependent flow fields," in *Topology-Based Methods in Visualization II*, ser. Mathematics and Visualization. Springer, 2009, pp. 75–88.
[29] A. Pobitzer, A. Lež, K. Matković, and H. Hauser, "A Statistics-based Dimension Reduction of the Space of Path Line Attributes for Interactive Visual Flow Analysis," in *Pacific Vis*, 2012, pp. 113–120.
[30] R. van Pelt, J. Olivan Bescos, M. Breeuwer, R. Clough, M. Groller, B. ter Haar Romeny, and A. Vilanova, "Interactive Virtual Probing of 4D MRI Blood-Flow," *IEEE Trans. Vis. Comput. Graphics*, vol. 17, no. 12, pp. 2153–2162, 2011.
[31] R. Gasteiger, D. J. Lehmann, R. van Pelt, G. Janiga, O. Beuing, A. Vilanova, H. Theisel, and B. Preim, "Automatic Detection and Visualization of Qualitative Hemodynamic Characteristics in Cerebral Aneurysms," *IEEE Trans. Vis. Comput. Graphics*, vol. 18, no. 12, pp. 2178–2187, 2012.
[32] L. J. O'Donnell and C.-F. Westin, "Automatic tractography segmentation using a high-dimensional white matter atlas," *IEEE Trans. Med. Imag.*, vol. 26, no. 11, pp. 1562–1575, 2007.
[33] L. O'Donnell, A. J. Golby, and C.-F. Westin, "Tract-based morphometry for white matter group analysis," *NeuroImage*, vol. 45, pp. 832–844, 2009.
[34] B. Moberts, A. Vilanova, and J. van Wijk, "Evaluation of Fiber Clustering Methods for Diffusion Tensor Imaging," in *IEEE Visualization*, 2005, pp. 65 – 72.
[35] S. Zhang, S. Correia, and D. Laidlaw, "Identifying white-matter fiber bundles in dti data using an automated proximity-based fiber-clustering method," *IEEE Trans. Vis. Comput. Graphics*, vol. 14, no. 5, pp. 1044 –1053, 2008.
[36] I. Corouge, S. Gouttard, and G. Gerig, "Towards a Shape Model of White Matter Fiber Bundles Using Diffusion Tensor MRI," in *ISBI*, 2004, pp. 344–347.
[37] P.-N. Tan, M. Steinbach, and V. Kumar, *Introduction to Data Mining*. Addison Wesley, 2005.
[38] U. von Luxburg, "A tutorial on spectral clustering," *Statistics and Computing*, vol. 17, no. 4, pp. 395–416, 2007.
[39] J. Handl, J. Knowles, and D. B. Kell, "Computational cluster validation in post-genomic data analysis." *Bioinformatics*, vol. 21, no. 15, pp. 3201–3212, Aug 2005.
[40] L. Ferreira and D. B. Hitchcock, "A comparison of hierarchical methods for clustering functional data," *Communications in Statistics - Simulation and Computation*, vol. 38, no. 9, pp. 1925–1949, 2009.
[41] J. Shi and J. Malik, "Normalized Cuts and Image Segmentation," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 22, no. 8, pp. 888–905, 2000.
[42] L. Zelnik-Manor and P. Perona, "Self-tuning Spectral Clustering," in *Advances in Neural Information Processing Systems 17*. MIT Press, 2004, pp. 1601–1608.
[43] L. Hubert and P. Arabie, "Comparing partitions," *Journal of classification*, vol. 2, no. 1, pp. 193–218, 1985.
[44] J. Klein, P. Bittihn, P. Ledochowitsch, H. K. Hahn, O. Konrad, J. Rexilius, and H.-O. Peitgen, "Grid-based Spectral Fiber Clustering," in *SPIE Medical Imaging: Visualization and Image-Guided Procedures*, vol. 6509, 2007, pp. 65 091E–65 091E–10.
[45] S. Salvador and P. Chan, "Determining the Number of Clusters/Segments in Hierarchical Clustering/Segmentation Algorithms." in *International Conference on Tools with Artificial Intelligence (ICTAI)*, 2004, pp. 576 – 584.
[46] M. H. Everts, H. Bekker, J. B. Roerdink, and T. Isenberg, "Depth-dependent halos: Illustrative rendering of dense line data," *IEEE Trans. Vis. Comput. Graphics*, vol. 15, no. 6, pp. 1299–1306, 2009.
[47] L. Zelnik-Manor and P. Perona, "Self-Tuning Spectral Clustering Code." [Online]. Available: www.vision.caltech.edu/lihi/Demos/SelfTuningClustering.html
[48] [Online]. Available: http://www.ics-meeting.net/ics_visc_past_.html
[49] G. Janiga, O. Beuing, S. Seshadhri, M. Neugebauer, R. Gasteiger, B. Preim, G. Rose, M. Skalej, and D. Thévenin, "Virtual Stenting Using Real Patient Data," in *International Conference on Fluid Flow Technologies*, 2009, pp. 111–117.
[50] M. Kim, D. B. Taulbee, M. Tremmel, and H. Meng, "Comparison of two stents in modifying cerebral aneurysm hemodynamics." *Ann Biomed Eng*, vol. 36, no. 5, pp. 726–41, 2008.

Steffen Oeltze is a Postdoctoral Research Fellow at the Computer Science Department at the University of Magdeburg, Germany. In 2004, he received the diploma in Computational Visualistics and in 2010, a Ph.D. in Computer Science from the University of Magdeburg. His research interests are in the visual analysis and exploration of medical, biological, and epidemiological data.

Dirk J. Lehmann is a Postdoctoral Research Fellow at the Computer Science Department at the University of Magdeburg, Germany. In 2008, he received the M.Sc. in Computational Visualistics and in 2012, a Ph.D. in Computer Science from the University of Magdeburg. His research interests focus on flow and volume visualization as well as information visualization and visual analytics.

Alexander Kuhn is a Postdoctoral Research Fellow at the Department of Visualization and Data Analysis at the Zuse Insitute Berlin (ZIB), Germany. In 2009, he received the M.Sc. in Computational Visualistics and in 2012, a Ph.D. in Computer Science from the University of Magdeburg. His research interests are in the visualization, segmentation, and analysis of flow data.

Gábor Janiga is researcher (Privatdozent) at the Laboratory of Fluid Dynamics and Technical Flows at the University of Magdeburg, Germany. In 1998, he received his diploma and in 2002 a Ph.D. from the University of Miskolc (Hungary), and a habilitation (venia legendi) in 2011 from the University of Magdeburg. His research interests are in the simulation of blood flow in vascular pathologies and in virtual stenting.

Holger Theisel is professor for Visual Computing at the Computer Science Department at the University of Magdeburg, Germany. In 1994, he received the diploma in Computer Science, in 1996 a Ph.D. in Computer Science, and a habilitation (venia legendi) in 2001 from the University of Rostock. His research interests focus on flow and volume visualization as well as on CAGD, geometry processing and information visualization.

Bernhard Preim is professor for visualization at the Computer Science Department at the University of Magdeburg, Germany. In 1994, he received the diploma in Computer Science and in 1998 a Ph.D. in Computer Science from the University of Magdeburg, and a habilitation (venia legendi) in 2002 from the University of Bremen. His research interests are in medical visualization and applications in diagnosis and treatment.

Cluster Analysis of Vortical Flow in Simulations of Cerebral Aneurysm Hemodynamics

Steffen Oeltze-Jafra, *Member, IEEE*, Juan R. Cebral, Gábor Janiga, and Bernhard Preim

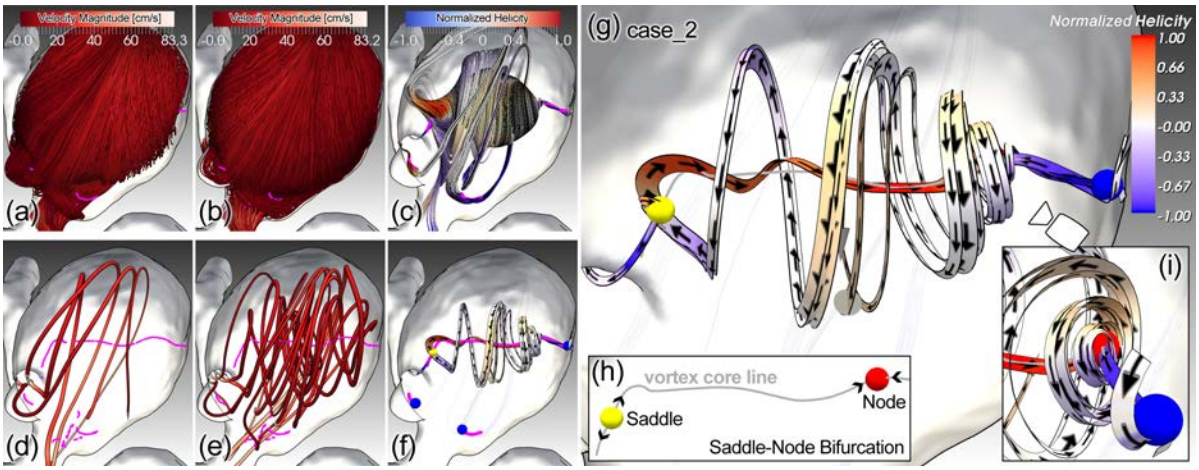


Fig. 1. Vortical flow in aneurysm with major center swirl (case #2, Fig. 9(b)). Vortex core lines are rendered magenta in *(a-f)*. Neither streamlines seeded at the aneurysm inlet *(a)* or inside *(b)* nor the representatives *(d-e)* of their clustering results [32] clearly convey the swirl. We propose a new approach to grouping streamlines seeded at vortex core lines *(c)* comprising custom group representatives *(f)*. It reveals an *embedded vortex (g)* — a small vortex enveloped by a larger one swirling in the opposite direction — forming around a *saddle-node bifurcation (h)*. In *(g-i)*, arrows indicate local flow direction and the vortex core line of the major swirl is rendered gray.

Abstract—Computational fluid dynamic (CFD) simulations of blood flow provide new insights into the hemodynamics of vascular pathologies such as cerebral aneurysms. Understanding the relations between hemodynamics and aneurysm initiation, progression, and risk of rupture is crucial in diagnosis and treatment. Recent studies link the existence of vortices in the blood flow pattern to aneurysm rupture and report observations of *embedded vortices* — a larger vortex encloses a smaller one flowing in the opposite direction — whose implications are unclear.

We present a clustering-based approach for the visual analysis of vortical flow in simulated cerebral aneurysm hemodynamics. We show how embedded vortices develop at *saddle-node bifurcations* on vortex core lines and convey the participating flow at full manifestation of the vortex by a fast and smart grouping of streamlines and the visualization of group representatives. The grouping result may be refined based on spectral clustering generating a more detailed visualization of the flow pattern, especially further off the core lines. We aim at supporting CFD engineers researching the biological implications of embedded vortices.

Index Terms—Blood Flow, Aneurysm, Clustering, Vortex Dynamics, Embedded Vortices

1 Introduction

Cerebral aneurysms represent a type of cerebrovascular disorder in which a weakening of the arterial wall leads to a balloon-like dilation (Fig. 2(a)). The prevalence of unruptured cerebral aneurysms in the general population has been estimated as 3.2% [48]. Their rupture is associated with a mortality rate of $\approx 50\%$.

Computational fluid dynamic (CFD) simulations of blood flow play a crucial role in understanding aneurysm rupture and evaluating its risk since they provide insights into the aneurysm hemodynamics [4, 7, 17]. Hemodynamic parameters are evaluated as predictors for aneurysm rupture together with geometrical descriptors, wall properties, inflammatory effects, genetic predispositions, and behavioral factors. The blood flow pattern and in particular, the formation of vortices are among the hemodynamic parameters that were linked to rupture, thus motivating a detailed investigation of swirling flow [4, 7].

Previous work reports the frequent observation of *embedded vortices* — one vortex enveloped by a second one swirling in the opposite direction — in unsteady CFD simulations [3] (Fig. 1 (g)). It relates their formation and collapse over the cardiac cycle to *saddle-node bifurcations* — the collision and annihilation of a saddle and a node located on the *vortex core line* (Fig. 1 (h)). We aim at supporting CFD engineers in revealing the biological implications of embedded vortices. Here, we study the structure of these vortices at full manifestation based on a representative point of the cardiac cycle.

The investigation of embedded vortical flow benefits from an extraction of the vortex core line and the seeding of integral curves in its vicinity (Fig. 1(a-f)). In [3], groups of streamlines are integrated from manually selected core line segments which were extracted by Sujudi and Haimes' method [45]. The segments are selected and integration and visualization parameters are adjusted until inner and outer vortex are properly conveyed by a separate group of streamlines.

- *S. Oeltze-Jafra and B. Preim are with the Department of Simulation and Graphics, University of Magdeburg, Germany. E-mail: {stoeltze|preim}@isg.cs.uni-magdeburg.de.*
- *J.R. Cebral is with the Center for Computational Fluid Dynamics, George Mason University, Fairfax, Virginia, USA. E-mail: jcebral@gmail.com.*
- *G. Janiga is with the Institute of Fluid Dynamics and Thermodynamics, University of Magdeburg, Germany. E-mail: janiga@ovgu.de.*

Manuscript received 31 Mar. 2015; accepted 1 Aug. 2015; date of publication xx Aug. 2015; date of current version 25 Oct. 2015. For information on obtaining reprints of this article, please send e-mail to: tvcg@computer.org.

We automate this tedious, error-prone process and contribute a fast generation of the streamline groups. The display of *group representatives* yields a comprehensive and comprehensible *visual summary* of the vortical flow (Fig. 1(g)). On demand, the groups are refined in a clustering step leading to a more detailed visualization, e.g., of flow further off the core line. We show that this two-step approach outperforms the direct clustering [32] in terms of speed and accuracy and illustrate our method by six (out of 17 investigated) case studies. We further contribute an expert evaluation of the visual summary, exploration techniques, and an automatic assembly of core line segments output by Sujudi and Haimes' algorithm [45] into continuous lines.

2 Hemodynamic Background

We introduce cerebral aneurysm hemodynamics with a focus on vortical flow and outline our hemodynamic data generation pipeline.

2.1 Hemodynamic Parameters

The role of quantitative parameters in characterizing the hemodynamic environment in ruptured and unruptured aneurysms is detailed in [6]. Here, we focus on qualitative parameters describing the blood flow pattern. They are derived from a visual inspection. If flow particles are moving mostly parallel or in a helical fashion along a common axis, the flow is considered *laminar* while *turbulent* flow is characterized by chaotic property changes. If flow from a forward stream reverses and flows back into a separation zone, *recirculation* is observed. The existence, number, and persistence of *vortices* over the cardiac cycle influence the flow *complexity* and *stability*. In complex flow, more than one recirculation zone or vortex are observed. Their disappearance and recreation during the cardiac cycle are characteristic for unstable flow.

2.2 Vortical Flow in Cerebral Aneurysms

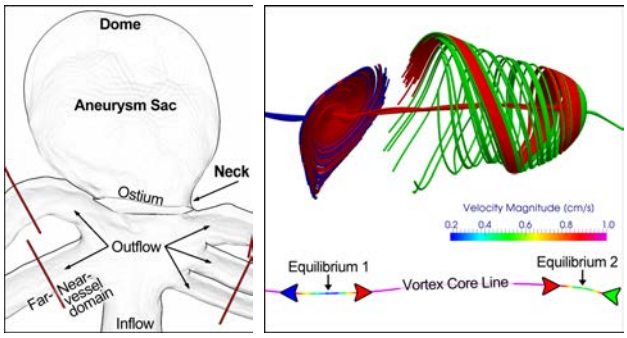
Cebral et al. introduced and later reviewed a flow pattern classification scheme comprising flow complexity and stability [7]. Blinded observers visually classified CFD simulation results of a large database comprising ruptured and unruptured cerebral aneurysms ($n = 210$). Simple, stable patterns were seen in unruptured aneurysms while complex, unstable patterns were observed in ruptured aneurysms. More than 95% of the aneurysms contained at least one vortex. Gambaruto and João link vortex cores anchored at the aneurysm wall to *wall shear stress* and monitor the persistence of vortex cores [14]. The most rigorous investigation of vortical flow has been accomplished by Byrne et al. based on vortex core line extraction [4]. They reviewed the database from Cebral et al. [7] and expressed flow complexity by core line length, with multiple core lines, i.e., vortices, resulting in a higher overall length, i.e. complexity. The persistence of core lines was related to flow stability. Unstable, complex vortical flow was more frequently seen in ruptured than in unruptured aneurysms.

2.3 Embedded Vortices

Byrne et al. observed the formation of embedded vortices in cerebral aneurysms [3]. These vortices are characterized by a vortex layer swirling in one direction along the vortex core line, and a second vortex layer swirling in the opposite direction and enveloping the first (Fig. 2(b), top). Their formation and collapse over the cardiac cycle has been related to saddle-node bifurcations — a concept from dynamical systems theory. It describes the collision and annihilation of two *equilibrium points* with alternating stability in a dynamical system [9].[1] In one-dimensional phase space, an unstable saddle and a stable node collide at the bifurcation point and annihilate each other.

Transferred to embedded vortices, two points at which the velocity magnitude is (near-)zero, i.e. two equilibria, emerge along the vortex core line causing the formation of embedded vortical flow (Fig. 2(b)). During the cardiac cycle, they converge along the core line, collide, and disappear causing a collapse of the embedding and a regression to uni-directional vortical flow [3]. Both equilibria are of type *focus saddle* since they simultaneously exhibit stable and unstable components

[1] In fluid dynamics, a saddle-node bifurcation is also referred to as *fold* or *blue skies bifurcation* and equilibria are known as *stagnation* or *critical points*.



(a) (b)

Fig. 2. *(a)* Morphological features (bold) of an aneurysm and subdivision of the vascular domain (red lines). *(b)* Embedded vortex (top) and its exposed core line (bottom). Two equilibria form a saddle-node bifurcation. Arrows indicate the saddle- (Equilibrium 1) and node-like (Equilibrium 2) flow transport along the core line. Vortex formation is detailed in the text.

— one 2D component describing the direction of swirling flow in a plane transverse to the core line (in or out relative to the equilibrium point, i.e. stable or unstable *focus*) and a 1D component describing the direction of flow transport along the core line (attracting or repelling from the point, i.e., stable node or unstable saddle). The direction of transport is expressed by stable and unstable *manifolds* also known as *stable insets* and *unstable outsets* (arrowheads in Fig. 2(b) bottom).

In Figure 2(b) top, streamlines are seeded at three manually defined core line parts separated by the equilibria. Red streamtubes spiral away from the stable focus (Equilibrium 1) along its unstable outset (left red arrow), flow along the length of the core line and spiral in towards the unstable focus (Equilibrium 2) along its stable inset (right red arrow). As they approach the unstable focus, they collide with green streamtubes flowing in along the inset on the opposite side (green arrow). Red and green streamtubes are repelled by the unstable focus, swirl outwards, and jointly flow back along the core line.

2.4 Hemodynamic Data Generation Pipeline

The pipeline employed for the cases shown in Figure 9 *(a-c)* is described in the following. The pipeline for cases *(d-f)*, which differs in steps tailored to virtual stenting, is outlined in [32].

CFD models of intracranial aneurysms are generated from 3D rotational angiography images [5]. The aneurysm and the surrounding vasculature are segmented by thresholding and region growing algorithms followed by iso-surface deformable models [49]. The resulting surface is smoothed with a non-shrinking algorithm [46]. Unstructured grids composed of tetrahedral elements are generated with an advancing front technique [25]. A maximum element size of 0.2 mm is used, resulting in meshes with approximately 1.5-3 million elements. Numerical solutions of the incompressible Navier-Stokes equations under unsteady pulsatile flows are obtained with an implicit finite element formulation and a deflated conjugate gradients solver to accelerate its convergence [28]. The *neck* of the aneurysm is manually traced on the 3D model and used to triangulate the *ostium* and subdivide the volume mesh into *aneurysm sac* and *parent vasculature* [29] (Fig. 2(a)). The latter is further subdivided into *near- and far-vessel domain* [6].

3 Related Work

Post et al. categorized flow visualization techniques into direct, geometric, texture-based, and feature-based [33]. Salzbrunn et al. added partition-based techniques [38], which decompose a flow field based on vector values, integral curve properties or contained features. Our approach is related to geometric, feature-based, and partition-based techniques since we employ streamribbons to convey flow dynamics, restrict the visualization to vortices, i.e. features, and decompose the flow field around a vortex core line. We restrict our literature overview to the visualization of vortical flow and the partition-based visualization of blood flow.

3.1 Visualization of Vortical Flow

A qualitative investigation of vortical flow does not require an explicit representation of the vortex. Basic techniques such as arrow glyphs and selectively seeded streamlines were compared in this context [24]. A quantitative investigation builds upon a vortex representation, which can be the vortex core line or the vortex region. While the latter is directly visualized by a surface [43], the former serves as a basis for visualizing the flow in its vicinity, mostly by integral curves [21].

We aim at a more quantifiable and reproducible analysis of vortical flow and hence, build upon a vortex representation. We favor the core line since embedded swirling motion is better conveyed by enveloping sets of integral curves than by nested, semi-transparent surfaces. Furthermore, we need the core line to detect the equilibria constituting a saddle-node bifurcation. In the following overview of visualization approaches, we omit a description of the respective core line extraction algorithm. An introduction to extraction methods is given in [21].

Based on Vortex Core Line. The core line is often rendered as polyline or tube. Integral curves are seeded in its vicinity or cutting planes showing the local 2D flow are erected along the line [21]. Embedded vortices are visualized by means of manually defined, colored sets of streamlines in [3] (similar to Fig. 2(b) top). Filled contour plots of *normalized helicity* [27] erected along the core line facilitate the visual separation of a primary and an embedded, secondary vortex in [11]. A streamsurface tightly enclosing the core line and a color mapping to convey local flow rotation are proposed in [16]. An iconic representation of core lines is employed in [36]. A tube is colored and/or scaled according to a vortex criterion and, e.g., colored stripes on the tube indicate flow rotation. Striped pathlines around core lines for conveying vorticity transport in unsteady flow are presented in [35]. The stripes convey scalar flow quantities and simulation quality. In [37], vortex skeletons are extracted from scalar fields indicating vortex activity and visualized by tapered tubes. An illustrative rendering conveying both, vortical flow and vortex extent is introduced in [42].

3.2 Partition-Based Visualization of Blood Flow

Most techniques decompose a blood flow field into regions of similar behavior based on integral curves [2, 17, 23, 32]. Others employ local vectors [47] or aneurysm wall properties [14, 18, 30]. As argued in [50], we favor integral curves over local flow information since they represent continuous flow patterns traced over the entire domain.

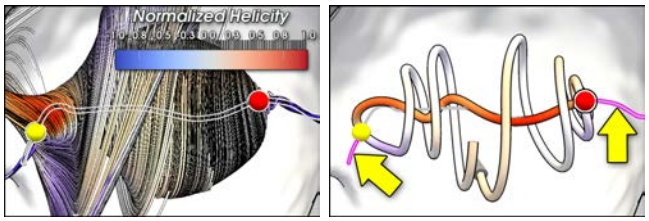
3.2.1 Partitioning Based on Integral Curve Properties

Streamline predicates represent and combine local integral curve properties [39]. The predicate-based grouping of streamlines which constitute the *inflow jet* in a cerebral aneurysm is proposed in [17]. Predicates for the visual analysis of measured cardiac and aortic blood flow can be defined such that, e.g., flow passing vortices is extracted [2]. A comparison of vortex criteria for defining *pathline predicates* tailored to vortical cardiac blood flow is provided in [23].

Predicate-based approaches require the user to define attributes of interest and thresholds on attribute values. *Clustering approaches* employ a data-driven strategy for grouping integral curves and are hence self-tuning with respect to differences in the flow across datasets. In [32], different integral curve properties and techniques for clustering streamlines were compared for analyzing the blood flow pattern in cerebral aneurysms. A spectral clustering approach [51] performed best and will hence, be applied in the refinement step of our approach.

3.2.2 Partition/Cluster Representatives

A condensed visualization of clustered integral curves or fiber tracts is achieved by displaying one or more *representatives* per partition/cluster. In [50], only streamlines located at cluster boundaries are shown adhering to a user-defined density. An interactive filtering mechanism iteratively removing the most similar lines from a cluster until the characteristic ones remain is suggested in [26]. In [2], representative streamlines are derived from the skeleton of a line



(a) (b)

Fig. 3. *(a)*: Streamlines seeded at core line points. The core line is superimposed and its equilibrium points are rendered as spheres (yellow: saddle, red: node). *(b)*: Streamline clustering and display of cluster representatives [32]. The result is lacking a representation of the flow along the magenta ends (arrows) of the gray shaded core line. The same streamlines were grouped using our new approach (Fig. 1(g)). The group representatives clearly convey also the missing flow.

predicate-based streamline bundle and visualized by illustratively rendered ribbon-like structures with arrowheads. We aim at a user-independent approach exploiting the good cluster cohesion and separation achieved by our clustering technique [32]. Related work displays down to a single representative per cluster. In [8], the streamline(s) closest to the cluster centroid are computed in cluster space and displayed by streamribbons with arrowheads. Streamtubes with arrowheads are employed in [32]. Here, the streamline traversing the densest cluster parts on average is computed in streamline space employing an adaptation of [31]. We adopt the density-based representative since it does not require expensive inter-streamline similarity computations.

4 Previous Approaches and New Analysis Pipeline

In previous work, uncluttered visualizations of the aneurysmal blood flow pattern were generated by clustering streamlines seeded at the inlet of the aneurysm, i.e. the ostium surface (Fig. 2(a)), or inside the aneurysm, and showing cluster representatives [32]. However, neither the lines (Fig. 1 *(a,b)*) nor their representatives (Fig. 1 *(d,e)*) are guaranteed to properly capture and convey (embedded) vortical flow and non-vortical flow causes visual clutter (vortices are indicated by their magenta core lines). Extracting vortex core lines and seeding streamlines in their vicinities resolves these issues (Fig. 1 *(c)*). However, conveying the structure of an embedded vortex requires additional efforts. In previous work, streamlines were seeded at manually selected core line parts such that inner and outer vortex were represented by a separate group of streamlines [3]. Integration length, seeding density, and transparency of the streamline groups were manually adjusted until the embedding was properly conveyed. In this tedious and error-prone process, often only trade-offs can be achieved between a clear visibility of the embedding structure, an indication of flow feeding and draining the vortex, and an uncluttered visualization also in case of multiple vortices in one aneurysm. An automated definition of the streamline groups based on the cluster representative approach in [32] and a dense set of lines seeded along the entire core line suffers from inaccuracies and performance issues. First, small parts of distinct vortical flow at core line ends are not always captured by a separate cluster (representative) hampering the assessment of local flow near equilibria (Fig. 3). This is due to the spectral clustering approach which is biased towards clusters of similar size and thus, may merge a very small cluster with a larger one [32]. Second, the clustering can take minutes depending on streamline count and length due to the employed inter-streamline similarity measure. We propose a solution that is fast, avoids visual clutter, and generates a complete visual summary of vortical flow.

Our pipeline for the cluster analysis of vortical flow is shown in Figure 4. It starts by extracting the vortex core lines from the hemodynamic data (Sec. 5.1). Since the extraction generates sets of line segments suffering from erroneous gaps and noise, continuous lines are formed in an enhancement step (Sec. 5.2). Then, streamlines are integrated from seed points along the core lines (Sec. 5.4). The subsequent clustering step comprises an initial grouping of streamlines

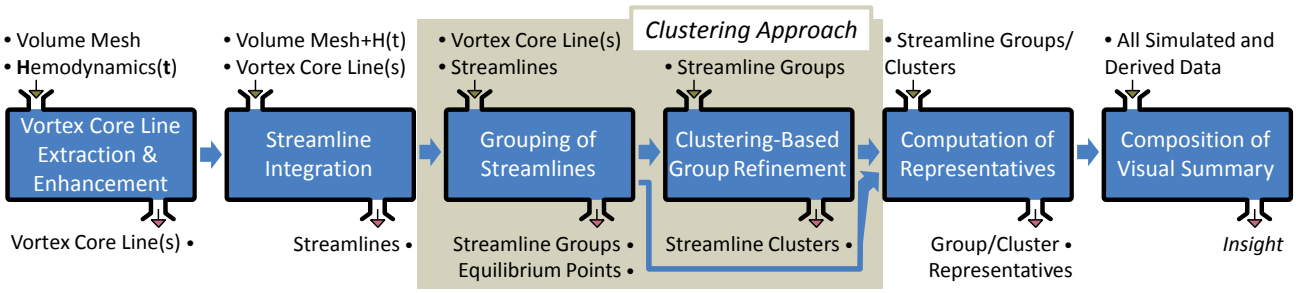


Fig. 4. Analysis pipeline. Vortex core line segments are extracted from time step t of the simulation data and continuous lines are formed in an enhancement step. Then, streamlines are integrated from core line points. Next, streamlines are grouped and the groups are optionally refined in a clustering step. Finally, group/cluster representatives are computed and integrated with other simulated or derived data in a visual summary.

and an optional refinement step. The grouping is sufficient to observe the global structure of an embedded vortex and to assess the local flow near equilibria (Fig. 1 *(f-i)*). It comprises detecting the equilibria along the core lines for defining core line partitions (Sec. 6.1) and associating each streamline with a partition (Sec. 6.2). This structured process ensures that all parts of embedded vortical flow are captured by a separate group of streamlines. It is fast since it does not require the computation of inter-streamline similarity. In an optional clustering-based group refinement, a more detailed representation of flow further off the core line and of vortical flow without an embedded structure can be achieved (Sec. 7.2). The refinement step employs the grouping result as constraint to improve accuracy and reduce computational time. After the clustering, group/cluster representatives are computed and aggregated in a visual summary of vortical flow (Sec. 8.1). The user may interact with the visual summary to explore the flow (Sec. 8.2).

5 Computation of a Vortex Representation

We describe vortex core line extraction and enhancement. The core line is part of our vortex representation which is completed by streamlines integrated from seed points on the core line. Note that we consider only vortices located inside the aneurysm sac (Fig. 2(a)).

5.1 Vortex Core Line Extraction

Our collaborating CFD engineers extracted the core lines partially within previous projects employing the eigenvector method by Sujudi and Haimes [45]. It returns a line segment per element of the volume mesh that is penetrated by the core line. The set of segments is known to suffer from discontinuities between neighboring segments due to the piecewise linearity of the Jacobian and from false positives, i.e. noise. Hence, filtering mechanisms were applied.

Despite careful fine-tuning, often only trade-offs between noise removal and maintenance of "real" core line segments were achieved. Gaps resulted from a too generous filtering and the filtered data still suffers from noise (Fig. 5(a)). More sophisticated algorithms for core line extraction were proposed [40]. However, no algorithm or vortex criterion are known to guarantee vortex detection [1]. Moreover, different algorithms generate different core line candidates. To find the best algorithm is out of scope of this paper. In preliminary tests, we visually compared the given segments to the results of two algorithms implemented in EnSight (CEI Inc., Apex, NC) — the method by Sujudi and Haimes and a vorticity-based approach. The agreement also between EnSight's methods was very high with the vorticity-based method producing more noise. A more thorough comparison involving more approaches would be required for implementing our approach in a clinical setting requiring maximum confidence. However, we focus on supporting the research endeavors of CFD engineers who are aware of the vortex extraction issues.

We stick with the available core line segments and propose an enhancement approach incorporating automatic removal of the remaining noise, gap closure, and the formation of continuous lines adhering to the segments (Sec. 5.2). In a sensitivity analysis, we visually verify all core line candidates and elaborate on the rate of false-positive and -negative (missing) core lines after enhancement (Sec. 5.3).

5.2 Vortex Core Line Enhancement

The enhancement step may neglect the flow field since we do not aim at the precise core line but at a solid base for seeding streamlines and computing equilibria. We identify curve-shaped clusters of segments, e.g., the green and red cluster in Figure 5(a), and per cluster, find the shortest path P along the segments from one end of the curve/cluster to the other (Fig. 5(c)). Gaps are closed by merging nearby, similarly oriented paths P. Stair-case artifacts caused by discontinuities between adjacent segments are resolved by smoothing (Fig. 5(b)). The proposed approach can handle multiple core lines but fails to handle branching ones. So far, we did not observe branching vortex cores.

Clustering. At first, the segments are clustered using a density-based technique (Fig. 5(a)). Such techniques are particularly suited for noisy data and non-spherical clusters. Instead of employing the line segments, we cluster their endpoints. This is feasible since the endpoints of a segment are very close together due to their location inside the same cell of the volume mesh. We employ a variant of the DBSCAN clustering algorithm that requires only a single parameter *minPts* steering the minimum number of points to be considered as a cluster [10]. The second parameter ε of the original DBSCAN algorithm is derived from *minPts*. It represents the distance between points up to which they are considered to lie in the same cluster. Small clusters can be tagged as outliers, i.e. false positives, by adjusting *minPts*. However, increasing *minPts* leads to larger values of ε which in turn leads to the merging of outliers with larger clusters in close proximity. We thus set $minPts = 1$ and skip small clusters when determining P.

Graph Representation. In a second step, the set of points is treated as undirected, edge-weighted graph. Each of the n points corresponds to a node. An edge exists between two nodes if their points are contained in the same cluster. The Euclidean distance is employed as edge weight unless the points belong to the same line segment, in which case a predefined value very close to zero is assigned. This guarantees that the search for the shortest path P prefers the line segments. For instance, P of a U-shaped set of segments would otherwise correspond to the direct connection of the U's ends. The graph is represented by its sparse $n \times n$ adjacency matrix A.

Shortest Path Search. A series of steps is carried out to determine P_c for each cluster c. Clusters smaller than a threshold T_1 are neglected since they are likely to represent false positives (too small clusters in Fig. 5(a)). *(1)* For the rest, the corresponding subgraph A_c is extracted from A. A straightforward but expensive approach to search for P_c would be the computation of all shortest paths between pairs of nodes in A_c with P_c being the longest of them. *(2)* To minimize the number of pairs to test, the *minimal spanning tree* MST_c of A_c is computed [34] (black structure in Fig. 5 (close-up)). It forms a natural skeleton of c suggesting its shape. *(3)* Then, the shortest paths between all pairs of nodes in MST_c are computed [22] and the longest of them $Ptmp_c$ is memorized (yellow path in Fig. 5 (close-up)). It provides a good estimate of P_c but its computation considers only a subset of A_c, namely the edges in MST_c. As a consequence, $Ptmp_c$ may not be optimal everywhere and make detours as for instance in the middle of Figure 5 (close-up) where the red line represents the shortest path.

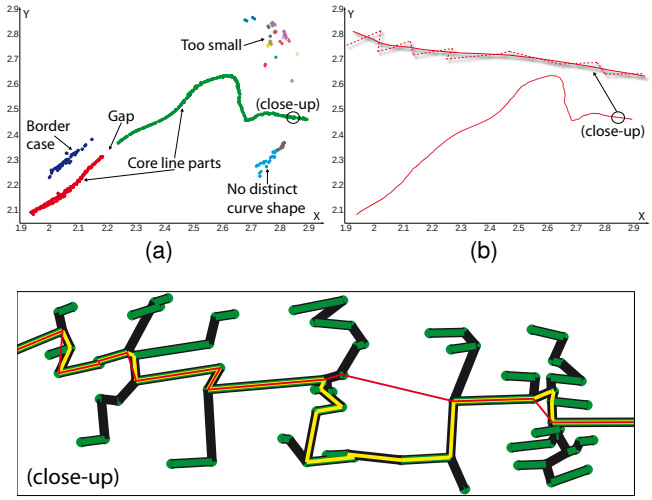


Fig. 5. Vortex core line enhancement. *(a)* Core line segments are clustered. *(close-up)* For each larger cluster, the shortest path (red) from end to end along the line segments (green) is computed based on the cluster's minimal spanning tree (black) and its longest shortest path (yellow). Note that segment overlaps result from 3D to 2D projection. *(b)* Gaps are closed, the path is smoothed, and small clusters representing noise and those without a distinct curve shape are neglected.

(4) Hence, the path search is extended to A_c and point pairs in the ε-neighborhood of $Ptmp_c$'s terminal vertices. Note that the two terminal vertices are located at either "end" of the cluster due to the definition of MST_c. For all point pairs, the shortest path is computed [12] and the longest of the shortest is chosen as P_c (red path in Fig. 5 (close-up)).

Verification Step. Clusters without a distinct curve shape, i.e. a high percentage of segments deviates strongly from P_c, shall be neglected from further processing since they are likely to represent false positives. At first, the median of all distances $\tilde{d}$ between two neighboring points on P_c is computed. Then, all points in c which are not part of P_c and further off P_c than $\tilde{d}$ are determined. For each of these points, its closest point on P_c is computed. If the share of unique closest points in total points of P_c exceeds a threshold T_2, c is neglected since it shows a high degree of dispersion which is not restricted to a small region along P_c. A locally high dispersion is tolerable. It is occasionally observed close to the vessel wall where core line extraction is error-prone due to small velocities. The light and the dark blue clusters in Figure 5(a) fail the verification step. The latter is a border case which becomes more apparent in a 3D view.

Path Merging and Smoothing. The merging step checks whether ends of paths are close together and oriented in a similar direction. If so, the paths are likely to belong to the same core line and they are merged (paths of red and green cluster in Fig. 5(a) are merged in Fig. 5(b)). The similarity of direction is checked based on the dot product of direction vectors extracted from the paths' ends. Two thresholds T_3 and T_4 are applied to proximity and similarity of direction, respectively. In a final step, a low-order smoothing spline is fit to each path based on an automatically determined amount of smoothing [15] (Fig. 5(b)). As argued in [19], this eliminates high frequency noise, i.e., the stair-case artifacts caused by the discontinuities between neighboring segments, while maintaining the core line's curvature.

5.3 Parameter Sensitivity Analysis

A parameter sensitivity analysis was conducted to determine the stability of the enhancement result as well as default values for $T_1 - T_4$. In a visual inspection of 17 aneurysm cases with differing number of vortices, streamlines were integrated from each cluster of line segments produced by DBSCAN. If the lines were swirling around the segments, this cluster was tagged as vortex. None of the clusters was seen to represent more than one vortex indicating a good separability of the data and performance of DBSCAN. From 94 clusters, 42 were tagged as vortex and 9 as embedded vortex. To initialize T_1 and T_2 properly, the number of points per cluster and the share of unique closest points from the verification step were opposed in a plot (Fig. 1 of the supplemental material). It reveals that vortices cannot perfectly be separated using T_1 and T_2. However, setting $T_1 = 26$ points and $T_2 = 15\%$ achieves a good trade-off between false-positives and -negatives and only neglects two small vortices. While the inclusion of embedded vortices, which only appear on longer core lines ($T_1 > 100$ points) with a very distinct curve shape ($T_2 < 5\%$), is insensitive to slight parameter changes, the inclusion of other vortices is sensitive. In future work, an automatic vortex verification will obviate T_1 and T_2 [20].

The stability of the core line merging step has been investigated by setting T_1 and T_2 to their default values and letting T_3 vary over multitudes of ε starting by a value close to zero and terminating at half the largest diameter of the aneurysm sac. With each value of T_3, the similarity of direction of core line ends (T_4) being closer together than T_3 was recorded. Only values in the range $[0,1]$ ($0 =$ orthogonal ends and $1 =$ collinear ends) were considered. The results are given in Figure 2 of the supplemental material. Merging only occurs for five cases and does not start until $T_3 = 1.5\varepsilon$. Multiple merges were observed for two datasets (pink and turquoise). The merges at values of $T_4 < 0.9$ were visually identified as false merges fusing unique vortices. The data at hand suggests that a combined thresholding of distance and direction at $T_3 = 1.5\varepsilon$ and $T_4 = 0.9$ (equal to $\approx 26°$ maximum deviation) yields satisfying results. Small changes of T_3 and T_4 do not effect the results.

5.4 Streamline Integration

We seeded streamlines in the vicinity as well as directly on the core line and found that the latter sufficiently conveys the structure of embedded vortices. Hence, seed points are computed along each core line by equidistant resampling employing half the minimum distance between any two consecutive points in P_c. A 5^{th} order Runge-Kutta method is employed for streamline integration. The maximum integration length is set to three times the maximum edge length of the volume mesh's bounding box. This high value accounts for the high vorticity of the investigated flow. The integration is carried out in backward and forward direction and each resulting pair of lines is merged.

6 Grouping of Streamlines

The grouping requires the computation of all equilibrium points along a core line (Sec. 6.1) and then, classifies streamlines based on the position of their respective seed point relative to the equilibrium locations (Sec. 6.2). Since equilibria subdivide a core line into regions of coherent flow transport, this structured processing ensures that all parts of embedded vortical flow are captured by a separate group of streamlines. This was confirmed by three observations in a visual investigation of 9 embedded vortices. Streamlines seeded on (see Fig. 6)

1. either side of an equilibrium point differ strongly in shape.
2. the same side exhibit a very similar shape in the vicinity of the core line and may only diverge further off.
3. two different vortex core lines differ strongly in shape unless they are traced for a long time and participate in both vortices.

Here, "shape of a streamline" refers to its progression along the vortex core line. For instance, in Figure 6(b), all green streamlines swirl in from the left of the node and at the node (between green and red arrow), jointly spiral outwards. Therefore, their shape is said to be similar. Note that for the grouping, similarity considerations are restricted to the vicinity of the core line. In the clustering-based refinement step (Sec. 7), they are extended to the entire aneurysm to capture differences further off the core line (yellow arrows in Fig. 6).

6.1 Detection of Equilibrium Points

So far, equilibria were detected visually based on color mapping the velocity magnitude onto the core line, clamping values > 1 cm/s, and searching for local minima [3]. We maximize comparability with [3]

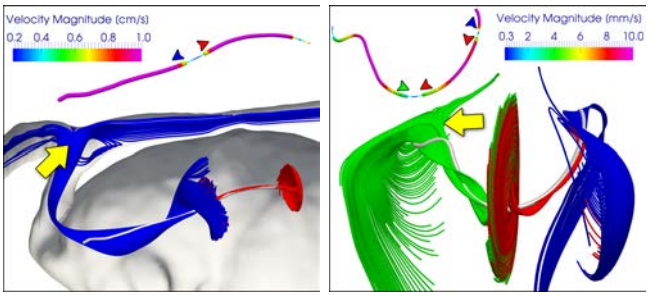


Fig. 6. Sets of streamlines seeded at two *(a)* and three *(b)* different core line parts. Lines within a set exhibit a similar shape and only diverge further off the core line (yellow arrows). The core lines are shown at their real location (gray tubes) as well as in a shifted display. In the latter, the in-/outflow direction at equilibria is depicted by arrows whose color corresponds to streamline color. Furthermore, the core line parts employed for seeding are rendered thicker.

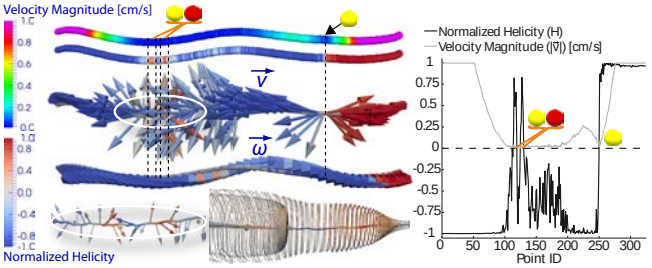


Fig. 7. Detection of nodes and saddles (red and yellow spheres) based on normalized helicity H and velocity magnitude $|\vec{v}|$ along the vortex core line. H is the normalized dot product of $\vec{v}$ and vorticity vector $\vec{\omega}$. All vectors are scaled to unit length to improve readability. Equilibria are more reliably indicated by zero crossings of H than by minima of $|\vec{v}|$.

by adopting their color mapping. The classification of equilibria, i.e., saddle or node, was achieved by integrating streamlines from manually selected core line segments only in forward or backward direction.

The strict calculation and classification of equilibria/critical points in vector fields with a finite precision suffers from numerical instabilities. We suggest a more robust alternative that is based on analyzing the local flow transport (insets/outsets) along the core line. Before we describe the approach, we characterize the types of equilibria that were observed in our case studies.

Types of Equilibrium Points. The following types are indicated by colored spherical glyphs in all visualizations (e.g., Fig. 9):

- <— o —>, saddle point (o), two unstable outsets (—>)
- —> o <—, stable node, two stable insets
- | o —>——, end of core line near wall (|), one unstable outset
- o —>——, end of core line offside wall, one unstable outset

In accordance with [14], we observed equilibria at the ends of vortex cores, particularly, at those anchored to the aneurysm wall. Here, wall-near flow is advected in a swirling motion, leaves the wall along the outset, and spirals along the core line. Detecting these equilibria is desirable since the inner vortex of an embedded vortex may originate here and not as typical, at a saddle (e.g., left blue sphere in Fig. 9(a,g)).

Detection based on Normalized Helicity. While detecting equilibria as local minima in a function of clamped velocity magnitude $|\vec{v}|$ over core line points (plot in Fig. 7) may seem straightforward, it fails in two situations (topmost core line in Fig. 7). First, equilibria within larger "stagnation zones", such as the left large dark blue zone, may not be reflected by distinct minima (plot). However, the velocity vectors $\vec{v}$ in this zone point in alternating directions of flow transport indicating the presence of equilibria (encircled close-up). Second, multiple equilibria, i.e. local minima, may exist within a region of very low velocity magnitude, such as those within the stagnation zone plus the one within the smaller dark blue zone to its right.

Our solution is inspired by the visual separation of a primary and an embedded secondary vortex based on the switching sign of *normalized helicity* in [11]. Helicity is defined as the knottedness of *vortex lines* [27]. These lines are everywhere tangent to the vorticity vector $\vec{\omega} = \nabla \times \vec{v}$ which describes the curl of the flow velocity. Normalized helicity H is computed as:

$$H = \vec{\omega} \cdot \vec{v} / (|\vec{\omega}|\,|\vec{v}|), \in [-1, +1] \qquad (1)$$

The sign of H is negative if $\vec{v}$ and $\vec{\omega}$ point in an opposite direction and positive if they point in the same direction. This property is exploited in detecting the equilibria. Since the direction of $\vec{\omega}$ is stable along the core line and the direction of $\vec{v}$ switches at both, saddles and nodes (Fig. 7), equilibria are indicated by zero crossings in a function of H over core line points (plot in Fig. 7). The precise positions of the equilibria are gained by interpolation based on the respective two points "enclosing" a crossing and their values of H. Finally, the equilibria are classified by comparing the average direction of $\vec{v}$ on either side in a local neighborhood. For a saddle, the dot product of the two average vectors is negative while it is positive for a node.

The detection of equilibria based on H and $|\vec{v}|$ is compared in Figure 7. Two saddle-node pairs within the stagnation zone and the nearby saddle are identified using H. The pairs would have been missed employing $|\vec{v}|$. They form two adjacent embedded vortices with one enveloping the other (illustrated by streamlines). The analysis of H is also more robust since zero crossings are global features of a function and hence, less sensitive to low frequency jags than local minima.

Detection at Vortex Core Line Ends. Equilibria at core line ends are not reflected by a switching sign of H due to the unidirectional flow there. However, they are located inside small regions with $|\vec{v}| < 1$ cm/s (clamping value). If such a region exists, its point with minimum $|\vec{v}|$, which in all cases corresponded to the terminal vertex of the core line, is taken as equilibrium point. Depending on its distance to the aneurysm wall, it is classified as "near wall" or "offside" wall (blue or green sphere).

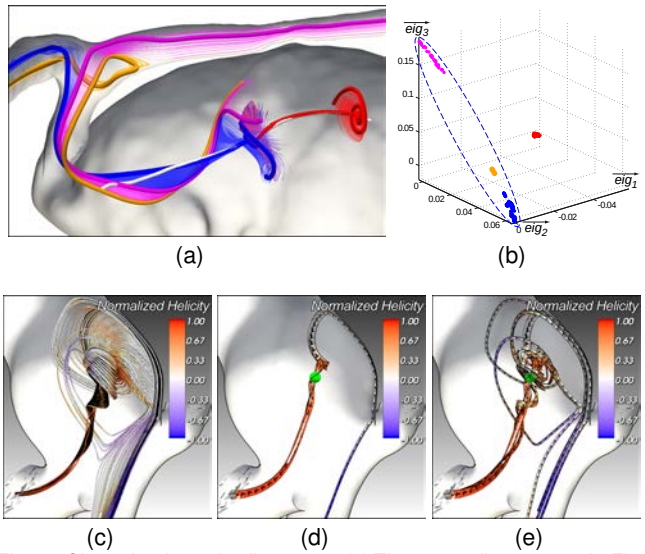
6.2 Grouping

This step iterates over all vortex core lines, partitions them according to the respective equilibria, and assigns each streamline to a partition.

Grouping at Saddles and Nodes. Only saddles and nodes are considered for the partitioning since only streamlines seeded on either side of these equilibria differ strongly in shape (Fig. 6). Accordingly, the core line in Figure 7, is divided by the five equilibria into six partitions. Before streamlines are assigned to them, *safety margins* are defined around saddles and nodes (thin parts of exposed core lines in Fig. 6). They account for the potential inaccuracies involved in core line extraction and enhancement and comprise all core line points within the next five penetrated volume elements to either side.

Streamlines seeded outside the safety margins are assigned to the partition containing their respective seed point thereby creating initial streamline groups. For each group, a *representative* is determined employing our variant [32] of *density-based representatives* [31]. In short, the longest streamline traversing the densest parts of the group is computed. For streamlines seeded inside the margins, the similarity to the neighboring group representatives is calculated based on minimum closest point distances. Each line is then assigned to the group of the most similar representative and finally, all representatives are updated.

Grouping at Near-Wall Equilibria. Group representatives shall convey each distinct streamline shape occurring in the vicinity of the core line. To guarantee this at near-wall equilibria, a refinement step is necessary. At the corresponding core line end, two types of streamlines are observed (arrow in upper inset of Fig. 9(g)): longer lines whose backward integration indicates the inflow of the vortex and



(a) (b) (c) (d) (e)

Fig. 8. Clustering-based refinement. *(a)* The streamline groups in Figure 6(a) were refined based on their spectral embedding *(b)*. The encircled clusters constitute the blue group in Figure 6(a). *(c)* Streamlines at a vortex without an embedded structure were grouped in *(d)* and the groups were refined in *(e)* yielding a comprehensive representation.

short lines whose backward integration terminated early due to near-zero velocities close to the aneurysm wall. To ensure that representatives convey both types, the streamline group that has been assigned to the partition containing the core line end is split according to the length of the backward integrated streamline part. The set of lengths forms two crisp clusters that can be detected by k-means with $k = 2$. The two corresponding subgroups of streamlines replace the current group and their representatives are computed. In the lower inset of Figure 9(g), the upper and the lower arrow point at the representative indicating the inflow and the early terminated streamlines, respectively.

Grouping at Core Lines with a Single Equilibrium Point. We observed core lines with a single equilibrium located at one of their ends, mostly near the wall and once, offside the wall (Fig. 9(b) and 9(e)). We chose to generate two representatives per such core line so that their intertwining conveys the vortical nature of the flow. For core lines with a near-wall equilibrium, k-means clustering as described above is employed. For the rest, the density-based representative of all associated streamlines is first computed. Next, the similarity of all lines to the representative is calculated and the similarities are clustered using k-means with $k = 2$. Finally, the representatives of both clusters are computed, one conveying the densest part of the streamline group and one conveying the part that deviates most from it.

7 Clustering-Based Group Refinement

The clustering-based refinement of streamline groups allows for a more detailed visualization of flow further off the core line and of vortical flow without an embedded structure. In previous work, we showed that *spectral clustering* of streamlines generates expressive blood flow summaries [32]. Here, we recapitulate the approach and explain its adaptation to the refinement of an initial grouping result.

7.1 Spectral Clustering

Spectral clustering (SC) maps all streamlines to a *spectral embedding space* where each line is represented by a point (Fig. 8 (a,b)). This mapping preserves local distance relations and increases cluster separability. SC starts by constructing a $n \times n$ symmetric *distance matrix* $\mathbf{M}$ that contains all pair-wise distances/dissimilarities between n streamlines. Then, the $n \times n$ *affinity matrix* $\mathbf{W}$ is constructed by applying a function f to the entries of $\mathbf{M}$. As f, a Gaussian is used assigning high affinities to low distances and vice versa. Based on $\mathbf{W}$, the $n \times n$ *Graph Laplacian* $\mathbf{L}$ is computed [51]. Then, the k largest eigenvalues and their corresponding eigenvectors of $\mathbf{L}$ are determined with k representing the number of clusters. These eigenvectors span the k-dimensional spectral embedding space (Fig. 8(b)). Instead of providing k, a range of values $[a,b]$ is provided by the user. The optimum number of clusters within $[a,b]$ and a cluster label per streamline are then returned by an *eigenvector rotation* approach [51].

7.2 Group Refinement

Our adaptation of SC employs the grouping result to constrain the clustering-based refinement. Since streamlines from different groups exhibit a considerably different shape, they must not fall into the same cluster. As a consequence, distances in $\mathbf{M}$ must only be computed for pairs of streamlines from the same group. To speed up distance computation, *hierarchical signatures* [26] are employed. Distances between lines from different groups are set to a predefined very high value such that the Gaussian f is evaluated to zero (affinity). The remaining steps are identical to our previous approach except for the initialization of $[a,b]$. Instead of assigning an arbitrary minimum to a, it is set to the number of streamline groups. Assuming that each group will not be partitioned more than the number of draining vessels dv, b is set to $a * dv$. Constraining SC yields a performance gain due to fewer distance computations. In the best case, the g streamline groups are of equal size reducing the number of computations by factor $\approx g$.

The streamline groups in Figure 6(a) were refined in Figure 8(a), yielding a separate cluster for each unique streamline shape. Representatives now well convey the fact that flow leaving the vortex proceeds through different draining vessels. The spectral embedding of the streamlines reflects the four well-separable clusters (Fig. 8(b)). In Figure 8(c), streamlines convey vortical flow without embedded structure. The corresponding core line exhibits a single equilibrium at one of its ends offside the aneurysm wall (green sphere) and the streamlines were grouped accordingly (Fig. 8(d)). While the two group representatives partially convey the vortex, the group refinement yields a more comprehensive representation (Fig. 8(e)).

8 Visual Summary of Vortical Flow

We describe the aggregation of group/cluster representatives in a visual summary of vortical flow, present exploration techniques, and demonstrate the summary in six case studies. We use the term cluster to refer to both, group and cluster. Exemplary summaries are shown in Figure 9(g-j). In Section 2 of the supplemental material, we report on timings of generating the summary.

8.1 Composition of the Visual Summary

Stream ribbons and tubes have been tested for visualizing the cluster representatives. While tubes allow for a good readability of color-mapped parameters independent of the viewpoint, we favor ribbons since they additionally encode the local curl of the flow. A narrow strip of polygons is added to each representative streamline. Its curl is either determined by the local vorticity vectors or by normal vectors, which are forced to minimally rotate along the line. The former yields strongly twisted ribbons in regions of high vorticity hampering the readability of color and textures. The latter resolves this problem but indicates the torsion of the streamline instead of the flow's local curl. Hence, both opportunities were integrated. The latter is shown in the paper to avoid readability issues.

The width of a ribbon encodes the number of streamlines in its associated cluster. Black halos improve the visual separability of the ribbons. An arrowhead glyph is attached to the ribbon's end that points in outflow direction. To visualize the local flow direction at equilibria, arrow textures are mapped onto the ribbons inspired by [44]. Normalized helicity H (Eq. 1) is mapped along the ribbons by a diverging blue-to-red color scale. The display of representatives is embedded in a surface rendering of the aneurysm. The front faces are culled to provide an unobstructed view. A silhouette drawing lets the surface stand out from the background. The streamlines of a cluster can be displayed to indicate its extent. They are rendered with GPU support, black halos are added to improve spatial perception [13], and normalized helicity

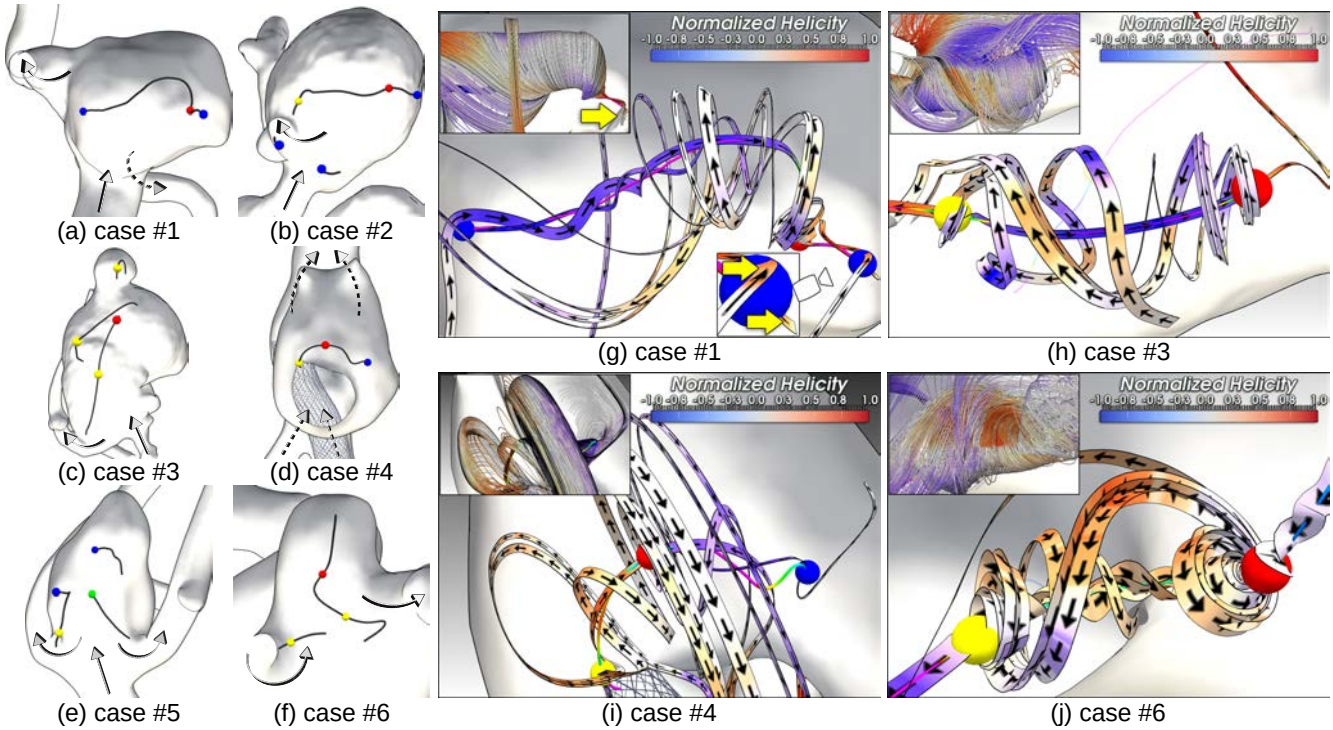


Fig. 9. The cases analyzed in this work are presented in the first two columns. Aneurysmal morphology including feeding and draining arteries, inflow and outflow direction (arrows), vortex core lines (dark gray tubes), equilibrium points (colored spheres), and a stent (wire mesh in *(d)*) are depicted. Embedded vortices are shown in detail views of selected visual summaries of vortical flow in the last two columns. Insets depict the original streamlines. The lower inset of *(g)* shows a near-wall equilibrium (blue sphere) and the representatives of the two types of streamlines (arrows) which are typically observed here (Sec. 6.2). In the corresponding upper inset, an arrow points at the represented streamlines.

is color-coded (insets of Figure 9(g-j)). Inspired by [14], equilibria are represented by spherical glyphs whose color encodes the type of equilibrium (Sec. 6.1). Vortex core lines are rendered as tubes and colored according to velocity magnitude as described in Section 6.1.

8.2 Exploration of the Visual Summary

The exploration techniques discussed in [32] and two new techniques are integrated. First, an inspection of the flow in the immediate vicinity of an equilibrium point is supported. Clicking on the spherical glyph causes an emphasis of the two representatives associated with the adjacent core line partitions. Second, investigating flow in the presence of multiple vortices benefits from a smart visibility strategy. Please, see the supplemental video for a demonstration. In an overview visualization, the representatives of all vortices are displayed. As the user zooms in on a vortex, all other vortices gradually disappear. This is realized by coupling opacity with distance to the camera. A proper functioning requires that the vortex of interest is always closest to the camera and in the approximate center of the viewport. Once the distance is < 4 mm, the representatives of the vortex in focus also start to fade-out and its spherical glyphs gradually shrink to reveal the core line and facilitate an inspection of its velocity magnitude profile.

8.3 Case Studies

We processed 17 cerebral aneurysm cases from which we chose six representative ones for illustration (Fig. 9(a-f)). Visual summaries of selected additional cases are shown in the supplemental material (Sec. 4). The six cases differ with respect to location at the circle of Willis, type: side-wall (cases #4,case #6) vs. basilar tip (the rest), size: small (case #6) vs. giant (case #3) vs. big (the rest), and number of vortices. Furthermore, case #4 was virtually stented and case #2 exhibits a local outpouching (*bleb*) at the top of the aneurysm sac.

For each of the first three cases, at least one embedded vortex had been observed in unsteady CFD simulation data. To study the structure of the vortex at full manifestation, a representative point of the cardiac cycle was selected. For the last three cases, only steady-state CFD simulation data existed. A candidate embedded vortex was observed in case #6. For case #4, no clear statement could be made since the outer layers of two adjacent candidate embedded vortices only briefly envelope their inner counterpart before leaving the aneurysm. No embedded vortex seems to exist in case #5. A conclusive assessment of the three cases yet requires unsteady data covering the cardiac cycle.

Detail views of the visual summaries of vortical flow are presented for case #2 in Figure 1(g), for case #5 in Figure 8(c-e), and for the remaining cases in Figure 9(g-j). Each summary has been visually validated by our collaborators. The correctness of the groups/clusters and their representatives was assessed based on a display of the corresponding original streamlines and on additional arrow glyphs conveying the flow direction along the core line as well as along the streamlines. The glyphs helped in validating the determined type of equilibrium as well as the orientation of the ribbon textures. All visual summaries were assessed as correct. Note that this assessment neglects vortices potentially missed during core line extraction (Sec. 5.1). The false-positive and -negative core line candidate clusters described in Section 5.3 do not affect the six demonstration cases.

We restrict our discussion to cases exhibiting a definite embedded vortex (candidate). While case #2, case #3, and case #6 show the typical saddle-node configuration (yellow and red sphere), the inner vortex of case #1 originates at a near-wall equilibrium (left blue sphere). The latter configuration must be considered in embedded vortex detection since core lines of strong vortices are frequently attached to the aneurysm wall and end their in an equilibrium [14]. A single embedded vortex had been reported for the ruptured *giant aneurysm* case #3 [3]. We found two more embedded vortices (Sec. 6.1, Fig. 7). However, both together reach their full extent at a different point of the cardiac cycle and are small-scale phenomena compared to the known vortex (Fig. 9(h)). Hence, they are neglected here and their equilibria are not shown in Figure 9(c). Case #6 in Figure 9(f) had been investigated in previous work [32]. While the swirl in the center of

Table 1. The visual summary *VS* of vortical flow was compared to interactively seeding *IS* streamlines at selected core line parts and spectral clustering *SC* of streamlines and display of cluster representatives [32].

Analysis Time [m:s]			Correct Equilibrium Classification [%]			Embedded Vortex Detection [%]		
IS	SC	VS	IS	SC	VS	IS	SC	VS
9:24	5:36	2:06	90	76	100	95	100	100
Convenience of Analysis [-2,+2]			Confidence in Analysis Results [-1,+1]			Clarity of Visualization [-2,+2]		
-0.4	-0.1	1.3	0.0	0.1	0.8	-0.6	0.5	1.3

the aneurysm was observed, details of the swirling motion remained hidden. Accordingly, the embedded vortex along the upper core line (Fig. 9(j)) had been described just as "complex flow".

9 Evaluation

We carried out an evaluation comparing our new and two prior approaches with regard to the quality and speed of detecting embedded vortices and equilibria as well as of characterizing the flow near equilibria. We compared *(VS)* our visual summary without the clustering-based group refinement, *(IS)* the interactive seeding of streamlines at selected core line parts as in [3] (Sec. 4), and *(SC)* the spectral clustering of streamlines and display of cluster representatives [32]. *IS* is employed by our collaborators and has been reimplemented in ParaView (Kitware, Clifton Park, NY). For *SC*, we used our in-house implementation. Since the core line enhancement is not part of *IS* and *SC*, the originally extracted core line segments were employed as seeding geometry for both (Sec. 5.1). They were rendered and colored according to velocity magnitude in order to indicate equilibria [3]. Seeding regions as well as integration direction and length could be modified in *IS*. In *SC*, individual representatives could be turned on/off and streamlines could be displayed for each cluster. Please consult Section 3 of the supplemental material for illustrating screenshots.

Two CFD engineers (no co-authors) with long-term experience in analyzing blood flow and eight computer scientists with a background in flow visualization participated in the evaluation. Each session started with a briefing including the definition of embedded vortices and equilibria and a short introduction. Then, each participant employed each approach twice to analyze all six cases (Fig. 9(a-f)). A questionnaire had to be filled in per case showing the vortex core lines inside a drawing of the aneurysm. Participants were asked to encircle core lines exhibiting embedded flow, sketch inner and outer vortex, indicate the equilibria and determine their type (Sec. 6.1). Since spherical glyphs represent equilibria in *VS*, the participants were asked here to validate their location and classification. Finally, they rated the convenience of the analysis process, their confidence in the results, and the visual clarity of the visualization. The analysis process was timed.

The evaluation results are summarized in Table 1. On average, *VS* reduces the analysis time by a factor of ≈ 4.5 and ≈ 2.7 as compared to *IS* and *SC*, respectively. Participants spent plenty of time in *SC* with deducing local flow direction at an equilibrium from the arrowheads' direction of the corresponding representatives (Fig. 3(b)). In *VS*, this information is locally encoded by arrow textures (Fig. 1(g)). The participants detected all equilibria in all cases independent of the applied approach. Hence, the detection rates are omitted and only the classification rates are given. In *VS*, all equilibria were confirmed as properly localized and classified. The classification rate in *IS* was also high while only 76% of the equilibria were classified correctly in *SC*. This is due to an inadequate visualization of flow (direction) at some equilibria (Fig. 3). Participants were explicitly asked not to guess the flow (direction) but only to report what they actually see. All embedded vortices were detected in *VS* and *SC*. One participant did not observe the embedding in case #3 (Fig. 9(h)) based on *IS*. Some participants saw an embedded vortex in the ambiguous case #4, while others did not (Fig. 9(i)). Both answers were counted as correct. The convenience of the analysis was rated highest for *VS* with a substantial lead on *SC* and *IS*. The participants had great trust in their analysis results based on *VS* but felt less confident when employing *IS* and *SC*. Furthermore, the visual clarity of *VS* was rated highest. In *IS*, participants complained about the cluttered streamline display while in *SC*, they found parts of vortical flow not properly conveyed (Fig. 3).

Limitations. The evaluation has the character of first preliminary tests. A prospective user study will be based on a larger database of aneurysms ($n = 210$) capturing the variety of location, shape, and hemodynamics [7]. It will involve more and a wider variety of users, e.g., neurosurgeons and neuroradiologists. Since expert users will employ our approach, participants should be familiar with flow visualization and analysis. So far, the evaluation focused on the streamline grouping step whose results are sufficient for conveying the global structure of embedded flow and the location and type of equilibria. The prospective study will include the group refinement step.

A reworked study design will ensure that metrics such as analysis time are not biased towards *VS*, which is designed to highlight vortices and equilibria. Instead, users will be asked to accomplish the tasks employing *IS* and *SC* and their results will be compared to those of our new pipeline (Fig. 4). Our case studies indicate that the pipeline can batch the database and solve the tasks automatically. The generated quantitative results will be compared to the manually derived ones and the visual summaries *VS* will be rated by the users.

Given a positive evaluation of *VS*, an insight-based study will be conducted with expert users. For this purpose, *VS* will be integrated in an analysis system [30] and embedded in a more realistic data analysis context. More general analysis tasks will aim at understanding the effects of (embedded) vortical flow on quantitative hemodynamic parameters, such as pressure and *wall shear stress*, and the relation to other qualitative parameters, such as wall-near flow and the *inflow jet* [17]. In pair analytics sessions, an expert user and a developer will analyze the data and try to generate application-specific insight and hypotheses. Both will be counted yielding metrics of success.

10 Summary and Outlook

We presented a pipeline for the cluster-based visual analysis of vortical flow in simulated cerebral aneurysm hemodynamics. Segments of vortex core lines affected by artifacts are transformed into continuous core lines serving as a basis for seeding streamlines that convey the vortical flow. Streamlines are grouped and group representatives are computed such that each distinct flow behavior in the vicinity of the core line is properly captured. On demand, the groups are refined yielding a more detailed representation of more distant flow. The group/cluster representatives are aggregated in a visual summary of vortical flow.

With the focus on embedded vortices, the pipeline was applied to 17 aneurysm cases from which six were chosen for demonstration here. The corresponding visual summaries were positively evaluated by experts. They outperform the summaries generated by our previous approach [32] in terms of production time and accuracy. Our collaborating CFD engineers so far managed to manually investigate a few cases a day. Assuming an automated, reliable vortex core line extraction, the pipeline facilitates a batch processing of their database ($n = 210$) [7].

Relating embedded vortices to aneurysm rupture and hemodynamic parameters may contribute to an understanding of their implications, which may range from thrombosis initiation to high/low wall shear stress. However, this requires extending the considered time-window from one point in time to the full cardiac cycle. This will pose further challenges such as tracking saddle-node bifurcations over time [41].

Our results strongly depend on the vortex core line extraction and enhancement. While for the extraction no perfect criterion or algorithm delivering a ground truth exist [1], we carried out a parameter sensitivity analysis for the enhancement step. The resulting default values of all involved parameters yielded satisfactory results for the 17 cases. The parameters of streamline grouping and clustering-based group refinement influence the summaries rather slightly. No adjustment of their proposed default values was required.

Acknowledgments

The authors wish to thank G. Byrne, T. Günther, and H. Theisel for the fruitful discussions on flow topology. This work was partly funded by the Federal Ministry of Education and Research in Germany within the Research Campus STIMULATE under grant number 03FO16102A.

REFERENCES

[1] A. Biswas, D. Thompson, W. He, Q. Deng, C.-M. Chen, H.-W. Shen, R. Machiraju, and A. Rangarajan. An uncertainty-driven approach to vortex analysis using oracle consensus and spatial proximity. In *PacificVis, accepted*, 2015.

[2] S. Born, M. Markl, M. Gutberlet, and G. Scheuermann. Illustrative visualization of cardiac and aortic blood flow from 4D MRI data. In *PacificVis*, pages 129–136, 2013.

[3] G. Byrne and J. R. Cebral. Vortex dynamics in cerebral aneurysms. *arXiv:1309.7875v1*, 2013.

[4] G. Byrne, F. Mut, and J. R. Cebral. Quantifying the large-scale hemodynamics of intracranial aneurysms. *AJNR Am J Neuroradiol*, 35(2):333–338, 2014.

[5] J. R. Cebral, M. A. Castro, S. Appanaboyina, C. M. Putman, D. Millan, and A. F. Frangi. Efficient pipeline for image-based patient-specific analysis of cerebral aneurysm hemodynamics: Technique and sensitivity. *IEEE Trans Med Imaging*, 24(4):457–467, 2005.

[6] J. R. Cebral, F. Mut, J. Weir, and C. Putman. Quantitative characterization of the hemodynamic environment in ruptured and unruptured brain aneurysms. *AJNR Am J Neuroradiol*, 32(1):145–151, 2011.

[7] J. R. Cebral, F. Mut, J. Weir, and C. M. Putman. Association of hemodynamic characteristics and cerebral aneurysm rupture. *AJNR Am J Neuroradiol*, 32(2):264–270, 2011.

[8] C.-K. Chen, S. Yan, H. Yu, N. Max, and K.-L. Ma. An illustrative visualization framework for 3D vector fields. *Comput Graph Forum*, 30(7):1941–1951, 2011.

[9] J. Crawford. Introduction to bifurcation theory. *Rev. Mod. Phys.*, 63:991–1037, 1991.

[10] M. Daszykowski, B. Walczak, and D. Massart. Looking for natural patterns in data: Part 1. Density-based approach. *Chemometr Intell Lab Syst*, 56(2):83–92, 2001.

[11] D. Degani, A. Seginer, and Y. Levy. Graphical visualization of vortical flows by means of helicity. *AIAA journal*, 28(8):1347–1352, 1990.

[12] E. W. Dijkstra. A note on two problems in connexion with graphs. *Numerische Mathematik*, 1(1):269–271, 1959.

[13] M. H. Everts, H. Bekker, J. B. Roerdink, and T. Isenberg. Depth-dependent halos: Illustrative rendering of dense line data. *IEEE Trans Vis Comput Graphics*, 15(6):1299–1306, 2009.

[14] A. Gambaruto and A. João. Flow structures in cerebral aneurysms. *Computers & Fluids*, 65:56–65, 2012.

[15] D. Garcia. Robust smoothing of gridded data in one and higher dimensions with missing values. *Comput Stat Data An*, 54(4):1167–1178, 2010.

[16] C. Garth, X. Tricoche, T. Salzbrunn, T. Bobach, and G. Scheuermann. Surface techniques for vortex visualization. In *Joint Eurographics-IEEE TCVG Conference on Visualization*, pages 155–164, 2004.

[17] R. Gasteiger, D. J. Lehmann, R. van Pelt, G. Janiga, O. Beuing, A. Vilanova, H. Theisel, and B. Preim. Automatic detection and visualization of qualitative hemodynamic characteristics in cerebral aneurysms. *IEEE Trans Vis Comput Graphics*, 18(12):2178–2187, 2012.

[18] S. Glaßer, K. Lawonn, T. Hoffmann, M. Skalej, and B. Preim. Combined visualization of wall thickness and wall shear stress for the evaluation of aneurysms. *IEEE Trans Vis Comput Graphics*, pages 2506–2515, 2014.

[19] M. Jankun-Kelly, M. Jiang, D. Thompson, and R. Machiraju. Vortex visualization for practical engineering applications. *IEEE Trans Vis Comput Graphics*, 12(5):957–964, 2006.

[20] M. Jiang, R. Machiraju, and D. Thompson. Geometric verification of swirling features in flow fields. In *IEEE Visualization*, pages 307–314, 2002.

[21] M. Jiang, R. Machiraju, and D. Thompson. *The Visualization Handbook*, chapter Detection and Visualization of Vortices, pages 295–309. Elsevier, Amsterdam, 2005.

[22] D. B. Johnson. Efficient algorithms for shortest paths in sparse networks. *Journal of the ACM*, 24(1):1–13, 1977.

[23] B. Köhler, R. Gasteiger, U. Preim, H. Theisel, M. Gutberlet, and B. Preim. Semi-automatic vortex extraction in 4D PC-MRI cardiac blood flow data using line predicates. *IEEE Trans Vis Comput Graphics*, 19(12):2773–2782, 2013.

[24] R. S. Laramee, D. Weiskopf, J. Schneider, and H. Hauser. Investigating swirl and tumble flow with a comparison of visualization techniques. In *IEEE Visualization*, pages 51–58, 2004.

[25] R. Löhner. Automatic unstructured grid generators. *Finite Elem Anal Des*, 25(12):111–134, 1997. Adaptive Meshing, Part 1.

[26] T. McLoughlin, M. W. Jones, R. S. Laramee, R. Malki, I. Masters, and C. D. Hansen. Similarity measures for enhancing interactive streamline seeding. *IEEE Trans Vis Comput Graphics*, 19(8):1342–53, 2013.

[27] H. K. Moffatt. The degree of knottedness of tangled vortex lines. *J Fluid Mech*, 35(1):117–129, 1969.

[28] F. Mut, R. Aubry, R. Löhner, and J. R. Cebral. Fast numerical solutions of patient-specific blood flows in 3D arterial systems. *Int J Numer Method Biomed Eng*, 26(1):73–85, 2010.

[29] F. Mut, R. Löhner, A. Chien, S. Tateshima, F. Viuela, C. Putman, and J. R. Cebral. Computational hemodynamics framework for the analysis of cerebral aneurysms. *Int J Numer Meth Bio*, 27(6):822–839, 2011.

[30] M. Neugebauer, K. Lawonn, O. Beuing, P. Berg, G. Janiga, and B. Preim. Amnivis–a system for qualitative exploration of near-wall hemodynamics in cerebral aneurysms. *Comp Graph Forum*, 32(3pt3):251–260, 2013.

[31] L. O'Donnell, A. J. Golby, and C.-F. Westin. Tract-based morphometry for white matter group analysis. *NeuroImage*, 45:832–844, 2009.

[32] S. Oeltze, D. J. Lehmann, A. Kuhn, G. Janiga, H. Theisel, and B. Preim. Blood flow clustering and applications in virtual stenting of intracranial aneurysms. *IEEE Trans Vis Comput Graphics*, 20(5):686–701, 2014.

[33] F. H. Post, B. Vrolijk, H. Hauser, R. S. Laramee, and H. Doleisch. The state of the art in flow visualisation: Feature extraction and tracking. *Comput Graph Forum*, 22(4):775–792, 2003.

[34] R. C. Prim. Shortest connection networks and some generalizations. *AT&T Tech J*, 36(6):1389–1401, 1957.

[35] F. Sadlo, R. Peikert, and M. Sick. Visualization tools for vorticity transport analysis in incompressible flow. *IEEE Trans Vis Comput Graphics*, 12(5):949–956, 2006.

[36] J. Sahner, T. Weinkauf, and H.-C. Hege. Galilean invariant extraction and iconic representation of vortex core lines. In *EuroVis/IEEE VGTC Symposium on Visualization*, pages 151–160, 2005.

[37] J. Sahner, T. Weinkauf, N. Teuber, and H.-C. Hege. Vortex and strain skeletons in Eulerian and Lagrangian frames. *IEEE Trans Vis Comput Graphics*, 13(5):980–990, 2007.

[38] T. Salzbrunn, H. Jänicke, T. Wischgoll, and G. Scheuermann. The state of the art in flow visualization: Partition-based techniques. In *Simulation and Visualization (SimVis)*, pages 75–92, 2008.

[39] T. Salzbrunn and G. Scheuermann. Streamline predicates. *IEEE Trans Vis Comput Graphics*, 12(6):1601–1612, 2006.

[40] T. Schafhitzel, J. E. Vollrath, J. P. Gois, D. Weiskopf, A. Castelo, and T. Ertl. Topology-preserving λ_2-based vortex core line detection for flow visualization. *Comput Graph Forum*, 27(3):1023–1030, 2008.

[41] G. Scheuermann, W. Kollmann, X. Tricoche, and T. Wischgoll. *Numerical Simulations of Incompressible Flows*, chapter Evolution of topology in axi-symmetric and 3-d viscous flows, pages 622–643. World Scientific Publishing, 2003.

[42] S. Shafii, H. Obermaier, V. Kolár, M. Hlawitschka, C. Garth, B. Hamann, and K. I. Joy. Illustrative rendering of vortex cores. In *Eurographics Conference on Visualization (EuroVis)*, pages 61–65, 2013.

[43] S. Stegmaier, U. Rist, and T. Ertl. Opening the can of worms: An exploration tool for vortical flows. In *IEEE Visualization*, pages 463–470, 2005.

[44] C. Stoll, S. Gumhold, and H.-P. Seidel. Visualization with stylized line primitives. In *IEEE Visualization*, pages 695–702, 2005.

[45] D. Sujudi and R. Haimes. Identification of swirling flow in 3d vector fields. Technical Report AIAA 95-1715, Department of Aeronautics and Astronautics, MIT, Cambridge, MA, 1995.

[46] G. Taubin. A signal processing approach to fair surface design. In *ACM SIGGRAPH*, pages 351–358, 1995.

[47] R. F. P. van Pelt, S. S. A. M. Jacobs, B. M. ter Haar Romeny, and A. Vilanova. Visualization of 4D blood-flow fields by spatiotemporal hierarchical clustering. *Comp Graph Forum*, 31(3pt2):1065–1074, 2012.

[48] M. H. Vlak, A. Algra, R. Brandenburg, and G. J. Rinkel. Prevalence of unruptured intracranial aneurysms, with emphasis on sex, age, comorbidity, country, and time period: A systematic review and meta-analysis. *Lancet Neurol*, 10(7):626–636, 2011.

[49] P. J. Yim, G. B. C. Vasbinder, V. B. Ho, and P. L. Choyke. Isosurfaces as deformable models for magnetic resonance angiography. *IEEE Trans Med Imaging*, 22(7):875–881, 2003.

[50] H. Yu, C. Wang, C.-K. Shene, and J. H. Chen. Hierarchical streamline bundles. *IEEE Trans Vis Comput Graphics*, 18(8):1353–1367, 2012.

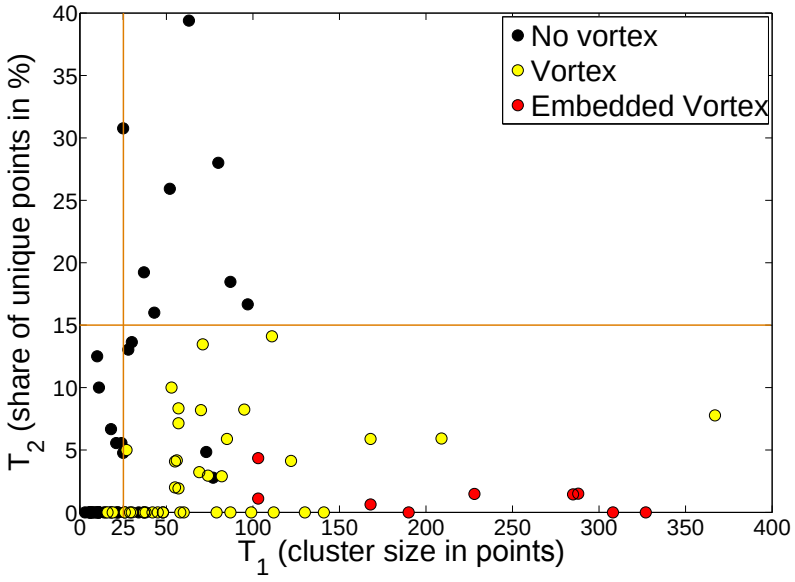
[51] L. Zelnik-Manor and P. Perona. Self-tuning spectral clustering. In *Advances in Neural Information Processing Systems*, pages 1601–08, 2004.

SUPPLEMENTAL MATERIAL

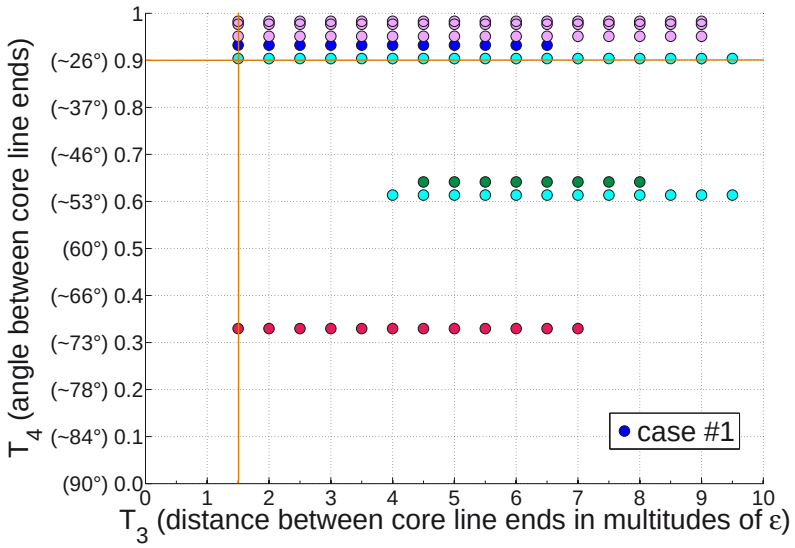
Paper Title: Cluster Analysis of Vortical Flow in Simulations of Cerebral Aneurysm Hemodynamics
Authors: Steffen Oeltze-Jafra, Juan R. Cebral, Gábor Janiga, and Bernhard Preim

1 PARAMETER SENSITIVITY ANALYSIS IN CORE LINE ENHANCEMENT

The parameter sensitivity analysis based on 17 aneurysm cases is detailed in Section 5.3 of the paper. Supplemental Figures 1 and 2 illustrate the sensitivity of the core line enhancement result with regard to the adjustable parameters T_1 to T_4 as well as good parameter default values.



Suppl.Fig. 1. Each dot in the plot represents a cluster. Cluster size (measured in points) and the deviation of a cluster from a distinct curve shape (a higher % indicates a higher deviation) are opposed on the x- and y axes. Orange lines are drawn at good default values for parameters T_1 and T_2. The set of clusters with $T_1 > 26$ points and $T_2 < 15\%$ exhibits the best trade-off between false-positive and -negative clusters representing vortices.



Suppl.Fig. 2. Each dot in the plot represents a merge event of two vortex core line ends. With an increasing distance (x-axis) and an increasing angle between core line ends (y-axis; measured as dot product between direction vectors) more merge events occur. All events occurring for the first time at $T_3 > 1.5\varepsilon$ (green and neighboring turquoise points) and $T_4 < 0.9$ (green and neighboring turquoise as well as red points) were identified as erroneous merges suggesting these values as the default setting (orange lines).

2 PERFORMANCE

Timings of selected steps of our analysis pipeline (Fig. 4 of the paper), the total time, and characteristics of our demonstration cases (Fig. 9(a-f) of the paper) are given in Supplemental Table 1. Note that vortex core line enhancement was timed but not core line extraction. Timings were taken on a 3.07 GHz Intel 8-core PC with 8 GB RAM and a 64 bit Windows operating system.

The duration of the core line enhancement (column 5) depends on the number of core line segments (c2) extracted before (Sec. 5.1 of the paper). The bottleneck is the extended shortest path search in a graph representing the segment endpoints (Sec. 5.2 of the paper). This search is based on an approximated path, which led to very similar visual summaries of vortical flow at a first glance. Applying this path would significantly reduce the computation time to the fraction of a second, e.g., 0.58 s for case #3.

Grouping streamlines (c6) takes longer than a second only for the giant aneurysm (case #3). The bottleneck are inter-streamline similarity computations based on closest point distances between all vertices. These are carried out for all lines seeded inside safety margins around saddles and nodes (Sec. 6.2 of the paper). Hence, the duration depends on the number of saddles and nodes (Fig. 9(a-f) of the paper), on safety margin size, and on the average number of streamline vertices (c4). The streamline count (c3) has a low impact since the majority of streamlines is seeded outside the safety margins (compare, e.g., case #2 and case #6).

The duration of group refinement (c7) depends on the number of streamlines (c3) and their average number of vertices (c4), both influencing the duration of inter-streamline similarity computations (Sec. 7.2 of the paper). It further depends on group count and group size (Suppl. Eq. 1) since the similarity computations are restricted to streamline-pairs from identical groups (compare, e.g., case #1 and case #2). In the best case, the g streamline groups are of equal size reducing the number of inter-streamline similarity computations from

$$\frac{n(n-1)}{2} \text{ to } \frac{n(-g+n)}{2g} \text{ by a factor of } \frac{g(n-1)}{(n-g)} \approx g. \tag{1}$$

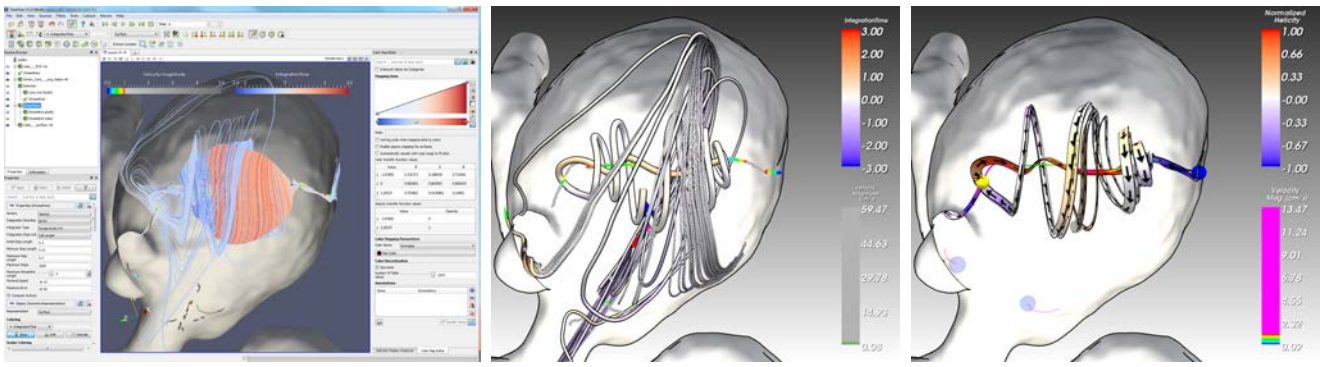
The total time (c8) represents the duration of our entire pipeline except the core line extraction and the group refinement step (c7). The streamline integration takes between 3.36 s for case #6 and 26.29 s for case #3 (Sec. 5.4 of the paper). The computation of group/cluster representatives takes the fraction of a second for all cases (Sec. 6.2 of the paper). The time for the composition of a visual summary ranges from 1.55 s for case #1 to 17.41 s for case #3 (Sec. 8.1 of the paper). Here, the geometry computation for the GPU-based streamline rendering is most time-consuming.

Suppl.Table 1. Characteristics of six demonstration cases (Fig. 9(a-f) of the paper) and timings $[s]$ of selected analysis pipeline steps (Fig. 4 of the paper). Number of vortex core line segments *(#Seg.)*, streamline count *(#Sl.)*, and average number of streamline vertices *(⌀#Vt.)* are given. Vortex core line enhancement *(1)*, grouping of streamlines *(3)*, optional clustering-based group refinement (4_{opt}), and the total duration have been timed.

	Case Characteristics			Time of Analysis Steps [s]			Total
Case	#Seg.	#Sl.	⌀#Vt.	1	3	4_{opt}	Time
#1	500	608	312	2.68	0.44	9.7	9.51
#2	674	859	299	2.80	0.46	10.0	13.78
#3	2469	1493	904	10.53	1.47	26.8	56.21
#4	628	735	444	2.63	0.56	12.7	13.77
#5	500	703	301	2.52	0.48	6.1	9.59
#6	350	587	533	1.84	0.57	6.8	11.65

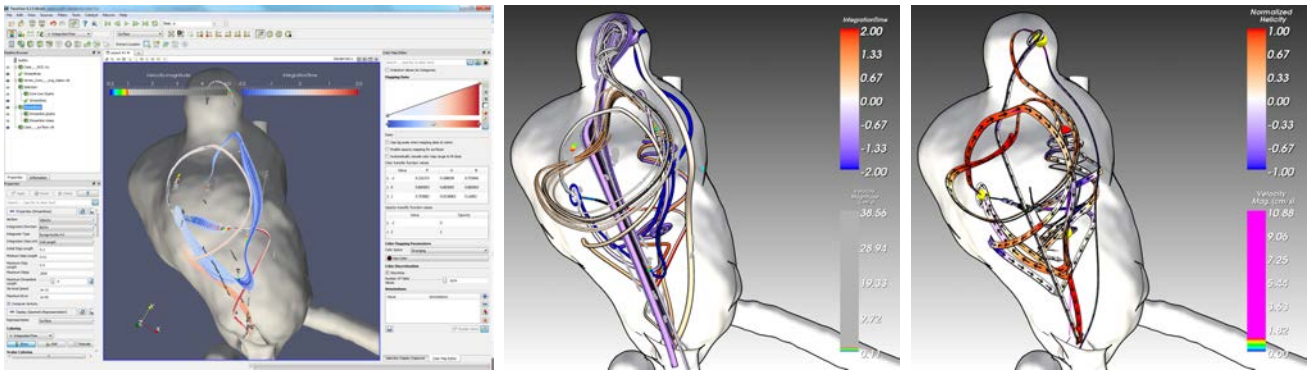
3 Interactive Visualizations Employed in the Evaluation

The evaluation is described in Section 9 of the paper. Supplemental Figures 3 and 4 show for two cases screenshots of the set of interactive visualizations on which the evaluation was based. All cases are surveyed in Figure 9 of the paper. The following abbreviations are used in the Supplemental Figure captions: *(VS)* our visual summary without the clustering-based group refinement, *(IS)* interactive seeding of streamlines at selected core line parts as in [3] (Sec. 4 of the paper), and *(SC)* spectral clustering of streamlines and display of cluster representatives [32].



(a) (b) (c)

Suppl.Fig. 3. Interactive visualizations of case #2 employed for the evaluation in Section 9: *IS* (left), *SC* (middle), and *VS* (right).

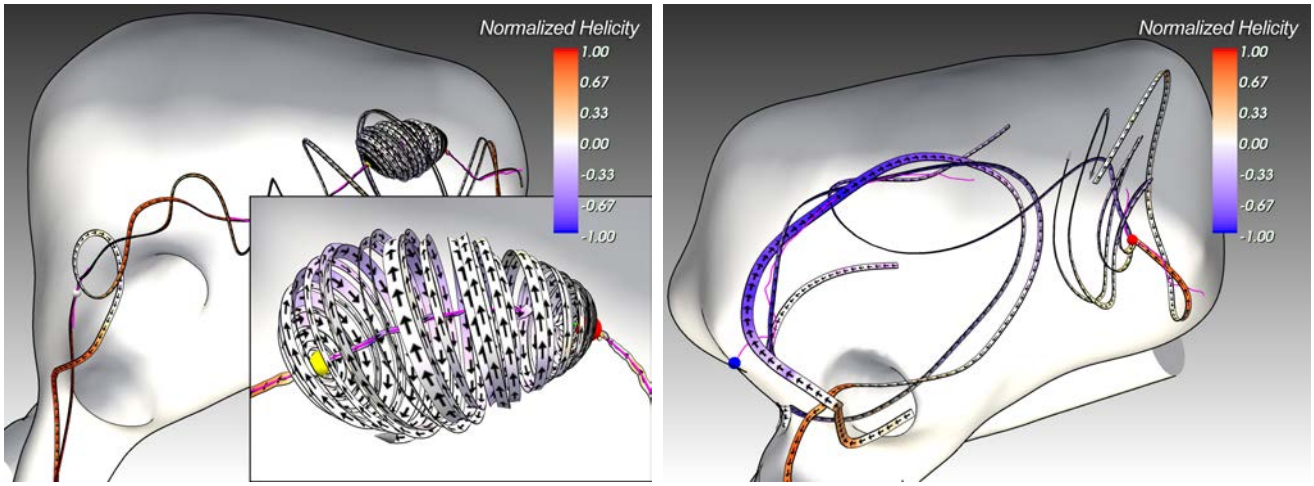


(a) (b) (c)

Suppl.Fig. 4. Interactive visualizations of case #3 employed for the evaluation in Section 9: *IS* (left), *SC* (middle), and *VS* (right).

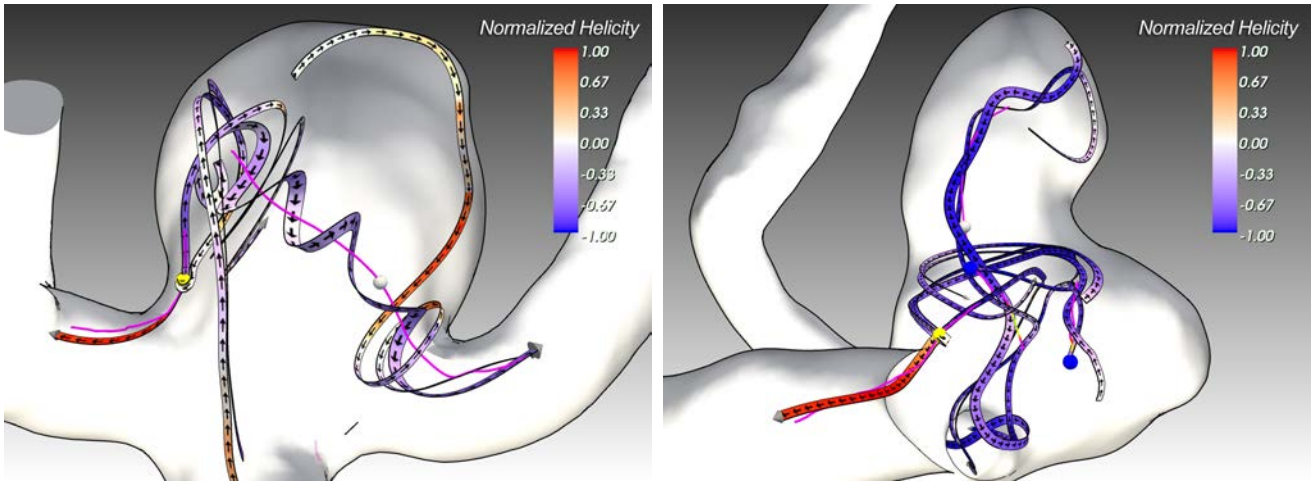
4 Visual Summaries of Vortical Flow of Additional Aneurysm Cases

Apart from the six demonstration cases in the paper, we investigated 11 additional cases. Visual summaries of a selection are shown in the following. Note that all of them could be generated using the default parameter settings proposed in the paper. Vortex core lines exhibiting no equilibrium point (not seen in the demonstration cases) are indicated by a gray sphere.



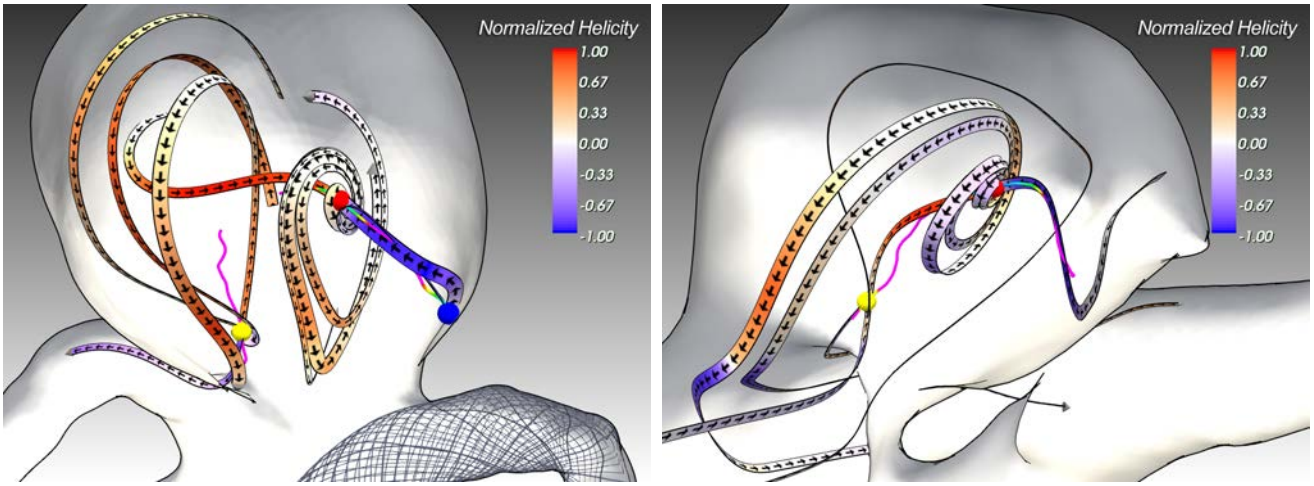
(a) (b)

Suppl.Fig. 5. The hemodynamic data of these two cases was generated within the scope of the CFD rupture challenge in 2013 (see Janiga et al., AJNR Am J Neuroradiol; 36(3):530-6, 2015 for details). The case in *(a)* exhibits an embedded vortex (inset). The case in *(b)* shows one core line attached not the neck of the aneurysm (blue sphere) and another core line attached to its dome and exhibiting a node (red sphere).



(a) (b)

Suppl.Fig. 6. The basilar-tip aneurysm shown in *(a)* exhibits a major swirl in the aneurysm's center (core line with gray sphere) and a saddle (yellow sphere) close to the neck of the aneurysm. Three vortices are observed for the case shown in *(b)*. The aneurysm exhibits a strong vortex (core line with gray sphere) located in a large local outpouching (*bleb*) at the top of the aneurysm sac. Blebs are associated with an increased risk of rupture. A saddle (yellow sphere) is located close to the neck of the aneurysm and one core line is attached to the aneurysm wall (blue sphere).

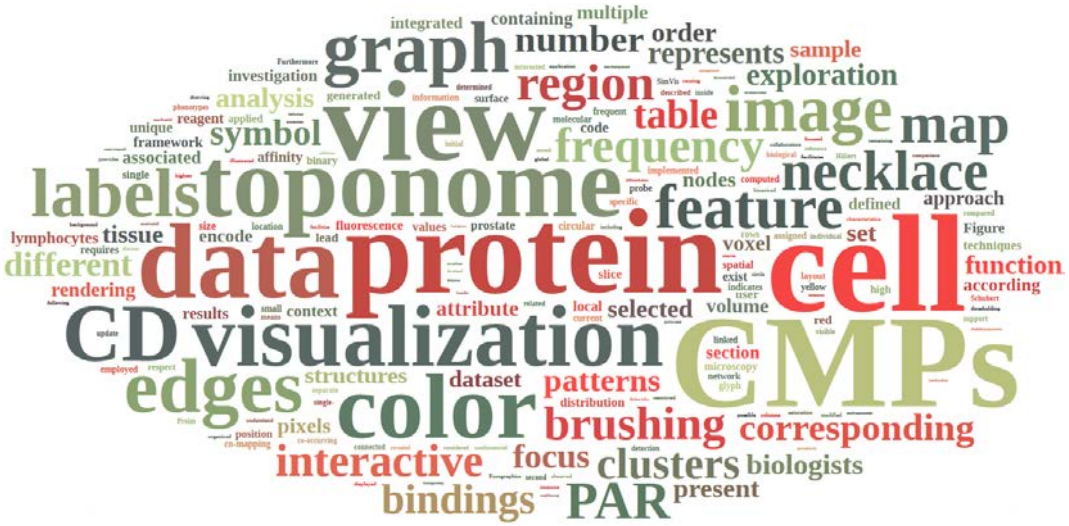


(a) (b)

Suppl.Fig. 7. The virtually stented case in *(a)* exhibits are core line attached to the aneurysm wall (blue sphere) and a node (red sphere) along this core line. A saddle (yellow sphere) is located close to the neck of the aneurysm. The stent is represented by a wire mesh in the lower right of the Figure. The case in *(b)* corresponds to case #6 in Figure 9(d) of the paper. It represents the aneurysm before the virtual stenting procedure. Note that the vortex (core line) is located at a very similar position and also exhibits a node (red sphere) and a saddle (yellow sphere).

Part II

Microscopy Data of Cells and Tissues



This part of the postdoctoral thesis cumulates the following publications:

Chapter 4 [166] **S. Oeltze**, W. Freiler, R. Hillert, H. Doleisch, B. Preim, and W. Schubert, "Interactive, Graph-Based Visual Analysis of High-Dimensional, Multi-Parameter Fluorescence Microscopy Data in Toponomics", *IEEE Trans. Vis. Comput. Graphics (IEEE TVCG)*, vol. 17, no. 12, pp. 1882-1891, 2011.

Chapter 5 [167] **S. Oeltze**, P. Klemm, R. Hillert, B. Preim, and W. Schubert, "Visualization and Exploration of 3D Toponome Data", *Eurographics Workshop on Visual Computing for Biology and Medicine (EG VCBM)*, pp. 115-122, 2012.

Chapter 6 [172] **S. Oeltze-Jafra**, F. Pieper, R. Hillert, B. Preim, and W. Schubert, "Interactive Labeling of Toponome Data", *Eurographics Workshop on Visual Computing for Biology and Medicine (EG VCBM)*, pp. 79-88, 2014.

Interactive, Graph-Based Visual Analysis of High-Dimensional, Multi-Parameter Fluorescence Microscopy Data in Toponomics

Steffen Oeltze, *Member, IEEE*, Wolfgang Freiler, Reyk Hillert, Helmut Doleisch, *Member, IEEE*, Bernhard Preim, and Walter Schubert

Abstract—In *Toponomics*, the function protein pattern in cells or tissue (the *toponome*) is imaged and analyzed for applications in toxicology, new drug development and patient-drug-interaction. The most advanced imaging technique is robot-driven multi-parameter fluorescence microscopy. This technique is capable of co-mapping hundreds of proteins and their distribution and assembly in protein clusters across a cell or tissue sample by running cycles of fluorescence tagging with monoclonal antibodies or other affinity reagents, imaging, and bleaching *in situ*. The imaging results in complex multi-parameter data composed of one slice or a 3D volume per affinity reagent. Biologists are particularly interested in the localization of co-occurring proteins, the frequency of co-occurrence and the distribution of co-occurring proteins across the cell.

We present an interactive visual analysis approach for the evaluation of multi-parameter fluorescence microscopy data in toponomics. Multiple, linked views facilitate the definition of features by brushing multiple dimensions. The feature specification result is linked to all views establishing a focus+context visualization in 3D. In a new attribute view, we integrate techniques from graph visualization. Each node in the graph represents an affinity reagent while each edge represents two co-occurring affinity reagent bindings. The graph visualization is enhanced by glyphs which encode specific properties of the binding. The graph view is equipped with brushing facilities. By brushing in the spatial and attribute domain, the biologist achieves a better understanding of the function protein patterns of a cell. Furthermore, an interactive table view is integrated which summarizes unique fluorescence patterns. We discuss our approach with respect to a cell probe containing lymphocytes and a prostate tissue section.

Index Terms—Visual Analytics, Fluorescence Microscopy, Toponomics, Protein Interaction, Graph Visualization.

1 Introduction

While the human genome project has revealed, among other things, the code for all proteins, the next big challenge is to understand how proteins cooperate in cells and tissues in time and space [30]. The *toponome* of a cell describes its function protein pattern, i.e. the location and topological distribution of proteins. In *Toponomics*, the toponome is imaged, explored and analyzed for applications in toxicology, drug development and patient-drug-interaction. In the traditional fluorescence microscopy, a maximum of five proteins may be mapped concurrently. With a sophisticated flow cytometer up to 17 proteins may be mapped [23]. However, this number is still insufficient for mapping protein network features. The most advanced imaging technique is robot-driven multi-parameter fluorescence microscopy [28]. It is capable of co-mapping hundreds of proteins and their distribution across a cell or tissue sample by running cycles of fluorescence tagging with monoclonal antibodies or other affinity reagents, imaging, and bleaching *in situ*. The imaging results in complex multi-parameter data composed of one slice or a 3D volume per affinity reagent. In a preprocessing step, the data is binarized such that 1 encodes protein present and 0 encodes protein absent. Biologists are particularly interested in the localization of co-occurring proteins, i.e. co-occurring affinity reagent bindings, the frequency of co-occurrence and the distribution of co-occurring proteins across the cell. This information is crucial in order to understand how proteins cooperate in cells and tissues in time and space.

We present an interactive visual analysis approach for the evaluation of multi-parameter fluorescence microscopy data in toponomics. Multiple, linked views facilitate the definition of features by brushing multiple dimensions. The feature specification result is linked to all views establishing a focus+context style of visualization in 3D. In a new attribute view, we integrate techniques from graph visualization. Each node in the graph represents an affinity reagent while each edge represents two co-occurring affinity reagent bindings. The graph visualization is enhanced by glyphs which encode specific properties of the binding. The graph view is equipped with brushing facilities. By brushing in the spatial and attribute domain, the biologist achieves a better understanding of the function protein patterns of a cell. Additional attributes, such as local binding frequency, are derived from the data and made accessible from other attribute views, e.g., histograms and scatter plots. Furthermore, an interactive table view summarizes unique fluorescence patterns existing in the data. We discuss our approach with respect to a cell probe containing lymphocytes and a prostate tissue section. The probe has been stained with 32 protein affinity reagents (PARs) and imaged at 20 different slice locations. The tissue section has been stained with 17 PARs and imaged at 1 slice location. So far, we could test our approach only for the investigation of these two datasets since datasets from other studies were not available in the present cooperation and the temporal and financial costs which are involved in the imaging are high (sample preparation and imaging together often require several days [10]). However, at the end of this paper, we will elaborate on the generality of our approach including its scalability.

- *Steffen Oeltze, Reyk Hillert, Bernhard Preim and Walter Schubert are with the University of Magdeburg, E-mail: {oeltze|preim}@ovgu.de, {reyk.hillert|walter.schubert}@med.ovgu.de.*
- *Wolfgang Freiler and Helmut Doleisch are with SimVis GmbH, Vienna, E-mail: {freiler|doleisch}@simvis.at.*

Manuscript received 31 March 2011; accepted 1 August 2011; posted online 23 October 2011; mailed on 14 October 2011.
For information on obtaining reprints of this article, please send email to: tvcg@computer.org.

2 Biological and Technical Background

This section gives an overview on the biological background and familiarizes the reader with the imaging technique that we apply for mapping the toponome.

2.1 Toponomics

Toponomics requires new technologies which are able to co-localize a quasi random number of different proteins in the one biological sample in order to map what any cell does in reality: it forms functional protein patterns (assemblies of clusters of different proteins) to generate concrete cell functions. Hence, the cell is a protein pattern formation apparatus [28, 30]. This machinery, the whole functional plan of the cell, still poses many open questions. Neither the really existing pro-

tein clusters in given functional states of a cell *in vivo/in situ*, nor the rules of their formation can be derived from genomic or pure molecular protein data. Although many details on the molecular function and structure of many proteins are known, we cannot simply derive the corresponding cellular functions of these proteins, because the latter are dependent on the contextual position of a given protein within a protein network inside the cell [28, 29, 30, 31, 32, 33]. It is therefore essential to distinguish the molecular function from the cellular function of a protein, defined by the spatial protein context cell by cell. The entirety of all protein networks, in which proteins are defined by their protein-to-protein context in any given cell, is defined as the *toponome* [10, 28, 32]. The toponome is a system of proteins. Its inner structure, its biological code, and its semantics are investigated in *toponomics*.

The toponome in cells is hierarchically organized: protein clusters interlocked as a network contain *lead protein(s)* that control the topology of the protein clusters and their function as a whole network [32]. This has been clearly demonstrated by knocking down the lead protein or inhibiting it by a chemical agent which result in a disassembly of the whole protein cluster network and an essential alteration of the cellular function that is encoded by this network [32]. Neither a lead protein nor the protein cluster which is controlled by it can be predicted from molecular data. Thus, the toponome must be mapped in human tissues and human or animal cells to understand how cells encode the myriads of different functionalities both in health and disease. Many investigations have shown that mapping the toponome is essential for finding new drugs in cancer and for finding protein clusters that can be regarded as a new system of biomarkers in disease [2, 30, 34].

2.2 Imaging the Toponome

Imaging the toponome is based on a cyclical imaging procedure in which a *tag library* (specific affinity reagents recognizing proteins) is conjugated to one and the same dye. A toponome imaging system (TIS) robot [10] applies these tags sequentially in the following way: the first dye-conjugated tag is applied to a fixed cell or tissue sample, the resulting fluorescence image is registered by an epifluorescence microscope which is connected to a CCD camera, and the dye is then bleached gently to avoid any energy transfer into the remaining proteins [27, 28, 32]. This first incubation imaging bleaching cycle is followed by a second incubation imaging bleaching cycle using a second tag reacting with the second protein of interest, and so on. This imaging procedure can be performed in 2D or 3D. Up to 100 cycles to label 100 proteins and co-map these proteins at any given sub-cellular data point in a cell have been demonstrated [32].

A wide-spread method for the rapid analysis of toponome data is thresholding each protein fluorescence signal thereby generating a combinatorial binary code, where 0 indicates protein absent and 1 indicates protein present [1, 9, 28, 32]. In case of a 3D dataset, this results in a binary code at each voxel. The size of the code equals the number of applied tags. All binary codes that exist in the data, out of all possible combinations of 0 and 1, can be assembled in a *toponome map* and be referred to as combinatorial molecular phenotypes (CMPs). In a toponome map, a unique color is assigned to each CMP. Since the corresponding binary code frequently occurs at several locations, the map visualizes the location of given protein clusters present in groups of pixels or voxels. We introduce the term *single-1-CMP* which refers to a CMP which contains only a single element equal 1.

3 Prior and Related Work

This section describes prior and related work on the application of graph visualization techniques for investigating biological data and on the visual analysis of such data by employing multiple coordinated views. It starts with a brief discussion on how our collaborating biologists have been analyzing their data so far.

After imaging the toponome and thresholding the data, the CMPs are determined and added as rows to a table. Each row then represents a group of pixels/voxels having the same binary code. The table may be sorted according to the columns which represent the employed affinity reagents. The table is linked to a toponome map such that the user may select individual CMPs and observe their location in the data or define a region of interest in the toponome map thereby restricting the table to the corresponding CMPs. A limitation of this setup is the ability to explore a 3D dataset as a whole volume instead of slice by slice. Furthermore, it is impossible to select individual co-occurring PAR bindings across different CMPs and the frequency of co-occurrence is very difficult to infer from just the table. The distribution of CMPs across the cell may be recognized in the toponome map whereas the distribution of individual co-occurring proteins remains hidden in the data.

3.1 Graph Visualization of Biological Data

The most frequent use of graphs in biology is the visualization of biological networks representing the metabolism of cells, the regulation of genes and the interaction of proteins. Graph visualizations of such networks encode biochemical processes and provide an important means to understand their complex nature. A survey of visualization tools for biological network analysis is given in [14]. One prominent example is *Cytoscape* [36]. A central component of the software is the graph view which represents molecular species as nodes and intermolecular interactions as edges. Severals graph layout algorithms have been integrated including a circular layout. Another popular software for studying pathways, gene regulation and systems biology is *VisANT* [18]. It is implemented as an on-line tool and integrated with standard databases for organized annotation. All information gathered in VisANT may be visualized in a graph. Among others, a circular layout has been integrated. The usefulness of circular graph layouts for visualizing expressions of genes across a set of patients has been investigated in [37]. Genes are clustered and the resulting clusters are visualized as a set of circles. Edges reflecting the correlation between two genes are drawn inside a cluster and across clusters. Special care is taken to generate a visually pleasing layout of the set of circles. An extension of circular graph layouts to 3D is integrated in the software *Arena3D* [22]. Proteins or pathways can be iteratively grouped and the result of each iteration can be arranged on a separate higher layer. This results in a hierarchical 3D layout where for each layer, e.g., a circular layout may be employed. The described tools are very diverse, but each of them provides among others a circular graph layout since this results in a compact view of the investigated network. Furthermore, the modification of a node's and an edge's visual properties according to a biological aspect as well as the ability to interact on the graph, e.g., select a subset of nodes, are essential.

3.2 Visual Analysis of Biological Data

DeLeeuw et al. present the *ARGOS* system for interactive exploration and batch processing of confocal laser microscopy data [5]. They combine views of the image data with attribute views, e.g. parallel coordinates, scatter plots and histograms. The views are realized as separate windows which leads to maximum flexibility at the expense of guidance. Weber et al. present results of a long-term effort of exploring gene regulatory networks by means of relating them to 3D gene expression data [38]. Thus they had to find visualization techniques which convey the spatial position of gene expression data and to relate these visualizations to attribute views. For this purpose, they created different views on the spatial data including those where they unroll the data to a cylindrical shape and project it in 2D. To relate the spatial data to the attribute views, they developed a query-mechanism. There are a number of other techniques and systems where multiple views are coordinated to investigate biological data. However, most of them do not incorporate image data. In [19], a multiscale synteny browser (MizBee) for the comparison of two genomes is presented. MizBee incorporates linked genome, chromosome, and block views. The genome view applies a double circular graph layout that opposes source and destination chromosomes. Schulz et al. developed special graph visualization techniques for bipartite biological networks, a special kind of network which is frequently used to study findings in biology [35]. The *SpRay*-system supports the exploration of gene expression data with various InfoVis views [6]. To the best of our knowledge, no system exists that can readily process our data and link the image data to the derived information.

4 METHODS

In this section, we provide details on our visual analysis framework and on the integration of the graph and the table view. We put special emphasis on the graph view, in particular, the graph layout, the modification of an edge's visual attributes, the integration with glyph drawings, and the equipment with brushing facilities.

4.1 The Visual Analysis Framework

The interactive visual analysis is carried out in a framework employing the SimVis technology [7]. SimVis has been developed for the analysis of multi-parameter data. It has originally been engineered for the analysis of 3D time-dependent flow simulation data but has been extended to cope with data measured on structured grids, e.g., 3D/4D medical data [20, 21]. Recently, a structured grid ray-casting renderer has been integrated. In the framework, multiple views from scientific and information visualization may be linked together, e.g., 3D view, histogram, scatter plot, and parallel coordinates. By linking these views, the framework supports the concurrent exploration of the observation space (3D view) and the attribute space (InfoVis views). The user may choose an attribute for representation in one of the InfoVis views and then, brush an interesting range of attribute values, i.e. specify a feature. The result of such a brushing operation is reintegrated into the data in form of a synthetic data attribute $DOI_j \in [0,1]$ (*degree of interest (DOI)*, compare to [12]). This DOI attribution is used in the views of the analysis setup to visually discriminate the interactively specified features from the rest of the data in a focus+context visualization style which is consistent in all (linked) views [15].

There are several interactive visual analysis tasks. In a *feature localization*, the user searches for places in the 3D/4D domain where certain feature characteristics are present. The user can brush features in attribute views and concurrently localize the respective feature in the observation space. In a *multi-variate analysis*, multi-variate data properties are investigated by specifying a feature in one attribute view and at the same time analyzing the DOI distribution with respect to other data attributes in other attribute views (through view linking). In a *local investigation*, the user inspects the values of selected data attributes with respect to certain spatiotemporal subsets of the 3D volume domain. In SimVis, the user can also load spatial as well as temporal data references into attribute views—brushing these kinds of data attributes then yields features which are specified according to their spatiotemporal extents. Each feature is described in SimVis by a feature characteristic and organized in a feature set:

- *feature set*: subsumes an arbitrary number of features (logical OR combination of features), only one feature set is used by the framework at the time
- *feature*: specified by one or more feature characteristics (logical AND combination of DOI functions of all feature characteristics)
- *feature characteristic*: simple (attribute+bounds) or complex (AND/OR/NOT of several brushes) feature characteristic defined on an attribute view

Each attribute view is associated to a feature characteristic and hence, to a feature. SimVis uses standard colors for the geometric primitives in an attribute view, e.g., scatters in a scatter plot, thereby indicating the belonging of the corresponding data points to features:

- *red*: belong to the associated feature and all other features in the current feature set
- *green*: do not belong to the associated feature but to at least one other feature in the current feature set
- *yellow*: do neither belong to the associated feature nor to any other feature in the current feature set
- *dark gray*: only exists if ≥ 1 brush is defined for the associated attribute view; then, corresponding data points do not belong to the associated feature

A mixing of color (averaged DOI) may occur when a geometric primitive in an attribute view represents multiple data points. For example, a simple range of values has been brushed on one attribute in a histogram. In a linked scatter plot, two other attributes are opposed. As a result of the brushing, pure red, pure yellow as well as scatters in various shades of orange appear. A mixed color scatter indicates that some of the corresponding data points exhibit values within the brushed range while others do not. The percentage of each separate color is proportional to the respective number of data points. In all attribute views, color mixing may be turned on and off. If it is off, the drawing priority in decreasing order is red, green, yellow, gray (see [20] for more details on the colors and their mixing).

4.2 The Graph View

The co-occurrence of proteins across the cell and the frequency of co-occurrence are of high interest for biologists. Also the spatial distribution of co-occurring proteins and a comparison of this distribution across cells are crucial. The co-occurrence is represented in the data by pairs of 1s in the binary codes and each such pair may be considered as a relation. The relation is not directional and since PARs are not grouped together, no hierarchy exists in the data. Graphs offer a comprehensive and comprehensible representation of relational data and are frequently used in bioinformatics [14]. Transferred to our binary data, each node in an undirected graph may represent a PAR and each edge may represent two co-occurring PAR bindings, i.e. a pair of 1s in a binary code. The frequency of co-occurrence should be encoded by a visual property of the edge, e.g., edge color or edge width. Understanding the spatial distribution of co-occurring PAR bindings would be supported by integrating the graph into the visual analysis framework as an additional attribute view. Furthermore, the new view may be equipped with brushing facilities, e.g., for node and edge selection, and the brushing result should be reintegrated into the data as DOI thereby linking the graph view to all other views. This allows to select an edge with an associated high co-occurrence frequency and then, show all sites in the 3D view where the two corresponding PARs are concurrently binding. Further, by also updating the graph view according to the modified DOI values, PAR bindings which coincide with the selected PARs are highlighted, i.e., edges connecting these PARs with the selected PARs are emphasized.

Requirements. The requirements on the graph view are derived from many discussions with the biologists. The graph layout should:

- enable a fast recognition of a PAR of interest
- enable a quick inference of which other PARs are connected by an edge to a PAR of interest, i.e., co-occurring PAR bindings
- be static, i.e. node positions must not change during a DOI update in order to guarantee a fast PAR recognition
- avoid visual clutter by minimizing the number of edge crossings

The following information should be communicated by the graph:

- PAR bindings do/do not co-occur
- Co-occurrence frequency in the feature and in the context
- SimVis standard colors (see Sec. 4.1)
- PAR name and frequency of the PAR binding in the feature
- frequency of single-1-CMPs with a 1 for this PAR in the feature
- degree of the node corresponding to the PAR

With regard to interactivity, the graph view should support:

- a repaint of the graph at interactive frame rates when the feature specification has been updated
- brushing in order to include PARs (nodes) in the feature (AND/OR) and to exclude PARs (NOT)

Development Environment. The graph view is implemented in C++ and added to SimVis as a plug-in. The graphical user interface of SimVis and hence, of the plug-in are realized in Qt. The information visualization classes of VTK are used for implementing the graph-related functionality. VTK provides most of the relevant graph features, such as different layout strategies. The most important VTK classes which have been employed are `vtkMutableUndirectedGraph` and `vtkGraphLayoutView`. While the first implements a data structure representing an undirected graph, the latter implements functionality for laying out and drawing a graph. Moreover, a plenty of new functionality has been added for customizing the graph drawing and augmenting its expressiveness by integrating additional information. A crucial step in SimVis is the continuous traversal of all voxels while the feature specification is being changed in order to determine those which do/do not adhere to the specification. In order to offer interactive frame rates, we have parallelized the voxel traversal using OpenMP [3]. Since SimVis processes the data block-wise, the parallelization of the traversal is relatively straightforward by letting each thread process a block. For the graph and the table view, a thread-safe computation of a graph and a list of CMPs per block and a final merging step have been implemented.

Preprocessing. Preprocessing includes the determination of the edges that exist in the graph out of all theoretically possible edges, the determination of the frequency of each PAR binding, i.e., the *edge-weight*, and for each PAR, the computation of the number of single-1-CMPs. The latter information is not inherently represented by the graph and is therefore added to the view.

At first, the PARs are determined which bind at neither position in the cell probe, i.e. whose corresponding binary images are zero everywhere. These PARs are excluded from further processing. They are presented in a text inset of the graph view (Fig. 3 (b)).

Next, several arrays are initialized. The array $[EW]$ is going to contain the edge-weights. Its size is set to the number of possible graph edges: $n(n-1)/2$, where n is the number of PARs. The arrays $[PAR]$ and $[PAR1]$ are going to contain the PAR binding-related information. The size of the arrays is set to the number of PARs. Finally, an array $[CN]$ is initialized which is going to contain indexes into the nodes connected by an edge in $[EW]$. $[CN]$ is filled with all possible index pairs and its size is set to two times the size of $[EW]$. Next, a nested `for`-loop begins. The outer loop iterates over all voxels (data points will be referred to as voxels in the following) whereas the inner loop iterates over all binary attributes, i.e. PARs. Within the inner loop, the current binary code is reconstructed and all pairs of 1s are determined without repetition and inversion of CMP elements. The elements of $[PAR]$ are incremented at those positions where the elements of the code equal 1. If the code contains just a single 1, $[PAR1]$ is updated accordingly. Next, the index into $[EW]$ is computed for each pair of ones and the corresponding element of $[EW]$ is incremented. At the end of the `for`-loop, $[EW]$ and $[CN]$ are squeezed such that zero elements are deleted.

An additional step, which is carried out within the inner part of the `for`-loop, is the filling of a data structure which will speed up the repaint of the graph on a feature specification update. The new data structure is realized as a Vector, a special type of dynamic array. It is initialized before the outer loop starts and its size is set to the size of the dataset. Within the inner loop, all indexes into $[EW]$ which have been computed for the current voxel, are stored at the corresponding voxel position in the Vector. They are again stored in the form of a Vector v_{sub}. The majority of voxels in the lymphocytes dataset represent protein absent. Here, no indexes are computed resulting in a sparse Vector containing many Vectors v_{sub} of size zero. On the repaint of the graph, all voxels are traversed again and for the non-background voxels, the corresponding indexes into $[EW]$ have to be determined. By means of the Vector, only a readout of each v_{sub} is required. Without the Vector, this would require recomputing the indexes which in turn involves touching every binary attribute as described above. This is particularly ineffective when the voxel represents background which may easily be inferred from v_{sub} by just checking its size.

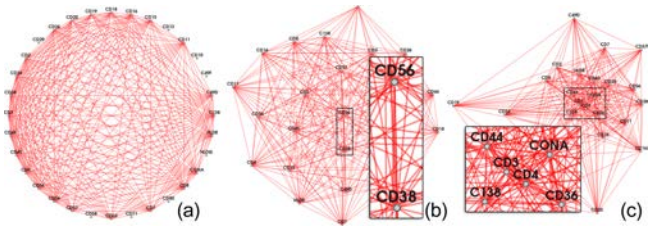


Fig. 1. Different graph layouts: circular (a), force-directed (b), and force-directed under consideration of edge frequency (c). The insets illustrate the strong overlap of edges in (b) and the aggregation of nodes which are connected by edges with a high frequency in (c).

Graph Layout. Two strategies have been tested for laying out the graph in 2D: a circular layout and a force-directed layout [11]. In a circular layout, all nodes are uniformly arranged on a circle (Fig. 1 (a)). In a force-directed layout, the nodes are arranged such that the variation in edge length as well as the number of edge crossings are minimized and a symmetric layout is achieved (Fig. 1 (b)). In addition, edge-weights may be considered such that the nodes connected by an edge with a high weight are positioned close together, e.g., CD3 and CD4 which are connected by the most frequent edge in Figure 1 (c).

Considering the above mentioned requirements on a graph layout, the circular layout performs better than the force-directed layout except for the number of edge crossings. The PARs are easier to find since their position follows an obvious rule and they are ordered alphabetically starting at 3 o'clock. In the force-directed layout, some edges strongly overlap which hampers a separate tracing (inset of Fig. 1 (b)). In an interactive analysis, edge-weights are constantly recomputed with respect to the updated feature characteristics. If edge-weights influence the positioning of nodes, the force-directed layout then constantly changes. This hampers a fast recognition and a visual tracking of nodes representing PARs of interest.

Edge Color. After the nodes of the graph have been laid out in a circular fashion, edges need to be constructed. For each edge in $[EW]$, the corresponding nodes are retrieved from $[CN]$ and a line is drawn between them. The edge color is modified according to the SimVis standard colors (Fig. 2 (a), recall Sec. 4.1). The selected color depends on the voxels which contain this edge and their belonging to the features in the feature set. As can be seen, it is difficult to visually group edges of equal color due to the large number of edge crossings. However, early feedback from the biologists indicated that mostly only edges being part of the feature set are of interest, i.e. red and green edges. Furthermore, green edges are only interesting in a comparison of different cells or different cell parts which must be described by different features. In order to account for this feedback, the opacity of the edges is initially modulated and may be further adjusted by the user. Red is assigned full opacity, while green and yellow are assigned an opacity of 50% and 20%, respectively (Fig. 2 (b)).

Edge Width. We decided to map the edge-weights in $[EW]$ to the width of the graph edges. Width is chosen instead of color or opacity since the perception of quantities depicted by area is more accurate [4]. Unfortunately, VTK only offers a uniform change of edge width which requires a workaround. Each line representing an edge is replaced by a ribbon that is generated in 2D by means of `vtkRibbonFilter` (Fig. 2 (c)). The width of the ribbons is modified according to the values in $[EW]$. All values are normalized with respect to an empirically determined minimum and maximum width. Width modulation and normalization are only applied to edges being part of the feature. All other edges are assigned a uniform width being slightly below the minimum. In Figure 2 ((a) and (c)), two problems become obvious. At first, yellow edges are often drawn in front of red feature edges since the drawing order does neither consider importance nor edge width. Second, edges with a strong overlap are sometimes difficult to visually separate when being assigned the same color (inset of Fig. 2 (c)). The first problem could be solved by adapting the drawing order each

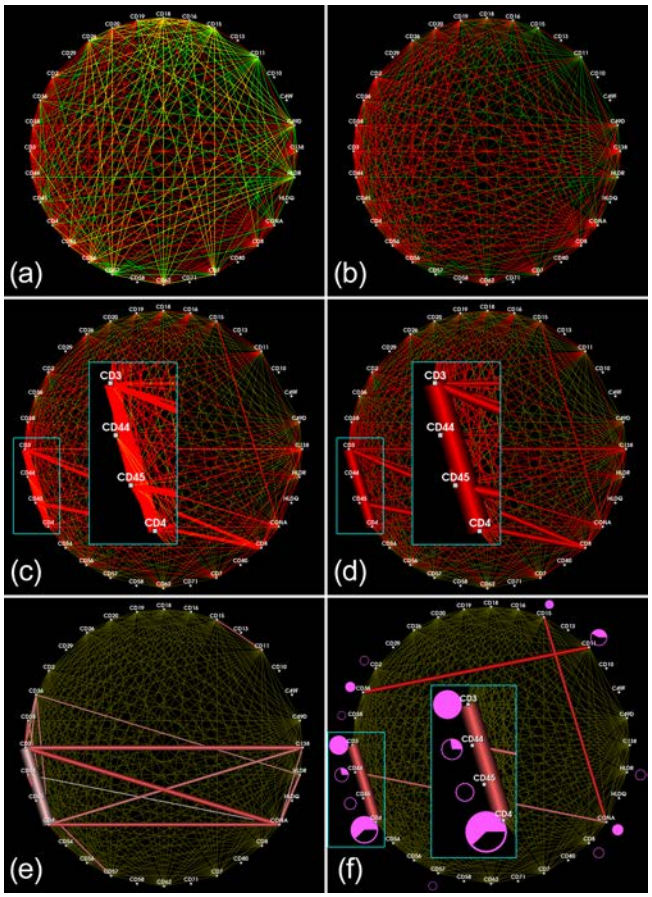


Fig. 2. Development of the graph view. *(a)*: Edge color is modified according to the SimVis standard colors. *(b)*: Opacity modulation to emphasize edges that are part of features. *(c)*: Ribbons encode the edge-weights. *(d)*: Ribbons are replaced by tubes thereby resolving the dependency between edge drawing order and good visibility. *(e)*: A feature has been defined. The different shades of red encode the frequency of each edge inside and outside of the feature. *(f)*: Circular glyphs encode node-specific properties such as binding frequency inside the feature. Please, note the close-up views in *(c)*,*(d)*, and *(f)*.

time the feature specification changes and draw red edges on top of the yellow edges. However, a less expensive solution also solves the second problem as a side effect. The ribbons have been replaced by tubes constructed of four faces by means of `vtkTubeFilter` (Fig. 2 (d)). The tube radius is modified according to the edge-weights. To ensure a rounded appearance, surface normals are computed and the tubes are oriented such that a salient edge faces the viewer. Without any additional sorting, edges now automatically appear in increasing order of their associated edge-weight from back to front. The thickest edge, i.e. the tube representing the highest weight is drawn in front of all other edges. Also, feature edges are drawn in front of context edges (yellow) since the latter are guaranteed to have a lower radius. With regard to the second problem, partially overlapping edges are now easier to separate due to shading (inset of Fig. 2 (d)).

Feature Edge Saturation. In a local investigation, the user inspects a certain subset of the observation space, e.g., a single cell. From the current graph visualization it remains unclear whether a PAR binding (edge) only occurs in this cell. In other attribute views of SimVis, such information is conveyed by percentaged color mixing of red and yellow (Sec. 4.1). However, a color map with a varying hue component is less effective in encoding variations in magnitude than isomorpic color maps, i.e. maps where saturation or luminance are increased monotonously [25]. Hence, we modify the saturation of the feature edges (Fig. 2 (e)). If the corresponding PAR binding is fully contained in the feature, the edge is assigned a fully saturated red. If it often occurs in the context, the edge is assigned a very light red. In between, the saturation is interpolated. From our meetings with the biologists, we learned that for a pure exploration of the data, the definition of one feature set including a single feature is sufficient. The biologists then browse through the data by constantly modifying the feature's characteristics. Hence, only mixtures of red and yellow may occur which are replaced here by a varying saturation. The definition of multiple features in a feature set becomes interesting if two cells or two parts of a single cell, i.e. two features shall be compared. In that case, green is used to convey the difference between the two features and other color mixtures, e.g., of green and yellow may occur. However, since the feature edges are of primary interest, we decided to neglect these mixtures and consider only the green color component.

Glyph Drawings. The frequency of an individual PAR binding and the existence and frequency of single-1-CMPs may not be inferred from the graph. In order to integrate this information, the graph view is augmented by glyph drawings (Fig. 2 (f)). Disc sections (`vtkSectorSource`) are employed as a base for the glyph shape and customized such that the outer construction circle is shown as a frame of reference. This simplifies the estimation of sector circumference with respect to the full circumference and better conveys glyph size. One glyph is created per PAR and placed close to the corresponding node. The glyph's inner and outer radius as well as its start and end angle may be modified. All modifications are carried out with respect to the associated feature and neglecting the context. The user is provided with two settings of glyph attribute modification. In the default setting, the inner radius is set to zero and the outer radius is adapted according to the binding frequency of the PAR (values of $[PAR]$). The start angle is set to zero and the end angle is modified according to the percentage of CMPs which exhibit a 1 for this PAR but are not single-1-CMPs (computed based on values of $[PAR]$ and $[PAR1]$). Hence, the missing piece of the full circle represents the percentage of single-1-CMPs.

In the feature represented by Figure 2f (see inset), CD4 is the most frequently binding PAR. About 35% of the CMPs with a 1 for CD4 are single-1-CMPs. CD3 has the second highest binding frequency. The associated filled circle indicates that no single-1-CMPs exists with a 1 for CD3. A special case from a visualization point of view is the existence of only single-1-CMPs for a specific PAR. Without special care, no glyph would be visible here since start and end angle would be equal in this case. To treat those cases, only the outline of the circle is drawn, as for CD45. In a second glyph modification setting, the inner radius and the start angle are set to zero. The outer radius is assigned a uniform value. The end angle is modified according to the percentage of CMPs with a 1 for the associated PAR with respect to the overall number of non-background CMPs in the feature. A very high percentage may indicate that the protein which is bound by this PAR is a lead protein. The glyphs are colored according to the degree of their corresponding node, i.e. the number of incident edges. Finally, the nodes are labeled with the name of the associated PAR.

Brushing Facilities. The graph view is equipped with a brushing facility. Each user-defined brush modifies the DOI attribute which is associated with the view. The modified DOI is then merged with the DOIs of all other views thereby linking the graph to these views. For simplification, we decided to restrict the brushing to nodes. Edge brushing would be hampered by the large number of edge crossings and can also easily be replaced by multiple node brushes. AND, OR, and NOT brushes are provided for the definition of simple and complex brushes. Complex brushes facilitate feature specifications such as 'Show all sites where CD3 AND CD4 bind but CD36 does NOT bind.'. Once a (complex) brush has been defined, a rule is constructed based on the involved brush types and the involved binary attributes corresponding to the selected nodes. Then, the DOI attribute is modified according to this rule thereby triggering a new merge with the DOI attributes of all other views. This again updates these views and also the graph view itself.

The brush type may be chosen from the graphical user interface. A rubber band brushing which is accomplished by pressing the left mouse button and moving the mouse allows for the selection of multiple nodes. The small squares which are drawn at the corresponding node positions are then colored according to the brush type. The colors adhere to the SimVis style: AND = yellow, OR = turquoise, NOT = magenta (not selected = gray). The deselection of nodes is accomplished by rubber band brushing after pressing the right mouse button. An extension of brushing is *Smooth brushing* facilitating the definition of non-discrete DOIs [8]. It has been neglected here since the underlying binarized fluorescence data does not exhibit smooth features.

4.3 The Table View

We integrate a table view in the framework in order to represent the unique CMPs being part of the current feature. Thereby, we take into account the familiarity of the biologists with a table view. Furthermore, the graph view is not able to convey individual CMPs. This information may only be inferred from the graph by a tedious search for individual cliques. A clique is a complete subgraph, i.e. each edge is connected to all other edges. This is inherent to the graph representation of each CMP. The table view has been implemented as a derivation of `vtkQtTableView` which is based on Qt's `QTableView`. It contains one column per PAR and three additional columns showing the absolute frequency of the CMP within the current feature, the percentaged frequency and the number of elements which equal 1 (Fig. 3 (e)). Each column may be applied for sorting the table. Alternating row colors simplify the differentiation of adjacent CMPs. Elements which equal 1, are highlighted in the SimVis standard color red since the table view shows only CMPs which are part of the current feature. If only a subset of PARs is of interest, other columns may be hidden by decreasing their width to zero. The table view is equipped with a brushing facility that allows for an OR selection being defined on the table rows (CMPs). This facilitates feature specifications such as 'Show all voxels with CMP_1 or CMP_2.'. AND and NOT brushes are needless here since only a single CMP may exist per voxel and unwanted CMPs may simply not be selected. The table supports the brushing of separate rows, a range of rows or multiple ranges. The brushed rows are highlighted in blue and white.

5 Application

Our analysis framework has been applied to one cell probe containing lymphocytes and a prostate tissue section. In the following, we discuss how a biologist uses the framework. Please note that for local investigation (Sec. 5.3), the biologist decreased the opacity of the context edges (yellow) to 0% such that visual clutter is reduced and switched off the saturation modulation of the feature edges. When investigating a cell of a specific type, it is less important to know which edges exist in other cells in the data likely being of another type. The situation changes when only cells of the same type are in the probe or can be extracted from the probe and shall be compared or in a sub-cellular examination of a single cell.

5.1 Case Studies

Lymphocytes. In this example, a large PAR library (monoclonal antibodies directed against cluster of differentiation (CD) marker proteins) was used to co-map the cell surface toponome of lymphocytes in a healthy subject. Lymphocytes are a special type of white blood cells. CD marker proteins are proteins that are expressed on the surface of immune cells, such as lymphocytes, but are also expressed in many other cell types of the human and animal organism. We have chosen lymphocytes, since this cell type in the human blood is frequently related to causing chronic inflammatory diseases by entering healthy tissues. Hence, a major biological challenge is to decipher the cell surface toponome code of lymphocytes to detect disease-specific codes, e.g., by a systematic co-mapping of a large number of CD surface proteins [29]. In the present study, we have co-mapped 32 CD surface proteins on a cell probe by using a TIS robot system [2, 10]. The probe has been imaged at 20 different slice locations with a matrix of 658×517 pixels, an in-plane-resolution of 216×216 nm and a slice distance of 200 nm. The data has been binarized by an expert and imported into our framework (see [10] for detailed information on the probe preparation, the data acquisition, and thresholding). The initial graph visualization of the entire dataset results in 24 nodes (24 PARs) and 271 edges (out of 276 theoretically possible co-occurring PAR bindings) as illustrated by Figure 3 (a-b). Only 24 of 32 PARs are included in the graph since the remaining PARs, e.g., CD49F and CD10, bind at neither position in the cell probe. Their names are instead displayed in a text inset of the graph view (Fig. 3 (b)). The colors in the 3D view encode the CMP frequency with red corresponding to the highest frequency. The cells appear as ring-shaped structures. In the graph view, it can be seen that CD3 and CD4 co-map most frequently since they are connected by the thickest edge. Also, CD3, CD4 and CONA individually map with a high frequency as indicated by the large, associated glyphs. The color of the glyphs encodes the number of incident edges per node (degree). The highest possible degree in case of 24 binding PARs is 23 (magenta). The high frequency of magenta-colored glyphs reflects the high number of edges. The table view summarizes all existing CMPs and sorts them from top to bottom in order of descending frequency.

Prostate Cancer. In this example, a tissue section has been investigated that was cut from a prostate tissue block of radical prostatectomy. A library of 17 PARs, among which 16 were cell surface proteins, was used to co-map the cell surface toponome in prostate tissue. This example is highly relevant since prostate cancer is the most common noncutaneous malignant neoplasm in men in western countries and its pathogenesis is still unclear [34]. A major biological challenge is to detect disease-specific codes, including the identification of lead proteins which may be candidates for therapeutic intervention: when lead proteins are inhibited, the corresponding protein clusters disassemble and loose their function [30, 34]. The tissue section has been imaged at a single slice location with a matrix of 658×517 pixels and an in-plane-resolution of 216×216 nm. Its analysis results are presented in Section 5.3. For a detailed discussion of the tissue section, see [34].

5.2 CMP-Guided Analysis

In Figure 3 (e), the CMPs with the five highest frequencies have been brushed in the table (bluish rows). The highest frequency has been computed for a single-1-CMP with a 1 at CD36. The corresponding regions appear red in the 3D view (Fig. 3 (c)). Amongst others, CD36 is found on *platelets* (small, solid structures in between the ring-shaped cells). Platelets are crucial in hemostasis since their aggregation causes a bleeding to stop. Since CD36 does not bind to lymphocytes, it may be excluded from the analysis by a NOT-brush in the graph view. Please note that the blank circle, which is drawn at CD36, indicates that this PAR does not co-occur with neither of the other PARs within the feature (Fig. 3 (d)). The CMP with the second highest frequency is again a single-1-CMP with a 1 at CD3 (yellow regions). CD3 may be found in the membrane of *T-lymphocytes* which are a special subgroup of lymphocytes. Different types of T-lymphocytes exist such as *T helper* and *Natural killer T* cells which all participate in the cell-mediated immunity. The third highest frequency is computed for a single-1-CMP with a 1 at CONA (Concanavalin A) which is a sugar-binding protein (green regions). The fourth highest frequency is computed again for a single-1-CMP with a 1 at CD15 (turquoise regions). CD15 binds to *mononuclear immune cells* and *granulocytes* with the latter also being a type of white blood cell. Here, the two ring-shaped, turquoise structures represent mononuclear immune cells since the cell nucleus (missing, inner part of the ring) of granulocytes has a very distinct shape. Mononuclear immune cells represent another type of white blood cell (just as lymphocytes) and are important in immune function since they respond to inflammation signals. The fifth highest frequency is computed for a CMP with two elements equal 1: CD3 and CD4 (blue regions). Hence, an edge exists that connects CD3 and CD4. CD4 binds to *T4 cells* whose function is to activate and direct other immune cells. The depletion of T4 cells is an important characteristic of an infection with the human immunodeficiency virus (HIV). The edge between CD3 and CD4 is colored in

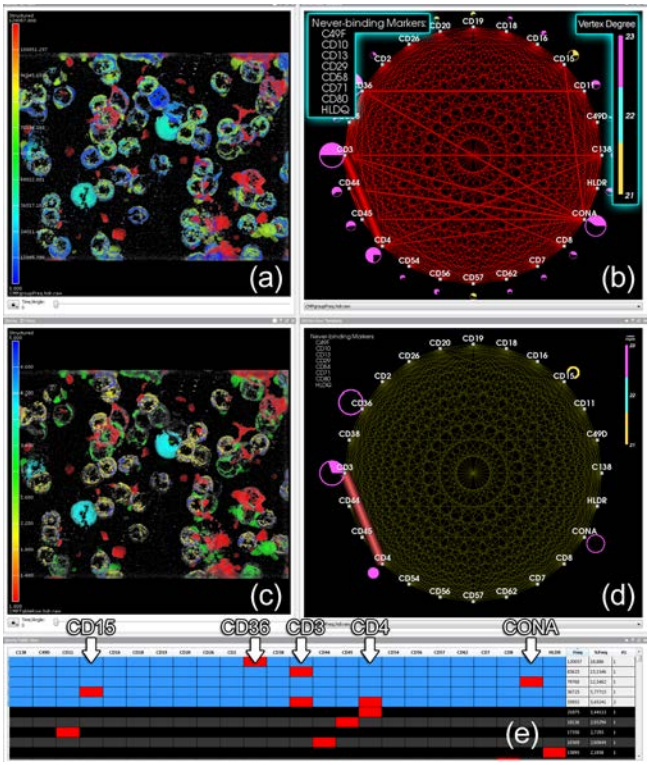


Fig. 3. CMP-Guided Analysis. *(a)*: The entire probe is visualized and colored according to CMP frequency. *(b)*: The graph represents all co-occurring PAR bindings. *(c)*: The visualization is restricted to the CMPs with the five highest frequencies (blue rows in *(e)*) and colored according to table row (first row = red). *(d)*: The corresponding graph reveals single-1-CMPs, e.g., for CD36, as well as one co-occurring PAR binding: CD3 and CD4.

a medium saturated red which indicates that the combination of these two PARs also often occurs outside the feature. The missing part of the filled, circular glyph at CD3 reflects that this PAR also exists in a single-1-marker within the feature. This has already been discussed above. Another type of T-lymphocyte, that has been identified in the data, is T8 cells (not illustrated by Fig. 3 (c-e)). T8 cells are characterized by a concurrent binding of CD8 and CD3. Their function is to destroy, e.g., cells which are effected by a virus.

5.3 Local Investigation

Biologists frequently inspect a single cell in the probe. From a visual analysis point of view this corresponds to a local investigation since the observation space is restricted to a subset and then the attribute distribution in this subset is investigated.

Lymphocytes. In Figure 4, a mononuclear immune cell is analyzed. The cell is focused by coloring the data in the 3D view according to CD15 and applying a rectangular brush to the x-and y-coordinates of the probe in a scatter plot (not illustrated). The glyph encoding in the graph view has been switched to the non-default setting (see Sec. 4.2, paragraph *Glyph Drawings*). All circles have the same radius and the filled portion of the circle is proportional to the number of non-background voxels in the feature that exhibit a binding of the PAR. Accordingly, CD15 binds in almost every part of the cell. If CD15 co-maps with another PAR, this is CONA in the majority of cases (thick edge). The modulation of edge saturation has been turned off since the focus is just on this single cell. It can be observed from the table that a single-1-CMP with a 1 at CD15 and a CMP with 1s at CD15 and CONA are the most frequent CMPs (together ≈86% of all non-background voxels). In the histogram, the number of PAR bindings per voxel is plotted and scaled logarithmically. At most, three

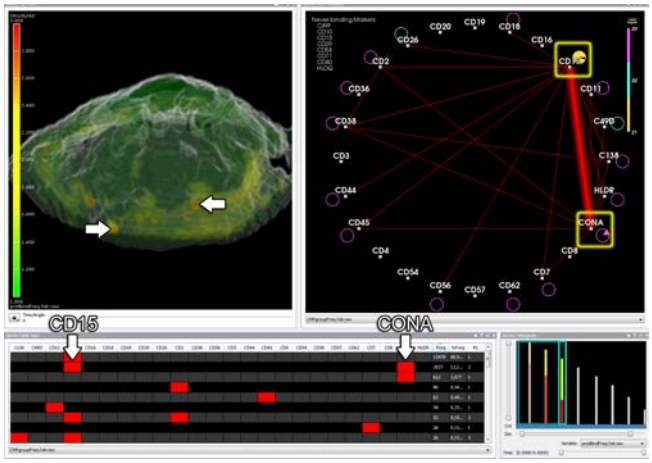


Fig. 4. Local investigation of a mononuclear immune cell toponome. CD15 binds to almost every part of the cell's surface (almost fully filled glyph). If it co-maps with another PAR, this is CONA in the majority of cases. A smooth brush is applied to the histogram which plots the number of PAR bindings per voxel (logarithmic scaling). The focus is on a high number. Thereby, corresponding regions are assigned a higher opacity in the 3D view (red regions, see arrows). Green indicates single-1-CMPs, yellow indicates two and red three concurrently binding PARs.

bindings co-occur (rightmost red bar). A smooth brush is defined such that the focus region includes the highest number of bindings and the near-focus region spans the remaining numbers. The 3D view is colored according to the number of bindings with red representing three co-occurring bindings (arrows). As can be seen, separate clusters exist within the cell that may correspond to individual cell parts and functions. The grayish context represents those parts of the cell which are not part of the feature. The structured grid ray-caster of SimVis supports focus and context visualization in one pass. In addition to the accumulation of color values, the context visualization is created using a second selection (feature set) which includes, e.g., all parts of the dataset with a specific PAR binding. This selection is represented by a 3D-texture containing 0 for unselected cells and 1 for selected cells. Due to interpolation, texture lookups between those cells can return values between 0 and 1. The distance to 0.5 is then used as a transparency value of the context, which leads to an opaque visualization of the context's silhouette. The interior of a selected region is not occluded, because a ray which intersects the volume perpendicular to the selection border collects only very few opaque values. In Figure 4, all voxels with a binding to CD15 have been selected for context generation.

In order to illustrate the striking difference between cells with regard to their surface protein pattern, we repeat the local investigation for a T8 cell (Fig. 5). The cell is focused in a similar way as described for the mononuclear immune cell. An initial rendering of the graph reveals the edge between CD3 and CD8 as having the highest frequency. Hence, it is selected by brushing the two corresponding nodes using an AND-brush (the small, filled squares representing the nodes are colored in yellow). Two thick edges connecting both nodes to CD45 show that this triple co-mapping frequently occurs across the cell surface. The histogram indicates that within the feature, between two and five PARs are co-mapped (red bars). The CMPs in the table are sorted according to the number of elements equal 1 in a decreasing order from top to bottom. Most of the higher binding numbers occur only at low magnitudes which may however not be related to their significance. To localize the co-mappings within the cell, the 3D view is colored according to the number of bindings per voxel. All voxels with a binding to CD8 have been selected for context generation.

Prostate Cancer. In the analysis of this tissue section, we follow the procedure described by our collaborators in [34]. We use

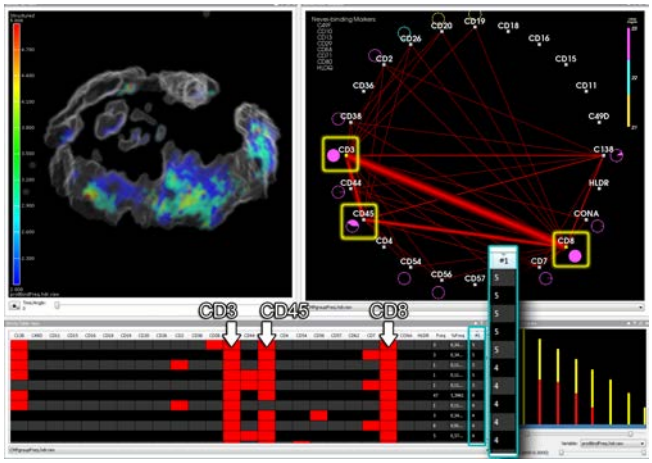


Fig. 5. Local investigation of a T8-cell toponome. CD3 and CD8 have been brushed in the graph by means of an AND-brush. The resulting cell parts are colored in the 3D view according to the number of PAR bindings per voxel (blue represents 2, red 5). The remaining parts of the cell are rendered in gray and serve as context information. The histogram plots the number of PAR bindings (logarithmic scaling). The CMPs in the table are sorted according to their number of elements equal 1 in a decreasing order from top to bottom.

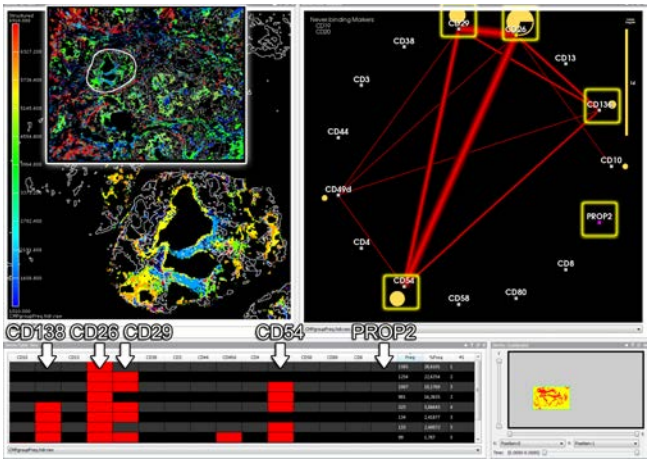


Fig. 6. Local investigation of a prostate cell toponome. The analysis has been restricted to the cell by brushing the x- and -y coordinates of the dataset in the scatter plot (lower right view). The cell is colored in the 3D view according to CMP frequency. The PAR CD138 is superimposed as gray isolines for providing context information. The inset shows the location of the cell in the tissue section. CD26 has been brushed in the graph by means of an AND-brush restricting the visualization to prostate epithelium. PROP2 has been excluded from the analysis by means of a NOT-brush. The table shows all CMPs that adhere to the final feature specification.

our new system and illustrate how the analysis may benefit from the linked views (Fig. 6). At first, we exclude the marker propidium iodide (PROP2) from the analysis by means of a NOT brush on the graph view. PROP2 is excluded since it is no cell surface marker, but instead binds to the cell nucleus. Then, we select the 30 most frequent CMPs in the table (not illustrated here) and superimpose a context visualization of the PAR CD138 (inset of the 3D view). CD138 has been chosen since its is a marker for prostate cancer progression. The 3D view is colored according to CMP frequency. The context appears as a gray silhouette. Next, the focus is on a single cell by brushing x- and y-coordinates of the dataset in the scatter plot (compare to Figs. 4-5 in [34]). This lens function was not available to the biologists in [34], requiring a manual cropping of the dataset. The inset of the 3D view shows the location of the cell within the entire tissue section. Then, we apply an AND brush on CD26 in the graph since CD26 is known to recognize prostate epithelium [34]. Two very frequent co-mappings are revealed: CD54 and CD26 as well as CD26 and CD29. In [34], the biologists had to extract this information from a table with 17 columns and 4217 rows (number of unique CMPs) which was a cumbersome task. After further investigation of other cells in the dataset, they concluded that CD26 and CD29 are lead proteins in a *motif* that is specific to epithelial cells with features of neoplasia inside prostate acini.

5.4 Comparative Analysis

The comparison of cell samples from healthy subjects and patients is crucial in detecting and understanding disease-related protein patterns. We mimic this procedure by comparing two cells in a sample, in particular, we compare two T8-cells: $T8_1$ and $T8_2$ (Fig. 7). Two features are defined in a feature set, one for each cell. The separation of each cell is carried out as described in Section 5.3. The result is visualized in the 3D view and colored according to CMP frequency with red representing the highest frequency. For each cell, a graph has been generated. The glyph encoding in the graph views has been switched to the default setting (see Sec. 4.2, paragraph *Glyph Drawings*). Apart from red edges, green edges appear indicating that the corresponding PAR pair is not co-mapped in this cell but in the other cell (see Sec. 4.1 for a detailed discussion of the edge colors). Thus, similarities and differences between the graphs can easily be inferred. At a first glance, the graphs and the glyphs look similar which is also reflected by similar colors in the 3D view. However, an interesting observation is the binding of CD36 at $T8_2$ in a single-1-CMP. Essentially, CD36 only binds to platelets and mononuclear immune cells. For further investigation, CD36 has been brushed in the graph view leading to a separation of the small, solid, red region in $T8_2$ (not illustrated here). It was concluded that this region represents a platelet which superimposes $T8_2$. Another interesting observation is that CD19 is not binding at $T8_1$ but at $T8_2$. This may indicate that $T8_2$ is in a different transition stage. Living cells undergo different stages during their life span. The two histograms plot the number of PAR bindings per voxel. In $T8_2$, the highest number of co-mappings is five whereas it is four in $T8_1$.

5.5 User Feedback

We gathered anecdotal feedback from two users, an experienced biologist and a computer scientist who has been working in the biology domain for several years. Both appreciated the possibility to explore the 3D data as a whole instead of slice by slice. The latter is a serious limitation of their homemade system which requires tedious manual post-processing in order to merge the analysis results from individual slices. They commended the interactive selection of parts of the cell probe and the interactive update of graph and table view which greatly simplifies the data exploration. Also, the graph brushing was appreciated as a fast method for defining a template CMP that can be searched for in the data. The graph provides a fast and easy to comprehend overview of a cell's toponome and clearly outperforms the table. As shown in Section 5.4, it can be used as a symbolic description of a cell's toponome in a cell-by-cell comparison. Information that is cumbersome to infer from a table but can easily be observed in the graph is the frequency of individual PAR bindings and co-occurring PAR bindings. Another advantage of the new system is the possibility to integrate histograms and scatter plots displaying derived parameters, e.g., the number of PAR bindings per voxel.

A drawback of the system is the way of defining a region of interest. For this purpose, a scatter plot opposes the x- and y-coordinates of the data points and the user may define a region by means of a rectangular brush. This requires the user to focus on the 3D view while moving the brush in order to select the desired cell. However, guidance is provided by a context visualization of the entire dataset. The biologist commented that a more flexible brush shape is desirable in order to adhere to more complex cell shapes. Furthermore, measurement units and the current scale should be included in the 3D view.

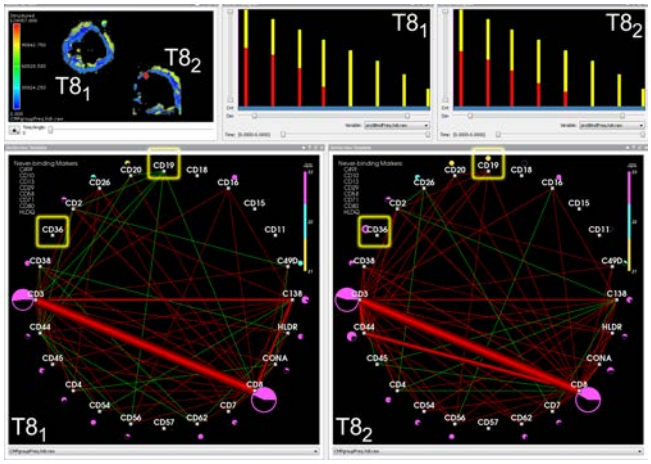


Fig. 7. Comparative analysis of two T8-cells: $T8_1$ and $T8_2$. Red edges in the graph represent PAR bindings that exist in both cells whereas green edges represent bindings that exist only in the other cell. For example, CD19 binds only to $T8_2$. The histograms plot the number of PAR bindings per voxel (logarithmic scaling). In the 3D view, the cells are colored according to CMP frequency.

Feature requests of our users include a picking facility in the 3D view that allows the selection of cell parts thereby inducing an update of all other views with respect to the selected part. Furthermore, it should be possible to crop the data such that only a single cell is considered in the analysis. At the moment, a cell can be focused but the entire dataset is considered as context. This hampers, e.g., a sub-cellular examination in which cell parts shall be focused while the rest of the cell serves as context. If this was possible, the users would fade in again the yellow context edges and apply the modulation of feature edge saturation (Sec. 4.2, paragraphs *Edge Color* and *Feature Edge Saturation*).

5.6 Scalability

In order to assess the scalability of our approach, we created two synthetic datasets: D_1 and D_2. D_1 has the characteristics: 1056×1027×32(slices)×31(PARs). At the moment of writing, a dataset of this size is being acquired by the biologists and represents their largest dataset with respect to voxel count. D_2 resembles the skin biopsy data analyzed in [32] which contains the highest number of PARs that have ever been applied: 658×517×1(slices)×100(PARs). The original data could not be provided for the present collaboration. Hence, we randomly distributed 155,000 unique CMPs over both datasets while treating 50% of the voxels as background. This is the highest number of CMPs that the biologists ever observed in one of their datasets. The maximum number of 1's in a CMP has been set to 37% of the PAR number corresponding to the maximum for the lymphocytes dataset.

All tests were carried out on a 3.07GHz Intel 8-core PC with 8GB RAM and a 64bit Windows operating system. Our view setup consisted of a 3D view, a table view, a graph view and a scatter plot. Both datasets were rendered at interactive frame rates in the 3D view. After defining the entire dataset D_1 (values of D_2 in parentheses) as a feature, the table view with 155,000 rows and 31 columns (155,000 rows, 100 columns) needed ≈240ms and ≈360ms (≈80ms and ≈200ms) rendering time with and without sorting according to one of its columns. The graph view with 32 nodes and 496 edges (100 nodes, 4950 edges) needed ≈2.44s (395ms) rendering time, 2.4s (51ms) for the voxel traversal and the update of the data structures and 40ms (344ms) for repainting the graph and the glyphs. For comparison, the lymphocytes dataset which contains 2228 unique CMPs needed ≈10ms and ≈11ms for table update and ≈70ms (32ms+38ms) for graph update. The long time for voxel traversal of D_1 when updating the graph view does not primarily result from the high voxel count. It is related to the high number of edges (10) which are stored on average per non-background voxel (17,352K voxels) in the custom data structure described in Section 4.2 (paragraph *Preprocessing*). For comparison, 1 edge is stored on average for the lymphocytes dataset and 3 edges are stored for the prostate dataset with the latter being significantly smaller then D_1. Further investigation and more real world examples are required to study this effect more thoroughly. From the timings, it can be concluded that even data with a high voxel count and a high number of PARs may be processed. However, interactive frame rates with 25 updates per second are not achieved anymore.

We also explored the visual scalability. The table of D_2 is very wide requiring ineffective horizontal scrolling. In a two screen solution, the table may span both screens. A more appropriate solution could be the application of focus+context techniques, e.g., table lenses [24]. The graph of D_2 is very cluttered and the glyphs may cover only a few pixels which hampers their readability. The situation is improved slightly when the graph is displayed in full screen mode on a second screen. In addition, focus+context techniques such as graphical fish eye views may be applied [26]. A more sophisticated solution to edge cluttering would be *edge bundling* [16]. The original method requires a hierarchy being defined on the nodes. However, such a hierarchy is not defined on the PARs we applied. An extension proposed by the same authors does not require a hierarchy but applies a flexible spring model [17]. An additional advantage of this approach is a parameter that controls the bundling strength. By reducing the bundling strength from full to a lower value, edges become slightly separated which is a prerequisite for the recognition of a varying edge width. Additional solutions for reducing edge cluttering are the reordering of nodes (in contrast to dynamic node arrangements, node positions on the circle are fixed but the assignment of PARs to nodes is flexible) and the exterior routing of edges [13, 37]. It should be noted that the visual clutter is already significantly reduced once only individual cells are investigated and context edges are fade out. However, in future scenarios fading out the edges may not be an option (recall Sec. 5.5).

6 Summary and Discussion

We presented the integration of techniques from graph visualization in an interactive visual analysis framework for the investigation of toponome data. Each node in the graph represents an affinity reagent while each edge represents two co-occurring affinity reagent bindings. The frequency of co-occurring affinity reagent bindings is encoded in the edge's width. The graph visualization is enhanced by glyphs which encode specific properties of the binding. The graph view is equipped with brushing facilities and linked to all other views of the framework. Furthermore, an interactive table view is integrated which summarizes unique fluorescence patterns existing in the data.

We applied the framework to a cell probe containing lymphocytes and to a prostate tissue section. By brushing in the spatial and attribute domain of the corresponding datasets, the biologist achieves a better understanding of the function protein patterns of a cell. In a local investigation, a single cell may be separated from the probe and inspected. By browsing the table view, individual combinatorial molecular phenotypes may be localized. Interesting co-mappings of individual affinity reagents may be localized by brushing the corresponding nodes in the graph view. Clusters of protein mappings with a different number of protein bindings per voxel were observed in a mononuclear immune and a T8 cell. By defining two features (cells), the biologist is able to compare two cells. Edges in the graph of one cell are then colored differently if they occur only in the other cells.

In the future, we will implement more advanced methods and brush shapes for cell separation which will more strongly adhere to the actual cell shape. So far, a reliable analysis of the cell depends on its separability from other cells in the probe.

Acknowledgments

Technological and biological toponome studies were supported by BMBF grants Biochance, CELLECT, NBL3, NGFN2, NGFNplus, and through DFGschu627/10-1, the Innovationskolleg INK15, and the Klaus Tschira foundation (project toponome atlas).

REFERENCES

[1] A. Barysenka, A. W. M. Dress, and W. Schubert. An information theoretic thresholding method for detecting protein colocalizations in stacks of fluorescence images. *J Biotechnol*, 149(3):127–131, 2010.

[2] S. Bhattacharya, G. Mathew, E. Ruban, D. B. A. Epstein, A. Krusche, R. Hillert, W. Schubert, and M. Khan. Toponome imaging system: in situ protein network mapping in normal and cancerous colon from the same patient reveals more than five-thousand cancer specific protein clusters and their subcellular annotation by using a three symbol code. *J Proteome Res*, 9(12):6112–6125, 2010.

[3] B. Chapman, G. Jost, and R. v. d. Pas. *Using OpenMP: Portable shared memory parallel programming*. The MIT Press, 2007.

[4] W. S. Cleveland and R. McGill. Graphical perception: Theory, experimentation, and application to the development of graphical methods. *J Am Stat Assoc*, 79(387):531–554, 1984.

[5] W. de Leeuw, P. J. Verschure, and R. van Liere. Visualization and analysis of large data collections: a case study applied to confocal microscopy data. *IEEE Trans Visual Comput Graph*, 12(5):1251–1258, 2006.

[6] J. Dietzsch, J. Heinrich, K. Nieselt, and D. Bartz. SpRay: A visual analytics approach for gene expression data. In *Proc. IEEE Symp. Visual Analytics Science and Technology VAST*, pages 179–186, 2009.

[7] H. Doleisch, M. Gasser, and H. Hauser. Interactive feature specification for focus+context visualization of complex simulation data. In *Proc. of IEEE TCVG - EUROGRAPHICS Symp. on Vis.*, pages 239–248, 2003.

[8] H. Doleisch and H. Hauser. Smooth brushing for focus+context visualization of simulation data in 3D. *Journal of WSCG*, 10(1):147–154, 2002.

[9] A. Dress, T. Lokot, L. Pustyl'nikov, and W. Schubert. Poisson numbers and poisson distributions in subset surprisology. *Annals of Combinatorics*, 8:473–485, 2005.

[10] M. Friedenberger, M. Bode, A. Krusche, and W. Schubert. Fluorescence detection of protein clusters in individual cells and tissue sections by using toponome imaging system: sample preparation and measuring procedures. *Nat Protoc*, 2(9):2285–2294, 2007.

[11] T. M. J. Fruchterman and E. M. Reingold. Graph drawing by force-directed placement. *Softw. Pract. Exper.*, 21:1129–1164, 1991.

[12] G. Furnas. Generalized fisheye views. In *Proc. of the ACM CHI Conf. on Human Factors in Computing Systems*, pages 16–23, 1986.

[13] E. Gansner and Y. Koren. Improved circular layouts. In M. Kaufmann and D. Wagner, editors, *Graph Drawing*, volume 4372 of *Lecture Notes in Computer Science*, pages 386–398. Springer, 2007.

[14] N. Gehlenborg, S. I. O'Donoghue, N. S. Baliga, A. Goesmann, M. A. Hibbs, H. Kitano, O. Kohlbacher, H. Neuweger, R. Schneider, D. Tenenbaum, and A.-C. Gavin. Visualization of omics data for systems biology. *Nat Methods*, 7(3 Suppl):S56–S68, 2010.

[15] H. Hauser. Generalizing focus+context visualization. In *Scientific Visualization: The Visual Extraction of Knowledge from Data*, pages 305–327, 2005.

[16] D. Holten. Hierarchical edge bundles: Visualization of adjacency relations in hierarchical data. *IEEE Trans Visual Comput Graph*, 12(5):741 –748, 2006.

[17] D. Holten and J. J. van Wijk. Force-directed edge bundling for graph visualization. *Comput Graph Forum*, 28(3):983–990, 2009.

[18] Z. Hu, J. Mellor, J. Wu, and C. DeLisi. VisANT: an online visualization and analysis tool for biological interaction data. *BMC Bioinformatics*, 5:17, 2004.

[19] M. Meyer, T. Munzner, and H. Pfister. MizBee: A multiscale synteny browser. *IEEE Trans Visual Comput Graph*, 15:897–904, 2009.

[20] P. Muigg, J. Kehrer, S. Oeltze, H. Piringer, H. Doleisch, B. Preim, and H. Hauser. A four-level focus+context approach to interactive visual analysis of temporal features in large scientific data. *Comput Graph Forum*, 27(3):775–782, 2008.

[21] S. Oeltze, H. Doleisch, H. Hauser, P. Muigg, and B. Preim. Interactive visual analysis of perfusion data. *IEEE Trans Visual Comput Graph*, 13(6):1392–1399, 2007.

[22] G. A. Pavlopoulos, S. I. O'Donoghue, V. P. Satagopam, T. G. Soldatos, E. Pafilis, and R. Schneider. Arena3D: visualization of biological networks in 3D. *BMC Syst Biol*, 2:104, 2008.

[23] S. P. Perfetto, P. K. Chattopadhyay, and M. Roederer. Seventeen-colour flow cytometry: unravelling the immune system. *Nat Rev Immunol*, 4(8):648–655, 2004.

[24] R. Rao and S. K. Card. The table lens: merging graphical and symbolic representations in an interactive focus + context visualization for tabular information. In *Proc. of the ACM CHI Conf. on Human Factors in Computing Systems*, pages 318–322, 1994.

[25] B. E. Rogowitz, L. A. Treinish, and S. Bryson. How not to lie with visualization. *Comput. Phys.*, 10:268–273, 1996.

[26] M. Sarkar and M. H. Brown. Graphical fisheye views of graphs. In *Proc. of the ACM CHI Conf. on Human Factors in Computing Systems*, pages 83–91, 1992.

[27] W. Schubert. Multiple antigen-mapping microscopy of human tissue. In *Advances in analytical cellular pathology*, Excerpta Medica, pages 97–98. Elsevier, 1990.

[28] W. Schubert. Topological proteomics, toponomics, melk-technology. *Adv Biochem Eng Biotechnol*, 83:189–209, 2003.

[29] W. Schubert. A three-symbol code for organized proteomes based on cyclical imaging of protein locations. *Cytometry A*, 71(6):352–360, 2007.

[30] W. Schubert. On the origin of cell functions encoded in the toponome. *J Biotechnol*, 149(4):252–259, 2010.

[31] W. Schubert, M. Bode, R. Hillert, A. Krusche, and M. Friedenberger. Toponomics and neurotoponomics: a new way to medical systems biology. *Expert Rev Proteomics*, 5(2):361–369, 2008.

[32] W. Schubert, B. Bonnekoh, A. J. Pommer, L. Philipsen, R. Bckelmann, Y. Malykh, H. Gollnick, M. Friedenberger, M. Bode, and A. W. M. Dress. Analyzing proteome topology and function by automated multidimensional fluorescence microscopy. *Nat Biotechnol*, 24(10):1270–1278, 2006.

[33] W. Schubert, M. Friedenberger, M. Bode, A. Krusche, and R. Hillert. Functional architecture of the cell nucleus: towards comprehensive toponome reference maps of apoptosis. *Biochim Biophys Acta*, 1783(11):2080–2088, 2008.

[34] W. Schubert, A. Gieseler, A. Krusche, and R. Hillert. Toponome mapping in prostate cancer: detection of 2000 cell surface protein clusters in a single tissue section and cell type specific annotation by using a three symbol code. *J Proteome Res*, 8(6):2696–2707, 2009.

[35] H.-J. Schulz, M. John, A. Unger, and H. Schumann. Visual analysis of bipartite biological networks. In *VCBM'08: Proc. of the Eurographics Workshop on Visual Computing for Biomedicine*, pages 135–142, 2008.

[36] P. Shannon, A. Markiel, O. Ozier, N. S. Baliga, J. T. Wang, D. Ramage, N. Amin, B. Schwikowski, and T. Ideker. Cytoscape: a software environment for integrated models of biomolecular interaction networks. *Genome Res*, 13(11):2498–2504, 2003.

[37] A. Symeonidis and I. Tollis. Visualization of biological information with circular drawings. In *Proc. Int'l Symp. on Biological and Medical Data Analysis (ISBMDA)*, pages 468–478, 2004.

[38] G. H. Weber, O. Rubel, M.-Y. Huang, A. H. DePace, C. C. Fowlkes, S. V. E. Keranen, C. L. Luengo Hendriks, H. Hagen, D. W. Knowles, J. Malik, M. D. Biggin, and B. Hamann. Visual exploration of three-dimensional gene expression using physical views and linked abstract views. *IEEE/ACM Trans Comput Biol Bioinformatics*, 6(2):296–309, 2009.

Visualization and Exploration of 3D Toponome Data

S. Oeltze[1] and P. Klemm[1] and R. Hillert[2] and B. Preim[1] and W. Schubert[2]

[1]Department of Simulation and Graphics, University of Magdeburg, Germany
[2]Molecular Pattern Recognition Research Group, University of Magdeburg, Germany

Abstract

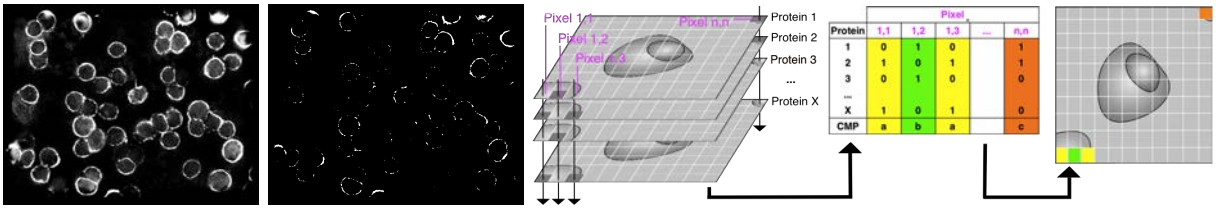
The toponome of a cell describes the location and topological distribution of proteins across the cell. In toponomics, the toponome is imaged and its inner structure and its semantics are investigated in order to understand how cells encode different functionalities both in health and disease. Toponome imaging results in complex multi-parameter data composed of a 3D volume per protein affinity reagent. After imaging, the data is binarized such that 1 encodes protein present and 0 encodes protein absent. Biologists are particularly interested in the clustering of these binary protein patterns and in the distribution of clusters across the cell.
We present a volume rendering approach for visualizing all unique protein patterns in 3D. A unique color is dynamically assigned to each pattern such that a sufficient perceptual difference between colors in the current view is guaranteed. We further present techniques for interacting with the view in an exploratory analysis. The biologist may for instance "peel of" clusters thereby revealing occluded cell structures. The 3D view is integrated in a multiple coordinated view system. Peeling off clusters or brushing protein patterns in the view updates all other views. We demonstrate the utility of the view with a cell sample containing lymphocytes.

Categories and Subject Descriptors (according to ACM CCS): J.3 [Computer Applications]: Life and Medical Sciences—Biology and genetics

1. Introduction

While the human genome project has revealed the code for all proteins, the next big challenge is to understand how proteins cooperate in cells and tissues in time and space [Sch10]. Although many details on the molecular function and structure of many proteins are known, their corresponding cellular functions cannot simply be derived. This is due to a dependence of the function on the contextual position of a given protein within a protein network. The *toponome* of a cell describes its functional protein pattern, i.e. the location and topological distribution of proteins. In *toponomics*, the toponome is imaged, explored and analyzed for applications in toxicology, drug development and patient-drug-interaction. Toponome imaging results in complex multi-parameter data composed of a 3D volume per protein affinity reagent. In a post-processing step, the data is binarized such that 1 encodes protein present and 0 encodes protein absent. For each data voxel, a binary vector can be constructed over all volumes, which then encodes the local protein co-mapping. Biologists are interested in answering questions such as: Which proteins co-map with which frequency?, Where across the cell surface do the binary protein patterns cluster?, and How does the clustering differ from cell to cell? These questions guide our research. In previous work, we developed a graph view which encodes co-mapping proteins and co-mapping frequency [OFH*11].

Here, we present a volume rendering approach that generates an integrated 3D visualization of all protein affinity reagent volumes. It is based on the assignment of a unique color to each unique binary protein pattern. Since several thousands of such patterns may exist in a dataset, we offer a dynamic color range distribution for the current view on the data, e.g., a close-up view of a cell. The algorithm then generates a set of unique, perceptually optimized colors for the currently visible patterns. Thus, the visual differentiability of these patterns is improved. To support an exploratory analysis of the toponome, the 3D view is equipped with several interaction techniques, e.g., *similarity brushing*. Furthermore, it is integrated via *linking* in a multiple coordinated view system. Brushing and linking supports the biologist in deciphering the toponome code. We demonstrate the utility of the 3D view with a cell sample containing lymphocytes.



(a) (b) (c)

Figure 1: *a: Original fluorescence signal in a single slice. b: The same slice as in (a) after thresholding. White pixels indicate protein present. c: Generation of combinatorial molecular phenotypes and display in a 2D toponome map. For each pixel, the binary fluorescence signal of all protein affinity reagents is collected in a combinatorial binary code (table columns). The set of unique binary codes (combinatorial molecular phenotypes) is computed and each code is assigned a color. A toponome map is generated by mapping each pixel to its corresponding color (right subimage).*

2. Biological Background

This section explains the toponome and how it is imaged. Further, it introduces *combinatorial molecular phenotypes*, which serve as the input for the 3D visualization.

2.1. The Toponome

The entirety of all protein networks, in which proteins are defined by their protein-to-protein context in any given cell, is defined as the *toponome* [FBKS07, SBP*06, SGK*11]. Its inner structure, its biological code, and its semantics are investigated in *toponomics*. It has been shown that the toponome is hierarchically organized [SBP*06]. It comprises protein clusters which contain *lead proteins* and are interlocked as a network. The lead proteins control the topology of the clusters and their function as a network. However, neither a lead protein nor the protein cluster which is controlled by it can be predicted from molecular data. Thus, the toponome must be imaged in human cells or tissues to understand how cells encode different functionalities both in health and disease. This is essential for finding new drugs in cancer and for detecting protein clusters that can be regarded as a new system of biomarkers in disease [Sch10, SGK*11].

2.2. Imaging the Toponome

The most advanced technique for imaging the toponome is robot-driven multi-parameter fluorescence microscopy TIS^{TM}. It is based on a cyclical procedure in which a *tag library* (specific affinity reagents recognizing proteins) is conjugated to one and the same dye. An imaging robot applies these tags to a fixed cell or tissue sample [FBKS07]. It starts with the first dye-conjugated tag and applies it to the sample. The resulting fluorescence image is then captured by a CCD camera, which is connected to an epifluorescence microscope. In the last step of the cycle, the dye is bleached gently to avoid any energy transfer into the remaining proteins [SBP*06]. Then, the second tag is applied and so on. The labeling of 100 proteins in 100 cycles has been demonstrated [SBP*06, SGK*11]. Imaging can be performed in 3D by modifying the microscope's focal plane. The acquisition of a 3D dataset is costly and may take longer than a day.

2.3. Combinatorial Molecular Phenotypes

After imaging the toponome, a thresholding algorithm is applied to the fluorescence signal (Fig. 1a-1b). This generates a combinatorial binary code for each voxel where 0 indicates protein absent and 1 indicates protein present [BDS10]. The unique binary codes that exist in the data, out of all possible codes, are referred to as *combinatorial molecular phenotypes* (CMPs). A simple technique for visualizing CMPs of a single volume slice is the assignment of a unique color to each CMP and the display of the slice as a colored image also known as *toponome map*. The generation of CMPs and of the toponome map are illustrated by Figure 1c. The binary code that corresponds to a certain CMP very often exists at several pixel/voxel positions. These positions are not randomly spread over the data but clustered at certain locations of the cells. The biologist is interested in these protein clusters since they correspond to functional units of the cells.

3. Related Work

3D microscopy data may consist of a single channel or multiple channels, with each channel showing a different aspect of the data. The most wide-spread visualization technique for 3D microscopy data is direct volume rendering.

3.1. Single-Channel 3D Microscopy Data

Sakas et al. presented a pipeline for visualizing confocal laser scan microscopy (CLSM) data [SVP96]. They integrated Maximum and Minimum Intensity Projection as well as full volume rendering into the rendering step. The volume visualization system VolVis was presented by Kaufman in the context of investigating CLSM datasets of nerve

cells [Kau98]. It supports surface as well as volume rendering and offers global rendering effects such as shadows and reflections. Fang et al. described an approach for microscopy data exploration with a focus on the intuitive design of transfer functions [FDM*00]. They employed an image-based model which defines a transfer function as a sequence of 3D image processing steps, e.g., intensity thresholding and boundary detection. An approach for the reconstruction of cellular structures from optical microscopy data was proposed by Mosaliganti et al. [MCS*08]. They employed sophisticated techniques for extracting cells and separating overlapping cells. The resulting, individually colored cells were then visualized through volume rendering. In recent work, Guo et al. presented a novel volume illustration technique [GYL*12]. It simulates an optical phenomenon in interference microscopy, which accounts light interference over transparent specimens, and thereby enhances the image contrast and structure details.

3.2. Multi-Channel 3D Microscopy Data

Razdan et al. worked on the visualization of multicolor laser confocal microscopy data [RPFC01]. In such data, three lasers providing light at different wavelengths reveal different substances in the same field of view. A RGB composite rendering was generated by means of ray-casting. A volume rendering system for confocal and two-photon fluorescence microscopy data was presented by Clendenon et al. [CPS*02]. They focused on a near real-time visualization of multichannel image stacks on standard PCs by exploiting 3D graphics processors. DeLeeuw et al. presented the ARGOS system for an interactive exploration and batch processing of confocal laser microscopy data [dLVvL06]. They linked a volume rendering of the image data to attribute views, e.g., parallel coordinates, which were equipped with brushing facilities. Bruckner et. al presented a system for an integrated visualization, exploration and annotation of anatomical brain microscopy data and molecular genetic data of fruit flies [BŠG*09]. The microscopy data was visualized through volume rendering with Maximum Intensity Difference Accumulation as a projection method. Wan et al. developed an interactive tool for visualizing multi-channel confocal microscopy data [WOCH09]. Multidimensional transfer functions and several compositing techniques were implemented for an integrated visualization of up to three channels. The approaches creating an integrated visualization of several data channels are closest to our work. However, none of these approaches can handle an arbitrary number of binary data channels (Sec. 2.3).

4. Analysis Framework and Prior Visualization

This section comprises a brief discussion of the in-house toponome analysis framework of our collaborators and their prior way of visualizing the 3D toponome data in an off-line process, i.e., outside the framework.

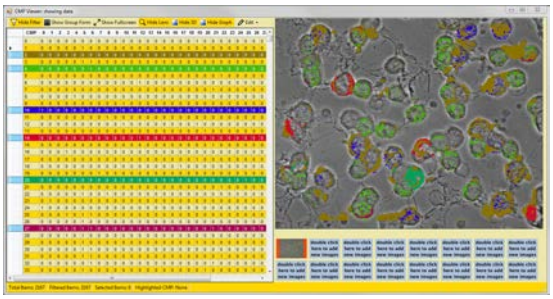


Figure 2: *Toponome analysis framework. Multiple views are implemented, such as a table view listing all CMPs as rows (left) and a toponome map showing the location of CMPs selected in the table (right). The toponome map is superimposed on a grayscale phase contrast image. Each ring-shaped structure represents a cell.*

4.1. Toponome Analysis Framework

The in-house toponome analysis framework implements multiple coordinated views: a filter view, a table view (Fig. 2, left), a toponome map (Fig. 2, right), and a recently added graph view [OFH*11]. All views are linked and equipped with brushing facilities. If a selection is brushed in one view, it is merged with the selections from all other views and the result is propagated to all views. The filter view facilitates the definition of a template CMP by setting the filter for each affinity reagent to zero, one or no filtering. The table view lists all CMPs as rows. Each selected row is assigned a unique color, which serves as an identifier for this CMP in all views. The table is linked, e.g., to the toponome map, such that the user may select individual CMPs and observe their spatial locations in the cell/tissue sample. The toponome map is often superimposed on a phase contrast image, which serves as a spatial reference (Fig. 2, right). In these images, a relatively clear distinction between cell surface, nucleus, and background is possible. The framework further contains a graph view which encodes the co-mapping of proteins (edges) and the co-mapping frequency (edge width).

4.2. Prior Visualization of 3D Toponome Data

Until now, our collaborators applied the commercial software IMARIS (Bitplane AG, Zurich, Switzerland) for generating volume renderings of their 3D toponome data in an off-line process. The entire procedure for visualizing a 3D toponome dataset took them $1-2$ workdays. Besides the high expenditure of time, a visualization at interactive frame rates in IMARIS is restricted to $\approx$ 30 different CMPs on a machine with 4 GB of working memory. Hence, our collaborators focused on the 30 most frequent CMPs out of a set that may comprise several 1000 CMPs. In the following, we describe the time-consuming visualization procedure. Since

the in-house framework could only handle 2D data so far, each volume slice was processed separately. At first, the 30 most frequent CMPs of the volume's middle slice were determined. Although these might not have been the most frequent ones in the remaining slices, they were treated as such in order to generate an inter-slice coherent visualization. In a next step, all n slices were processed sequentially: an image was generated per CMP, showing its locations in this specific slice, the image was screen captured and stored. For a common number of slices $n = 20$, this resulted in 600 screen captures which were then loaded into IMARIS. Each set of images showing the same CMP had to be loaded as a separate channel and perceptually well separated colors had to be defined manually and assigned to the channels.

5. Visualization and Exploration Methods

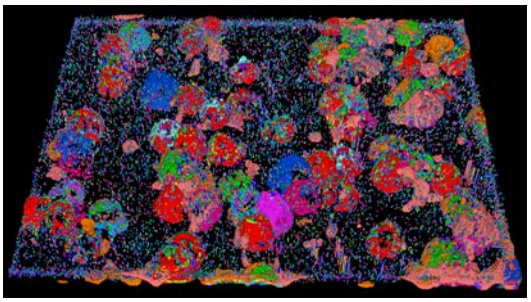
This section starts with a description of the toponome data, which we employ for demonstration purposes. Then, it elaborates on our volume rendering approach for visualizing the data, which has been integrated in a newly added 3D view of the toponome analysis framework. Finally, it discusses the equipment of the 3D view with interaction facilities.

5.1. Toponome Data

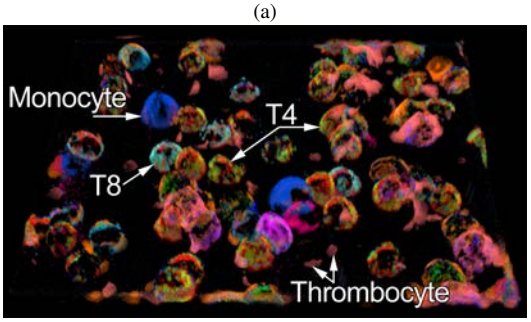
Our methods have been applied to a cell sample containing blood lymphocytes of a healthy subject. Lymphocytes are frequently related to causing chronic inflammatory diseases by entering healthy tissues. In imaging the sample, a large tag library, containing monoclonal antibodies directed against cluster of differentiation (CD) marker proteins, was used. CD marker proteins are expressed, among others, on the surface of immune cells such as lymphocytes. A major challenge is to decipher the cell surface toponome code of these cells in order to detect disease specifics. In this study, 32 CD surface proteins have been co-mapped on the sample (Sec. 2.2). The sample has been imaged at 20 slice locations with a matrix of 658×517 pixels, an in-plane-resolution of 216×216 nm and a slice distance of 200 nm. The data has been binarized and imported into the framework, which detected 2167 CMPs (see [FBKS07] for detailed information on sample preparation, data acquisition, and thresholding).

5.2. Volume Rendering of 3D Toponome Data

The combinatorial molecular phenotypes (CMPs) serve as an input for the volume rendering (recall Sec. 2.3). They are stored per slice in XML files. Each file contains the CMPs of the slice and the pixel positions per CMP. In order to construct a 3D dataset, the XML files are read in and an empty dataset is generated, whose dimensions correspond to those of the measured data. Then, each CMP is assigned a unique RGBA value and this value is stored at those voxels of the new dataset which are associated with the respective CMP (RGBA computation will be discussed in Section 5.3).



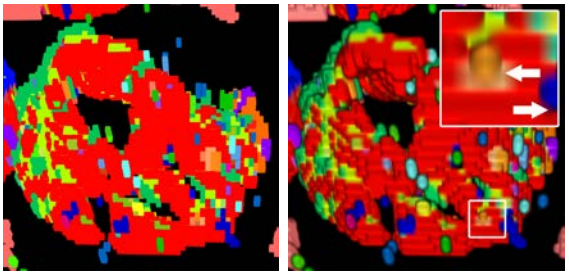
(a)



(b)

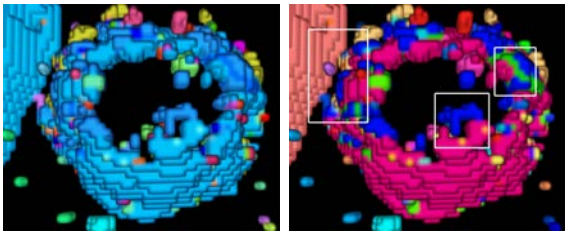
Figure 3: *(a) Overview of a cell sample containing lymphocytes (hemisphere-shaped structures). Each of the 2167 CMPs is assigned a unique color. Very small structures are likely to represent noise. (b) A globally assigned transparency visually suppresses these structures. A subset of the major cell types is annotated.*

A crucial aspect of volume rendering is the definition of transfer functions for mapping data values to color and transparency. Often, a color scale together with a range of transparency values between zero and one are defined and discretized for that purpose. In our case, a color has already been defined and associated to each data value which renders a mapping unnecessary. The application of transparency to visualizing toponome data is challenging. An increased transparency induces color mixing which hampers the identification of CMPs by their unique color. This scenario is very likely, since a cell or tissue section often contains heterogeneous regions with respect to CMP distribution. At a given pixel position, a different CMP may exist in each slice or all neighboring pixels in the same slice may exhibit different CMPs. Nevertheless, a semi-transparent visualization proved to be valuable in getting an initial overview of the data, in visually suppressing noise and in an interactive data analysis. The latter will be described in Section 5.4. The exploration of 3D toponome data starts with an overview visualization (Fig. 3a). Here, the biologist is particularly interested in the location of cells and in dominant CMP clusters. In case of noise or a high frequency of other small structures that visually overlap with the structures of interest, an increased global transparency simplifies the retrieval of this information (Fig. 3b). It visually suppresses small structures



(a) (b)

Figure 4: *(a) Ray casting using nearest neighbor interpolation at sample points. (b) Trilinear interpolation improves the recognition of depth information and cell surface morphology. However, mixed colors occur at transitions between different CMP clusters and between a cluster and the background (arrows in inset).*



(a) (b)

Figure 5: *(a) Close-up view of a cell and a neighboring cell (upper left). A large CMP cluster appears turquoise in both cells. (b) A color range redistribution for the same view reveals the true variety of CMPs (insets). It shows that the visible CMP cluster of the neighboring cell differs from the largest cluster of the focused cell.*

while large regions exhibiting identical CMPs remain visible due to a higher accumulated opacity. The global transparency is adjustable and initially set to zero.

The volume rendering has been implemented employing the open-source Visualization Toolkit (VTK, Kitware, Clifton Park, NY, U.S.). VTK offers fast volume rendering techniques exploiting graphics hardware, it implements 3D interaction techniques, e.g., picking, and it is freely available. Furthermore, it integrates well with the existing framework written in C#, by means of the ActiViz software (VTK, Kitware, Clifton Park, NY, U.S.). In VTK, two hardware-accelerated volume rendering techniques are integrated: 3D texture mapping and GPU-based ray casting. We employ the latter, since it is better suited for interactive applications where the input dataset is constantly updated, e.g., through brushing operations. These updates significantly slow down the texture mapper since it always resamples the data to be a power of two in each direction before rendering. In ray casting, the data is sampled along each ray. At each sample position, the corresponding data value is interpolated based on the neighboring values. VTK implements nearest neighbor and trilinear interpolation. We apply the latter since the resulting visualization better conveys depth information and cell surface morphology (Fig. 4). A drawback of trilinear interpolation is the generation of mixed colors along the border between two CMP clusters or a cluster and the background (arrows in inset of Fig. 4b). However, the corresponding regions are very narrow and hence can be distinguished from the real data. The sample distance along the ray as well as the image sampling density have a strong impact on the image quality. If either of them is too low, aliasing artifacts occur. VTK's default values for the sample distance and the sampling density are 1.0 and one ray per pixel, respectively. These settings led to an artifact-free visualization in all our experiments. However, a rendering at interactive frame rates is hampered for large render windows, e.g., 1600×1200 pixels. As a solution, VTK offers an automatic adjustment of the parameters to a desired update rate, which we set to 15 frames per second. This guarantees a fluent interaction but also causes aliasing artifacts. Hence, we turn the automatic adjustment on when the user is interacting with the scene and turn it off again when interaction stops.

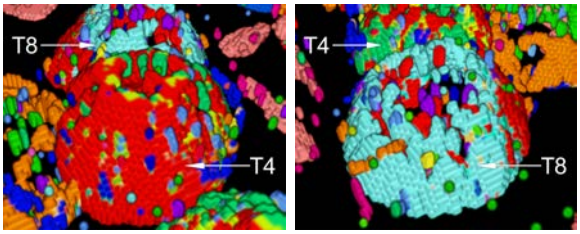
5.3. Perceptually Optimized Coloring

Color serves as CMP identifier across all views of the toponome analysis framework. It is crucial that different CMPs, i.e., their associated colors, may be well discriminated by the biologist. However, a toponome dataset is likely to contain several thousand CMPs. While generating a different color for each CMP is technically possible, visually discriminating these colors by far exceeds the capabilities of the human visual system. Hence, we offer a dynamic color range distribution, that can be activated for the current view on the data, e.g., a close-up view of a particular cell (Fig. 5). It then generates a set of colors with a sufficient perceptual difference for the currently visible CMPs. The colors are transfered to all other views for a coherent visualization. The color range distribution is carried out before the dataset is initially displayed. The user may trigger a redistribution for the current view at any time of the exploration process. In the following, we provide details on the implemented color range distribution algorithm. It strongly improves the former coloring algorithm, which was independent of the number of CMPs and generated 255 shades of color by uniformly sampling the perceptually non-linear RGB color space.

At first, an initial pool of colors is computed in HSV color space by a regular sampling of each component (Alg. 1). From this pool, the n colors, which will be associated to the n visible CMPs, will be drawn. When computing the pool, the hue component is sampled over the full range of 360° with a sample distance of $360^\circ/n$. For the saturation S and the value

Algorithm 1 Compute pool of colors in HSV space.

```
counter ← 0
degree_incr ← 360/n
for hue = 0 → 359 incr : degree_incr do
    for saturation = 60 → 100 incr : 10 do
        for value = 70 → 100 incr : 10 do
            colors[counter] ← [hue, saturation, value]
            counter ← counter + 1
        end for
    end for
end for
```



(a) (b)

Figure 6: *Close-up views of a T4 (a) and a neighboring T8 lymphocyte (b). Note the strikingly different toponomes.*

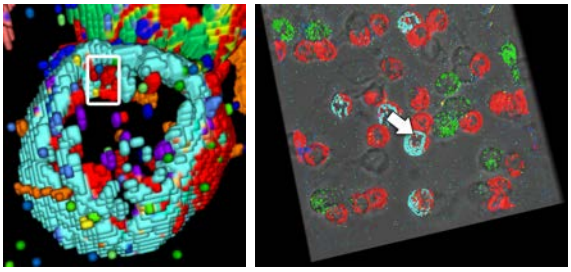
V component, only a subrange of $[0,100]$ is considered in order to avoid undersaturated and too dark colors. Both, S and V are iteratively incremented by 10, within the subranges $[60,100]$ and $[70,100]$, respectively. With this computation scheme, n hue samples, 5 saturation samples per hue sample and 4 value samples per saturation sample are employed resulting in a color pool size of $n \times 5 \times 4$. Less saturated and darker colors are computed first. They will later be assigned to CMPs with a high frequency of occurrence while highly saturated and bright colors will be assigned to CMPs with a low frequency of occurrence. Thus, large protein clusters remain well visible due to their size and small clusters are easier to perceive due to a striking coloring.

In the next step, the n visible CMPs of the current view are computed by casting a ray per pixel into the scene and determining each intersected voxel and the corresponding CMP. Finally, n colors with a sufficient perceptual difference are drawn from the pool. We compute the difference in the perceptually linear CIELab color space, since a difference between two colors there corresponds well to their perceived difference. As a difference measure, the Euclidean distance is employed [Sha03]. To facilitate distance computations, the color pool is transformed from HSV to RGB [Smi78] and from RGB to CIELab space [JL07]. Before the n colors are drawn, a distance threshold is defined beyond which two colors are considered as sufficiently different. If the threshold is too high, the number of colors that can be drawn is too small. If its is too small, the colors are not perceived as different. Hence, the threshold is determined in an iterative process. The iteration starts with a high value and decreases the value down to the *just noticeable difference* (JND) of two colors in CIELab space which is ≈ 2.3 [Sha03]. The iteration is terminated once n colors could be determined. The initial threshold as well as the step size for decreasing it have been determined empirically and set to $2.3+30$ and 0.5, respectively. A higher initial threshold led to very small color sets, e.g., containing just a single entry. During each step of the iteration, the algorithm tries to draw a set of n colors from the pool. It starts with the first color and adds it to the set. Thereafter, the distance between this color and the second color is computed. If it is higher than the current threshold, this color is also added. If not, the algorithms proceeds with the third color and measures its distance to the first one and so fourth. If more than a single color is contained in the already drawn set, distance tests are carried out between the new color and each of the set members. Only if all tests are passed, the new color is added.

The computation time of the color range distribution depends on the view size and the percentage of background pixels (black image regions). For a view size of 1600×1200 pixels and an overview of the dataset as in Figure 3, the time is 4.4s. For the same view size and a close-up view as in Figure 6a, the time is longer (6.2s) due to a lower percentage of background pixels. A drawback of the current approach is that the colors of the CMPs which are outside the current view are not taken into account. Hence, these colors may also be assigned to CMPs within the view. This poses a problem once the user changes the view, e.g., to an overview representation. In such a case, the color range distribution should be triggered again.

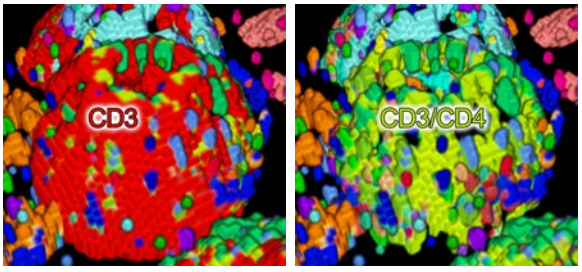
5.4. Interactive Exploration of 3D Toponome Data

In this section, we describe our methods for interacting with the 3D view in an exploratory analysis. After the dataset has been loaded, all CMPs are listed and automatically selected in the table view. Due to linked views, this generates an overview visualization of the cell sample in the 3D view (Fig. 3a). The biologist then modifies the global transparency in order to get an overview of the most prominent CMPs (Fig. 3b). Since the table is initially sorted from top to bottom according to decreasing CMP frequency, these CMPs are listed at its top. This together with the overview helps the biologist in identifying the cell types which are contained in the sample. For example, the cluster of differentiation (CD) marker protein CD4 binds to the surface of T4 lymphocytes. If now CMPs are listed in the table with an entry of 1 in the column representing CD4 and if the corresponding regions cover large parts of a cell's surface in the 3D view, this cell represents a T4 lymphocyte. In the present sample, for instance T4 and T8 lymphocytes as well as monocytes and thrombocytes may be observed. In Figure 3b, T4 lymphocytes, which activate and direct other immune cells, ap-



(a) (b)

Figure 7: *Similarity brushing of a T8 lymphocyte. (a) The biologist brushes an interesting part of the cell surface. (b) The corresponding CMPs are determined across all slices and the 3D visualization is restricted to these CMPs. A phase contrast image stack is volume rendered in grayscale mode and integrated into the scene as a spatial frame of reference.*



(a) (b)

Figure 8: *CMP peeling of a T4 lymphocyte. (a) Initial view showing all CMPs. The CMP with a single entry equal 1 for the cluster of differentiation (CD) marker protein CD3 shall be peeled off (CD3 is found in the membrane of T4 and T8 lymphocytes). (b) A CMP with entries equal 1 for CD3 and CD4 is revealed (CD4 is specific to T4 lymphocytes).*

pear as red and greenish sprinkled hemispherical structures. T8 lymphocytes destroy cells which are effected by a virus. Large parts of their surface are colored in turquoise. Monocytes are another type of white blood cells and appear as mostly blue structures. They are important in immune function and often serve as antigen presenting cells. Thrombocytes are the light red, small, solid structures in between the hemispherical cells. They play a role in blood coagulation. Please note that the coloring of the cells may change across the figures due to color range redistribution (Sec. 5.3).

For our collaborating biologist, the T4 and the T8 lymphocytes are of high interest. Hence, he zooms in to one of these cells and applies a color range redistribution (Fig. 6). Then, the proteins that contribute to the visible CMPs need to be identified. This may be accomplished sequentially by memorizing the color of each CMP and then searching for this color within the selected rows of the table view (Fig. 2). The entries in a detected row which are equal 1 then represent the co-mapped proteins. This approach has several drawbacks. It involves frequently changing the point of gaze between table and 3D view in order to recall the color of interest. This is further complicated as soon as scrolling the table is required. Furthermore, once the row of interest has been detected it needs to be scanned for its entries equal 1. We avoid these problems by implementing a CMP probing through mouse-over interaction. During mouse movement, a ray is cast into the scene starting from the current mouse pointer position. The first hit non-background voxel is reported, the corresponding CMP is extracted, and the names of its co-mapped CD marker proteins are rendered as text at the mouse pointer position. Further, the table row representing this CMP is moved to the top in order to avoid scrolling.

While the biologist samples the cell's surface, he mentally constructs a molecular "face" of the cell which he later compares to the "face" of other cells, e.g., for detecting (pathological) variations. We support this comparison by *similarity brushing* (Fig. 7). Here, a cell surface part is brushed in screen space, the corresponding CMPs are determined across all slices and the visualization in all views (including the 3D view) is restricted to these CMPs. Thus, regions containing at least one of the brushed CMPs are revealed. In practice, the biologist defines a rectangular ROI enclosing a cell surface part. Then, rays are cast into the scene for each ROI pixel. Finally, all hit non-background voxels and the corresponding CMPs are determined per ray. In order to augment the 3D visualization, which is restricted to a subset of CMPs and hence, to certain cell parts, a phase contrast image stack is volume rendered and integrated into the scene as a spatial reference (Fig. 7b). The reference is rendered in shades of gray and a global transparency is assigned such that a trade-off between a good visibility of the brushed data and a clear visibility of reference structures is achieved. The reference volume approach has been adopted from superimposing a toponome map on a phase contrast image (Sec. 4.1).

A common problem in 3D data visualization is occlusion. Transferred to toponome data, cell regions corresponding to one CMP occlude others that correspond to another CMP. Hence, we implemented a semi-transparent rendering of the occluding region which is triggered by a point-and-click interaction. After clicking, the desired CMP is determined as described for the CMP probing. This approach is very limited, since overlapping semi-transparent regions are difficult to distinguish and color mixing occurs. Hence, we implemented a CMP peeling interaction (Fig. 8). The desired CMP is selected by means of point-and-click and then, "peeled off" by rendering it fully transparent. Further, it is automatically deselected in all other views of the framework. The peeling step may be repeated any number of times. We enhance the usability of this technique by offering an undo/redo mechanism which is operated via the graphical user interface of the 3D view.

5.5. Anecdotal User Feedback

We gathered anecdotal feedback from our collaborating biologist and a computer scientist who has been working in his laboratory for several years. Both also co-authored this paper. They appreciated the fast generation of the 3D visualization, its completeness with regard to the number of displayed CMPs, and its integration into their framework. They used the color range redistribution extensively (Sec. 5.3). However, they criticized that it ignores the spatial arrangement of the differently colored cell surface parts. Thus, neighboring parts may be assigned rather similar colors as compared to further distant parts. The integration of a reference volume was commended (Fig. 7b). All interaction methods were assessed as useful. Particularly, CMP probing and CMP peeling were frequently used. The computer scientist requested an undo/redo mechanism for the peeling which we added.

6. Summary and Discussion

We presented a volume rendering approach for visualizing 3D toponome data, which significantly reduces the workload of our collaborators. Visualizing a 3D toponome dataset, which took them $1-2$ workdays before, now takes $\approx$ 30s. We integrated the 3D visualization into their analysis framework. In this framework, a unique color serves as CMP identifier. We improved the existing coloring scheme by computing a set of perceptually optimized colors. This set may be recomputed for varying views on the data in order to optimize the differentiability of the currently visible CMPs. To support an exploratory analysis, the 3D view was equipped with several interaction techniques. CMP probing supports the biologist in mentally forming the molecular "face" of a cell. CMP peeling provides insight into the composition of a cell's surface. For instance, a CD3 layer wrapping around a CD3/CD4 layer was frequently observed in T4 lymphocytes (Fig. 8). We further presented the brushing of interesting CMPs in 3D space, which restricts the visualization in all views to these CMPs. In future work, we will conduct a user study to evaluate the perceptually optimized coloring.

Acknowledgements

Technological and biological toponome studies were supported by BMBF grants Biochance, CELLECT, NBL3, NGFN2, NGFNplus, and through DFGschu627/10-1, the Innovationskolleg INK15, and the Klaus Tschira foundation (project toponome atlas).

References

[BDS10] BARYSENKA A., DRESS A. W. M., SCHUBERT W.: An information theoretic thresholding method for detecting protein colocalizations in stacks of fluorescence images. *J Biotechnol 149*, 3 (2010), 127–131.

[BŠG*09] BRUCKNER S., ŠOLTÉSZOVÁ V., GRÖLLER M. E., HLADŮVKA J., BÜHLER K., YU J., DICKSON B.: Braingazer - visual queries for neurobiology research. *IEEE Trans Visual Comput Graph 15*, 6 (2009), 1497–1504.

[CPS*02] CLENDENON J. L., PHILLIPS C. L., SANDOVAL R. M., FANG S., DUNN K. W.: Voxx: a pc-based, near real-time volume rendering system for biological microscopy. *American Journal of Physiology - Cell Physiology 282*, 1 (2002), C213–C218.

[dLVvL06] DE LEEUW W., VERSCHURE P. J., VAN LIERE R.: Visualization and Analysis of Large Data Collections: a Case Study Applied to Confocal Microscopy Data. *IEEE Trans Visual Comput Graph 12*, 5 (2006), 1251–1258.

[FBKS07] FRIEDENBERGER M., BODE M., KRUSCHE A., SCHUBERT W.: Fluorescence detection of protein clusters in individual cells and tissue sections by using toponome imaging system: sample preparation and measuring procedures. *Nat Protoc 2*, 9 (2007), 2285–2294.

[FDM*00] FANG S., DAI Y., MYERS F., TUCERYAN M., DUNN K.: Three-dimensional microscopy data exploration by interactive volume visualization. *Scanning 22*, 4 (2000), 218–226.

[GYL*12] GUO H., YUAN X., LIU J., SHAN G., CHI X., SUN F.: Interference microscopy volume illustration for biomedical data. In *Pacific Vis* (2012), pp. 177–184.

[JL07] JIN L., LI D.: A switching vector median filter based on the cielab color space for color image restoration. *Signal Processing 87*, 6 (2007), 1345 – 1354.

[Kau98] KAUFMAN A.: Visualizing Nerve Cells with VolVis. In *Pac Symp Biocomput* (1998), pp. 165–176.

[MCS*08] MOSALIGANTI K., COOPER L., SHARP R., MACHIRAJU R., LEONE G., HUANG K., SALTZ J.: Reconstruction of cellular biological structures from optical microscopy data. *IEEE Trans Visual Comput Graph 14*, 4 (2008), 863–876.

[OFH*11] OELTZE S., FREILER W., HILLERT R., DOLEISCH H., PREIM B., SCHUBERT W.: Interactive, Graph-Based Visual Analysis of High-Dimensional, Multi-Parameter Fluorescence Microscopy Data in Toponomics. *IEEE Trans Visual Comput Graph 17*, 12 (2011), 1882–1891.

[RPFC01] RAZDAN A., PATEL K., FARIN G. E., CAPCO D. G.: Volume visualization of multicolor laser confocal microscope data. *Computers & Graphics 25*, 3 (2001), 371–382.

[SBP*06] SCHUBERT W., BONNEKOH B., POMMER A. J., PHILIPSEN L., BÖCKELMANN R., MALYKH Y., GOLLNICK H., FRIEDENBERGER M., BODE M., DRESS A. W. M.: Analyzing proteome topology and function by automated multidimensional fluorescence microscopy. *Nat Biotechnol 24*, 10 (2006), 1270–1278.

[Sch10] SCHUBERT W.: On the origin of cell functions encoded in the toponome. *J Biotechnol 149*, 4 (2010), 252–259.

[SGK*11] SCHUBERT W., GIESELER A., KRUSCHE A., SEROCKA P., HILLERT R.: Next-generation biomarkers based on 100-parameter functional super-resolution microscopy TIS. *New Biotechnology* (2011), In Press, Available Online.

[Sha03] SHARMA G.: *Digital Color Imaging Handbook*. CRC Press, 2003.

[Smi78] SMITH A. R.: Color gamut transform pairs. *SIGGRAPH Comput. Graph. 12*, 3 (1978), 12–19.

[SVP96] SAKAS G., VICKER M. G., PLATH P. J.: Case study: visualization of laser confocal microscopy datasets. In *IEEE Visualization* (1996), pp. 375–379.

[WOCH09] WAN Y., OTSUNA H., CHIEN C.-B., HANSEN C.: An interactive visualization tool for multi-channel confocal microscopy data in neurobiology research. *IEEE Trans Visual Comput Graph 15*, 6 (2009), 1489–1496.

Interactive Labeling of Toponome Data

S. Oeltze-Jafra[1] and F. Pieper[1] and R. Hillert[2] and B. Preim[1] and W. Schubert[2]

[1]Department of Simulation and Graphics, University of Magdeburg, Germany
[2]Molecular Pattern Recognition Research Group, University of Magdeburg, Germany

Abstract

Biological multi-channel microscopy data are often characterized by a high local entropy and phenotypically identical structures covering only a few pixels and forming disjoint regions spread over, e.g., a cell or a tissue section. Toponome data *as an example, comprise a fluorescence image (channel) per protein affinity reagent, and capture the location and spatial distribution of proteins in cells and tissues. Biologists investigate such data using a region-of-interest in an image view and a linked view displaying information aggregated or derived from the channels. The cognitive effort of moving the attention back and forth between the views is immense.*
We present an approach for the in-place annotation of multi-channel microscopy data in 2D views. We combine dynamic excentric labeling *and static* necklace maps *to cope with the special characteristics of these data. The generated annotations support the biologists in visually exploring multi-channel information directly in its spatial context. A label is generated per unique phenotype included in a flexible, moveable focus region. The labels are organized in a circular fashion around the focus region. On demand, a nested labeling can be generated by displaying a second ring of labels which represents the channels characterizing the focused phenotypes. We demonstrate our approach by toponome data of a rhabdomyosarcoma cell line and a prostate tissue section.*

Categories and Subject Descriptors (according to ACM CCS): J.3 [Computer Applications]: Life and Medical Sciences—Biology and genetics

1. Introduction

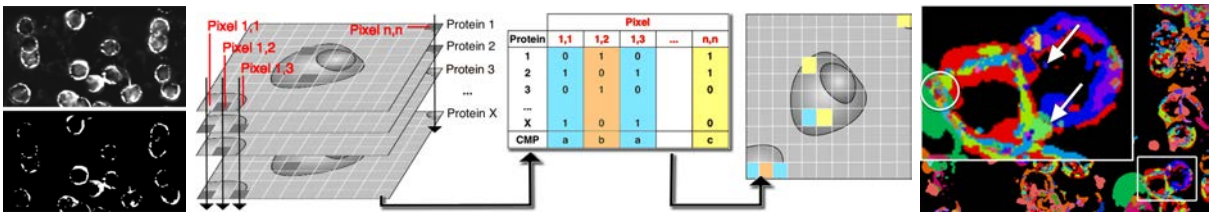
Proteins are the basic modules of cells performing a huge variety of functions in living organisms. A major challenge in biology is to understand how proteins cooperate in cells and tissues in time and space [Sch10]. The *toponome* of a cell describes its functional protein pattern, i.e. the location and spatial distribution of proteins. In *toponomics*, the toponome is investigated in order to understand how cells encode different functionalities both in health and disease. Robot-driven multi-parameter fluorescence microscopy is employed for imaging the toponome [Sch03]. The imaging may be carried out in 2D or 3D and results in a fluorescence image or volume per protein. Here, we focus on a 2D slice-based analysis of toponome data.

In a post-processing step, the fluorescence data is binarized. For each pixel, a binary code (protein pattern) is constructed over all images, i.e. proteins, which then encodes the local protein co-mapping. Finally, all unique protein patterns are determined and each is assigned a unique color. Biologists are interested in the natural clustering of protein patterns across a cell, in the difference in clustering between cells or healthy and pathologic tissue, and in the frequencies of proteins and protein patterns. Hence, they visually explore the toponome data piece by piece. They repeatedly define a region-of-interest in an image view and inspect the corresponding unique patterns in a separate table view. The cognitive effort of moving the attention back and forth between the views is immense.

We present an approach to interactively label toponome data in image views facilitating an exploration of the toponome in its spatial context. Labeling the clusters of protein patterns is challenging since (Fig. 1c):

- very small clusters cover only a few pixels,
- the local entropy, i.e. variety of clusters, is high, and
- phenotypically identical clusters form disjoint regions.

To cope with the high local entropy and to account for the piece-wise exploration of the data, we adopt dynamic *excentric labeling* of a focus region [FP99]. Phenotypically identical but disjoint regions, such as the turquoise or red clusters in Figure 1c, require either multiple converging lines (*leaders*) connecting the regions with a single label (*many-to-one*



(a) (b) (c)

Figure 1: *(a) Fluorescence signal of a protein affinity reagent as measured (top) and after binarization (bottom). White pixels indicate protein present. (b) Generation of Combinatorial Molecular Phenotypes (CMPs). For each pixel, the binarized fluorescence signal of all protein affinity reagents is collected in a combinatorial binary code. The set of unique codes, i.e. the CMPs, is computed and visualized in a toponome map (right). Image adapted from [OKH*12]. (c) Inset of an exemplary toponome map illustrating the challenges on labeling. Very small clusters of protein patterns exist (arrows). The local variety of clusters, e.g., inside the circle, is high. Phenotypically identical clusters form disjoint regions, e.g., the turquoise and the red regions.*

labeling) or also multiple labels. In order to avoid visual clutter, we combine excentric labeling with static leader-free *necklace maps* [SV10], which line up a single label per unique protein pattern on a curve surrounding the focus region and relate labels to regions, e.g., by matching colors. To the best of our knowledge, we are the first to present a dynamic variant of necklace maps posing special requirements on label update during exploration.

We support multiple labeled focus regions facilitating cell-to-cell comparisons, which so far required the tedious comparison of individual tables. Copies of the labelings are organized in a management view to structure and log the exploration. We demonstrate our approach by a rhabdomyosarcoma cell line and a prostate tissue section. It may be transferred to similar image data, e.g., light microscopy images of differently stained tissue, or maps of geospatial data, e.g., the world-wide distribution of mineral resources.

2. Biological Background

The *toponome* of a cell is defined as the entirety of all protein networks, in which proteins are defined by their protein-to-protein context [Sch03]. It is hierarchically organized and comprises protein clusters which in turn contain *lead proteins* and are interlocked as a network [SBP*06]. The lead proteins control the topology of the clusters and their function as a network. The most advanced toponome imaging technique is robot-driven multi-parameter fluorescence microscopy TIS^{TM} [FBKS07]. It is capable of co-mapping hundreds of proteins and their distribution across a cell or tissue sample in situ [SBP*06, SGK*12]. Imaging and analyzing the toponome are essential in finding new drugs, e.g., for cancer treatment, and for detecting protein clusters that can be regarded as a new system of biomarkers in disease [Sch10, SGK*12].

Combinatorial Molecular Phenotypes. After imaging the toponome, the fluorescence data is binarized [BDS10] (Fig. 1a). This generates a combinatorial binary code (protein pattern) for each pixel where 0 indicates protein absent and 1 protein present. The unique binary codes in the data are referred to as *Combinatorial Molecular Phenotypes* (CMPs). A simple technique for visualizing CMPs is their color-coded representation in a *toponome map*. The computation of a unique color per CMP is described in [OKH*12]. The generation of binary codes, the concept of CMPs, and the toponome map are illustrated by Figure 1(b,c). The binary code corresponding to a CMP very often exists at many pixel positions, which are clustered at several locations of a cell or tissue sample. These protein clusters correspond to functional cell units and are of crucial interest.

3. Biological Workflow and Requirement Analysis

The analysis of toponome data starts with a hypothesis-free visual exploration of the CMPs and involves the following biological tasks:

(1) detection of selective CMP patterns,
(2) comparison of CMP patterns, and
(3) identification of lead proteins.

In (1), patterns characteristic for a particular cell type, a developmental stage of cells or a pathology are searched for. Such patterns support an understanding of cell composition and function, protein interaction, and may serve as biomarkers in disease. The comparison of patterns (2) is crucial, e.g., in comparing healthy and pathologic tissue or cells in different developmental stages for understanding stage transition. The detection of lead proteins (3) may be the first step in drug development. Inhibiting a lead protein causes a disassembly and function loss of the associated protein network, which may eventually stop the disease [SBP*06].

Workflow. The biologists perform these tasks following a specific workflow implemented by their in-house, multiple coordinated view framework (see [OFH*11, OKH*12] for

details on the framework). Here, we focus on the 2D view showing the image data and the toponome map and on the table view listing the CMPs as rows and the proteins as columns. Together, they are the main vehicles of initial toponome exploration (Fig. 2a,b).

After toponome data have been acquired, the biologists browse the morphology in the 2D view to orient themselves in the spatial domain of the data. This step is carried out, e.g., based on a phase contrast image facilitating a good visual separation between cells and background (Fig. 2b). Next, the biologists investigate the CMP data at morphologically interesting locations and search for selective CMP patterns. For this purpose, a focus region is defined on the morphology image. This region is neither draggable nor resizable. After its definition, the corresponding part of the toponome map is superimposed. Note that the corresponding CMPs are not only superimposed on the focus region but on the entire image (Fig. 2b). This is necessary to assess whether a CMP pattern is selective or appears anywhere in the data.

Once a selective pattern has been detected, its CMPs and their contributing proteins are investigated in the table view, which is often shown on a second screen (Fig. 2a). The table lists the CMPs of the entire dataset sorted according to each CMP's overall frequency. The rows corresponding to the focused CMPs are colored. A CMP's unique color is employed to establish visual correspondence between table and toponome map. Comparing cells or cell parts regarding their CMP pattern and proteins requires multiple focus regions. Since this was not supported so far, multiple instances of the framework were created or screenshots were compared.

Requirement Analysis. To investigate the CMPs of an interesting pattern, the user browses the table, which may list hundreds or thousands of CMPs. *All* columns must be checked to retrieve the present proteins. This is essential, e.g., for detecting lead proteins. If all CMPs of a pattern contain a specific protein, it represents a lead protein candidate.

The exploration requires the user to constantly move the focus of attention back and forth between table and 2D view. The static focus region prevents a fluent sampling of the toponome and a comprehension of pattern changes between neighboring image regions. The missing support for multiple focus regions hampers the comparison of CMP patterns.

The primary requirement of our collaborators on a novel approach is the embedding of information derived from the table into the 2D view such that the toponome may be explored directly in its spatial context. Further requirements are the support of multiple focus regions and the management of these regions and their respective CMP pattern, e.g., capture, show, hide, and store.

4. Related Work

This section is based on a survey of labeling techniques in medical visualizations [OJP14]. Ali et al. studied handmade illustrations in scientific and technical textbooks and identified two types of labels: internal and external [AHS05].

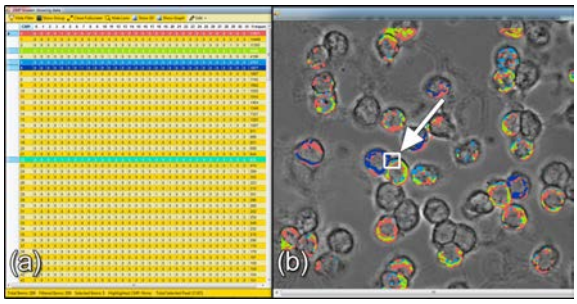


Figure 2: *Toponome analysis framework. (a) The table view lists all CMPs as rows and proteins as columns. (b) The 2D view shows a grayscale phase contrast image as spatial context. The ring-shaped structures represent cells. Each CMP within a user-defined focus region (arrow), i.e., the corresponding part of the toponome map, is superimposed in color and the respective table row is colored likewise.*

Internal Labels. Labels being superimposed on the structure of interest are referred to as internal labels. They have been applied, e.g., to virtual bronchoscopy images [MHST00], medical surface [RPRH07] and volume rendered data [JNH*13]. Their application to toponome data is challenging since clusters of the same CMP do not form a single, continuous region in the 2D toponome map (e.g., the turquoise or red clusters in Fig. 1c). The problem might be tackled by multiple identical labels as shown for annotating vascular structures in volume rendered images [JNH*13]. Here, a vessel is often partially occluded by other vessels or organs. However, another problem prevents the application of internal labels. Often, CMP clusters cover only a few pixels, which would be largely occluded by the label.

External Labels. The occlusion problem is solved by external labels. They are positioned outside the structure of interest and connected to it by a line. This so-called *leader* connects an *anchor point* on the structure and a point on the label box holding the label's textual representation. Ali et al. proposed a variety of real-time label layout algorithms for anatomical 3D models [AHS05]. Labels are arranged in a circular fashion around the model or along its silhouette. Mühler et al. demonstrated the labeling of 3D medical structures located inside a transparent structure or being currently hidden but still of importance for surgical planning [MP09]. Mogalle et al. presented the automatic optimal placement of external labels representing findings in 2D radiological slice data [MTSP12]. They focused on avoiding leader crossings and labels occluding crucial image parts. Their approach is limited to ≈ 10 labels, which is realistic for radiological data. However, the number of CMPs even in a small subregion of the toponome map is often higher.

Boundary Labeling. In early work, Preim et al. presented a system for exploring anatomical models which combines zooming techniques, fisheye views, and interactive labels [PRS97]. The labels are aligned on the left and right boundary of a virtual rectangle enclosing the model. Bekos et al. later coined the term "boundary labeling" in the context of annotating static maps [BKSW07]. A virtual rectangle containing the map is constructed and external labels are placed outside the rectangle. They are connected by leaders to the map areas of interest. Crossings of leaders are avoided and total leader length is minimized. Boundary labeling is generally applied to the entirety of data. Labeling the entire toponome map is however, neither feasible due to the hundreds or thousand of CMPs nor required by the biologists who explore the data piece by piece.

Excentric Labeling. The cell-wise or subcellular piece-wise exploration of the data is very well related to excentric labeling by Fekete and Plaisant [FP99]. Their dynamic approach aims at labeling dense maps interactively by means of a moveable, flexible focus region. The labels are displayed in stacks to the left or right of the focus region and connected to the structure of interest inside the region by a leader. Fink et al. extended the approach by various techniques for creating a visually pleasing annotation, e.g., the use of straight lines or Bézier curves instead of zigzagging polylines [FHS*12]. Luboschik et al. presented a fast point-feature labeling approach, which avoids the placement of labels over other labels or visual representatives such as leaders and icons [LSC08]. They coupled the approach with a moveable label lens. Transferring excentric labeling to toponome data is not straightforward. Several leaders originating either from a single label (*many-to-one labeling* [Lin10]) or from multiple identical labels would be necessary to annotate multiple clusters of the same CMP. Even with minimized leader crossings, this would cause a cluttered visualization for a larger number of CMPs.

Necklace Maps. A static labeling approach abandoning leaders has been proposed by Speckmann and Verbeek for visualizing statistical data on geographical maps [SV10]. Glaßer et al. have applied necklace maps to labeling clusters of breast tumor tissue with cluster-specific perfusion information [GLP14]. In a necklace map, the labels are related to structures of interest by matching colors – the unique CMP color in our case – and spatial proximity. They are organized on a one-dimensional curve (the necklace) that surrounds the map or a subregion.

Consequences. We choose *external labels* over internal ones since the latter would occlude very small CMP clusters. To cope with the high local entropy of toponome data and to account for its piece-wise exploration, we adopt *excentric labeling* of a focus region [FP99]. Disjoint regions, such as the turquoise clusters in Figure 1c, require either multiple converging leaders connecting the regions with a single label (*many-to-one labeling*) or also multiple labels. In order to avoid visual clutter, we adopt the leader-free *necklace maps* [SV10], which line up a single label per CMP or protein on a curve surrounding the focus region. The combination of excentric labeling and necklace maps meet our requirements on a visual exploration of toponome data (Sec. 3).

5. Interactive Labeling of Toponome Data

We discuss our visual encoding, aspects of label position, order, and count, and we emphasize modifications to the original static necklace map approach. After describing the necklace composition, we elaborate on interaction facilities and introduce a view for managing multiple necklaces.

5.1. Basic Approach

Initialization. At first, the user defines a focus region (region-of-interest, abbrev. ROI) on the toponome map by means of a flexible lens. We have implemented three lens shapes: circle, rectangle, and lasso. Circular and rectangular lenses are adjustable with respect to size and position. Both are meant for a quick inspection of the CMP distribution. The lasso is employed for a more targeted inspection of separate cellular subregions. It does not need to be adjustable since it is aligned with a particular shape. In an early prototypical implementation, our collaborators favored the circular lens since it adheres to the metaphor of exploring a dark room by means of a flashlight. For a recent survey on interactive lenses in visualization, see [TGK*14].

Nested Necklaces. After ROI definition, all pixel positions within the ROI and their associated CMPs are determined. Then, a one-dimensional curve (the necklace) surrounding the ROI is constructed. Currently, our implementation is restricted to a circular necklace since it best matches the circular lens shape (see [SV10] for arbitrary necklace shapes). The CMPs are represented by graphical symbols, which are strung on the necklace (inner necklace in Fig. 3). Following Speckmann and Verbeek [SV10], we provide circular and bar-shaped symbols (Fig. 4). In the remainder, we use the terms *symbol* and *label* interchangeably.

On demand, a second necklace enclosing the former is displayed. One symbol per protein present in the focused CMPs is drawn (outer necklace in Fig. 3). This nested labeling facilitates the concurrent exploration of CMPs and proteins. While dragging the focus region, the protein necklace is hidden by default to avoid mental overload.

5.2. Visual Encoding

Label Text. When the CMPs of a new toponome dataset have been determined, each is assigned a unique name which simply equals its place in a frequency ranking of all CMPs. This name is typeset within the corresponding symbol. The

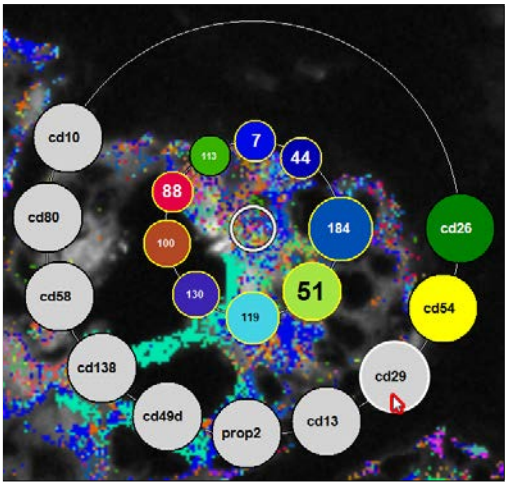


Figure 3: *Nested necklace map. Two one-dimensional curves (the necklaces) surround a focus region (white center circle). The CMPs in the focus region and their present proteins are represented by circular symbols strung on the inner and the outer necklace, respectively. The CMP symbol colors match the unique CMP colors while the colors of the protein symbols indicate lead protein likelihood. Please see the text for all other encodings and interaction facilities.*

name of the protein affinity reagent is typeset in the symbol of the corresponding protein. The names relate the symbols to the table view since the latter consists of columns listing the ranking place and the proteins (Fig. 2a).

Symbol Size. The relative frequency of a CMP inside the ROI f_{cmp} is of particular interest to the biologists. It is defined as the number of ROI pixels being associated with the CMP normalized by the overall number of ROI pixels. We map the CMP frequencies to the area of the circular symbols and to the length of the bar-shaped symbols, respectively. In accordance with [SV10] and Tufte who demands to "tell the truth about data" [Tuf01], we employ *mathematical scaling*, which directly relates the symbol area/length to the underlying data. However, for the circular symbols, we offer *perceptual scaling* by Flannery's compensation which aims at compensating for the non-linear relationship between an increase in circular area and the perceived increase [Fla71]. Our collaborators prefer circular symbols due to their orientation-independent encoding of frequency and the more symmetric and aesthetic appearance of the resulting necklaces (Fig. 4). Hence, we show circular symbols in the remainder.

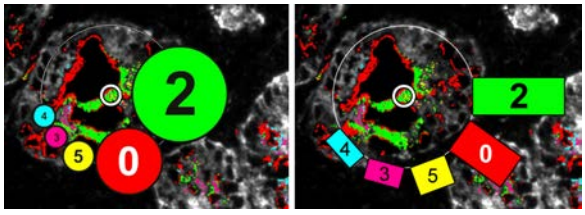


Figure 4: *Circular and bar-shaped labels are implemented. Circles encode CMP frequency by area and bars by length.*

The biologists are also interested in the relative frequencies of the proteins inside the ROI f_{prot}. A protein's relative frequency is independent of the number of pixels. It is defined as the number of focused CMPs with this protein present normalized by the overall number of focused CMPs (except for the background zero-CMP). The biologists categorize the frequencies rather than considering individual values. For the detection of lead proteins, it is sufficient to know whether a protein is present in (nearly) all CMPs inside the ROI or only in a small subset. Hence, we assign a uniform size to the protein symbols and employ color to encode the frequency category (outer necklace in Fig. 3).

Symbol Color. A necklace map communicates the relation between a symbol and its corresponding pixels within the ROI by matching colors and spatial proximity. Hence, we color each symbol on the CMP necklace according to the CMP's unique color in the toponome map (Fig. 3). For the symbols on the protein necklace, we use a segmenting color scale. Symbols of proteins with a relative frequency $f_{prot} < 80\%$ are shaded in gray, $80\% \leq f_{prot} < 100\%$ in yellow, and $f_{prot} = 100\%$ in green. This facilitates an easy detection of lead protein candidates (green; recall Sec. 3) and of such near the mark (yellow).

5.3. Label Position, Order, and Count

The following methods are straightforward to implement based on simple trigonometry facilitating an update of the necklaces at interactive frame rates during exploration.

Position and Order. Besides color, necklace maps employ spatial proximity to relate image or map regions and their corresponding symbol. Optimizing spatial proximity is a hard problem having received special attention in [SV10]. For toponome data, this problem is even aggravated. Often, multiple clusters of the same CMP exist in a focus region (Fig. 3,5) and also a protein may be scattered across the entire region. Optimization with respect to one cluster is not reasonable in particular for similar-sized, equally distributed clusters. Generating multiple symbols would require their mental integration during exploration. The integration is particularly cumbersome if symbol attributes encode data variables, e.g. size encoding CMP frequency. Finally, very small clusters may exist in the center of a focus region where spatial proximity is hard to achieve by means of a standard convex necklace shape. Discussing these problems with our collaborators revealed that in an initial exploration of toponome data, they are rather interested in the relative frequency of

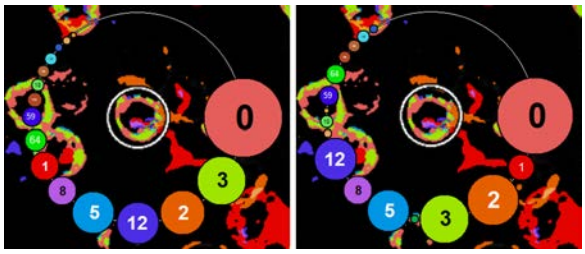


Figure 5: *Local vs. global sorting of CMP symbols. (Left) In local mode, the symbols are sorted clockwise according to their CMP's frequency inside the focus region (inner circle). Note that symbol size encodes local CMP frequency while the label text equals the CMP's place in a ranking of global frequencies. (Right) In global mode, the frequency inside the entire dataset is employed for sorting. The symbols are not ordered anymore according to size, but the label texts are ordered now.*

the CMPs than in their exact location inside the focus region. Hence, we decided to sacrifice the spatial proximity criterion in favor of a sorted symbol line-up along the necklace starting at 3'o clock with the most frequent CMP and proceeding in clockwise order. Due to the sorted line-up, simple comparisons of CMP frequency within a necklace are even possible when symbols sizes are visually not distinguishable.

The symbols on the CMP necklace may be sorted according to the CMP frequency inside the ROI (*local*) or the total frequency in the dataset (*global*). An exploration in local mode supports the detection and tracking of a CMP's place in a local frequency ranking (Fig. 5a). An exploration in global mode simplifies the tracking of a CMP's presence and frequency inside the focus region (Fig. 5b). This is due to the rather stable place of its corresponding symbol in the order of symbols, which is fix as long as the more frequent CMPs also remain in focus. Note that in global mode, the symbol sizes are not ordered since they still represent the local frequency which often differs from the global one. Please also see our supplemental video for an illustration of the modes.

In order to simplify the search for a specific protein, the symbols on the protein necklace may be arranged alphabetically. Alternatively, the symbol order may be chosen to reflect each protein's place in a ranking of the number of associated ROI pixels. The latter is set by default and also shown in all figures of the remainder.

The necklace radii and the arc length distance between neighboring symbols are chosen such that labels do neither overlap the focus region nor each other. The latter is guaranteed along the necklace and across inner and outer necklace.

Count. In a toponome dataset, hundreds to thousands of CMPs may exist depending on the investigated biology and the number of applied protein affinity reagents. Even in a small focus region, the number of CMPs can be quite high. However, the number of labels that can be drawn on the CMP necklace is restricted by the minimum size of a symbol down to which it is readable and by the necklace perimeter. Since the necklace should closely adhere to the focus region rather than exploiting the entire available screen space, its perimeter is bounded above. Instead of predefining the perimeter, we first map each CMP's relative frequency f_{cmp_i} to symbol size. Based on $f_{cmp_i} \in [0,1]$, the diameter $\emptyset_{s_i}$ of the corresponding symbol s of the necklace map is computed:

$$\emptyset_{s_i} = \emptyset_{base} \cdot \sqrt{f_{cmp_i}}, i \in [1, n_{cmp}] \quad (1)$$

The number of CMPs inside the ROI is denoted by n_{cmp}. The global scaling factor $\emptyset_{base}$ corresponds to an adjustable maximum symbol size which is initially set to 150 pixels. Note that this high value is only achieved in the rare case of a single CMP covering the entire ROI ($f_{cmp_i} = 1$). If necessary, $\emptyset_{s_i}$ is clamped to the minimum value of four pixels to guarantee the readability of its symbol color. The mathematical scaling in Equation 1 directly relates the symbol area – not the radius/diameter – to the underlying data by employing the square root.

Based on the maximum of $\emptyset_{s_i}$, we then compute the necklace diameter such that this symbol does no overlap the focus region. We then draw the symbols starting at 3'o clock and proceeding clockwise until a new symbol would intersect the first one. Following this strategy, the most frequent CMPs inside the ROI are labeled. This has been agreed upon with the biologists, since very small CMP clusters might represent noise not being eliminated in the course of binarization (Fig. 1a,b). However, special care must be taken when the labels shall be ordered according to global CMP frequency. If for instance only 20 out of 30 CMPs can be labeled, the 20 locally most frequent CMPs do not necessarily coincide with the 20 globally most frequent ones. To guarantee that always the former are labeled, we first determine them and then, sort only these in descending order according to global CMP frequency. A more fine-granular inspection of the CMP distribution can be accomplished by capturing the necklace and labeling all CMPs in an enlarged separate widget (Sec. 5.5).

The number of labels that can be drawn on the protein necklace is also limited by the same factors but the number of proteins is small as compared to the number of CMPs. The most comprehensive toponome study hereof, employed 100 protein affinity reagents [SBP*06]. Furthermore, only a subset of all proteins is included in a reasonably sized focus region. So far, we have been able to draw a label for each protein inside a ROI employing a symbol size that guarantees good readability and at the same time a perimeter that is not far off the perimeter of the CMP necklace. Drawing all symbols is crucial here since otherwise lead protein candidates may remain unnoticed.

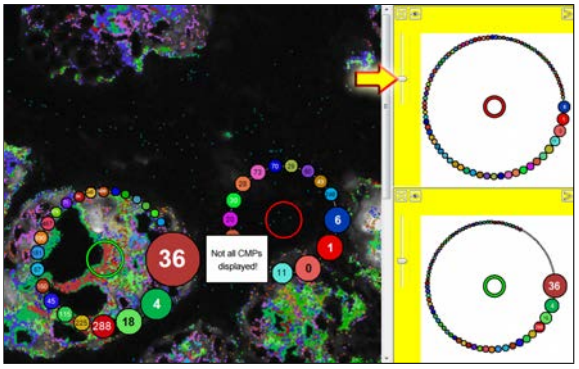


Figure 6: *2D view of image data and toponome map (left) and necklace management view (right). The management view organizes the necklaces of the two focus regions as widgets. Both widgets have been enlarged by means of a slider (arrow) to gain space for more CMP symbols.*

5.4. Necklace Interaction

The user can drag the focus region across the toponome map and modify its size by scrolling the mouse wheel. The necklace of a selective CMP pattern can be captured via mouse-click causing an interactive copy to be added to the necklace management view (Sec. 5.5). Another necklace map may be initialized, causing a fade-out of the old map. For orientation purposes, the old focus region remains visible. If multiple necklaces have been defined, any of them can be reactivated by clicking the respective focus region. Note that during interaction, only the CMPs inside the focus region of the active necklace map are colored in the toponome map.

A tooltip listing the relative and the absolute CMP frequency inside the ROI is shown during mouse hover of a CMP's symbol. If the symbol is clicked, the CMP's pixels are highlighted by a temporary blinking. This is particularly useful in cases of CMPs with barely distinguishable colors. Furthermore, if the protein necklace is visible, the symbols of the proteins present in the CMP are highlighted.

The protein necklace is by default only visible on demand. Hovering the mouse pointer over a symbol causes an emphasis of the symbols of all CMPs with this protein present by means of a yellow contour. On clicking the symbol, the CMPs' pixels are highlighted by a temporary blinking.

5.5. Necklace Management View

The necklace management view facilitates the organization of multiple necklaces and helps to structure the exploration. It is attached to the 2D view of the toponome map (Fig. 6). The view is based on requests by the biologists for having a means to record their exploration results. Such records illustrate the daily work and are integrated in the laboratory book. They support scientific reporting of research results and simplify the communication with other biologists. Furthermore, the management view arranges the necklaces in a non-overlapping fashion thereby simplifying a comparison of the associated CMPs. Superimposing all necklace maps on the toponome map would lead to overlapping necklaces and considerable occlusions of the image data.

In the management view, each necklace is presented in a resizable widget. If a widget is enlarged, the necklace diameter is increased causing previously neglected CMPs to be displayed (recall paragraph "Count" in Sec. 5.3). This facilitates a more fine-granular inspection of the CMP distribution. The background of the widget may be set to the corresponding part of the toponome map. For comparing necklaces, a plain color background causes less distraction.

A necklace map may be shown/hidden in the toponome map by selecting/deselecting its widget. Note that the focus region of a hidden map remains visible. The background color of a selected widget switches from white to yellow. Multiple selections are supported. For each necklace map, the user may choose whether the corresponding CMPs, i.e. their pixels in the toponome map, are shown in color. In Figure 6, the coloring is restricted to the left necklace.

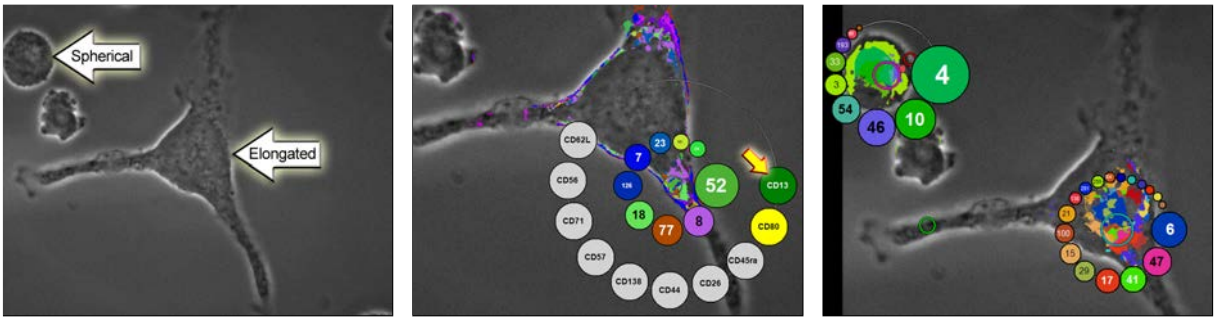
6. Application

We demonstrate our approach by a rhabdomyosarcoma cell line and a prostate tissue section. Both probes have been imaged by means of the TIS robot system with an in-plane-resolution of 216×216 *nm* (Sec. 2). Protein affinity reagents, more precisely, monoclonal antibodies directed against cluster of differentiation (CD) surface marker proteins, were co-mapped on the probes. The resulting fluorescence images were binarized according to [BDS10] (Sec. 2). We conclude the section by providing anecdotal user feedback.

6.1. Rhabdomyosarcoma Cell Line

Rhabdomyosarcoma (RMS) is the most common peripheral malignant tumor of soft tissue in children and adolescents and its causes are unclear [HJC*13]. RMS is made up of cells which normally develop into skeletal muscles. To research RMS, muscle cells were extracted from the RMS cell line TE671. Cell lines are populations of cells which have been cultivated from a single cell thus held to contain the same genetic makeup. The cell sample has been imaged in a single transection with a matrix of 693×552 pixels employing 23 protein affinity reagents. 958 CMPs were derived from the binarized data. Sample preparation, data acquisition, and binarization are detailed in [SBP*06].

RMS cells enter two different evolutionary states characterized by a specific cell shape: spherical and elongated with spindle-shape extensions [SBP*06] (Fig.7a). Spherical cells spontaneously enter an exploratory state in which they form three spindle-shaped extensions. Once a promising direction



(a) (b) (c)

Figure 7: *Necklace maps for visually exploring the toponome of Rhabdomyosarcoma (RMS) cells. (a) Phase-contrast image of RMS cells in two different states of their evolution: spherical and elongated with spindle-shaped extensions. (b) A necklace map at one of the extensions confirms CD13's function as a lead protein [SBP*06] as indicated by the green symbol (arrow). (c) Two focus regions have been defined in the cell bodies. Note the strikingly different toponome despite the same cell type.*

has been detected by the cell, it proceeds to a migratory state characterized by a withdrawal of one of the extensions. The whole process is targeted at metastasis formation.

Previous toponome decoding work has shown that the proteolytic enzyme CD13 functions as a lead protein driving and directing the formation of the cell extensions [SBP*06]. Based on the same cell type and a similar dataset, we recapitulate this finding (Fig. 7b). While previous work required a time-consuming investigation of the CMP table view, the necklace facilitates a quick identification of CD13 as a lead protein. Its corresponding symbol is colored in green and appears at the starting position of symbol drawing (arrow).

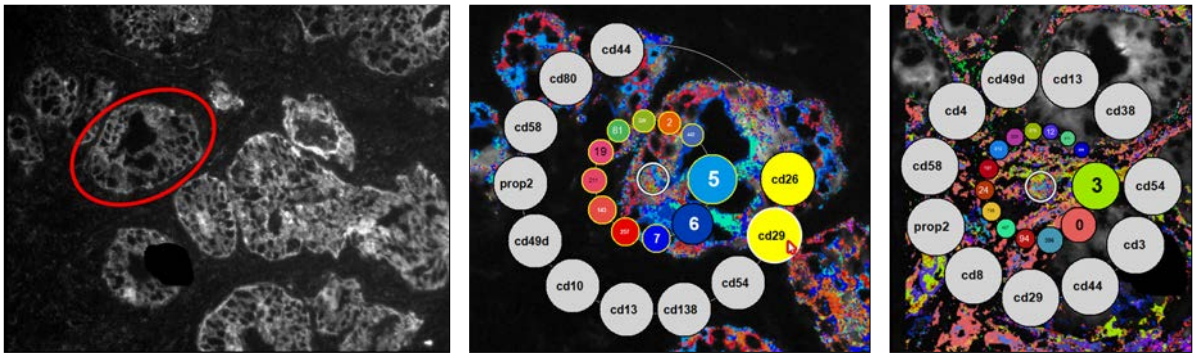
Furthermore, we show that the protein network controlled by CD13 across the cell body shows strikingly different variations for cells in the spherical as compared to the exploratory state (Fig.7c). Two focus regions were placed within the cell bodies. The toponomes represented by the corresponding necklaces are completely disjoint. Furthermore, the CMPs included in the focus region of the spherical cell barely occur in the elongated cell and vice versa. An investigation of the protein necklaces of both focus regions revealed an omnipresence of CD13 (not illustrated here to simplify a comparison of the CMP patterns). This provides further evidence that CD13 functions as a control element steering the transformation from the spherical to the exploratory state by a recombination with other proteins. It was shown in [SBP*06], that inhibiting CD13 prevents the transformation from the spherical to the exploratory state.

6.2. Prostate Tissue Section

The tissue section was cut from a prostate tissue block of radical prostatectomy — the surgical removal of the entire prostate gland in the therapy of prostate cancer. This type of cancer is the most common noncutaneous malignant neoplasm in men in western countries and its pathogenesis is still unclear [SGKH09]. The tissue section has been imaged in a single transection with a matrix of 658×517 pixels employing 17 protein affinity reagents. 2100 CMPs were derived from the binarized data. Sample preparation, data acquisition, and binarization are detailed in [SGKH09].

The tissue section contains several *prostate acini* — many-lobed, berry-shaped terminations of the prostate glands lined by secretory epithelial cells — and the fibromuscular *stroma* between the acini. The protein affinity reagent CD138, which is a marker for prostate cancer progression, singles out the acini in its fluorescence image (Fig. 8a). For clarification, one acinus has been encircled. Its epithelial cells appear white in the image while their nuclei and the lumen of the acinus show no response to CD138 and hence, appear as small black circular and large black centered regions, respectively. The encircled acinus drew the interest of the biologists since a fraction of its epithelial cells exhibits features of *prostate intraepithelial neoplasia (PIN)* [SGKH09].

Researching PIN is crucial since it is considered to be a pre-malignancy of the prostatic glands. In order to investigate the toponome of PIN, we have dragged a focus region across the epithelial cells. A representative necklace map including the protein necklace is shown in Figure 8b. The CMP pattern is selective for epithelial cells since none of the CMPs appear in the stroma surrounding the acini. The protein necklace reveals CD26 and CD29 as lead protein candidates indicated by the yellow colored symbols. Both contribute to all but one CMP, which in both cases is the one with only the respective other protein present. For instance, only CMP 6 does not exhibit CD29 but instead solely contains CD26 (Fig. 8b). Similar to the role of CD13 in tackling rhabdomyosarcoma (Sec. 6.1), inhibiting CD26 and CD29 may contribute to preventing the transformation of PIN to prostate adenocarcinoma [SGKH09]. CD26 and CD29 were already identified as lead proteins in [SGKH09]



(a) (b) (c)

Figure 8: *Necklace maps for visually exploring the toponome of a prostate tissue section. (a) Fluorescence image of protein affinity reagent CD138 with one acinus encircled. (b) A necklace map at epithelial cells of the acinus from (a) indicates CD26 and CD29 as lead protein candidates (yellow circles). CD29 is mouse hovered causing all symbols of CMPs containing CD29 to be highlighted (yellow border). (c) A focus region is defined below the acinus in the stroma. Note the strikingly different CMP pattern compared to (b) despite the overlap of contributing proteins (7 out of 11).*

and [OFH*11] however, by means of a more complex and time-consuming pipeline of analysis and interaction steps involving additional views.

A second necklace has been positioned over a part of the stroma (Fig. 8c). The corresponding CMP pattern is selective for the stroma and considerably differs from the one in the acinus (Fig. 8b). In Figure 8c, the acinus is located in the upper right corner. The protein necklace reveals again a high frequency of CD29 but also no mapping of CD26. Since the latter specifically recognizes prostate epithelium, this may be seen as a validation of our labeling algorithm. Furthermore, the necklace shows a mapping of CD4 and CD8 indicating the presence of T4 and T8 lymphocytes both participating in the cell-mediated immunity. This in turn, substantiates the presence of inflammatory cells.

6.3. User Feedback

We gathered anecdotal feedback from a biologist with a long-term, strong background in oncology and a computer scientist who has been working in his laboratory for many years. Both are co-authors of the paper. They used our necklace map approach and we simultaneously recorded their comments. They appreciated the in-place annotation of CMPs and proteins as a great cognitive relief since it avoids the tiresome shifting of attention back and forth between table and 2D view (Sec. 3). The comprehensive and sorted display of CMPs along the necklace obviates the search for the focused CMPs in the table. The display of the protein necklace and the interaction with it simplify the identification of present proteins, the detection of lead proteins, and the determination of cell types. Retrieving this information from the table view requires scrolling through the rows and examining each selected row for 1s (Fig. 2a).

The interaction with the necklace map was considered simple and effective. Merely, the temporary blinking of CMP pixels after clicking a symbol causes distraction and should be replaced by a less discomposing highlighting technique. The necklace management view was considered useful. It was heavily used for hiding and showing individual necklace maps. In contrast, the scalability of the necklace widgets was barely utilized due to a common focus of the CMP analysis on the most frequent ones, which were always visible.

7. Summary and Discussion

We have presented an approach to interactively label toponome data in 2D views thereby supporting biologists in visually exploring the data. The approach may be readily transferred to other image data exhibiting a very high local entropy, phenotypically identical structures forming multiple disjoint regions, and very small structures.

We have combined the dynamic excentric labeling of a focus region [FP99] with the static leader-free labeling of necklace maps [SV10]. The user may place a single or multiple focus regions in the image view causing the contained protein patterns to be displayed as symbols strung on a necklace surrounding each focus region. On demand a second necklace illustrating the proteins present in the focused patterns can be displayed. A focus region may be dragged and adjusted causing an update of the necklace(s) at interactive frame rates. For the use cases in Section 6 and larger test images (1600×1200 pixels), no restricted interactivity even for unreasonably large focus regions was observed.

A necklace management view has been implemented for organizing multiple necklaces and structuring the exploration. While necklaces may overlap in the toponome map, the management view arranges them in a non-overlapping

fashion subserving a comparison of the represented toponomes. We have demonstrated our approach for the visual exploration of a rhabdomyosarcoma cell line and a prostate tissue section. We plan to integrate the approach into volume rendered views of 3D toponome data [OKH*12].

Acknowledgements

Technological and biological toponome studies were supported by the Klaus Tschira foundation (project toponome atlas), the BMBF grants Biochance, CELLECT, NBL3, NGFN2, NGFNplus, and through DFGschu627/10-1, and the Innovationskolleg INK15.

References

[AHS05] ALI K., HARTMANN K., STROTHOTTE T.: Label Layout for Interactive 3D Illustrations. *Journal of WSCG 13*, 1 (2005), 1–8.

[BDS10] BARYSENKA A., DRESS A. W. M., SCHUBERT W.: An Information Theoretic Thresholding Method for Detecting Protein Colocalizations in Stacks of Fluorescence Images. *J Biotechnol 149*, 3 (2010), 127–131.

[BKSW07] BEKOS M. A., KAUFMANN M., SYMVONIS A., WOLFF A.: Boundary Labeling: Models and Efficient Algorithms for Rectangular Maps. *Comp Geom-Theor Appl 36*, 3 (2007), 215–236.

[FBKS07] FRIEDENBERGER M., BODE M., KRUSCHE A., SCHUBERT W.: Fluorescence Detection of Protein Clusters in Individual Cells and Tissue Sections by Using Toponome Imaging System: Sample Preparation and Measuring Procedures. *Nat Protoc 2*, 9 (2007), 2285–2294.

[FHS*12] FINK M., HAUNERT J.-H., SCHULZ A., SPOERHASE J., WOLFF A.: Algorithms for Labeling Focus Regions. *IEEE Trans. Vis. Comput. Graphics 18*, 12 (2012), 2583–2592.

[Fla71] FLANNERY J. J.: The Relative Effectiveness of Some Common Graduated Point Symbols in the Presentation of Quantitative Data. *Cartographica 8*, 2 (1971), 96–109.

[FP99] FEKETE J.-D., PLAISANT C.: Excentric Labeling: Dynamic Neighborhood Labeling for Data Visualization. In *SIGCHI Conference on Human Factors in Computing Systems* (1999), pp. 512–519.

[GLP14] GLASSER S., LAWONN K., PREIM B.: Visualization of 3D Cluster Results for Medical Tomographic Image Data. In *Conference on Computer Graphics Theory and Applications (GRAPP)* (2014), pp. 169–176.

[HJC*13] HINSON A. R., JONES R., CROSE L. E., BELYEA B., BARR F. G., LINARDIC C. M.: Human Rhabdomyosarcoma Cell Lines for Rhabdomyosarcoma Research: Utility and Pitfalls. *Frontiers in Oncology 3*, 183 (2013), eCollection.

[JNH*13] JIANG Z., NIMURA Y., HAYASHI Y., KITASAKA T., MISAWA K., FUJIWARA M., KAJITA Y., WAKABAYASHI T., MORI K.: Anatomical annotation on vascular structure in volume rendered images. *Comput. Med. Imag. Grap. 37*, 2 (2013), 131–141.

[Lin10] LIN C.-C.: Crossing-Free Many-to-One Boundary Labeling With Hyperleaders. In *Pacific Visualization Symposium (PacificVis)* (2010), pp. 185–192.

[LSC08] LUBOSCHIK M., SCHUMANN H., CORDS H.: Particle-Based Labeling: Fast Point-Feature Labeling Without Obscuring Other Visual Features. *IEEE Trans. Vis. Comput. Graphics 14*, 6 (2008), 1237–1244.

[MHST00] MORI K., HASEGAWA J., SUENAGA Y., TORIWAKI J.: Automated Anatomical Labeling of the Bronchial Branch and its Application to the Virtual Bronchoscopy System. *IEEE Trans. Med. Imag. 19*, 2 (2000), 103–114.

[MP09] MÜHLER K., PREIM B.: Automatic Textual Annotation for Surgical Planning. In *Vision, Modeling, and Visualization (VMV)* (2009), pp. 277–284.

[MTSP12] MOGALLE K., TIETJEN C., SOZA G., PREIM B.: Constrained Labeling of 2D Slice Data for Reading Images in Radiology. In *Eurographics Workshop on Visual Computing for Biology and Medicine (VCBM)* (2012), pp. 131–138.

[OFH*11] OELTZE S., FREILER W., HILLERT R., DOLEISCH H., PREIM B., SCHUBERT W.: Interactive, Graph-Based Visual Analysis of High-Dimensional, Multi-Parameter Fluorescence Microscopy Data in Toponomics. *IEEE Trans. Vis. Comput. Graphics 17*, 12 (2011), 1882–1891.

[OJP14] OELTZE-JAFRA S., PREIM B.: Survey of Labeling Techniques in Medical Visualizations. In *Eurographics Workshop on Visual Computing for Biology and Medicine (VCBM)* (2014), p. this volume.

[OKH*12] OELTZE S., KLEMM P., HILLERT R., PREIM B., SCHUBERT W.: Visualization and Exploration of 3D Toponome Data. In *Eurographics Workshop on Visual Computing for Biology and Medicine (VCBM)* (2012), pp. 115–122.

[PRS97] PREIM B., RAAB A., STROTHOTTE T.: Coherent zooming of illustrations with 3d-graphics and text. In *Graphics Interface* (1997), pp. 105–113.

[RPRH07] ROPINSKI T., PRASSNI J.-S., ROTERS J., HINRICHS K.: Internal Labels as Shape Cues for Medical Illustration. In *Vision, Modeling, and Visualization (VMV)* (2007), pp. 203–212.

[SBP*06] SCHUBERT W., BONNEKOH B., POMMER A. J., PHILIPSEN L., BÖCKELMANN R., MALYKH Y., GOLLNICK H., FRIEDENBERGER M., BODE M., DRESS A. W. M.: Analyzing Proteome Topology and Function by Automated Multidimensional Fluorescence Microscopy. *Nat Biotechnol 24*, 10 (2006), 1270–1278.

[Sch03] SCHUBERT W.: Topological proteomics, toponomics, MELK-technology. *Adv Biochem Eng Biotechnol 83* (2003), 189–209.

[Sch10] SCHUBERT W.: On the origin of cell functions encoded in the toponome. *J Biotechnol 149*, 4 (2010), 252–259.

[SGK*12] SCHUBERT W., GIESELER A., KRUSCHE A., SEROCKA P., HILLERT R.: Next-generation biomarkers based on 100-parameter functional super-resolution microscopy TIS. *New Biotechnology 29*, 5 (2012), 599–610.

[SGKH09] SCHUBERT W., GIESELER A., KRUSCHE A., HILLERT R.: Toponome mapping in prostate cancer: detection of 2000 cell surface protein clusters in a single tissue section and cell type specific annotation by using a three symbol code. *J Proteome Res 8*, 6 (2009), 2696–2707.

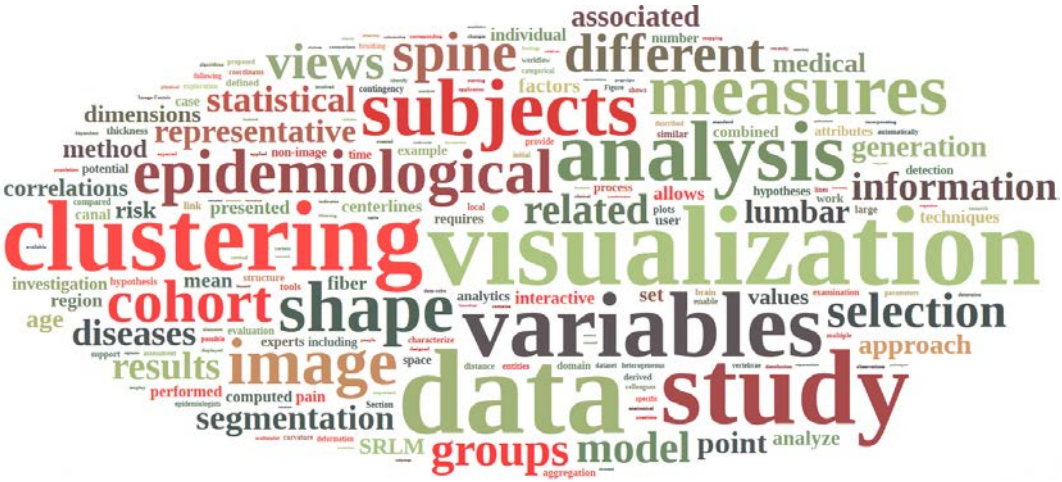
[SV10] SPECKMANN B., VERBEEK K.: Necklace Maps. *IEEE Trans. Vis. Comput. Graphics 16*, 6 (2010), 881–889.

[TGK*14] TOMINSKI C., GLADISCH S., KISTER U., DACHSELT R., SCHUMANN H.: A Survey on Interactive Lenses in Visualization. In *EuroVis State-of-the-Art Reports* (2014), pp. 43–62.

[Tuf01] TUFTE E. R.: *The Visual Display of Quantitative Information*, 2nd ed. Graphics Press, 2001.

Part III

Epidemiological Population Study Data



This part of the postdoctoral thesis cumulates the following publications:

Chapter 7 [181] B. Preim, P. Klemm, H. Hauser, K. Hegenscheid, **S. Oeltze**, K. Toennies, and H. Völzke, *Visual Analytics of Image-Centric Cohort Studies in Epidemiology*, vol. 3, ch. Visualization in Medicine in Life Sciences, in print. Springer, 2016.

Chapter 8 [112] P. Klemm, K. Lawonn, M. Rak, B. Preim, K. D. Tönnies, K. Hegenscheid, H. Völzke, and **S. Oeltze**, "Visualization and Analysis of Lumbar Spine Canal Variability in Cohort Study Data," in *Vision, Modeling, and Visualization (VMV)* (M. Bronstein, J. Favre, and K. Hormann, eds.), pp. 121-128, 2013.

Chapter 9 [113] P. Klemm, **S. Oeltze-Jafra**, K. Lawonn, K. Hegenscheid, H. Völzke, and B. Preim, "Interactive Visual Analysis of Image-Centric Cohort Study Data," *IEEE Trans. Vis. Comput. Graphics (TVCG)*, vol. 20, no. 12, pp. 1673-1682, 2014.

Chapter 10 [4] P. Angelelli, **S. Oeltze**, C. Turkay, J. Haász, E. Hodneland, A. Lundervold, A. J. Lundervold, B. Preim, and H. Hauser, "Interactive Visual Analysis of Heterogeneous Cohort Study Data," *IEEE Comput. Graph. Appl. Mag. (CG&A)*, vol. 34, no. 5, pp. 70-82, 2014.

Chapter 11 [170] **S. Oeltze**, H. Schütze, A. Maaß, E. Düzel, and B. Preim, "Measurement of the Stratum Radiatum/Lacunosum-Moleculare (SRLM)," in Bildverarbeitung für die Medizin (BVM), pp. 264-269, 2014.

Visual Analytics of Image-Centric Cohort Studies in Epidemiology

Bernhard Preim, Paul Klemm, Helwig Hauser, Katrin Hegenscheid, Steffen Oeltze, Klaus Toennies, and Henry Völzke

Abstract Epidemiology characterizes the influence of causes to disease and health conditions of defined populations. Cohort studies are population-based studies involving usually large numbers of randomly selected individuals and comprising numerous attributes, ranging from self-reported interview data to results from various medical examinations, e.g., blood and urine samples. Since recently, medical imaging has been used as an additional instrument to assess risk factors and potential prognostic information. In this chapter, we discuss such studies and how the evaluation may benefit from visual analytics. Cluster analysis to define groups, reliable image analysis of organs in medical imaging data and shape space exploration to characterize anatomical shapes are among the visual analytics tools that may enable epidemiologists to fully exploit the potential of their huge and complex data. To gain acceptance, visual analytics tools need to complement more classical epidemiologic tools, primarily hypothesis-driven statistical analysis.

1 Introduction

Epidemiology is a scientific discipline that provides reliable knowledge for clinical medicine focusing on prevention, diagnosis and treatment of diseases [14]. Research in epidemiology aims at characterizing *risk factors* for the outbreak of diseases and at evaluating the efficiency of certain treatment strategies, e.g., to compare a new treatment with an established gold standard. This research is strongly hypothesis-driven and statistical analysis is the major tool for epidemiologists so far. Correlations between genetic factors, environmental factors, life style-related parameters, age and diseases are analyzed. The data are acquired by a mixture of interviews (self-reported data, e.g., about nutrition and previous infections) and clinical ex-

Bernhard Preim
Otto-von-Guericke University Magdeburg, 39106 Magdeburg, e-mail: bernhard.preim@ovgu.de

aminations, such as measurement of blood pressure. Statistical correlations, even if they are strong, may be misleading because they do not represent *causal relations*. As an example, the slightly reduced risk of heart infarct and cardiac mortality for elderly people reporting to drink one glass of wine every evening (compared to people drinking no alcohol at all) may be due to the involved low level of alcohol but may also be a consequence of a very regular and stress-free lifestyle [14]. When something happened, *before* an event, it is an indicator for a causal relationship. However, care is necessary, since many things happen in the life of an individual before, e.g., a heart attack, but do *not* cause it.

Thus, statistical correlations are the starting point for investigating *why* certain factors increase the risk of getting diseases. Epidemiology is not a purely academic endeavor but has huge consequences for establishing and evaluating preventive measures even outside of medicine. The protection of people from passive smoking, recommendations for various vaccinations and the introduction of early cancer detection strategies, e.g., mammography screening, are all based on large-scale epidemiological studies. Also the official guidelines for the treatment of widespread diseases, such as diabetes, are based on *evidence* from epidemiological studies [14]. While this all may sound obvious, it is a rather recent development. *Evidence-based* medicine often still has to "fight" against recommendations of a few opinion leaders arguing based on their personal experience only.

The analysis techniques used so far are limited to investigating hypotheses based on known or suspected relations, e.g. hypotheses related to observations or previous publications. The available tools support the analysis of a few dimensions, but not of the hundreds of attributes acquired per individual in a cohort study. Both typical visualization techniques as well as analysis techniques, e.g., support vector machines, do not scale well for hundreds of attributes [41]. While we are not able to describe *solutions* for these challenging problems, we give a survey on recent approaches aiming also at *hypothesis generation*.

Organization. This chapter is organized as follows. In Sect. 2 we describe important concepts and terms of epidemiology including observations from epidemiologic workflows. This discussion is restricted to those terms that are crucial for communicating with epidemiologists, understanding requirements and for designing solutions that fit in their process. In Sect. 3, we discuss how (general) information visualization and data analysis techniques may be used for epidemiologic data. Section 4 describes the analysis of image data from cohort studies and how this analysis is combined with the exploration of non-image attribute data. This section represents the core of the chapter and employs a case study where MRI data of the lumbar spine are analyzed along with attributes characterizing life-style, working habits, and back pain history.

2 Background in Epidemiology

Population-based studies. Epidemiological studies are based on a *sample* of the population. The reliability of the results obviously depends on the size of that sample but also strongly on the selection criteria. Often, data from patients treated in one hospital are analyzed. While this may be a large number of patients, the selection may be heavily biased, e.g., since the hospital is highly specialized and diseases are often more severe or in a later stage compared to the general population.

Population-based studies, where representative portions of a population (without known diseases) are examined, have the potential to yield highly reliable results. The source population may be from a city, a region or a country. Individuals are randomly selected, e.g., approaching data bases of population registries. The higher the percentage of people who accept the invitation and actually take part in the study, the more reliable the results are.

In this chapter, we focus on longitudinal population-based studies. The sheer amount and diversity in terms of type of data makes it difficult to fully identify and analyze interesting relations. We will show that information visualization and visual analytics techniques may provide substantial support that complements the statistical tools with their rather simple statistical graphics. Most epidemiological studies were restricted to nominal (often called categorical) and scalar data, e.g., related to alcohol consumption, and body mass index as one measure of obesity.

Image-centric epidemiological studies. More recently, for example, in the Rotterdam study [22], also non-invasive imaging data, primarily ultrasound and MRI data, are employed. Petersen and colleagues [32] report on six studies involving cardiac MRI from at least 1000 individuals in population-based studies. These high-dimensional data enable to answer analysis questions, e.g., how does the shape of the spine changes as a consequence of age, life style and diseases? We focus on such *image-centric* epidemiological studies.

Epidemiology and public health. There are different branches of epidemiology. One branch deals with predictions to inform public health activities. These include measures in case of an epidemic – an acute *public health* problem, mostly related to infectious diseases. The recent article ”computational epidemiology” [29] was focussed on this branch of epidemiology. Another branch of epidemiology aims at long-term studies and at findings primarily essential for prevention. Image-centric cohort studies, the focus of this article, belong to this second branch. The target user group consists of epidemiologists who can be expected to have a high level of expertise in statistics. Thus, their findings involve statistical significance, confidence intervals and other measures of statistical power.

Healthy aging and pathologic changes. An essential problem in the daily clinical routine is the discrimination between healthy age-related modifications (that may not be reversed by treatment) and early stage diseases (that may benefit from immediate treatment). As a consequence, elderly people are often not adequately treated.

As a general goal for epidemiological studies, better and more reliable markers for early stage diseases are searched for. The cardiovascular branch of the Rotterdam study, for example, aims at an understanding of atherosclerosis, coronary heart disease and "cardiovascular conditions at older age" [22].

Modern epidemiology. Epidemiology faces new challenges due to the rapid progress, e.g., in genetics and sequencing technology as well as medical imaging. Acquisition of health data thus becomes cheaper and more precise. In cohort studies, as much potentially relevant data as possible are acquired as a basis for an as broad as possible spectrum of analysis questions. This includes blood, urine and tissue samples, information about environmental conditions and the social milieu.

Visual analytics for modern epidemiology. In the past, epidemiology primarily dealt with hypotheses aiming to prove them, e.g., the efficiency of early cancer detection programs in terms of mortality and long term survival [14]. Since recently, more and more data mining is performed to identify correlations. Results of such analyses, however, need to be very carefully interpreted. If thousands of potential correlations are analyzed automatically, just by chance some of them will reach a high level of statistical significance.

An essential support for epidemiology research is to define relevant subgroups. To perform separate analyses for women and men as well as for different age groups is a common practice in epidemiology. However, relevant subgroups may be defined by a non-obvious combination of several attributes that may be detected by a combination of cluster analysis and appropriate visualization.

Since the information space is growing with each examination cycle, Pearce and Merletti [31] pointed out in 2006 that methods are needed which can cope with this complexity and enable the analysis of underlying causes of a certain disease. Visual analytics (VA) methods can support epidemiological data assessment in different ways, e.g. by defining subgroups based on a multitude of attributes that exhibit a certain characteristic. For the analysis of scalar and categorical data, established information visualization techniques combined with clustering and dimension reduction are a good starting point, but need to be tightly integrated with statistic tools epidemiologists that are more familiar with. For image-centric studies, however, new visualization, (image) analysis and interaction techniques are needed.

In the following, we define essential terms in epidemiology and give an overview on cohort studies that employ medical image data as an essential element. Finally, we describe how image data, derived information and other data complement each other to identify and characterize risks.

2.1 Important Terms

Prevalence and incidence. Epidemiology investigates how often certain diseases or *clinical events*, such as a cerebral stroke or sudden heart death, occur in the pop-

ulation. Two terms are important to characterize this frequency. The *prevalence* indicates the portion of people suffering from a disease at a given point in time. The *incidence* represents how many people suffer from a disease or event in a certain interval, usually one year. High prevalence is usually associated with high economic costs. Population-based studies focus on diseases with a high prevalence, such as diabetes, coronary heart disease or neurodegenerative diseases. Even these diseases do not occur frequently in a random population including many younger people (where the prevalence of these diseases is low). A rare disease, such as amyotrophic lateral sclerosis, may have a prevalence of 5 from 100,000. Thus, even in a large population-based study probably no individual suffers from this disease.

Absolute and relative risks. Another essential epidemiological term is the *risk* for a clinical event, such as outbreak of a certain disease, severity (stage) or death. As an example, a study related to cardiac risk may investigate angina pectoris, myocardial infarction, atrial fibrillation depending on attributes such as age and sex. The *absolute risk* characterizes the likelihood of getting a disease in life time. The absolute risk for a woman to develop breast cancer in the Western world is particularly high for women aged 50-60 (2.6%) and 60-70 (3.7%). Therefore, for these age groups, mammography screening – aiming at early detection and thus optimal treatment – was introduced.

The *relative risk* (RR) characterizes the increased risk if an individual is exposed to a certain risk factor, e.g., smoking, excessive weight, or alcohol abuse. It is based on a comparison with a control group not exposed to that risk factor. A value of $RR < 1$ represents a factor that protects, e.g., moderate physical activity. Exciting observations are often the combined effects of several parameters. A certain factor may be protective for some people (younger, slim women) and is involved with an increased risk for others. The combined risk may be significantly smaller or larger than could be expected from individual factors.

Moreover, relationships are often distinctly non-linear or even non-monotonic. Dose-response relationships are often non-linear. RR increases slowly (almost no effect for a small dose) and increases much faster for higher levels of a dose, e.g., exposure to toxicity. A typical non-monotonic relation is *U-shaped*, that is both very low and very high instances of an attribute involve an increased risk, whereas values in between are associated with a reduced risk. Examples are weight (both very low and very high weight are associated with an increased risk for mortality) and sleeping time (both very short and very long sleepers have an increased risk for developing psychiatric disorders [22]). Such relations cannot be characterized by a global RR value. Instead, tools are necessary that support the hypothesis of a U-shaped relation by estimating their parameters with some kind of best-fit algorithm.

2.2 Image-Centric Cohort Studies

Image data in epidemiology. The acquisition of image data is determined by the available time, by financial resources, by the epidemiological importance and by ethic considerations. Epidemiological studies require approval by a local ethics committee. As a consequence, healthy individuals in a cohort study should not be exhibited to a risk associated to the examinations carried out. Thus, MRI should be preferred over X-ray or CT imaging for its non-radiation nature. Petersen and colleagues [32] explain why cardiac CT is less feasible in a cohort study and even MR is only used without a contrast agent in their study due to ethical reasons. MRI data and ultrasound data are the prevailing modalities in both the SHIP as well as the Rotterdam study. Unfortunately, MRI and ultrasound data do not exhibit standardized intensity values (in contrast to CT data). Moreover, MRI and ultrasound data suffer from inhomogeneities and various artifacts. Thus, they are more difficult to interpret for humans and more difficult to analyze with computational means. These data are used to measure, e.g., the thickness of vessel walls, the abdominal aorta diameter and plaque vulnerability in the coronary vessels [22]. The intensive use of MRI in epidemiological research also explains to some extent which questions are analyzed: MRI is the best modality for the analysis of brain structures and thus serves to explore early signs of Parkinson's, Alzheimer's and other neurodegenerative diseases. Epidemiological research aims at identifying such brain pathologies in a pre-symptomatic stage. Among the sources for such investigations are MR Diffusion Tensor Imaging data that enable an assessment of white matter integrity [22].

The selection of imaging parameters is always a trade-off between conflicting goals related to quality, e.g., image resolution, signal-to-noise ratio, patient comfort, e.g., examination time and associated costs. As a consequence, to shorten overall examination times in cohort study examinations, not the highest possible quality is available, i.e., a slice distance of 4 mm is more typical than 1 mm. A great advantage of MRI is that this method is very flexible and enables to display different structures in different sequences, such as T1-, T2- and proton density-weighted imaging. MRI data in cohort studies often comprise more than ten different sequences.

Standardization in image acquisition. Due to the rapid progress in medical imaging, sequences, protocols and even (MR) scanners are frequently updated in clinical routine (similar to the update frequency on a computer). These updates would severely hamper the comparison of imaging results and thus the assessment of natural changes and disease outbreak. Thus, differences in acquisition parameters are essential *confounding variables*. Therefore, for one cohort and examination cycle that may last up to several years, no updates are allowed. Moreover, all involved physicians and radiology technicians are carefully instructed to use the same standardized imaging parameters. This point is even more important for longitudinal studies with repeated imaging examinations. Even if MR scanners and protocols are not updated, the life cycle of MR coils leads to changes of image quality that need to be monitored and compensated.

2.3 Examples for Image-Centric Cohort Study Data

In the following, we describe selected comprehensive and on-going longitudinal cohort studies. Both use a number of (epidemiologic) *instruments* that are innovative in cohort studies and thus lead already to a large number of insights documented in hundreds of (medical) publications. A considerable portion of these publications employ results from imaging data. However, the full potential of analyzing organ shapes, textures and spatial relations quantitatively is not exploited so far.

The Rotterdam study. A prominent example is the *Rotterdam* Study[1], initiated in 1990 in the city of Rotterdam, in the Netherlands. Similar to later studies, it was motivated by the demographic change with more and more elderly people suffering from different diseases and their interactions. After the initial study involving almost 8,000 men and women, follow-ups at four points in time were performed—the most recent examinations took place in the 2009-2011 period. In the later examination cycles, also new individuals were involved leading to datasets from almost 15,000 patients [22].

The original focus of the Rotterdam Study was on neurological diseases, but meanwhile it has been extended to other common diseases including cardiovascular and metabolic diseases. The study has an enormous impact on epidemiological and related medical research, documented in 797 journal publications registered in the pubmed database (search with keyword "Rotterdam Study", January 30, 2014). Among them are predictions for the future prevalence of heart diseases and many studies on potential risk factors for neurodegenerative diseases. For a comprehensive overview of the findings, see [22] that summarizes the findings of more than 240 papers related to the Rotterdam Study. In a similar way, [23] is a significant update of these findings with more recent data.

Norwegian Aging Study. A long-term study in Norway investigates the relations between brain anatomy (as well as brain function), cognitive function, and genetics in normally aging people.[2] In total 170 individuals (120 of them female), aged between 46 and 77 (mean 62), were examined in Bergen and Oslo in by now three waves (1st wave in 2004/2005, next in 2008/2009, and most recently in 2011/2012) [48]. While naturally not all of these subjects could be followed through all three waves, still most of them were subjected to an extensive combination of

1. neuropsychological tests, including tests of the intellectual, language (memory), sensory/motor, and attention/executive function,
2. MRI data, including co-registered T1-weighted anatomical imaging, diffusion tensor imaging, and – from the 2nd wave on – also resting-state functional MRI, as well as
3. genotyping (1st wave only) [46].

[1] `http://www.erasmus-epidemiology.nl/research/ergo.htm`, accessed: 1/31/2014

[2] `http://org.UiB.no/aldringsprosjektet/`, accessed: 1/31/2014

The substantially heterogeneous imaging and test data are used to study aging-related questions about the modern Norwegian population, for example, how anatomical and functional changes in the human brain possibly relate to the later development of Alzheimer and dementia. Important findings include the relation between hippocampal volumes and memory function in elderly women [48] and the relation between subcortical functional connectivity and verbal episodic memory function in healthy elderly [47].

SHIP. The Study of Health in Pommerania (SHIP) is another cohort study broadly investigating findings and their potential prognostic value for a wide range of diseases. The SHIP tries to explain health-related differences after the German reunion between East and West Germany. It was initiated in the extreme northeast of Germany, a region with high unemployment and a relatively low life expectancy.

In the first examination cycle (1997-2001) 4,308 adults of all age groups were examined, followed by a second and a third cycle that was finished at the end of 2012. The instruments used changed over time with some initial image data (liver and gallbladder ultrasound) available already in the first cycle and others, in particular whole body MRI, added later. The use of whole body MRI was unique in 2008 when the third examination cycle started. Breast MRI for women is performed, whereas for men MR angiography data are acquired, since men suffer from cardiovascular diseases significantly earlier than women [42]. In addition, a second cohort (SHIP-Trend) was established comprising 4,420 adult participants.

Diagnostic reports are created by two independent radiologists who follow strict guidelines to report their findings in a standardized manner. The pilot study to discuss the viability and potential of such a comprehensive MR exam is described by [20]. The overall time for the investigation is two (complete) days with 90 minutes for the MR exam. The SHIP helped to reliably determine the prevalence of risk factors, such as obesity, and diseases. Major findings of the SHIP are increased levels of obesity and high blood pressure (compared to the German population) in the cohort. The MR exams alone identified pathological findings in 35% of the sample population. More than 400 publications in peer-reviewed journals are based on SHIP data (January 2014).

UK Biobank. The UK Biobank started recently and represents a comprehensive approach to study diseases with a high prevalence in an aging society, such as hearing loss, diabetes and lung diseases. Half a million individuals will be investigated in one examination cycle from which 100,000 receive an MRI from 2014 onwards. The rationale for the number of individuals to be included is explained by Peterson and colleagues [32]: they aim at a reliable identification of even moderate risk factors (RR between 1,3 and 1,5) for diseases with a prevalence of 5%. The prospective study should have a comprehensive protocol of cardiac MRI, brain MRI and abdominal MRI. This prospective cohort study also involves genetic information.[3]

[3] `http://www.ukbiobank.ac.uk`, accessed: 1/31/2014

The German National Cohort. The recently started "German National Cohort" in Germany is based on experiences with a number of moderate-size studies, such as SHIP, and examines some 200,000 individuals over a period of 10-20 years. Individuals will be invited in three waves to characterize changes. Due to the large-scale character, imaging is distributed over five cities. Thus, the subtle differences in imaging within different scanners have to be considered. [4] It explicitly aims at improvements in the treatment of chronic diseases and involves a variety of tissue samples, e.g., lymphocytes. Imaging in 30,000 individuals is again performed with MRI, comprising whole body, brain and heart.

2.4 Epidemiological Data

Epidemiological data are huge and very heterogeneous. As an example, in the UK biobank 329 attributes relate to physical measures, such as pulse rate, systolic and diastolic blood pressure, and various measures relate to vision or hearing. 471 attributes relate to interviews (socio-demographics, health history, lifestyle, ...).

The data that are stored per individual is standardized but not completely the same, e.g., childbirth status and menstrual period are available for women only. Image data and derived information, e.g., segmentation results, significantly increased both the amount and complexity of data. Longitudinal cohort study data are time-dependent. While some *instruments*, such as blood pressure measurements, are available for all examination cycles, others were added later or removed. Individuals drop out, because they move, die or just do not accept the invitation to a second or third examination cycle. It is important to consider also such incomplete data but to be aware of potentially misleading conclusions.

The great potential of image-centric studies is that image data and associated laboratory data as well as data from interviews are available. An epidemiological study, such as the SHIP, has a large *data dictionary* that precisely defines all attributes and their ranges. While laboratory data are scalar values, most data from interviews are nominal or ordinal values. In particular, data from interviews exhibit an essential amount of uncertainty. Self-reports with respect to alcohol and drug use, cigarette smoking and sexual practices may be biased towards "expected or socially accepted" answers. Epidemiologists are not only aware of these problems but developed strategies to minimize the negative effects, e.g., by asking redundant questions. After data collection, experts spend a lot of effort to improve the quality of the data. Despite these efforts, visual analytics techniques have to consider outliers, missing and erroneous data.

Geographic data. Geographic data play a central role in public health where the dynamics of local infections are visualized and analyzed (*disease mapping*). Chui and colleagues [9] presented a visual analytics solution directly addressing this

[4] `http://www.nationale-kohorte.de/`, accessed: 1/31/2014

problem by combining three dedicated views. Also in cohort study data, geographic data are potentially interesting to understand local differences in the frequency and severity of diseases as an interaction between environmental factors and genetic differences. This branch of epidemiology is referred to as *spatial epidemiology*. Beale and colleagues [5] investigated differences between rural and urban populations. In their comprehensive survey, Jerrett and colleagues [24] considered spatial epidemiology as an emerging area. However, we do not focus on spatial epidemiology since cohort study data typically comprise rather narrow regions and thus may not fully support such analysis questions.

2.5 Analysis of Epidemiological Workflow

The following discussion of observations and requirements for computer support is largely based on discussions with epidemiologists as well as the inspiring publication by Thew and colleagues. According to [40]

- epidemiological hypotheses are mostly observations made by physicians in clinical routine,
- corresponding attributes are chosen based on the observations and further experience, and
- regression analysis is frequently used to determine whether the investigated attribute is a risk factor or not.

Major requirements for an epidemiological workflow (again based on [40]) are:

- Results have to be reproducible. Due to the iterative data assessment, methods need to be applied to new data sets as well and the results need to be comparable between different assessment times to characterize the change. User input needs to be monitored all the time to enable reproducible results.
- A major result of an epidemiological analysis is whether certain factors influence a disease significantly. Relative risk (as a measure of effect size) and p-values as statistical significance level are particularly important.

Although these requirements neither consider image data nor visual analytics, they have to be considered also in these more innovative settings. Reproducibility, for example, means that clustering with random initialization is not feasible. Moreover, reports must be generated that clearly reveal all settings, e.g., parameters of clustering algorithms that were used for generating the results.

Since statistical analysis plays such an important role, statistics packages, such as SPSS[5], R[6] and STATA[7] dominate in epidemiology. They provide various statistical tests also in cases where assumptions, such as a normal distribution, are not valid.

[5] `http://www-01.ibm.com/software/analytics/spss/products/statistics/`, accessed: 1/31/2014

[6] `http://www.r-project.org/`, accessed: 1/31/2014

[7] `http://www.stata.com/`, accessed: 1/31/2014

Also the peculiarities of categorical data are considered. Visual analysis, so far, plays a minor role. As an example, Figures 1 and 2 illustrate two graphical representations frequently used in epidemiology: *Kaplan-Meier curves* and *interaction terms*. A Kaplan-Meier curve shows the survival of patients, often as a comparison between different treatment options.

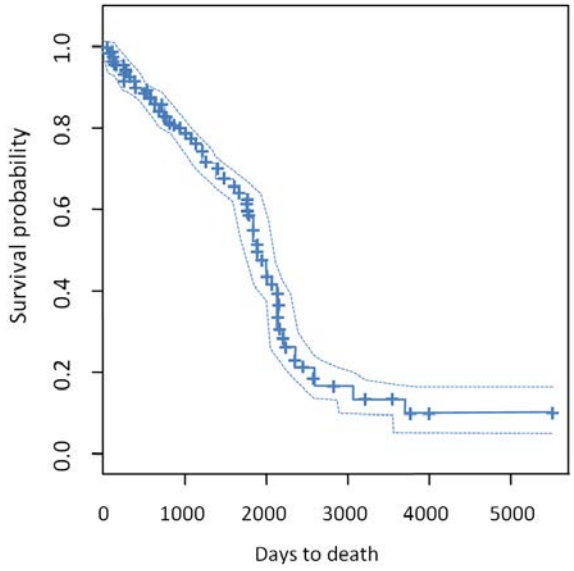


Fig. 1 A Kaplan-Meier curve indicates how many patients survive at least a certain time. The more patients pass away, the larger is the confidence interval indicated by the dotted lines. The crosses mark each time a patient dies to further provide information on the reliability of the data that decreases over time (Courtesy of Petra Specht, University of Magdeburg).

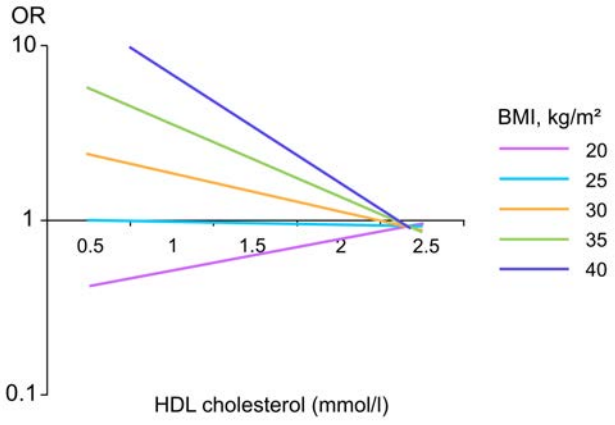


Fig. 2 The relative risk for cholelithiasis in men associated with a high level of a certain type of cholesterol slightly increases with a low BMI, but decreases for individuals with high or very high BMI. This multifactorial situation is depicted in an *interaction term* (Inspired by [43]).

3 Visual Analytics in Epidemiology

The visualization of correlations in the epidemiological routine is largely restricted to scatterplots with regression lines and box plots to convey a distribution. Scatter-

plots may be enhanced, e.g., by coloring items according to certain characteristics, e.g., a diagnosis or by adding results from cluster or Principal Component Analysis (PCA) [38]. The frequency related to a particular combination of values is often encoded by adapting saturation or darkness of colors.

Visual analytics methods can complement the statistical analysis and provide methods to *explore* the data. Efficient methods are essential to cope with the large amount of data and provide rapid feedback that is essential for any exploration process.

One of the first attempts to employ information visualization techniques for medical (image and non-image) data was realized in the WEAVE system [19]. The system incorporated parallel coordinates as well as real time synchronization between different views. Another essential tool, inspired by the WEAVE system, was presented by Blaas and colleagues [7]. They enabled feature derivation techniques and incorporated segmentation techniques providing a powerful framework for heterogeneous medical data. Later work by Steenwijk and colleagues was more focussed on epidemiology. They provided an exploratory approach to analyze heterogeneous epidemiological data sets, including MRI [38]. They consider parameters on normalized and not normalized domains, while only normalized domains are comparable between subjects. Normalization means, for example, to register MRI brain data to an atlas to compare individual differences.

Mappers are used to project data into normalized domains. As an example, in brain analysis, a mapper defines the relation between an individual brain and a brain atlas that contains normalized and averaged information derived from many individual data. Feature extraction pipelines can be build visually by using a pipeline of mappers. The visualization is realized through multiple coordinated views which either represent scalar data or volumetric images. Different techniques to color code data, to align them and add further information are provided to enhance scatterplots. Steenwijk and colleagues evaluated their tools with specific examples from neuroimaging and questions related to a neurological disease where relations between clinical data (anxiety-depresssion scales, mental state scales) and MR-related data are analyzed (Fig. 3). Normalized data domains are represented using scatterplots and parallel coordinated views. Dynamic changes are visualized using a time plot. The selection is linked between views and allows for multi-parameter comparison of clusters.

Zhang and colleagues [49] build a web-based information visualization framework for epidemiological analysis through different views. They divide the analytics process in *batch analytics* and *on-demand analytics*. Batch analytics steps are performed automatically as a new subject is added to the data set and aim to create groups by means of a certain condition. On-demand analytics are performed by user requests. Subjects are visualized using treemaps, histograms, radial visualizations and list views. However, neither filtering and grouping nor the interaction between the views are explained.

Recently, Turkay and colleagues [41] described a framework to analyze the data of the Norwegian aging study, aiming particularly at *hypothesis generation*. For this purpose, they give an overview on the dimensions in their dataset that conveys statis-

tical properties, such as mean, standard deviation, skewness and kurtosis. The two latter measures characterize how asymmetric the data distribution is. Scatterplots display pairwise measures related to *all* dimensions. *Deviation plots*, a new technique, enables grouping and supports a comparison of measures for a subgroup to the whole cohort.

The possibilities of VA tools can be summarized as follows:

- Manual/automatic definition (brushing) of interesting parameters and ranges of values in attribute views,
- Linking of attribute views for identifying relations,
- Analysis across aggregation levels, parameters and subjects,
- Definition of groups either interactively by means of (complex) brushes, or semi-automatically by means of clustering, and
- Visual queries and direct feedback enable easy exploration

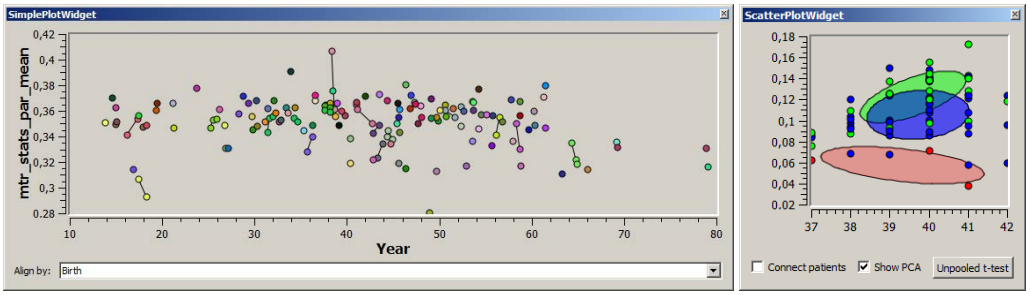


Fig. 3 Left: A scatterplot relates magnetic transfer ratios to age. Items relating to the same patient (over time) are connected via a line. Colors indicate a diagnosis. Right: An enhanced scatterplot with PCA results for three subgroups overlaid (Courtesy of Martijn Steenwijk, VU University Medical Center Amsterdam).

While having different applications in mind, the Polaris system of Stolte and colleagues also employs multiple coordinated views to validate hypotheses [39]. It uses a variety of different information visualization techniques to map ordinal/nominal or quantitative data of a relational data base. The system itself formulates data base queries and the mapping to create visualizations for the requested attributes. They choose the visualization mapping automatically based on the attribute types that are viewed in context with each other. This allows for a fast visualization of different attribute combinations in order to drill down to the information of interest. General visual analytics tools, such as Polaris and Weaver, in principle support some of the requirements for epidemiology. The use of coordinated views, brushes and switches is advantageous [44]. However, they are not designed to cope with the special requirements of cohort study data and do not directly support epidemiological workflows. In particular, no support for a combined assessment of image data and other epidemiological data is available.

Commercial Visual Analytics Systems. There are a number of commercial software tools specialized on data visualization. As an example, Tableau [28] provides an interface for creating different visualizations based on attribute drag-and-drop.[8] These approaches deliver fast results with respect to visualizing different attributes. However, it is not supported to derive new data, such as scores for diseases, body mass index and other data that is relevant in epidemiology. QLikView is a similar tool for visualizing data associations. The user can design a frontend where associations can be assessed using multiple information visualizations. Thus, the user can drill down to the desired information. Statistics features regarding epidemiological key figures are limited.

Spotfire/IVEE [2] is able to handle more complex analysis of data sets and allows for interactive filtering of attributes. It can be linked to the statistical computing programming language `R`, which makes it versatile in comparison with its competitors. However, users need to be familiar with the `R` syntax.

Commercial systems cannot be enhanced or embedded in another system with hassle-free data exchange. The focus of commercial data visualization tools is business intelligence yielding a focus on quantitative data sources. At the same time they excel at incorporating user collaboration by including comment sections and share filters or entire setups of a dashboard.

4 Analysis of Medical Image Data for Epidemiology

Medical images are not by themselves useful for epidemiological analysis, since the semantics of image elements (pixel, voxels) is too low. The resulting extremely high-dimensional feature space would be unsuitable for visual analysis. Hence, image data is sequentially aggregated and reduced. The different steps of this process i.e. image analysis, shape analysis of extracted objects and subsequent clustering are characterized and discussed in this section. Throughout the section, we often refer to one case study, where MRI data from the lumbar spine is analyzed. The representation of assumptions on the lumbar spine shape and location as well as the object detection scheme used are examples of viable and common approaches. We do not claim that these techniques are better than any other approaches.

4.1 Medical Image Analysis

One of the major purposes of image analysis for a cohort study is to quantify anatomy, e.g., by volume, shape or spatial relations between structures. Quantification may be used to establish a range of *normal values* for different age groups and to characterize variations. Such variations may also confirm a disease and thus

[8] `http://www.tableausoftware.com/`, accessed: 1/31/2014

add to data derived from clinical tests. Thus, the use of image data enables more reliable conclusions w.r.t. the incidence and prevalence of diseases. As discussed in Sect. 2.2, MRI is particularly interesting because of the wide range of different information represented in these data. Thus, the examples discussed in the following relate to MRI data although the principles are more general.

Due to the wide range of image analysis tasks, the techniques should be adaptable to different analysis goals. The parameterization of an image analysis technique should be intuitive and the interaction should be kept to a minimum. The latter aspect is particularly important, since often several thousand datasets need to be analyzed. While largely manual approaches are acceptable in some clinical settings, such as radiation treatment planning, they are not feasible in the evaluation of cohort study data. The reduction of interactivity is not only a matter of effort, but also to meet the essential goal of *reproducible approaches*.

Detection and segmentation of anatomical structures. A modular system is a possible means to meet the central requirement of image analysis in cohort study data. An example of a cohort study is the liver segmentation of [17], where concurrent detection and localization processes are combined for initial segmentation that is then fine-tuned in a model-driven segmentation step and finalized by a data-driven correction process. It has been shown that processes can be re-used and re-combined to solve a different segmentation task on similar data (kidney segmentation in MRI [18]). Alternatively, the necessary *domain knowledge*, related to expected size, basic shape, position and grey values, can be separated from the detection and segmentation module. This strategy is attractive, since the user has not to care about the detection process when changing the application. Two problems have to be addressed in this case:

- What is the expectation about the data support integrated to fit a model?
- How is the with-class variation of the object in search separated from the between-class variation?

Point distribution models (PDM) [30] address the second question by training on sample segmentations. Model fitting is realized by a registration step. When training is not feasible, a prototypical model may be used instead. It is associated with restricted input about variation (a few parameters only) and qualitative knowledge about configuration or part-relationship.

In the following, we describe the linear elastic deformation of a finite element model (FEM) as a common method to model shape variation. The user specifies the average shape and two elasticity parameters: Young's modulus defines how much external force is needed for a deformation and Poisson's ratio describes how the deformation is transferred orthogonal to the direction of an incident force [35]. The decomposition of the prototypical shape into finite elements bounded by nodes and specification of the elasticity parameters results in a stiffness matrix K that relates the node displacement u to incident forces f (Eq. 1):

$$Ku = f \tag{1}$$

Different kinds of nodes may be specified that are attracted by different kinds of forces. Boundary nodes are attracted by the intensity gradient and inner nodes are attracted from expected intensity or texture. For letting an FEM move and deform into an object in an image, deformation is made dependent on time t. Behavior then also depends on mass M of the FEM and object-specific damping D (Eq. 2).

$$M\ddot{u}(t) + D\dot{u}(t) + Ku(t) = f(t) \tag{2}$$

M represents the resistance of the moving FEM to external forces and allows the model to move over spurious image detail (e.g., gradients caused by noise). Damping D avoids oscillation of the FEM. The system of differential equations is decoupled by solving the following generalized eigenproblem.

$$KE = ME\Lambda \text{ with } E^T KE = \Lambda \text{ and } E^T ME = I$$

where Λ is the diagonal matrix with real-valued eigenvalues and I is the identity matrix.

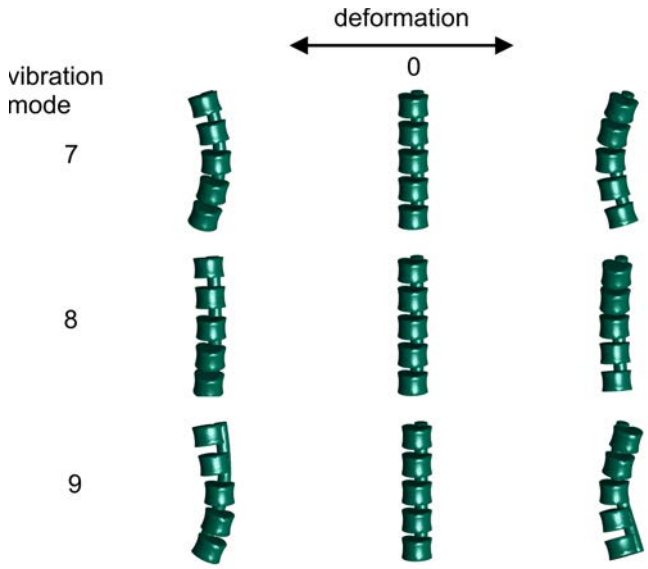


Fig. 4 Vibration modes 7 to 9 of a lower spine model. Vibration modes 1 to 6 represent rigid transformations (Courtesy of Marko Rak, University of Magdeburg).

After projecting the data on the eigenvector matrix, the differential equations can be solved fast and in a stable manner. Moreover, the projection on the eigenvectors (called *modes of vibration*, see Fig. 4) separates deformation into components representing rigid transformation, major deformation modes and remaining minor deformation modes. The vibration modes can be used similar to the variation modes of an ASM to derive a quality-of-fit formulation for a fitted model instance. Since only a few anatomical objects have such a specific shape that it can be described by

a simple deformation model and since training of additional information should be avoided, it is useful to complement simple deformation with pre-specified information on part-relationships. Part-relationships may describe the configuration of the object of interest w.r.t neighboring structures or may represent decomposition into parts (see [33]).

Extending the FEM to a hierarchical model requires the introduction of a second layer FEM. Each sub-shape is represented by an FEM on the first layer and sub-shape FEMs are connected to the second layer. The type of connection regulates dependencies between sub-shapes and may range from distance constraints to co-deformation. FEMs for the first and second layer are created and assembled in the same fashion than elements are assembled for the sub-shape FEM [35].

Case study: Analysis of vertebrae. Back pain and related diseases exhibit a high prevalence and are thus a focus of the SHIP (recall [42]). Specific goals are

- to define the prevalence of degenerative changes of the spine,
- to identify risk factors for these changes,
- to correlate degenerative changes with actual symptoms, and
- to better understand the progress from minor disease to a severe problem that requires medical treatment.

Epidemiologists hypothesize that smoking, heavy physical activity and a number of drugs that are frequently used are risk factors for back pain. Based on clinical observations, epidemiologists suggested to focus on the lumbar spine – the lowest part of the spine comprising five vertebrae. As a first step, the spine and lumbar vertebrae should be detected in T1-weighted and T2-weighted MRI data from SHIP.

Although local optimization could be complemented by stochastic global optimization [11], Rak and colleagues used only local optimization, since the initialization is simple for the given data. The user places a model instance in a sagittal view on the middle slice of the image sequence which is then transformed based on local image attributes. The model is constructed according to the appearance of vertebrae and spine in a sample image sequence. Vertebra sub-shapes were connected with a spine sub-shape by a structural model on the second level.

The spine model supported proper localization of the vertebrae. Since its most discriminate aspect was the cylindrical shape, it was represented by a deformable cylinder consisting of inner nodes only. The vertebrae shape was represented indirectly by inner nodes as well, since reliability of the intensity gradient was low. For each of the two shape models, the vertebra and the spine, a weighted combination of the T1-weighted and the T2-weighted image was computed as appearance input. Weights for each of the two models were determined a priori and produced a clearly recognizable local minimum for vertebra and spine appearance, respectively. The user placed a model instance in the vicinity of the object on a sagittal slice. Computation time until convergence was between 1.1 and 2.6 seconds per case. The method was evaluated on 49 data sets from the SHIP. The detection was considered successful if the center of each vertebra sub-shape was in the corresponding vertebra in the

image data, which was achieved in 48 of the 49 cases (see Fig. 5 for examples and [35] for further details).

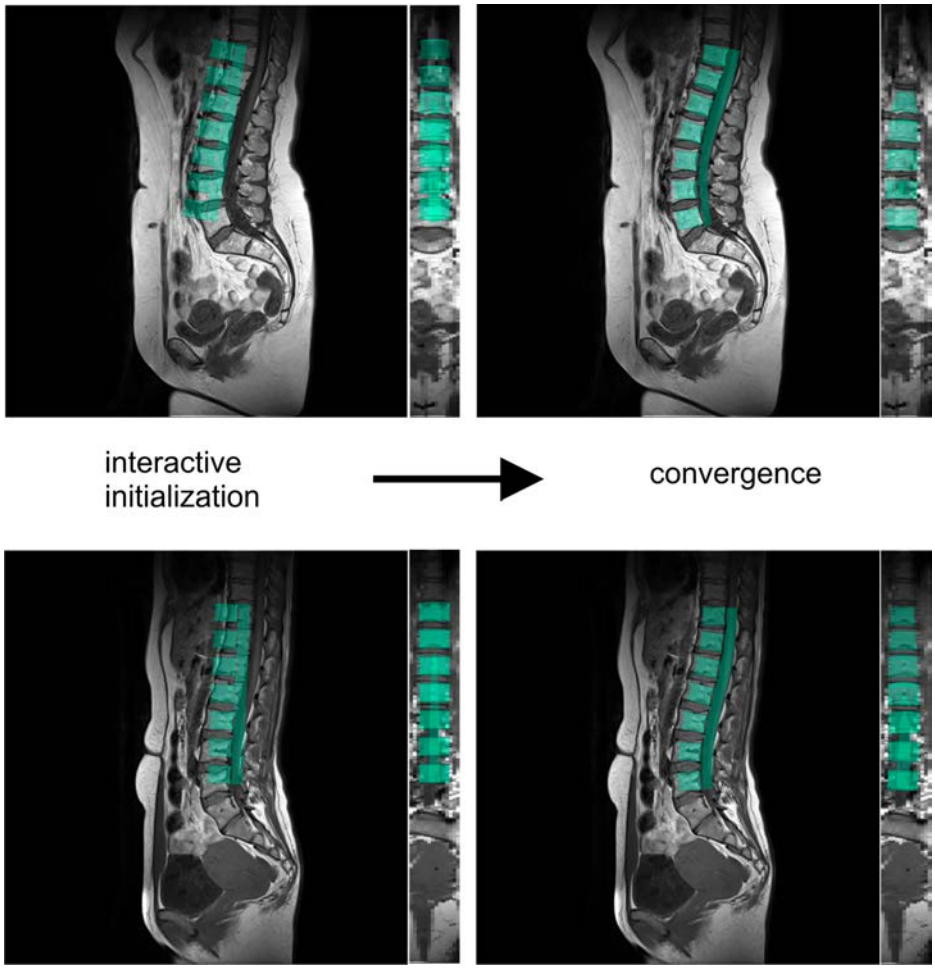


Fig. 5 Examples for initialization and convergence of model instances applied to the MRI data (Courtesy of Marko Rak, University of Magdeburg).

4.2 Shape Analysis

Epidemiologists are used to work with numerical and categorical data which then is tested for statistic validity. Medical image data also allows to consider characteristic object shapes. As an example, the shape of the liver may depend in a characteristic manner on infections (hepatitis), alcohol consumption, or obesity. Eventually, shape characteristics may change even before a disease becomes symptomatic. If this turns out, shape changes may be employed as an early stage indicator.

While the quantitative analysis of shapes or parts thereof (*morphometrics*) is a recent trend in epidemiology due to the availability of image data, it is established in anatomy and evolutionary biology.

Shape analysis requires that different shapes are transformed in a common space, typically by a rigid transformation (translation and rotation). The parameters of this rigid transformation are determined in an optimization process that minimizes the distances of corresponding points. A major challenge is to determine these corresponding points that serve as landmarks. In particular for soft tissue structures there are not sufficient recognizable landmarks and therefore a parameterization is necessary to define these points. Without going into detail, we assume that this process is applied to many individual shapes S_i, say livers in a cohort study. Then, for each S_i an optimal non-rigid transformation to a reference shape R defines a deformation with displacement vectors for each landmark.

For use in epidemiology, a large set of displacement vectors is not the right level of granularity. Instead, a few dimensions are desirable that characterize major differences. Thus, typically a dimension reduction technique, such as PCA, is employed to characterize the directions that represent the major differences. This process may be adapted to specific analysis questions by assigning individual weights to the landmarks expressing a strong interest in particular displacements [21]. Thus, epidemiological hypotheses may be incorporated.

While the establishment of point correspondences is often a major challenge, recently alternative approaches were developed. The GAMES algorithm (Growing and adaptive meshes) [13] creates a data structure to represent the shape variance if no pairwise correspondence between points is given. However, it can be prone to errors since it requires a prior registration of segmentation masks.

Shape analysis, of course, may also be supported by appropriate visualizations that enable pairwise comparisons and emphasize differences. In this vein, [8] presented a system for shape space exploration based on carefully designed multiple coordinated views.

4.3 Analysis of Lumbar Spine Canal Variability

In Sect. 4.1, we introduced a case study related to the analysis of the lumbar spine in cohort data and explained how the spine and the vertebrae are detected in MRI data from the SHIP. Here, we extend this discussion by the analysis of the spinal canal and non-imaging attributes related to back pain. In the SHIP, attributes related to back pain history, e.g., working habits, physical activities, size and weight, are available to identify and analyze potential correlations with findings from the MRI data. After careful discussions, we selected 77 attributes (60 are ordinal or nominal and 17 scalar) to investigate back pain [25]. Ordinal data are primarily results of multiple choice questions. The epidemiologists suggested to focus on the overall shape and curvature of the spine in that region instead of individual vertebrae. This overall shape is well characterized by the lumbar canal. Thus, correlations between

the shape of the lumbar spine, attributes of back pain history and activities both in leisure and working time may be analyzed.

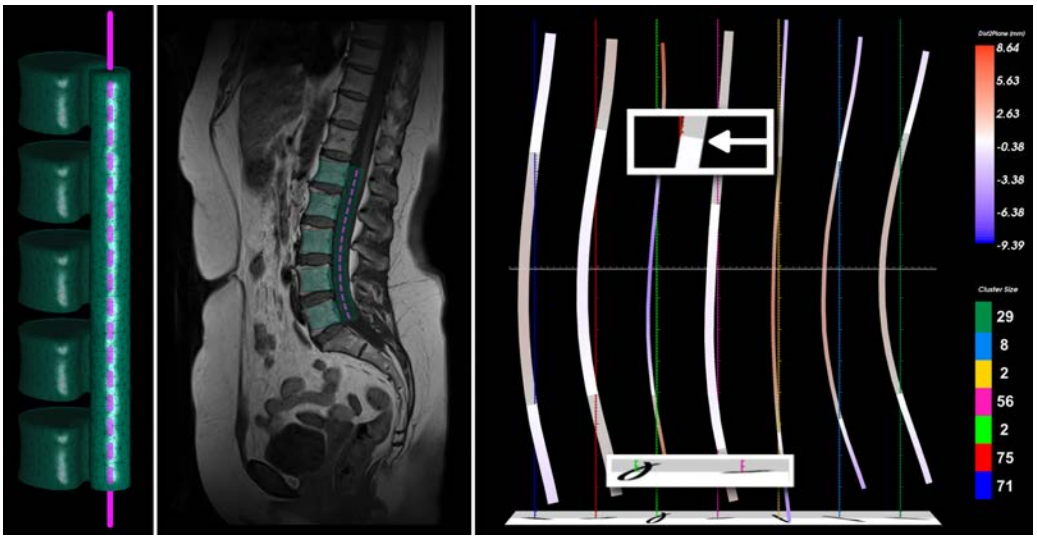


Fig. 6 Lumbar spine visualization of 243 female subjects. *Left:* Tetrahedron-based finite element model from Rak and colleagues [35]. The dashed purple line indicates the lumbar spine canal centerline *Middle:* Model used to detect lumbar spine canal in an MRI scan *Right:* Agglomerative hierarchical clustering of 243 centerlines yields seven clusters. Their representatives are visualized as ribbons mapping cluster size to width. The ribbon color encodes the distance to the semi-transparent plane orthogonal to the view direction (lower inset). Shadow projections (upper inset) provide an additional visual hint on the curvature extent [26].

Klemm and colleagues [26] extracted, clustered and visualized spine canal centerlines. Image segmentation of 493 MRI data sets was carried out automatically using tetrahedron-based finite element models of vertebrae and spinal canal [35]. Using barycentric coordinates of the tetrahedrons, a centerline consisting of 93 discrete points was extracted for each segmentation, as seen in Figure 6 (left, middle). They served as input for an agglomerative hierarchical clustering which created groups of subjects based on differences in shape. This special clustering technique was chosen, since it produces meaningful results in the clustering of similar structures, such as fiber tracts derived from MR-Diffusion Tensor Imaging data [26].

The cluster visualization in Figure 6 (right) displays each cluster representative as ribbon in a sagittal plane. The representative is the centerline with the smallest sum of distances to all other centerlines, i.e., the centroid line of the cluster. The width of the ribbon encodes the cluster count and the color encodes the distance to the sagittal plane. Shadow projections also (redundantly) convey the distance to the sagittal plane. This allows to assess the 3D shape in a 2D projection. The results of the clustering can be used in different ways.

- **Outlier detection**: Extraordinary shapes yield clusters of small size that differ strongly from the global mean shape. This can point to pathologies or errors in the segmentation process.

- **Hypothesis generation**: Shape groups serve as a starting point for an exploratory analysis to analyze disease-related correlations. The usual workflow requires epidemiologists to define groups, e.g., age ranges. Groups calculated solely using shape information can be analyzed to detect statistically relevant associations in other expositions which can lead to new hypotheses.
- **Hypothesis validation**: Clustering based on non-image related features can be used to analyze if these clusters are correlated to characteristic shapes, e.g., a strongly bent spine canal representative.

Calculating curvature on groups created according to body height starting from 150 cm in 10 cm steps was performed. Klemm and colleagues found that taller people have a more straight spine compared to small people. They also found multiple clusters of people 10 years above average age across all groups that exhibit a strong ”S” shape of the spine, which was the starting point for new investigations using expert chosen spine-related attributes. This method was extended to integrate the relevant information for identifying correlations. Thus, for a selected cluster, information related to the distribution of attributes, such as the back pain history (frequency and intensity of back pain), may be displayed as a tool tip. The initial observations show, that a box plot summarizing the distribution is more suitable than the full histogram. For routine use in epidemiology, the lumbar spine visualization (Fig. 6) has to be complemented with at least simple statistics to answer questions, such as: Is there a statistically significant difference between the curvature of the lumbar spine canal and back pain frequency? If so, what is the effect size?

4.4 Cluster Analysis and Information Space Reduction

A crucial task in epidemiology is the definition of groups of subjects. Differences and similarities among groups are investigated and control groups are defined to detect and assess the impact and interaction of risk factors to define the relative risk. A straightforward approach is the manual definition based on study variables and ranges of interest. A data-driven extension is the automatic detection of potentially relevant subgroups in the often high-dimensional data by means of clustering algorithms [25]. In particular, the generation of new hypotheses, which may be tightly connected to the identification of new groups, benefits from the latter. In clustering, subjects, being similar with respect to a certain similarity metric, are grouped in clusters with a low intra-cluster and a high inter-cluster variance. Particular challenges in the cluster analysis of cohort study data are:

- Missing data, e.g., denied answers to inconvenient questions [1]
- Mixture of scala and categorical study variables [3]
- Time-varying variables in longitudinal studies [15]

As a consequence of the first problem, it should be reported to the user how many datasets were actually used for clustering. Depending on the chosen attributes, this

may be a subset of the overall amount of data. Incomplete datasets may and should be used when the relevant data is available. It is essential to use similarity metrics that consider also ordinal or categorical data. Usually, the following convention is used: The distance between datasets equals 1 if their categorical data is different and 0 if it is the same. With ordinal data, more care is necessary. The difference between ”strongly agree” and ”strongly disagree” on a Likert scale is larger than the difference between ”agree” and ”disagree”. However, the precise quantification is not straightforward. As a first step to explore the SHIP data, a parallel coordinate view is combined with scatterplots and clustering (Fig. 7).

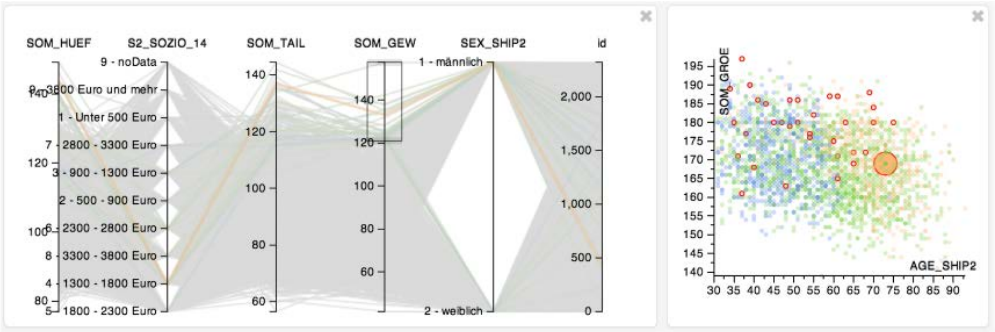


Fig. 7 A parallel coordinate view enables the selection of a relevant subset (here persons with a weight larger than 120 kg). The scatterplots represent correlations between age, size and weight. The elements are color-coded according to a clustering result that yields three clusters. The encircled elements correspond to the selection in the parallel coordinates view.

With respect to clustering algorithms, it is essential that the number of resulting groups has not to be specified in advance. Moreover, algorithms are preferred that allow outliers instead of forcing all elements to be part of a cluster. Outliers may be particularly interesting and thus serve as a starting point for further investigation. Of course, they may also indicate a bad quality of some data. In our experiments, density-based clustering with DBSCAN [12] produced plausible results when applied to non-image data of the SHIP study. The DBSCAN result is sensitive to the *minPoints* parameter that determines the minimum size of a cluster. Some cycles of clustering are necessary to make a suitable choice. In this process, an appropriate visualization is essential to easily understand the results. A visualization that conveys the location, size and shape of clusters is difficult in case that clustering is applied to more than two-dimensional data. A recently developed approach enables 3D visualizations of clustering results with very low levels of occlusion [16].

The majority of cohort-studies are restricted to non-image data, i.e., categorical and scalar data, which may directly serve as input for a clustering algorithm. In [4], 176 patients with lower back pain have been monitored over six months via text messages describing their bothersomeness. All patients received chiropractic treatment. A hierarchical clustering of the individual temporal courses of bothersomeness revealed groups of patients who responded differently to the therapy, which may im-

prove the optimal individual treatment selection. Hypotheses about the differences between paternal age-related schizophrenia (PARS) and other cases of schizophrenia were generated in [27]. A k-means clustering of demographic variables, symptoms, cognitive tests and olfaction for 136 subjects (34 with PARS) delivered clusters containing a high concentration of PARS cases. Significant characteristics of these clusters may give a hint on features of PARS improving its dissociation of other cases of schizophrenia.

In analyzing image-centric cohort study data, besides categorical and scalar data, images may be clustered for group definition. Image intensities, segmentation results, e.g., the surface of the segmented liver or the centerline of the spinal canal (recall Section 4.3) and derived information, e.g., liver tissue texture, liver volume and spinal canal centerline geometry, may serve as input. In [37], the anatomical variation of the mandibles is assessed across a population. For treating mandible fractures, subjects with similar characteristics are grouped in clusters and a suitable implant is designed per cluster. The clustering algorithm k-means is applied to transformation parameters of a locally affine registration between all mandible surfaces segmented in CT data. A cohort of 50 patients with suspicious breast lesions was investigated by means of dynamic contrast enhanced MRI (DCE-MRI) in [34]. Each lesion was clustered according to its perfusion characteristics by means of a region merging approach. Perfusion is represented by the temporal course of the DCE-MRI signal intensities. The clustering itself did not generate groups of patients here, but based on each individual number of clusters and their perfusion characteristics, two groups of lesions could be defined: benign and malignant. These groups were then compared to histological results from core needle biopsy.

Investigating high-dimensional non-image cohort study data often benefits from an information space reduction, e.g., by means of PCA. Plotting the data for inspection in a lower dimensional space while capturing the greatest level of variation, e.g., a scatterplot of the first two principal components, as well as detecting trends in the data and ordering them according to the variance they describe are important applications of PCA. In [36], symptom data gathered in interviews of 410 people with Turret syndrome was investigated in order to specify homogeneous symptom categories for a better characterization of the disease's phenotype. First, clusters of symptom variables were generated using *agglomerative hierarchical clustering*. Then, for each cluster and each participant a score was computed equal to the sum of present symptom variables in the cluster. These scores were the input for PCA, which produced homogeneous symptom categories, sorted according to their percentage of represented symptomatic variance. In [38], scatterplots of variables from cohort study data are extended by superimposing PCA ellipses. The ellipses are computed per group of subjects and illustrate its global distribution with respect to the two opposed variables. They are spanned by the principal component axes of a groups data points and centered at their mean. Optionally, their transparency is adjusted with respect to the groups confidence, i.e. the number of contained subjects.

4.5 Categorical Data

Cohort study data sets comprise many categorical data such as answers to questionnaires or categorizations that may result from binning continuous data, such as intervals of income or age groups. These data are *discrete* and exhibit a *low range*. While data resulting from binning or from answers marked at a Likert scale have an inherent order (ordinal data), often no inherent order exists. Standard information visualization methods like scatterplots and parallel coordinates are designed for continuous data and thus not ideal for displaying categorical data, since many data points occlude each other.

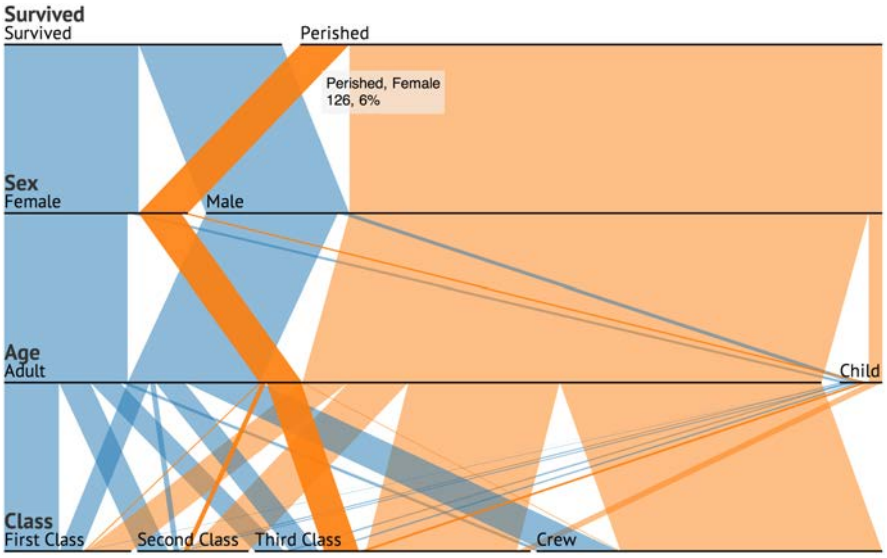


Fig. 8 Parallel sets are designed to explore categorical data. Parallel sets are based on the parallel coordinates layout, but it maps frequency of the data instead of rendering them just as data points. The user can interactively remap the data to new categorizations as well as highlight entries to examine their distribution along other mapped dimensions. The displayed data set shows the distribution of passengers of the RMS Titanic and whether they survived the sinking of the ship [10].

Categorical data are often visualized using boxes where the width is scaled to frequency [45]. These approaches use much space and are also not well suited for encoding multidimensional relationships. *Parallel sets*, introduced by Bendix and colleagues, comprise the same layout as parallel coordinates, ”but the continuous axes were replaced with sets of boxes . . . scaled to the frequency of the category” [6] (see Fig. 8). Selecting an attribute will map each category to a distinct color so that they can be traced through all visualized dimensions. Highlighting a category may be realized by drawing the selected category in a higher saturation leading to a pop-out effect. It is also useful to display a histogram for the selected category annotated with statistical information. Selecting a category will only display the particular box on an axis and make more room for connections to other axes.

This is especially helpful if the number of categories for one dimension is large. In case of data without inherent order, reordering of axes is possible and supports the exploration by providing more comprehensible layouts.

5 Concluding Remarks

A variety of large prospective cohort studies are established and ongoing. They generate a wealth of potentially relevant information for assessing risks, for estimating costs involved in treatment and thus inform health policy makers with respect to potentially preventive measures or cost-limiting initiatives. Despite the great potential of data mining and interactive visualization, none of these studies included such activities in their original planning. The role of computer science, so far, was limited to database management and data security. Visual analytics has a great potential for exploring complex health-related data, as recently shown by Zhang and colleagues [50] for clinical applications, such as treatment planning. Similar techniques may be employed to address epidemiology research. In contrast to clinical applications, where a severe time pressure leads to strongly guided workflows, epidemiology research benefits from powerful and flexible tools that enable and support *exploration*. Currently, existing and widespread information visualization and analytics techniques are employed and adapted to epidemiologic data. The high-dimensional nature of these data, however, also requires to develop new techniques.

The specific and new aspect discussed in this paper was the integration of image data, information derived from image data, such as spine curvature-related measures, and more traditional socio-demographics data.

Future work. So far, visual analytics research and software was rarely focused on epidemiology. Thus, to adapt visual analytics to epidemiology and to integrate the solutions with tools familiar to epidemiologists is necessary. In epidemiology, national studies are prevailing, which is due to the large amount of legislative conditions to be considered. International studies would enable to explore diseases with lower prevalence, subtypes of diseases that occur rarely, e.g., cancer in early age, and specific questions of spatial epidemiology. The SHIP Brazil study will provide such information. It was recently initiated to perform a study in Brazil according to the standards and experiences gained in the German SHIP study. A nasty but essential problem of image-based epidemiologic study is quality control. Research efforts are necessary to automatically check whether image data fully cover the target region, whether the alignment of slices, e.g., in heart imaging, is correct and whether severe artifacts appear, e.g., motion artifacts. This chapter discussed exciting developments related to the combined use of radiologic image and more classical epidemiological data. The next wave is already clearly recognizable: genetics information will be integrated in the search for early markers associated with risks for diseases.

Acknowledgements We want to thank Lisa Fraunstein, David Kilias and David Perlich who supported our analysis of the SHIP data as student workers as well as Marko Rak who provided the detection algorithm for the vertabrae and Myra Spilopoulou for fruitful discussions on clustering and data mining (all University of Magdeburg). We thank Martijn Steenwijk for providing images from his work and Charl Botha for fruitful discussions. Matthias Günther (Fraunhofer MEVIS) explained us quality aspects of MR imaging in epidemiologic studies. This work was supported by the DFG Priority Program 1335: Scalable Visual Analytics. SHIP is part of the Community Medicine Research net of the University of Greifswald, Germany, which is funded by the Federal Ministry of Education and Research (grant no. 03ZIK012), the Ministry of Cultural Affairs as well as the Social Ministry of the Federal State of Mecklenburg-West Pomerania.

References

1. Review: A Gentle Introduction to Imputation of Missing Values. Journal of Clinical Epidemiology **59**(10), 1087 – 1091 (2006)
2. Ahlberg, C.: Spotfire: an information exploration environment. SIGMOD Record **25**(4) (1996)
3. Ahmad, A., Dey, L.: A k-mean Clustering Algorithm for Mixed Numeric and Categorical Data. Data Knowl. Eng. **63**(2), 503–527 (2007)
4. Axén, I., Bodin, L., Bergström, G., Halasz, L., Lange, F., Lövgren, P.W., Rosenbaum, A., Leboeuf-Yde, C., Jensen, I.: Clustering Patients on the Basis of Their Individual Course of Low Back Pain Over a Six Month Period. B M C Musculoskeletal Disorders **12**, 99–108 (2011)
5. Beale, L.L., Abellan, J.J., Hodgson, S.S., Jarup, L.L.: Methodologic issues and approaches to spatial epidemiology. Environmental health perspectives **116**(8), 1105–1110 (2008)
6. Bendix, F., Kosara, R., Hauser, H.: Parallel sets: Visual analysis of categorical data pp. 133–140 (2005)
7. Blaas, J., Botha, C.P., Post, F.H.: Interactive Visualization of Multi-Field Medical Data Using Linked Physical and Feature-Space Views. In: Proc. of EuroVis, pp. 123–130 (2007)
8. Busking, S., Botha, C.P., Post, F.H.: Dynamic Multi-View Exploration of Shape Spaces. Computer Graphics Forum **29**(3), 973–982 (2010)
9. Chui, K.K., Wenger, J.B., Cohen, S.A., Naumova, E.N.: Visual analytics for epidemiologists: understanding the interactions between age, time, and disease with multi-panel graphs. PloS one **6**(2), e14,683 (2011)
10. Davies, J.: Parallel Set of the Titanic Data Set. `http://www.jasondavies.com/parallel-sets/` (2012). [Online; accessed 30-January-2014]
11. Engel, K., Toennies, K.D.: Hierarchical vibrations for part-based recognition of complex objects. Pattern Recognition **43**(8), 2681–2691 (2010)
12. Ester, M., Kriegel, H.P., Sander, J., Xu, X.: A Density-Based Algorithm for Discovering Clusters in Large Spatial Databases with Noise. In: Proc. of the Second International Conference on Knowledge Discovery and Data Mining (KDD), pp. 226–231 (1996)
13. Ferrarini, L., Olofsson, H., Palm, W., Vanbuchem, M., Reiber, J., Admiraalbehloul, F.: GAMEs: Growing and adaptive meshes for fully automatic shape modeling and analysis. Medical Image Analysis **11**(3), 302–314 (2007)
14. Fletcher, R.H., Fletcher, S.W.: Clinical Epidemiology. Lippincott Williams & Wilkins (2011)
15. Genolini, C., Falissard, B.: KmL: k-means for Longitudinal Data. Comput. Stat. **25**(2), 317–328 (2010)
16. Glaßer, S., Lawonn, K., Preim, B.: Visualization of 3D Cluster Results for Medical Tomographic Image Data. In: In Proc. of Conference on Computer Graphics Theory and Applications (VISIGRAPP/GRAPP), pp. 169–176 (2014)
17. Gloger, O., J., K., Stanski, A., Völzke, H., Puls, R.: A fully automatic three-step liver segmentation method on LDA-based probability maps for multiple contrast MR images. Magnetic Resonance Imaging **28**(6), 882–897 (2010)

18. Gloger, O., Toennies, K.D., Liebscher, V., Kugelmann, B., Laqua, R., H., V.: Prior shape level set segmentation on multistep generated probability maps of MR datasets for fully automatic kidney parenchyma volumetry. IEEE Transactions on Medical Imaging **31**(2), 312–325 (2012)
19. Gresh, D.L., Rogowitz, B.E., Winslow, R.L., Scollan, D.F., Yung, C.K.: Weave: a system for visually linking 3-d and statistical visualizations, applied to cardiac simulation and measurement data. In: Proc. of IEEE Visualization, pp. 489–492 (2000)
20. Hegenscheid, K., Kühn, J.P., Völzke, H., Biffar, R., Hosten, N., Puls, R.: Whole-body magnetic resonance imaging of healthy volunteers: pilot study results from the population-based SHIP study. Fortschritte auf dem Gebiet der R"øntgenstrahlen und der bildgebenden Verfahren (Röfo) **181 (8)**, 748–759 (2009)
21. Hermann, M., Schunke, A.C., Klein, R.: Semantically steered visual analysis of highly detailed morphometric shape spaces. In: Proc. of IEEE Symposium on Biological Data Visualization (BioVis), pp. 151–158 (2011)
22. Hofman, A., Breteler, M.M.B., van Duijn, C.M., Janssen, H.L.A., Krestin, G.P., Kuipers, E.J., Stricker, B.H.C., Tiemeier, H., Uitterlinden, A.G., Vingerling, J.R., Witteman, J.C.M.: The Rotterdam Study: 2010 objectives and design update. European Journal of Epidemiology **24**, 553–572 (2009)
23. Hofman, A., van Duijn, C.M., Franco, O.H., et al.: The Rotterdam Study: 2012 objectives and design update. European Journal of Epidemiology **26**, 657–686 (2011)
24. Jerrett, M., Gale, S., Kontgis, C.: Spatial Modeling in Environmental and Public Health Research. International Journal of Environmental Research in Public Health **7**(16), 1302–1329 (2010)
25. Klemm, P., Frauenstein, L., Perlich, D., Hegenscheid, K., Völzke, H., Preim, B.: Clustering Socio-demographic and Medical Attribute Data in Cohort Studies. In: Proc. of Bildverarbeitung fuer die Medizin (BVM) (2014)
26. Klemm, P., Lawonn, K., Rak, M., Preim, B., Tönnies, K., Hegenscheid, K., Völzke, H., Oeltze, S.: Visualization and Analysis of Lumbar Spine Canal Variability in Cohort Study Data. In: Proc. of Vision, Modeling, Visualization (VMV), pp. 121–128 (2013)
27. Lee, H., Malaspina, D., Ahn, H., Perrin, M., Opler, M.G., Kleinhaus, K., Harlap, S., Goetz, R., Antonius, D.: Paternal Age Related Schizophrenia (PARS): Latent Subgroups Detected by k-means Clustering Analysis. Schizophrenia Research **128**(13), 143 – 149 (2011)
28. Mackinlay, J., Hanrahan, P., Stolte, C.: Show me: automatic presentation for visual analysis. IEEE Transactions on Visualization and Computer Graphics **13**(6), 1137–1144 (2007)
29. Marathe, M., Vullikanti, A.K.S.: Computational epidemiology. Communications of the ACM **56**(7), 88–96 (2013)
30. McInerney, T., Terzopoulos, D.: Deformable models in medical image analysis: a survey. Medical Image Analysis **1**(2), 91–108 (1996)
31. Pearce, N., Merletti, F.: Complexity, simplicity, and epidemiology. International Journal of Epidemiology **35**(3), 515–519 (2006)
32. Petersen, S.E., Matthews, P.M., Bamberg, F., et al.: Imaging in population science: cardiovascular magnetic resonance in 100,000 participants of UK Biobank - rationale, challenges and approaches. J. Cardiovasc Magn Reson. **28**, 15–46 (2013)
33. Petyt, M.: Introduction to finite element vibration analysis. Cambridge University Press (1998)
34. Preim, U., Glaßer, S., Preim, B., Fischbach, F., Ricke, J.: Computer-aided Diagnosis in Breast DCE-MRI-Quantification of the Heterogeneity of Breast Lesions. European Journal of Radiology **81**(7), 1532–1538 (2012)
35. Rak, M., Engel, K., Tönnies, K.D.: Closed-Form Hierarchical Finite Element Models for Part-Based Object Detection. In: Proc. of Vision, Modeling, Visualization (VMV), pp. 137–144 (2013)
36. Robertson, M.M., Althoff, R.R., Hafez, A., Pauls, D.L.: Principal Components Analysis of a Large Cohort With Tourette Syndrome. The British Journal of Psychiatry **193**(1), 31–36 (2008)
37. Seiler, C., Pennec, X., Reyes, M.: Capturing the Multiscale Anatomical shape Variability with Polyaffine Transformation Trees. Medical Image Analysis **16**(7), 1371 – 1384 (2012)

38. Steenwijk, M.D., Milles, J., van Buchem, M.A., Reiber, J.H.C., Botha, C.P.: Integrated Visual Analysis for Heterogeneous Datasets in Cohort Studies. In: Proc. of IEEE VisWeek Workshop on Visual Analytics in Health Care (2010)
39. Stolte, C., Tang, D., Hanrahan, P.: Polaris: A system for query, analysis, and visualization of multidimensional relational databases. IEEE Transactions on Visualization and Computer Graphics **8**(1), 52–65 (2002)
40. Thew, S., Sutcliffe, A., Procter, R., de Bruijn, O., McNaught, J., Venters, C.C., Buchan, I.: Requirements engineering for e-Science: experiences in epidemiology. IEEE Software **26**(1), 80–87 (2009)
41. Turkay, C., Lundervold, A., Lundervold, A.J., Hauser, H.: Hypothesis generation by interactive visual exploration of heterogeneous medical data. In: Proc. of Human-Computer Interaction and Knowledge Discovery in Complex, Unstructured, Big Data, pp. 1–12 (2013)
42. Völzke, H., Alte, D., Schmidt, C., et al.: Cohort Profile: The Study of Health in Pomerania. International Journal of Epidemiology **40**(2), 294–307 (2011)
43. Völzke, H., Baumeister, S.E., Alte, D., Hoffmann, W., Schwahn, C., Simon, P., John, U., Lerch, M.M.: Independent Risk Factors for Gallstone Formation in a Region with High Cholelithiasis Prevalence. Digestion **71**, 97–105 (2005)
44. Weaver, C.: Cross-filtered views for multidimensional visual analysis. IEEE Transactions on Visualization and Computer Graphics **16**(2), 192–204 (2010)
45. Wittenburg, K., Lanning, T., Heinrichs, M., Stanton, M.: Parallel bargrams for consumer-based information exploration and choice. In: Proc. of the ACM symposium on User interface software and technology (UIST), pp. 51–60 (2001)
46. Ystad, M.: Quantitative structural and functional brain imaging in cognitive aging. Ph.D. thesis, University of Bergen (2010)
47. Ystad, M., Eichele, T., Lundervold, A.J., Lundervold, A.: Subcortical functional connectivity and verbal episodic memory in healthy elderly—resting state fmri study. NeuroImage **52**(1), 379–388 (2010)
48. Ystad, M., Lundervold, A.J., Wehling, E., Espeseth, T., Rootwelt, H., Westlye, L., Andersson, M., Adolfsdottir, S., Geitung, J., Fjell, A., Reinvang, I., Lundervold, A.: Hippocampal volumes are important predictors for memory function in elderly women. BMC Medical Imaging **9**(1), 1–15 (2009)
49. Zhang, Z., Gotz, D., Perer, A.: Interactive Visual Patient Cohort Analysis. In: Proc. of IEEE VisWeek Workshop on Visual Analytics in Healthcare (2012)
50. Zhang, Z., Wang, B., Ahmed, F., Ramakrishnan, I., Viccellio, A., Zhao, R., Mueller, K.: The Five Ws for Information Visualization with Application to Healthcare Informatics. IEEE Transactions on Visualization and Computer Graphics **19**(11), 379–388 (2013)

Visualization and Analysis of Lumbar Spine Canal Variability in Cohort Study Data

P. Klemm[1], K. Lawonn[1], M. Rak[1], B. Preim[1], K. Toennies[1], K. Hegenscheid[2], H. Völzke[2], S. Oeltze[1]

[1]Otto-von-Guericke University Magdeburg, Germany
[2]Ernst-Moritz-Arndt-University Greifswald, Germany

Abstract

Large-scale longitudinal epidemiological studies, such as the Study of Health in Pomerania (SHIP), investigate thousands of individuals with common characteristics or experiences (a cohort) including a multitude of socio-demographic and biological factors. Unique for SHIP is the inclusion of medical image data acquired via an extensive whole-body MRI protocol. Based on this data, we study the variability of the lumbar spine and its relation to a subset of socio-demographic and biological factors. We focus on the shape of the lumbar spinal canal which plays a crucial role in understanding the causes of lower back pain.
We propose an approach for the reproducible analysis of lumbar spine canal variability in a cohort. It is based on the centerline of each individual canal, which is derived from a semi-automatic, model-based detection of the lumbar spine. The centerlines are clustered by means of Agglomerative Hierarchical Clustering to form groups with low intra-group and high inter-group shape variability. The number of clusters is computed automatically. The clusters are visualized by means of representatives to reduce visual clutter and simplify a comparison between subgroups of the cohort. Special care is taken to convey the shape of the spinal canal also orthogonal to the view plane. We demonstrate our approach for 490 individuals drawn from the SHIP data. We present preliminary results of investigating the clusters with respect to their associated socio-demographic and biological factors.

Categories and Subject Descriptors (according to ACM CCS): J.3 [Computer Applications]: Life and Medical Sciences—Health

1. Introduction

Exploiting the full potential of huge information spaces created by cohort studies like the Study of Health in Pomerania (SHIP) is one of the major challenges in modern epidemiology. The SHIP [VAS*11] aims at characterizing health by assessing data relevant to prevalence and incidence of diseases and identifying their risk factors. With the recent incorporation of medical image data in cohort studies, shape and texture of organs may be characterized. Shape information linked to other medical or lifestyle data show great promise for better understanding of risk factors for certain diseases [WP03]. For example, how does a physically hard job influence the shape of the spine? Scientific findings yield in precise precautions for people who belong to risk groups.

Our focus is on the lumbar spine, which is most often the source of musculoskeletal disorders in clinical practice [vTKB02, WP03]. The whole-body MRI scans of the SHIP are the basis for our approach to enable a reproducible analysis of the lumbar spine canal variability. Our contributions are:

- generation of groups of individuals sharing a similar shape of the lumbar spine canal,
- visualization of these groups by means of representatives,
- illustration of 3D shape in a 2D view.

While the processing of the 490 data sets represents first results, we were able to observe the expected behavior like decreasing spine curvature with increasing subject body height. We also found unexpected clusters of unusual shape, which are now subject to further epidemiological analysis.

2. Related Work

To the best of our knowledge, only Steenwijk and colleagues concurrently query and visualize both image and non-image

data in a Visual Analytics framework [SMvB*10]. They put emphasis on a structured data organization and employ a relational database. Their work is closest to ours albeit our investigation of image and non-image data is at the moment still being performed sequentially.

Non-image Data. Cohort study data is often very heterogeneous. It consists of image and non-image data, different types of parameters, e.g. ordinal, nominal, and quantitaive, and parameters of the same type but having different domains, which may partially overlap. Schulze-Wollgasts [SWST03] work supports the data exploration process and hypotheses generation by dividing the information space into data cubes, which can then be understood as *n*-dimensional arrays. They are used to investigate normalized parameters of different modalities and individuals. Linking & brushing is used to investigate interesting details in the resulting spaces. Zhang and colleagues [ZGP12] extended this approach by a web-based system which allows for grouping of subjects based on associated data variables and feeding groups into a visualization system to support insight into complex correlations of the data attributes. Groups are precomputed by calculating common sets of risk factors. This can serve as starting point for an exploratory analysis. We adapt this approach by computing clusters based on shape.

Image Data. Caban and colleagues [CRY11] give an overview on how shape distribution models can be compared using different methods like deformation grids, likelihood volumes and glyphs. Their presented study favors a spherical glyph representation of variation modes. Busking and colleagues [BBP10] proposed a method which plots instances of a structure on a 2D plane. The user can then generate interpolated views in an object space view via mesh morphing on a reference structure together with a color-coded deformation field on the surface. In the shape evolution view, 2D projections of all structure instances can be compared. With pairwise corresponding data points their segmentation model is of the same type as our spine detection model. Their methods, however, focuses largely on local structural changes while we address curvature. Visualizing our data with their open source `ShapeSpaceExplorer` lead to a very cluttered view, since it is not suited for a large number of input objects. We do not use their approach of main variation modes, since they also display models by interpolating between standard deviation steps, which are not part of the data. Hermann and colleagues [HSK11] compared statistical deformation models to detect anatomically different individuals of the rodent mandibles. They propose a semantically driven user-centered pipeline that includes expert knowledge as region-of-interest selection via interactive volume deformation. This takes especially into account that not all shape information in a model is of equal interest to the user. Chou and colleagues [CLA*09] investigated the correlation of Alzheimer's disease for 240 subjects with ventricular expansion, clinical characteristics, cognitive values and related biomarker by statistically linking them together and ploting their p-values onto the ventricle surface. This way of directly mapping disease-related biomarkers is an example of how different data modalities can be expressively combined. A visual analytics approach for improving model based segmentation is presented by von Landesberger and colleagues [vLBK*13]. They introduced expert knowledge via visual analytics tools into every important step of segmentation from pre-processing to evaluation.

Using deformation fields that describe dense correspondences, Rueckert and colleagues [RFS03] constructed an atlas of average anatomy with variability across a population. Registration-based statistical deformation models are shown to be suitable for characterizing shape over many subjects.

3. Epidemiology of Back Disorders

Epidemiological cohort studies aim to identify factors which are associated with diseases and mortality risks. This includes socio-economic characteristics and medical parameters. While the understanding of genetic mutations regarding back disorders made progress, the correlations with different environmental factors as well as physical stress are not sufficiently understood [MM05]. Manek and colleagues reviewed the progress made in understanding causes of back pain and present influencing factors like age, gender, weight and different lifestyle aspects, such as smoking behavior and work conditions. Tucer and colleagues [TYO*09] conclude that depression is one of the independent risk factors for experiencing low back pain, although their analysis is based on surveys of the subjects and does not rest upon clinical analysis. Lang-Tapia and colleagues [LTERAC11] used a non-invasive method for analyzing spine curvature using a so-called "Spine-Mouse". They correlated spine curvature with age, gender, and weight-status. They did not find correlations between lumbar spine deformation and weight status. Van Tulder and colleagues [vTKB02] conclude that the value of such identified risk factors as prognostic value remains low. No factor arose as strong indication for back pain through many different studies.

These studies share the relation to socio-demographic and medical attribute data with most cohort studies that analyze back disorders. Many studies do not include shape information, only very few use medical imaging at all. One distinct feature of the SHIP are the whole-body MRI scans gathered for a large cohort of 3,368 subjects [HSS*13]. Radiation acting on subjects makes CT imaging ethically unjustifiable. Body-imaging allows for linking the spine shape to other attributes. Spines can be divided into groups to evaluate their potential to induce a pathology. Future cohort assessments even allow to determine change of spine shape.

4. Image Data Acquisition and Spine Detection

All whole-body MRI scans were acquired on a 1.5 Tesla scanner (Magnetom Avanto; Siemens Medical Solutions,

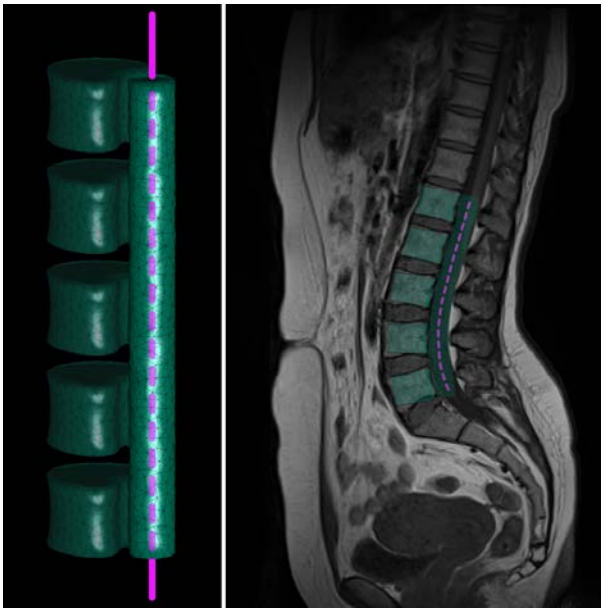


Figure 1: *The layered finite element model consists of more than 2,000 tetrahedrons (left). The spine canal center line is indicated by the dashed line. The model uses the image-induced potential field to align itself to find a local minimum after the initialization (right).*

Erlangen, Germany) by four trained technicians in a standardized way. Subjects were placed in the supine position. Five phased-array surface coils were placed to the head, neck, abdomen, pelvis, and lower extremities for whole-body imaging. The spine coil is embedded in the patient table. The spine protocol consisted of a sagittal T1-weighted turbo-spin-echo sequence (676 / 12 [repetition time msec / echo time msec]; 150° flip angle; 500 mm field of view; $1.1 \times 1.1 \times 4.0$ *mm* voxels) and a sagittal T2-weighted turbo-spin-echo sequence (3760 / 106 [repetition time msec / echo time msec]; 180° flip angle; 500 mm field of view; $1.1 \times 1.1 \times 4.0$ *mm* voxels). First, both sequences were placed over the cervical and upper thoracic spine. Then, they were placed over the lower thoracic and lumbar spine. The MRI software automatically composed a whole spine sequence from the two T1-weighted and T2-weighted sequences [HSS*13]. We were provided with 490 data sets.

Our work requires a detection of the lumbar spine in the MRI data. We employ a hierarchical finite element method according to [RET13]. Tetrahedron-based finite element models (FEM) of vertebrae and spinal canal are connected by a bar-shaped FEM (Fig. 1). The model comprises a fixed number of points which are pairwise relatable between instances of the model. Hence, correspondences between lumbar spine representations of different data sets can easily be established. The model is placed in the scene using an empirically chosen initialization point. The force acting on the model stems from aggregation of loads, which are derived from a potential field resulting from a weighted sum of the T1- and T2-weighted MRI images, see [RET13]. After detecting all spines, we register the models because in a later clustering step we only want to capture the local deformation of the lumbar spine, not different locations in world space. The models are registered using the Kabsch Algorithm [Kab76], which is designed to minimize the root mean squared deviation between paired sets of points. The model-based detection captures information about the spine canal curvature as well as the alignment of the vertebrae. It is not meant to capture information about vertebrae deformation and differences in spine canal extent.

5. Analysis of Lumbar Spine Canal Variability

We investigate the variability of the lumbar spine canal based on the deformed and registered models of the detection step. Since our primary interest is on the curvature of the spine, we focus on the spinal canal. Centerlines capture curvature and are easier to handle than the tetrahedral mesh. Clustering using Agglomerative Hierarchical Clustering is carried out to form groups that exhibit low intra-group and high inter-group shape variability. The clusters are visualized by means of representatives to reduce visual clutter and simplify a comparison between subgroups of the cohort.

5.1. Centerline Extraction

In this subsection, we describe how we compute the centerline $c_{\mathcal{S}}$ of the lumbar spine model $\mathcal{S}$. The model is given as a cylindrically shaped tetrahedral mesh. The axis of rotation is aligned to the z axis. Therefore, we use a parametric curve $c(t) = p_0 + t \cdot v_z$ where the z-component lies in $[h_{min}, h_{max}]$. Here, h_{min} and h_{max} are the minimal and maximal height of the mesh, respectively. We can write the parametric curve $c(t)$ as:

$$c(t) = \underbrace{\begin{pmatrix} 0 \\ 0 \\ h_{min} \end{pmatrix}}_{p_0} + t \cdot \underbrace{\begin{pmatrix} 0 \\ 0 \\ h_{max} - h_{min} \end{pmatrix}}_{v_z}, \qquad t \in [0,1]. \quad (1)$$

We determine the intersection points of the parametric curve with the faces of the tetrahedra $\tau \in \mathcal{S}$ of the undeformed lumbar spine model $\mathcal{S}_0$. Thus, we combine the vertices to obtain the triangles, faces and assess the intersection points with the curve. For this, we use vertices v_0, v_1, v_2, v_3 of every tetrahedra $\tau = \{v_0, v_1, v_2, v_3\}$ and solve the following matrix equation:

$$\begin{pmatrix} v_k & v_l & v_m & v_z \\ 1 & 1 & 1 & 0 \end{pmatrix} \cdot \begin{pmatrix} \alpha \\ \beta \\ \gamma \\ -t \end{pmatrix} = \begin{pmatrix} p_0 \\ 1 \end{pmatrix}, \quad (2)$$

with different permutated $k,l,m \in \{0,1,2,3\}$ for the four different faces of the tetrahedra. The equation combines the

parametric curve with the triangle face according to barycentric coordinates to obtain the intersection point. If we obtain a positive solution $\alpha, \beta, \gamma > 0$, the considered curve point lies in the interior of a triangle of τ. Thus, we assign the corresponding tetrahedron with its triangle and their barycentric coordinates to the curve point $p_i = p_0 + t \cdot v_z$. If one curve point lies on the boundary of a triangle, i.e., one of the coordinates is equal to zero, we assign only one tetrahedron to the curve point. Using these values, we obtain the centerline of every deformed lumbar spine model by applying the stored barycentric coordinates to the corresponding tetrahedron. Having one intersection point p_i of the undeformed lumbar spine model with the assigned tetrahedra τ, the coresponding triangle face v_k, v_l, v_m, and the assigned barycentric coordinates α, β, γ, we extract the new point p_i' of the deformd lumbar spine model by applying:

$$p_i' = \alpha v_k + \beta v_l + \gamma v_m. \tag{3}$$

Hence, we gain the new centerline.

5.2. Centerline Clustering

To cluster the centerlines, we employ an Agglomerative Hierarchical Clustering (AHC) approach. It has been demonstrated that AHC delivers meaningful results in the clustering of other plane and space curves, such as fiber tracts from Diffusion Tensor Imaging (DTI) data [MVvW05], streamlines from flow data [YWSC12], and brain activation curves (time-series) from functional Magnetic Resonance Imaging (fMRI) data [LCYL08]. Furthermore, it is flexible with regard to cluster shape and size. AHC relies on the difference/similarity between data entities. Thus, a definition of centerline similarity is the prerequisite for AHC of centerlines.

Similarity is often evaluated by a distance measure. General requirements for such a measure are positive definiteness and symmetry. A valid example, that has been successfully employed for clustering fiber tracts and streamlines [MVvW05, YWSC12], is the *mean of closest point distances* (MCPD) proposed in [CGG04]. For two centerlines c_i and c_j with points p, the MCPD is computed as:

$$\begin{aligned} d_M(c_i, c_j) &= mean(d_m(c_i, c_j), d_m(c_j, c_i)) \\ with\ d_m(c_i, c_j) &= mean_{p_l \in c_i} \min_{p_k \in c_j} \|p_k - p_l\| \end{aligned} \tag{4}$$

Cluster Proximity. AHC requires beforehand the computation of all pairwise centerline distances and their storage in a quadratic and symmetric distance matrix **M**. The algorithm operates in a bottom-up manner. Initially, each centerline is considered as a separate cluster. The algorithm then iteratively merges the two closest clusters until a single cluster remains. The merge step relies on **M** and a measure of cluster proximity. Various cluster proximity measures have been published, among which *single link*, *complete link*, *average link*, and *Ward's method* [TSK05] are the most popular. In single link, the proximity of two clusters is defined as the minimum distance between any two centerlines in the different clusters. Complete and average link employ the maximum and the average of these distances, respectively. Ward's method aims at minimizing the total within-cluster variance at each iteration. It defines the proximity of two clusters as the sum of squared distances between any two centerlines in the different clusters (SSE: sum of squared errors). Before we elaborate on the most suitable proximity measure for our application, we focus on automatically computing a reasonable number of clusters k. This computation helps us in providing a good initial visual summary of the variants in spinal canal shape and it facilitates a more reproducible analysis.

Number of Clusters. Salvador and Chan propose a method for automatically computing the number of clusters in hierarchical clustering algorithms [SC04]. Their *L-method* is based on determining the *knee/elbow*, i.e., the point of maximum curvature, in a graph that opposes the number of clusters and a cluster evaluation metric. The knee is detected by finding the two regression lines that best fit the evaluation graph, and then, the number of clusters that is closest to their point of intersection is returned. Locating the knee depends on the shape of the graph, which again depends on the number of tested cluster numbers k. Salvador and Chan recommend using a full evaluation graph, which ranges from two clusters to the number of data entities. Starting with the full graph, the L-method is carried out iteratively on a decreasing focus region until the current knee location is equal to or larger than the previous location. As evaluation metric, the proximity measure used by the different link versions of AHC is applied. Furthermore, the evaluation is not based on the entire dataset but only on the two clusters that are involved in the current merge step.

Evaluation of Cluster Proximity Measures. In an informal evaluation based on 16 datasets, we tested AHC with the four proximity measures and the L-method. The 16 datasets represent the complete set of centerlines ($n = 490$) and epidemiologically relevant subsets derived according to gender, age, e.g. 20-40, 41-60 and 61-80, body weight, and body height. For each dataset, we applied the four proximity measures and visualized all clustering results side-by-side. A visual inspection of the results confirmed textbook knowledge with regard to the strengths and weaknesses of the proximity measures [TSK05] (Fig. 2 shows an exemplary scenario). In single link clustering, the *chaining effect* could be observed for every dataset. Here, a single large cluster arises containing almost the entire set of centerlines. This cluster contains very dissimilar centerlines but they are connected by a chain of similar ones via some transitive relationship. For the majority of datasets, average link failed to avoid this effect. Instead, strong outliers were represented as individual clusters while the remaining centerlines, being dissimilar and still comprising outliers, were grouped in a single

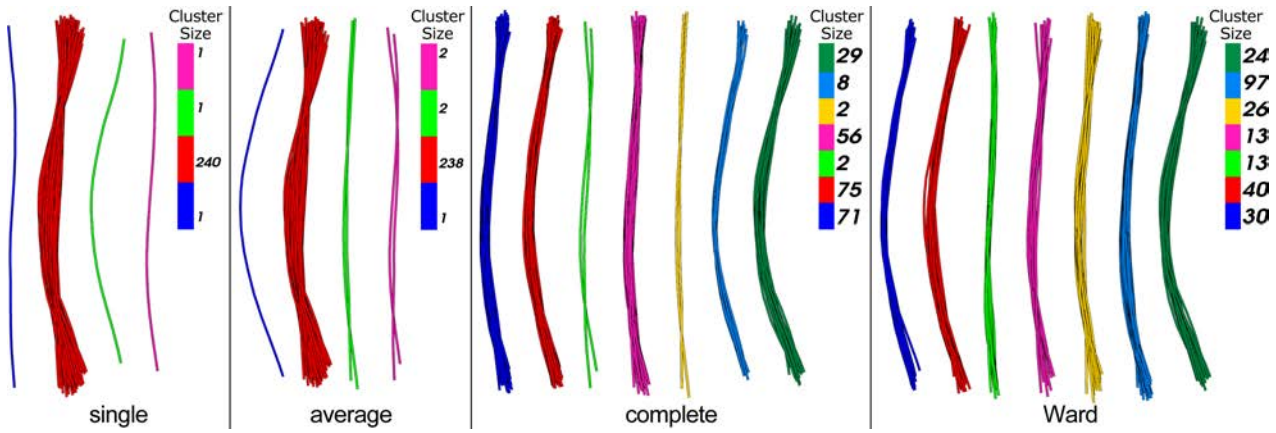


Figure 2: *Spinal canal centerlines of 242 female subjects clustered with Agglomerative Hierarchical Clustering using four different proximity measures and a technique for automatically computing the cluster count. Single link and average link suffer from the chaining effect (single large cluster), complete link produces compact, tightly bound clusters and Ward's method is biased towards generating clusters of similar size. The difference in centerline shape also occurs orthogonal to the view plane.*

large cluster. Complete link clustering produced small, compact, and tightly bound clusters. Ward's method was biased towards generating clusters with similar size. These clusters showed less diversity than the ones generated by means of complete link. In summary, due to the chaining effect of single link and average link, and the arbitrary assumption of similar cluster sizes in Ward's method, we favor complete link as a proximity measure.

The bottleneck of AHC in terms of time complexity is the computation of $\mathbf{M}$, in particular when a multitude of closest point distances must be calculated (Eq. 4). However, our total number of centerlines ($n = 490$) and the number of vertices per centerline ($v = 93$) are relatively small. Furthermore, we have parallelized the computation and the matrix must be computed only once and may be stored. The computation of $\mathbf{M}$ based on the complete set of centerlines, i.e. the entire population, can be considered as the worst case. On a 3.07 GHz Intel 8-core PC with 8 GB RAM and a 64 bit Windows operating system, the computation took 7.9 *s*. The L-method for determining the number of clusters took 24.2 *s* and represents the bottleneck in processing our data. This is due to the multitude of computations required for finding the two best fit regression lines but may be mitigated by cutting off unlikely high numbers of clusters from the full evaluation graph [SC04].

The clustering implementation is based on the AHC algorithm and the proximity measures being part of MATLAB's Statistics Toolbox (MathWorks, Natick, MA, U.S.). The source code of the L-method is provided by A. Zagouras as part of MATLAB Central's file exchange [Zag].

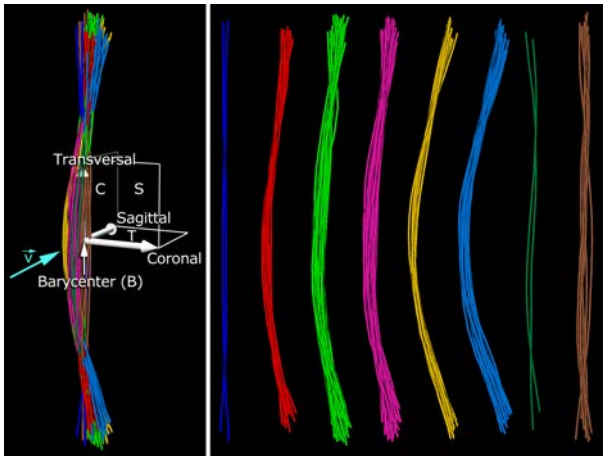


Figure 3: *Initially, all centerline clusters are closely intertwined (left). To simplify their interpretation, they are translated along the coronal axis and lined up at equidistant locations (right). The annotations illustrate typical medical view planes/axes: sagittal (S), coronal (C), and transversal (T). Our default viewing direction $\vec{v}$ is parallel to the sagittal axis (as can be seen in the right view).*

5.3. Visualization of Clustered Centerlines

A standard medical view for inspecting the spine in MR images is the sagittal view with the vertebrae located to the left of the spinal canal (Fig. 1, right). Hence, we choose it as the default view for the presentation of the clustering results. Initially, all centerlines and hence also the clusters, are closely intertwined in space due to the co-registration of all spine detection results (Sec. 4 and Fig. 3, left). In order to get a better overview of the individual clusters, they are translated

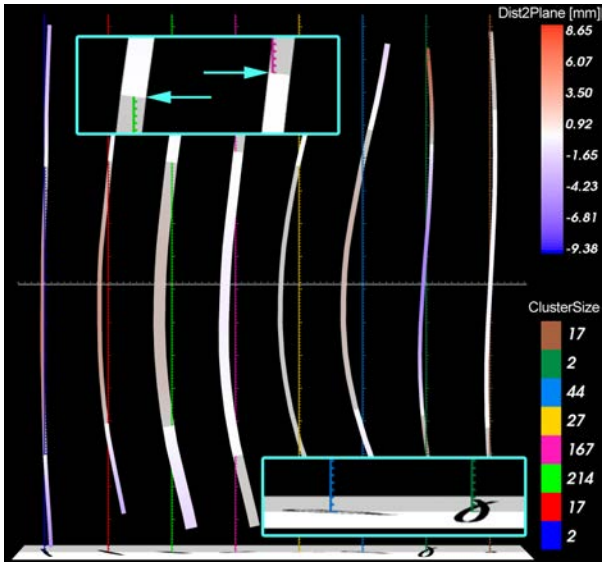


Figure 4: *Spinal canal centerlines of all subjects ($n = 490$) clustered with Agglomerative Hierarchical Clustering employing complete link. For each cluster, a representative centerline is visualized as a ribbon. Ribbon width encoded cluster size. Ribbon color encodes the distance to a view-aligned, highly transparent, sagittal plane passing through the barycenter B of the original centerline bundle (Fig. 3, left). The sequence of a ribbon's intersections with the plane supports an assessment of its curvature (upper inset). Shadow projections reveal how far a representative extends to either side of the plane (lower inset).*

along the coronal axis and lined up at equidistant locations (Fig. 3, right). The centerlines are visualized with GPU support as illuminated streamlines with halos [EBRI09]. The halos improve the visual separation of individual lines. Before the centerlines are translated, the barycenter B of the entire bundle of lines is computed (Fig. 3, left). It will be used for positioning visual hints in the scene.

Cluster Representatives. In order to simplify the interpretation of a cluster, to further reduce visual clutter, and to improve a visual comparison of clustering results between groups, e.g., younger and elder subjects, we compute a representative centerline for each cluster. This is inspired by the computation of a representative fiber tract for a bundle of fibers derived from DTI tractography data [BPHRA13]. Here, the fiber with the smallest sum of distances to all other fibers, i.e. the centroid fiber, of the bundle is chosen. Since all pairwise centerline distances are stored in $\mathbf{M}$, the selection of a centroid centerline is straightforward (Sec. 5.2). Each such centroid is then visualized by a ribbon whose width is scaled according to the size of the corresponding cluster (Fig. 4). Please note that the location of the vertebrae corresponding to this centroid centerline is intentionally not indicated since the ribbons are representative for the course of the spinal canal but not necessarily for the vertebrae location.

Visual Hints. The curvature of the spinal canal along the coronal axis is perceived well in the sagittal view. However, the curvature along the sagittal axis, i.e. the viewing direction, is only deducible by rotating the scene. Hence, we augment the sagittal view by three visual hints improving the curvature perception. (1) A highly transparent sagittal plane passing through B is added to the scene. The position of the ribbon parts with respect to the plane (in front/behind) and the visible intersections of ribbons and plane support the differentiation between spinal canals being mostly bended towards the viewer from those being bended away (Fig. 4, upper inset). (2) The ribbons are colored according to their distance to the sagittal plane. A diverging color scale is used to distinguish between parts in front of the plane (blue), close to the plane (white), and behind the plane (red). (3) A transversal plane is positioned below the ribbons and a light source is positioned above them. Shadow projections are computed and drawn on the plane. They provide an estimate of how far the representatives extend to either side of the plane (Fig. 4, lower inset). In some cases, the projections revealed subtle differences in shape, which could hardly be inferred from the other two hints.

Measurement and Interaction. In order to facilitate a more quantitative analysis of the centerlines and to support a comparison of individual representatives, a vertical and a horizontal axis including tick marks are added to each cluster representative (Fig. 4). All axes are located within the sagittal plane (1). An initial pair of axes running through B has been computed based on the entire set of centerlines and then copied and translated together with each cluster along the coronal axis (Fig. 3). The vertical axes are assigned a unique cluster color to interrelate the representatives and the cluster size legend. The interaction with the visualization exceeds standard 3D scene navigation. Individual representatives may be picked by the user and all centerlines of the corresponding cluster are visualized. The measurement of the spine based on neuralgic points is of crucial importance and has a long tradition in orthopedics. Hence, two measurement widgets have been added for measuring distances and angles (Fig. 5). Both widgets are bound to the geometry of the ribbons in order to simplify measurements in 3D space. The visualization has been implemented in C++ and the Visualization Toolkit. (Kitware, Inc., Clifton Park, NY, U.S.).

6. Results & Discussion

In this section, we present preliminary results combining our shape visualization with associated cohort study data. As seen in Fig. 4, the clustering step is a good way to detect outliers in the data as clusters with very few subjects that have an unusual shape. This can be utilized for finding pathological spine shapes–even for subjects, which do not have a diagnosed back disorder. The technique scales well

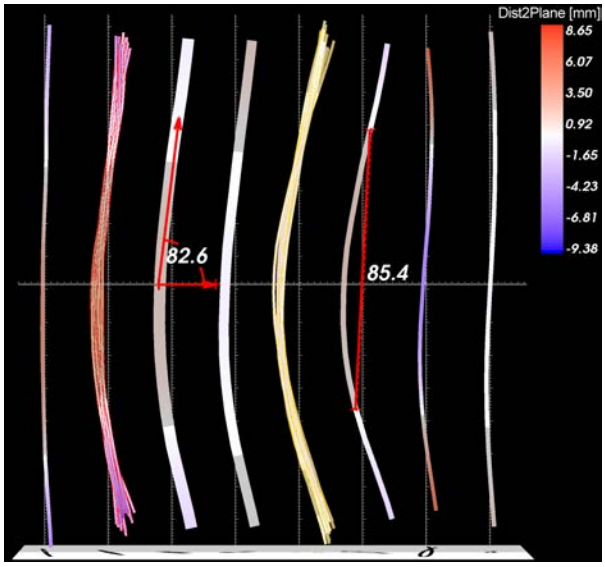


Figure 5: *Interaction facilities. The user may pick a cluster representative, i.e. a ribbon, causing the corresponding cluster to be visualized (centerlines with red and yellow halos). Widgets for measuring distances and angles facilitate a quantitative analysis of the spinal shape.*

regarding the number of input center lines. It is possible to generate an overview for hundreds of subjects as well as for smaller subsets, e.g. subjects which share certain similar attributes. A subset visualization can be applied to detect if the different shape clusters imply a significant difference in associated variables of interest. Does for example a physical demanding job correlate with an extraordinary curved spine?

Our clinical partners expected the lumbar spine to be more straight along the coronal axis for tall people, while being more sinuous ("*lordosis*") with decreasing body height. To check our results for medical plausibility, we created subsets of the data based on *body height*. For each cluster we calculated the distance to the arithmetic mean of *age*, *body height*, and *weight*. We computed the mean of the absolute lordosis curvatures K using the Frenet formulas [Fre52].

While the mean curvature K for people sized 150 – 160 cm is $38.99 \cdot 10^{-4}$ ($\sigma = 9.99 \cdot 10^{-4}$), it gets smaller the larger the subjects are, being at $34.59 \cdot 10^{-4}$ ($\sigma = 9.98 \cdot 10^{-4}$) for 160 – 170 cm and at $31.95 \cdot 10^{-4}$ ($\sigma = 8.88 \cdot 10^{-4}$) for 180 – 190 cm tall people. We could not only confirm the expected differences in the distinct groups, but also give clues for groups which share similar curvature. When looking at subject groups of body height 150 – 160 cm, 160 – 170 cm and 170 – 180 cm we always found a cluster of subjects which are about 10 years older than the rest of the group. They all presented a lordosis shape as well as an "S" shape in sagittal direction ("*scoleosis*"). Since a cluster showing the same characteristics was found in distinct subject groups, it is subject of further investigation.

This finding is an example of how a clustering result can create groups related by shape in order to find other correlations in the associated socio-economic and medical attribute parameters. It can also serve as starting point for a visual analytics tool to detect risk factors.

The visualization aims for at a visual comparability of the clusters. Additionally statistically reliable shape describing features would enhance the method by making statistical calculation applicable to deformation information. This can be achieved by storing the curvature and position of several fixed points in the FEM model. While the visualization allows for characterization of the lumbar spine curvature, it is currently not possible to predicate information about spinal canal narrowings, which can also be an indicator for pathologies like spinal stenosis. This is also the case for a vertebrae deformation, which is an indicator for osteoporosis. We plan to incorporate such information, e.g, based on an extension of the finite element model used for spine detection.

7. Conclusion & Future Work

Applying analysis of medical image data associated with non-image data in a cohort study context is both promising and challenging. The multitude of subjects requires robust yet precise and at least semi-automatic detection and segmentation algorithms which capture the shape of a structure of interest over a large space of subjects. Assessing the resulting information space demands visualizations, which map relevant information among large groups of subjects.

We aim to include more shape describing metrics and apply the technique to all cohort study subjects. This allows for a statistically reliable comparison of clusters. Currently, only the overall curvature and torsion is calculated. Those can be misleading metrics, since coronal as well as sagittal deformation can induce a large curvature. The deformation should be class-divided with the analyzed pathology in mind. Those and other morphology describing metrics can be transferred to the cohort study data dictionary. We also want to include information about unusual vertebrae alignment.

Our presented approach implements a pipeline for analyzing the lumbar spine canal in order to correlate its shape to other variables associated with the cohort study. This was done using an association to *body height*, *gender*, *age* and *weight*. While this was a first step to confirm the expected shape in different subject groups, it has to be enhanced to be applicable to all data variables measured in the cohort.

We plan a web-based visual analytics framework that allows for information visualization on non-image data in combination with complex data set queries including the shape of structures. This allows for possibilities to support queries which are not easy to make in classic statistics

software, like filtering by geographic location as closeness to the coast. We want to provide the epidemiologists with a fast and effective way to analyze their data sets exploiting the potential which lies beneath the numbers.

Acknowledgements: SHIP is part of the Community Medicine Research net of the University of Greifswald, Germany, which is funded by the Federal Ministry of Education and Research (grant no. 03ZIK012), the Ministry of Cultural Affairs as well as the Social Ministry of the Federal State of Mecklenburg-West Pomerania. Whole-body MR imaging was supported by a joint grant from Siemens Healthcare, Erlangen, Germany and the Federal State of Mecklenburg-Vorpommern. The University of Greifswald is a member of the 'Centre of Knowledge Interchange' program of the Siemens AG. This work was supported by the DFG Priority Program 1335: Scalable Visual Analytics.

References

[BBP10] BUSKING S., BOTHA C., POST F.: Dynamic Multi-View Exploration of Shape Spaces. *Computer Graphics Forum 29*, 3 (2010), 973–982.

[BPHRA13] BRECHEISEN R., PLATEL B., HAAR ROMENY B., A. V.: Illustrative uncertainty visualization of DTI fiber pathways. *The Visual Computer 29*, 4 (2013), 297–309.

[CGG04] COROUGE I., GOUTTARD S., GERIG G.: Towards a shape model of white matter fiber bundles using diffusion tensor mri. In *IEEE International Symposium on Biomedical Imaging: Nano to Macro, 2004.* (2004), pp. 344–347 Vol. 1.

[CLA*09] CHOU Y.-Y., LEPORÉ N., AVEDISSIAN C., MADSEN S. K., PARIKSHAK N., HUA X., SHAW L. M., TROJANOWSKI J. Q., WEINER M. W., TOGA A. W., THOMPSON P. M., ALZHEIMER'S DISEASE NEUROIMAGING INITIATIVE: Mapping correlations between ventricular expansion and CSF amyloid and tau biomarkers in 240 subjects with Alzheimer's disease, mild cognitive impairment and elderly controls. *NeuroImage 46*, 2 (June 2009), 394–410.

[CRY11] CABAN J. J., RHEINGANS P., YOO T.: An Evaluation of Visualization Techniques to Illustrate Statistical Deformation Models. *Computer Graphics Forum 30*, 3 (2011), 821–830.

[EBRI09] EVERTS M. H., BEKKER H., ROERDINK J. B., ISENBERG T.: Depth-dependent halos: Illustrative rendering of dense line data. *IEEE Trans. Vis. Comput. Graphics 15*, 6 (2009), 1299–1306.

[Fre52] FRENET F.: Sur les courbes à double courbure. *Journal de Mathématiques Pures et Appliquées* (1852), 437–447.

[HSK11] HERMANN M., SCHUNKE A. C., KLEIN R.: Semantically steered visual analysis of highly detailed morphometric shape spaces. In *BioVis 2011: 1st IEEE Symposium on biological data visualization* (Oct. 2011), pp. 151–158.

[HSS*13] HEGENSCHEID K., SEIPEL R., SCHMIDT C. O., VÖLZKE H., KÜHN J.-P., BIFFAR R., KROEMER H. K., HOSTEN N., PULS R.: Potentially relevant incidental findings on research whole-body MRI in the general adult population: frequencies and management. *European Radiology 23*, 3 (2013), 816–826.

[Kab76] KABSCH W.: A solution for the best rotation to relate two sets of vectors. *Acta Crystallographica Section A 32*, 5 (Sep 1976), 922–923.

[LCYL08] LIAO W., CHEN H., YANG Q., LEI X.: Analysis of fMRI Data Using Improved Self-Organizing Mapping and Spatio-Temporal Metric Hierarchical Clustering. *IEEE Transactions on Medical Imaging 27*, 10 (2008), 1472–1483.

[LTERAC11] LANG-TAPIA M., ESPAÑA-ROMERO V., ANELO J., CASTILLO M.: Differences on spinal curvature in standing position by gender, age and weight status using a noninvasive method. *J Appl Biomech 27*, 2 (2011), 143–50.

[MM05] MANEK N. J., MACGREGOR A. J.: Epidemiology of back disorders: prevalence, risk factors, and prognosis. *Current opinion in rheumatology 17*, 2 (Mar. 2005), 134–140.

[MVvW05] MOBERTS B., VILANOVA A., VAN WIJK J.: Evaluation of Fiber Clustering Methods for Diffusion Tensor Imaging. In *IEEE Visualization* (2005), pp. 65 – 72.

[RET13] RAK M., ENGEL K., TÖNNIES K. D.: Closed-Form Hierarchical Finite Element Models for Part-Based Object Detection. In *Vision, Modeling, Visualization* (2013).

[RFS03] RUECKERT D. D., FRANGI A. F. A., SCHNABEL J. A. J.: Automatic construction of 3-D statistical deformation models of the brain using nonrigid registration. *IEEE Transactions on Medical Imaging 22*, 8 (July 2003), 1014–1025.

[SC04] SALVADOR S., CHAN P.: Determining the Number of Clusters/Segments in Hierarchical Clustering/Segmentation Algorithms. In *Proc. of Tools with Artificial Intelligence. ICTAI* (2004), pp. 576 – 584.

[SMvB*10] STEENWIJK M., MILLES J., VAN BUCHEM M., REIBER J. H. C., BOTHA C.: Integrated Visual Analysis for Heterogeneous Datasets in Cohort Studies. *Proc. of IEEE VisWeek Workshop on Visual Analytics in Health Care* (2010).

[SWST03] SCHULZE-WOLLGAST P., SCHUMANN H., TOMINSKI C.: Visual analysis of human health data. *International Resource Management Association, Philadelphia* (2003).

[TSK05] TAN P.-N., STEINBACH M., KUMAR V.: *Introduction to Data Mining*. Addison Wesley, 2005.

[TYO*09] TUCER B. B., YALCIN B. M. B., OZTURK A. A., MAZICIOGLU M. M. M., YILMAZ Y. Y., KAYA M. M.: Risk factors for low back pain and its relation with pain related disability and depression in a Turkish sample. *Turkish Neurosurgery 19*, 4 (Sept. 2009), 327–332.

[VAS*11] VÖLZKE H., ALTE D., SCHMIDT C., ET AL.: Cohort Profile: The Study of Health in Pomerania. *International Journal of Epidemiology 40*, 2 (Mar. 2011), 294–307.

[vLBK*13] VON LANDESBERGER T., BREMM S., KIRSCHNER M., WESARG S., KUIJPER A.: Visual analytics for model-based medical image segmentation: Opportunities and challenges. *Expert Systems with Applications 40*, 12 (2013), 4934–4943.

[vTKB02] VAN TULDER M., KOES B., BOMBARDIER C.: Low back pain. *Best Practice & Research Clinical Rheumatology 16*, 5 (2002), 761 – 775.

[WP03] WOOLF A. D., PFLEGER B.: Burden of major musculoskeletal conditions. *Bulletin of the World Health Organization 81*, 9 (2003), 646–656.

[YWSC12] YU H., WANG C., SHENE C.-K., CHEN J. H.: Hierarchical Streamline Bundles. *IEEE Transactions on Visualization and Computer Graphics 18*, 8 (2012), 1353–67.

[Zag] A. Zagouras. L-method for Computing the Optimal Number of Clusters. www.mathworks.com/matlabcentral/fileexchange/37295-l-method/content/Lmethod.m.

[ZGP12] ZHANG Z., GOTZ D., PERER A.: Interactive Visual Patient Cohort Analysis. *Proc. of IEEE VisWeek Workshop on Visual Analytics in Healthcare, Seattle, Washington* (2012).

Interactive Visual Analysis of Image-Centric Cohort Study Data

Paul Klemm, Steffen Oeltze-Jafra, Kai Lawonn, Katrin Hegenscheid, Henry Völzke, Bernhard Preim

Abstract—Epidemiological population studies impose information about a set of subjects (a *cohort*) to characterize disease-specific risk factors. Cohort studies comprise heterogenous variables describing the medical condition as well as demographic and lifestyle factors and, more recently, medical image data. We propose an Interactive Visual Analysis (*IVA*) approach that enables epidemiologists to rapidly investigate the entire data pool for *hypothesis validation* and *generation*. We incorporate image data, which involves shape-based object detection and the derivation of attributes describing the object shape. The concurrent investigation of image-based and non-image data is realized in a web-based multiple coordinated view system, comprising standard views from information visualization and epidemiological data representations such as pivot tables. The views are equipped with brushing facilities and augmented by 3D shape renderings of the segmented objects, e.g., each bar in a histogram is overlaid with a mean shape of the associated subgroup of the cohort. We integrate an overview visualization, clustering of variables and object shape for data-driven subgroup definition and statistical key figures for measuring the association between variables. We demonstrate the *IVA* approach by validating and generating hypotheses related to lower back pain as part of a qualitative evaluation.

Index Terms—Interactive Visual Analysis, Epidemiology, Spine

1 Introduction

Epidemiology aims at characterizing health and disease by determining risk factors. Clinical problems, such as the selection of diagnostic tools and efficient treatment, are tackled using results of epidemiological research. The introduction of preventive measures in medicine and beyond is also based on epidemiological research, where, for example, subgroups with increased risk are identified [12]. Observations made by clinicians in the daily routine are translated into hypotheses for epidemiological research. These are used to determine environmental and lifestyle factors as well as medical examination results that may influence a disease. Potentially useful data variables are gathered using structured interviews and clinical examinations. Methods like regression analysis check the attribute list for statistical soundness.

Longitudinal population-based studies, such as the Study of Health in Pomerania (SHIP) [41], gather as much information as possible about a defined sample of people (a *cohort*). The cohort consists of several thousands of people, randomly selected to avoid any bias. The subjects are selected without focus on a certain disease. A large cohort size is essential to investigate differences between healthy and diseased people. Cohort studies often include medical image data. The concurrent analysis of image data and non-spatial epidemiological factors requires techniques that reach beyond standard statistical methods. For instance, segmentation of the image data is required for an analysis of anatomical structure and of possible correlations between this structure and epidemiological factors. Semi-automatic segmentation techniques are promising but also challenging, since the employed modalities, such as magnetic resonance imaging (MRI) and ultrasound, are subject to inhomogeneity and noise.

Compiling a list of variables for tests of statistical resilience based on experience-driven hypotheses leaves out other variables in the data which potentially interact with a disease. This also applies to the chosen landmarks used to quantify medical image data information. The standard workflow lacks methods for automatically identifying correlations possibly buried deep in the data or overseen by the expert. Also, only a small subset of factors can be concurrently analyzed.

We propose an Interactive Visual Analysis (*IVA*) approach [37] for the combined analysis of image and non-image data. Visual queries and direct feedback of Visual Analytics systems allow for a fast exploration of the data space incorporating many different variables. Intended as an extension to the well-established epidemiological tools it provides a way to rapidly validate hypotheses and to trigger *hypothesis generation* using data mining methods, such as clustering. *Hypothesis generation* gains importance since the number of epidemiological variables increases and the focus shifts towards more complex relations involving more than two variables. Our contributions are:

- an IVA workflow for cohort study data to allow both, hypothesis-driven analysis and hypothesis generation,
- visualization techniques, which incorporate both information visualization and 3D rendering of organ shapes as well as combining them with epidemiological graphics and key figures,
- highlighting subject groups and variable associations using shape-based clustering and statistical contingency measures.

We applied our approach to a data set compiled to analyze lower back pain and aim to determine variables, which indicate pathological changes. This data set comprises 127 variables and 2 sequences of MRI data from 6,753 subjects. The method implementation is web based to allow a fast feedback loop with domain experts.

2 Epidemiological Background

This section covers the epidemiological workflow and requirements.

2.1 Epidemiological Workflow

The diversity of epidemiology is reflected in the different experts who work at cohort studies, ranging from specialized doctors to medical computer scientists with focus on biometrics, and statisticians. Epidemiologists follow a workflow mainly driven by statistic tools to validate hypotheses about disease-specific risk factors. Following Thew et al. [36], the workflow can be characterized as follows:

1. A hypothesis is derived from observations made by clinicians in their daily routine.
2. A set of variables depicting conditions affected by the hypothesis is compiled accordingly.
3. Confounding variables are identified and taken into account (for example using stratification).
4. Statistical methods, such as regression analysis, assess the association of selected variables with the investigated disease.

- *Paul Klemm, Steffen Oeltze-Jafra, Kai Lawonn, Bernhard Preim are with Otto-von-Guericke University Magdeburg, Germany. E-mail: {klemm,oeltze,lawonn,preim}@ovgu.de*
- *Katrin Hegenscheid, Henry Völzke are with Ernst-Moritz-Arndt University Greifswald, Germany. E-mail: {katrin.hegenscheid,voelzke}@uni-greifswald.de*

Manuscript received 31 Mar. 2014; accepted 1 Aug. 2014; date of publication xx xxx 2014; date of current version xx xxx 2014. For information on obtaining reprints of this article, please send e-mail to: tvcg@computer.org.

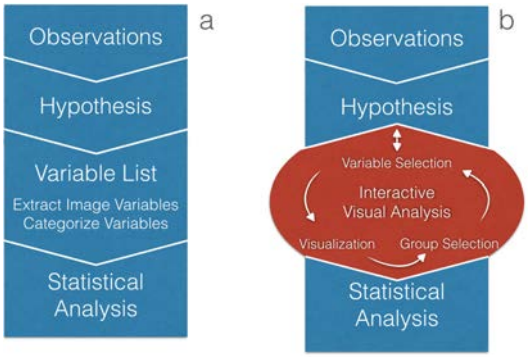


Fig. 1. (a) The standard epidemiology workflow consists of four steps. (b) *IVA* tools complement parts of this workflow instead of replacing them. The combination of statistical and interactive analysis shows promising potential to unveil information in the data. We call the iterative red highlighted part *IVA Loop*, described in detail in Figure 2.

The workflow is shown in Figure 1 (a) and serves as orientation for our approach. We focus on the potential of image data and attempt to support *hypothesis generation*.

Reproducibility of results is an epidemiological key requirement. It is difficult to achieve, since many physicians are involved when thousands of subjects are examined and interviewed. Thus, both intra- and inter-observer variability needs to be low for all aspects of a cohort study examination. Longitudinal studies require the acquired variables to be comparable for evaluation. Grouping subjects using epidemiological variables is essential in cohort studies to allow per-group risk determination. Grouping depends on the underlying hypothesis. Age, for example, is divided into groups (e.g. in 20 year steps) when investigating its influence. These groups strongly depend on the condition of interest and therefore there is no standard for their categorization.

Relative risks are determined to detect if a subject is prone to be affected by a certain disease. This includes confidence intervals indicating the certainty of that variable being a risk factor.

Statistical tools such as `SPSS`[1] play a major role for analyzing epidemiological data. Epidemiologists employ static graphical data representations primarily at the very end of an analysis session for presenting results or observing trends in the data.

2.2 Epidemiological Data

Epidemiological data are strongly heterogenous and incomplete. Information about medical history and examinations, genetic conditions, geographical data, questionnaire results and image data yield a complex data space for each subject. For ethical or medical reasons some variables cannot be gathered for each subject, e.g. women-specific questions about menstrual status or number of born children. Follow-up examinations or questions about conditions such as medications taken after a diagnosed disease also yield variables only available for a small amount of subjects.

Indicators for medical conditions as well as questions about a subject's lifestyle are often *dichotomous*–they have two manifestations (*Yes* or *No*). Dichotomous data can also be derived by aggregating variables to yield only two manifestations (e.g. subjects younger or older than 50 years). Medical examinations comprise categorical (e.g. levels of back pain) and continuous values (e.g. age or body size). Data analysis is usually carried out by calculating correlations, which is challenging due to the data type heterogeneity. Parameter correlation can also be associated with confounding, which cannot be automatically predicted. It has to be judged by a domain expert. Sparse populated variables are hard to assess statistically. Too few data samples may distort the real underlying distributions. Statistical correlations are prone to *confounding*, meaning that the association of two variables is influenced by a third variable, which needs to be isolated. A famous example is the association between shoe size and mortality, where it can be observed that people with larger shoe size have a smaller life expectation. The shoe size is actually associated with gender, where women have smaller feet and a longer life expectation.

Image Acquisition. Imaging techniques involving ionizing radiation for the subject are not suitable for ethical reasons. Therefore, MRI is the main method for collecting cohort study imaging data. The image quality is a tradeoff between accuracy and affordability [31]. This often yields image resolutions inferior to those of clinical practice.

Image Analysis. Decisions have to be made about *comparison* and *quantification* of image data. Segmentation masks representing the voxels of an anatomical structure would be ideal, since key figures, e.g., volume, largest diameter or aspect ratio, can be derived from them. Since reliable and efficient segmentation techniques are not available in general, epidemiologists are forced to measure the data by hand. Information derived by landmarks, such as top and bottom point of a vertebra, are by far not as expressive and versatile as segmentation masks describing its whole shape. They are also prone to a high inter-observer variability. Morphometric information from landmarks comprises thickness, diameter or length of a structure as well as grey value distribution in an area.

2.3 The Study of Health in Pomerania (SHIP)

After the pioneering Rotterdam study (started in 1990), several MR imaging study initiatives were initiated. They slightly differ in clinical focus, acquired data and epidemiological research questions. Starting in 1997 with a cohort of 4,308 subjects, the `SHIP`, located in Northern Germany, aims to characterize health and disease in the widest range possible [41]. Data are collected without focus on a group of diseases. This allows to query the data regarding many diseases and conditions. Subjects were examined in a 5-year time span, continuously adding new parameters including MRI scans in the last iteration [16].

3 Prior and Related Work

This section describes prior and related work and covers visual analysis methods incorporating both image and non-image data.

Visual Analysis of Image and Non-Image Data. Our work is closest to that of Steenwijk [35], Turkey [39], and Angelelli [1] and colleagues, who employ multiple coordinated view systems for the analysis of cohort study data.

Steenwijk et al. [35] propose a relational database to organize cohort study data for a visual analysis based on linked views such as parallel coordinates, scatter plots and time plots. Information about medical image data is incorporated via mappers, which extract comparable metrics about the data. Medical image data can be displayed individually for subjects, e.g., for analyzing outliers. While we use a similar approach when analyzing non-image data, our process also includes *overview visualizations* and statistical suggestions of potentially interesting variables.

Turkey et al. [39] present *hypothesis generation* based on descriptive statistics of the data dimensions. Key figures describing the distribution of data values, e.g., standard deviation and interquartile range, are computed per dimension and analyzed by pairs in a *deviation plot*. The *dual analysis* of data items and dimensions in multiple linked views led to several hypotheses in analysis sessions with domain experts. Hypotheses based on observations in the deviation plot may impose *overfitting* to the data because the measures highlight only parts of the statistical changes. Our approach uses information extracted from the segmented image data (such as 3D meshes) and variable associations with non-image epidemiological factors.

Angelelli et al. [1] focus on the data organization for an interactive visual analysis of heterogeneous cohort study data. The proposed data-cube model facilitates the seamless integration of image-based and non-image data. In a demonstration of the model, brain image data was integrated into the analysis by first segmenting brain regions and tracking neural pathways and then deriving attributes from both, e.g., volume and fractional anisotropy. A multiple coordinated view

[1] Product of IBM; `ibm.com/software/analytics/spss/`

framework then linked spatial and non-spatial data views. Our integration of image data into the analysis is similar to the work of Steenwijk [35] and Angelelli [1] and colleagues. While they offer a single spatial view for visualizing image-based information of one subgroup of the cohort, we provide multiple views showing the information of subgroups and their respective deviation from the entire cohort.

Gresh et al. [14] proposed WEAVE, one of the first systems which concurrently analyzed medical image and non-image data using linked views. Blaas et al. [2] presented a similar system, which also analyzed medical image data and variables derived from them using views from the variable and physical space. Both works are restricted to the analysis of one case at a time and to non-image data with a unique spatial reference, e.g., voltage simulated across the heart muscle. In epidemiology, multiple cases must be concurrently investigated and non-image data often lacks a spatial reference, such as gender or age.

Visual Analysis of Heterogenous Non-Image Data. Zhang et al. [44] provide a web-based system for analyzing subject groups with linked views and batch-processing capabilities for categorizing new subject entries into the data set. Their definition of a cohort differs from the understanding of the term in an epidemiological context by denoting every parameter-divided subject group as individual cohort. Due to the short paper length, detail is missing on the data types and their algorithms of identifying similar subjects or whether they employ statistical measures. We employ the idea of adding variables via drag and drop into a canvas area.

Generalized Pairs Plots (GPLOMS) are an information visualization technique comparing heterogenous variables pairwise using a plot-matrix grouped by type [10, 19]. They are useful to gain an overview over numerous variables and their distributions. Histograms, bar charts, scatter plots and heat maps are used to visualize variable combinations with regard to their type. The resulting matrix provides an *overview visualization*, but requires a lot of screen space for many variables (127 in our application scenario). We incorporate the idea of adaptive type-dependent visualizations. Dai et al. [9] explored risk factors by incorporating choropleth maps of epidemiological variables (e.g., mortality rates in a region) with parallel coordinates, bar charts and scatter plots with integrated regression lines. Their findings yielded a *Concept Map*, which linked cancer-related associations via graph edges. While their goal to identify possible risk factors using socio-economic and health data is similar to ours, they focus on iteratively refining defined hypotheses and on geographical data. We employ the use of small multiples for incorporating heterogenous data types for comparability. Chui et al. [6] visualized associations in time-dependent epidemiological data using time-series plots highlighting risk factor differences in age and gender. While the work shows how different visualization techniques provide insight into these data sets, it focuses on the time aspect, which is not present for our data.

Visualizing Shape Variance. Comparing tissue between many subjects requires shape variance visualizations. Caban et al. [5] investigated the suitability of variance visualizations of shape distribution models and concluded that users favor spherical glyph representations over deformation grids and likelihood volumes. The distribution of shapes in a space derived from a PCA is plotted by Busking et al. [4] in a 2D-projected plane of the space. Interpolated views can be created by the user in a separate view as well as comparisons in a contour view. Interpolation is carried out by mesh morphing. The distance to the mean shape is color-coded. We incorporate the idea of combining 3D shape rendering with information visualization techniques. Applying this technique to our data yielded a cluttered shape space due to the high subject count. The data needs to be abstracted to work in this context. Hermann et al. [18] identify local deformation changes by investigating shape-related differences on rodent mandibles. User-specified regions of interests are mapped to associated anatomic covariation using tensor visualization. This method enables rapid hypotheses validation and was able to reproduce textbook knowledge. It requires a spatial colocation of associated variables.

Prior Work. We visualized lumbar spine variabilities based on a semi-automatic shape detection algorithm of 490 participants of the SHIP-2 cohort [21]. Hierarchical agglomerative clustering divided the population into shape-related groups. As proof of concept, a relation between the size of the segmented shape and the measured size of the subjects was shown. This work focuses on incorporating these derived data as new variables, enabling to include it into the hypothesis validation and generation process. When applying clustering techniques to the non-image data it was found that k-Prototypes and DBSCAN are appropriate, but are strongly dependent on the chosen variables and distance measures [20]. Niemann et al. [26] presented an interactive data mining tool for the assessment of risk factors of hepatic steatosis, the fatty liver disease. Association rules created by data mining methods can be analyzed interactively with their tool and highlight potentially overlooked variables.

Interactive Visual Analysis. The strength of the *IVA* approach is its versatility with respect to the application field [22]. Oeltze et al. proposed a multiple coordinated view approach for the analysis of medical perfusion data [27] and biological multi-channel fluorescence microscopy data [28]. The approach is restricted to the investigation of a single subject at a time.

Lammarsch et al. [24] provided a workflow and terminology definition of Visual Analysis techniques. They define a model as a representation of system entities, phenomena, processes and hypotheses as models whose outcomes are not compared with real-world data (*validation*). The VA process is also reflected in our *IVA* loop.

Baldonado et al. [42] presented rules for designing multiple coordinated views. They point out the cost-benefit tradeoff introduced by the cognitive overhead by mentally connecting multiple views over more complex single views. Weaver et al. [43] extracted guidelines for cross-filtering multiple views by incorporating *views* mapping data to visual elements, *brushes* for selecting these elements and *switches* for linking brushing results between *views*. Our system follows the same rules for selecting subject groups, but our goal is to judge variable relations and potential outcomes.

The uniqueness of our workflow compared to the discussed work is threefold. (1) We incorporate 3D models abstracting shape information fused with non-image data visualizations, allowing to analyze local physiological changes related to non-image parameters. (2) We focus on hypothesis generation by discovering new relationships associated with shape information. (3) Overview visualizations using statistical abstractions aim to provide an unbiased variable relationship assessment.

4 Image-Centric Cohort Study Data in an Interactive Visual Analysis Context

We described the epidemiological workflow and emphasized the reproducibility and statistical integrity (recall Subsection 2.1). Introducing the *IVA* principle to the epidemiological domain aims to compensate the weaknesses of the existing workflow rather than replacing it (recall Fig. 1). In the current state, the workflow treats the data like a black box. Statistical tests on variables associated to a hypothesis yield a value for deciding whether the data supports the hypothesis. Variables not included in the analysis may potentially support the chosen hypothesis by discriminating the population in the expected way, but are not highlighted. This becomes even more important when the workflow is adapted to the analysis of the medical image data, where domain experts annotate landmarks tediously to derive measures, such as diameters. This leaves out the majority of information in the image data by abstracting it to single values. Considering more complex parts of the data would make those results more trustworthy and could also identify possible anatomical confounders–an epidemiolgical research result in itself.

IVA tries to illuminate the black box by making the domain experts part of an iterative variable selection process (see Fig. 1 b). It also aims to project back into the hypothesis formulation step to amplify hypothesis generation. This has to be handled with care, since *overfitting* of expectations to the data is an imminent danger [39].

Domain and Range Variables. In the *IVA* context, data are characterized by a combination of independent variables, such as space and/or time, and dependent variables, like temperature or pressure. Two kinds of views are employed to inspect the data:

- *physical views* [29], e.g. volume rendering, show information in the context of the spatio-temporal observation space [27], while
- *attribute views*, such as scatter plots and parallel coordinates, show relationships between multiple data attributes.

Transferred to epidemiological data, the residential area of cohort subjects could be interpreted as *space*, the different assessment cycles of a longitudinal study as *time*, and the image and non-image data as *dependent variables*. Our current work neglects geographical and temporal aspects. Instead we employ an abstract model and consider the subjects as living in a joint image space where each of them is represented by a segmented organ or structure. For instance, the lumbar spine is segmented over all subjects and all lumbar spines are co-registered spanning a joint space. Then, two types of dependent variables exist: the socio-demographic data and medical examination results, and variables derived from the segmented structures, e.g., spinal curvature or misalignment of the vertebrae. An alternative of the image space would be the shape space generated by extracting the major modes of variation from all segmentation results [4]. Based on our abstract model, the three analysis patterns of *IVA* can be employed.

Local Investigation refers to the inspection of dependent variables with respect to subsets of the image or shape space. For instance, the epidemiologist selects several lumbar spines with a common characteristic in the image or shape space and inspects the associated dependent variables in an attribute view [18]. The selection step requires dedicated interaction techniques for defining a subset. Alternatively, derived shape-related variables opposed in an attribute view or automatic techniques for shape clustering may be employed [21]. Clustering algorithms can be used to investigate associations between shape groups and other non-image variables. Analysis of outliers can indicate segmentation errors or a group of subjects sharing a pathology.

Feature Localization refers to the search for structures in the image or shape space with a defined characteristic. The epidemiologist may be interested in all female subjects with lower back pain and wishes to see the corresponding spines in a physical 3D view.

Multivariate Analysis refers to an investigation of multi-variate properties of the dependent data by specifying a variable in one attribute view while analyzing the value distribution with respect to other variables in other attribute views. Epidemiologists may define a variable in a scatter plot of the body mass index (BMI) and age to inspect the result in a histogram of body height. These associations may also be summarized using pivot tables, which are widely used in epidemiology.

4.1 Data Preprocessing

Non-Image Data. Data obtained using questionnaires or medical tests are often stored using statistical packages such as `SPSS`, which have a proprietary data format. Exporting the data in the respective tool to a `CSV` file and then converting it to file types that are easily manageable, such as `JSON`, makes it readable for modern programming languages. A data dictionary stores information about each manifestation of a variable. Detailed description of data variables, its meaning as well as unit of measurement are stored as a lookup table. Missing data are denoted using error codes indicating their cause ranging from ethical to medical and personal issues (recall Subsection 2.2).

Image Data. Information about anatomical structures, such as diameter or volumes, is extracted from the image data. This is either done manually by experts setting, landmarks or by a (semi-)automatic detection, registration and segmentation. These algorithms have to deal with a large inter-subject variability of the anatomical structure [31]. In principle, model-based approaches are effective for detection [32] and segmentation [13]. If a segmentation yields only binary masks, algorithms such as *Growing and Adaptive Shapes* can be applied for creating a surface grid where each point is comparable throughout the population [11]. Grey value comparison is used to measure the quantity of fat, water, and–application-specific–the iron content (liver) or the distribution of grey and white brain tissue. Morphometric variables are derived to allow for statistical comparison of the tissue, which incorporates mostly positions, diameters, volumes, relative distances and alignment to other structures.

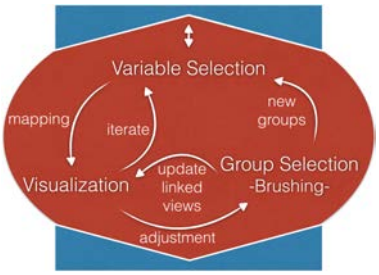


Fig. 2. Detailed *IVA Loop* as extension from Figure 1. Usually starting with a selection of a variable of interest (user-driven or via data mining techniques), the data are mapped using a visualization technique appropriate for the selected data types. The data are visualized in the range and domain space, which can be brushed, yielding new groups to be investigated using further variables. Note that adjacent steps are directly connected via feedback loops, allowing for an iterative refinement and giving as much freedom to the user as possible.

4.2 Analysis Workflow

Our proposed *IVA* workflow consists of three major steps, as illustrated in Figure 2: Variable selection, visualization and brushing. A hypothesis-driven analysis usually starts with the selection of variables or shape groups derived from a shape-based clustering. Hypothesis generation with focus on image data starts with a shape-based clustering or an *overview visualization* of all variables. The variable is mapped using an automatically chosen visualization appropriate for its data type (described in detail in the following section). The visualization techniques have to combine both image- and non-image data to set domain and range data in relation to each other. In our system, the visualization can either be brushed or new variables can be added to the analysis. Brushed regions are treated like categorical variables, as they divide the subject space in the same way. Selecting variables also triggers a *multivariate analysis* using contingency values (described in the following section) to highlight associated variables. A sample workflow using interaction and visualization techniques described in the next section can be seen in Figure 3.

5 System Design and Implementation

The suitability of visualization techniques for epidemiological data depends on their ability to compare multiple data variables while highlighting associations. Visual evaluations of data are therefore as important as methods allowing for numerical data analysis. In the following sections we present the different parts of our system.

5.1 Design and Visualization Techniques

Early it became clear that we have to rely a lot on online communication due to the large spatial distance towards each other. Hence, we built our system using web technologies. By running the prototypes on server machines, software exchange became as easy as sharing a weblink, giving us the opportunity to include the clinical experts in the development process with little effort. Incorporating the *IVA* workflow for image-centric cohort study data requires *overview visualizations* as well as *multivariate visualizations*, which bring image-derived information in context to non-image variables.

The focus on web technologies is not without tradeoffs. Classical UI elements, such as the menu bar or custom right-click menus, are technically possible, but not common in this domain. In favor of a clean layout, we designed the system without such components. Since

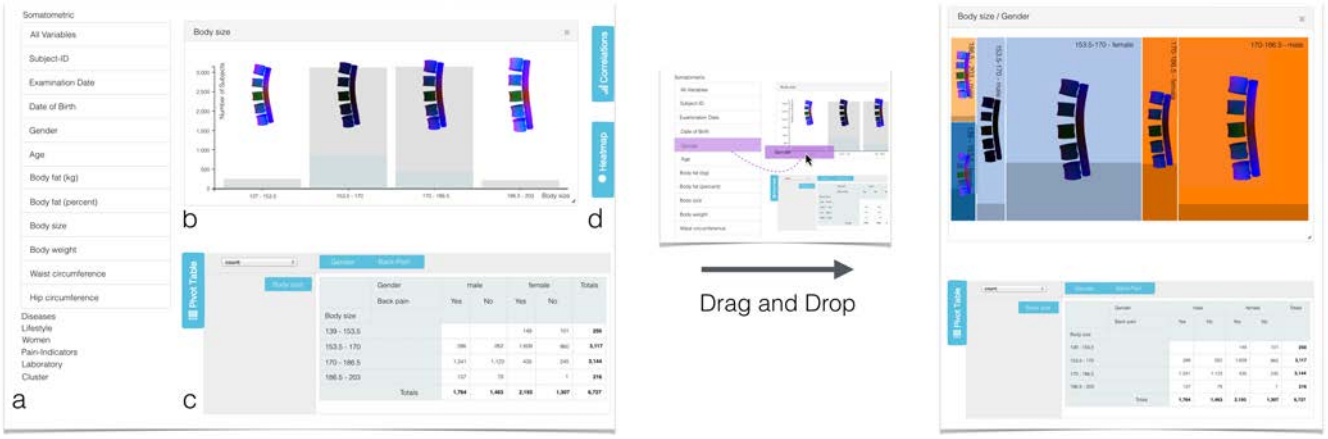


Fig. 3. (Left) Screenshot from the front-end, which is divided as follows: (a) The sidebar containing all variables as well as the groups defined in the analysis process; (b) the canvas area where variables can be added via drag and drop and the visualization is chosen automatically according to the data type; (c) the interactive pivot table showing the exact numbers for each displayed variable combination; (d) buttons to open panes containing the contingency matrix, contingency pane and pivot table. The data displayed is used to analyze the lumbar spine. Variables can be added freely on the canvas via drag and drop. Dropping the *gender* parameter on the already plotted *body size* container creates a mosaic plot combining both variables (right). In a prior step, the user selected all subjects with diagnosed thyroid disorder. These subjects are shown as shade in the visualizations, denoting their share. Subjects between 153.5-170 cm body size are more affected by thyroid disorder (box plot) and are mostly female (mosaic plot). Distance to the mean mesh of subjects with thyroid disorder is encoded as red for x axis, blue y axis and green z axis.

the previously described *IVA* workflow allows for many different ways to analyze the data, we designed the interface as minimalistic as possible, treating the resulting space as *canvas* for the data. We divided the workspace into four major parts, as illustrated in Figure 3 and 4.

- The *sidebar*, which contains all epidemiological variables. The cluster results group variables like categorical variables and are part of the sidebar as well (Fig. 3 a).
- The *canvas* holding all visualizations. Elements can be added, arranged, resized and removed freely (Fig. 3 b).
- The interactive *pivot table* gives detailed numerical information of the variables in the canvas view. This view on the data is familiar to epidemiologists (Fig. 3 c).
- The *contingency view* depicts relations for variables in the canvas in an contingency matrix (Fig. 4) and a *contingency list*.

System Layout. We experimented with several layouts. The initial idea was to make all components freely arrange- and resizable on a large *canvas* area. This idea was soon dropped, since domain experts reported a cluttered workspace, which required a lot of scrolling. The introduction of separate panes for the contingency matrix, pivot table and sidebar, displayed with a mouse click on the corresponding button and sliding on top of the *canvas* was considered more feasible (Figure 3 shows the system with reeled-out pivot table pane). All user-generated visualizations are part of the *canvas* and can be arranged freely.

Sidebar. Only the *sidebar* is visible at system start. It categorizes all variables into different types, such as somatometric (measurements of the human body dimensions), disease- or lifestyle-related, pain indicators and laboratory data (Fig. 3 a). It also contains subject groups defined by automated shape clustering. Groups are treated like dichotomous variables. Variables can be dragged from the sidebar into the canvas area for a *feature localization*, which works as follows.

Adaptive Variable Visualization. The visualization type, inspired by `GPLOMS` [10, 19], is dynamically chosen based on the variable types and number to allow for *multivariate analysis*. Categorical data are either mapped to bar charts (single variables) or mosaic plots (multiple variables). Figure 3 describes this dynamic adjustment. Continuous data can be visualized using scatter plots (two variables) or parallel coordinates (multiple variables), but in epidemiology, this data type is usually categorized into ordinal groups of *equal size*. Since the number of categories often depends on the hypothesis, the discretization steps can be adapted dynamically. Too many groups potentially generate sparse bins not suited for statistical evaluation. Not enough groups overgeneralize information. Adaptive discretization is an option, but imposes possible overfitting to the data. Conclusions based on statistical relationships derived from groups already biased by variable distribution are heavily influenced by the used discretization. Therefore, we follow the convention to use bins of equal size.

Following Tufte's concept of *small multiples* [38], information derived from the medical image data are incorporated into the plot by including color-coded mean shapes for each manifestation (Figure 3 b). The 3D plots can be navigated using standard mouse input, the camera is synchronized between all views to enable direct comparison. The distance from a group mean shape is mapped to the global mean using color. This allows to assess local shape changes (Fig. 3) and is an important information to the epidemiologist. Until now, epidemiologists were not able to inspect shape differences based on non-image variables. Dropping a variable on an existing plot adapts the visualization dynamically to allow for comparison (Fig. 3 right).

To support *feature localization*, subject groups can be brushed via a double-click on its representative in the visualizations. Holding down the shift key allows to select multiple manifestations. Brushed groups act as reference for the shape visualization, calculating distances based on the mean-shape of the brushed selection. The share of subjects of this subgroup is linked to all other views (Fig. 3 left). If the user selects all female subjects in a visualization of gender distribution, all other displayed meshes are color-coded with their distance to the female mean and the share of female subjects is highlighted in the information visualization.

Pivot Tables. Epidemiologists are used to perform *multivariate analysis* of groups based on table representations. Thus, we decided to introduce an interactive pivot table. These tables clearly convey the subject count in each group (see Figure 3 c). However, they quickly get confusing when they are divided into many subgroups. We tackled this problem by making the order and number of displayed variables adaptable. This also applies to the assignment of row or column variables. Another way to avoid clutter is the user-driven selection of displayed variables. To allow for better comparison with respect to variables, the values of each cell can also be displayed as percentage of the variable represented of either the row or column.

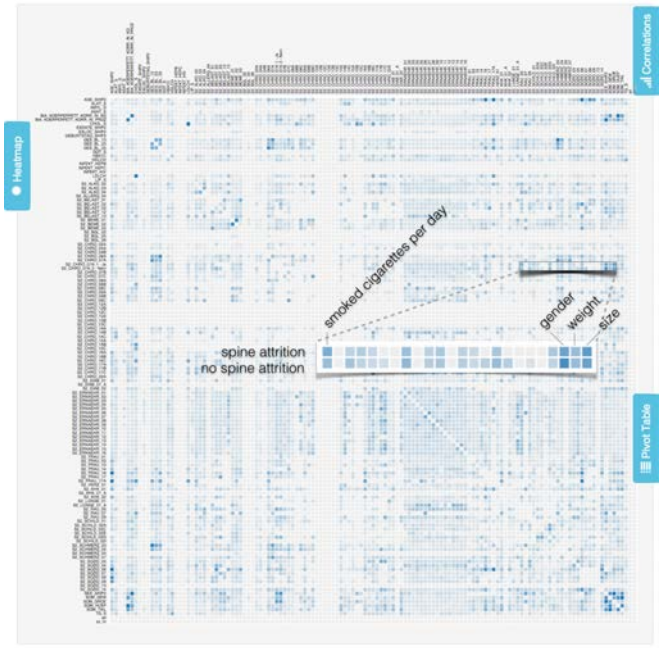


Fig. 4. Contingency matrix of 129 variables (127 data set variables, 2 cluster results) showing 16,641 combinations. Similarity is calculated using the *Cramér's V* contingency value. Color brightness encodes association strength. Moving the mouse over an entry enlarges the variable names for better readability. The enlarged excerpt shows associations for shape clusters of subjects with and without diagnosed spine attrition, which show associations between gender, weight, body height and smoking behavior. The contingency pane is not shown here.

Automated Variable Suggestion using a Contingency Matrix. Highlighting potentially interesting associations in the data set is one major benefits of the *IVA*-powered approach and is part of the *multivariate analysis* pattern for analyzing variables outside the shape space. Turkay et al. [39] used the approach to calculate statistical key figures based on the distribution functions of each variable derived from the image data. Since the majority of our data are categorical variables, we have to employ different solutions. The *Cramér's V* contingency coefficient can be used to calculate coherences between categorical variables [8]. It is based on *Pearson's* X^2 distribution test [30], which uses contingency tables holding the counts of subjects for all possible manifestations of two variables. *Cramér's V* is defined as:

$$V = \sqrt{\frac{X^2}{N(k-1)}}, \tag{1}$$

where X^2 equals *Pearson's chi squared*, N is the total number of observations and k is either the row or column count, depending on which one is lower. V yields values between 0, meaning that two variables are completely independent, and 1 indicates that they are the same. *Cramér's V* does not allow to infer the dependency direction.

It shares the same restrictions as *Pearson's* X^2. The expected counts in the contingency table have to be larger than five for 80% of the entries and no expected value must be smaller than one [7]. Some manifestations and variable combinations, which are only exposed by small subject groups, cannot be assessed with this technique. They cannot be included into the epidemiological analysis, since statistical validation needs a minimum count to be valid. The contingency matrix highlights correlations between all variables. This aims to highlight variables possibly associated with the focused hypothesis and to trigger new hypotheses. Contingency is visualized using an interactive contingency matrix with association power mapped to color brightness. The distinction whether an association is a confounder or an effect depends on the context defined by the hypothesis and is a decision to be made by the domain expert. The contingency matrix visualization is an *overview visualization*–something the epidemiological community lacks and is in great need of.

Contingency Pane. Dropping a variable into the canvas area adds an entry for each manifestation of it to the *contingency matrix*. Testing sessions revealed that it was tedious to open the matrix every time a new variable is added. As a consequence, the *contingency pane*, a table containing correlating variables for the last added visualizations in descending order of the *Cramér's V* value was added. *Contingency pane* entries can be dragged and dropped into the *canvas area* just like variables in the *sidebar*.

Initialization and Clustering. Using variable suggestion allows to initialize the system with a set of potentially interesting visualizations. After testing and domain expert feedback we dropped this idea. Reasons for this are twofold. Very often, high correlations are obvious, such as gender with menstrual status. Also, we observed that the variables of interest are dependent on the specialization of the domain expert (explained in detail in Section 6).

Subject clustering is triggered automatically as *local investigation* for a variable after it was added to the canvas by the user. A status indicator at the bottom of the screen keeps the user informed about the pending clustering result, since the process can take up to ten seconds. Clustering results are listed in their own category in the *sidbar*.

5.2 Implementation

Fig. 5. The front-end solution (left) uses *HTML5*/*CSS3*, *WebGL* and *SVG* to display the data. The *NodeJS* based back-end (right) stores all image and non-image data and transfers it to connected clients. All computation-heavy operations, such as calculation of mean shapes or distances, are performed on the server side. Client-server communication is accomplished via the *Websocket* protocol.

In this section, we discuss how we implemented the presented methods using open web standards. To provide a fast communication loop between method development and expert input, we decided to rely on modern web technologies. In addition to the obvious advantages of web technologies, the following aspects are crucial for our work:

- The client-server structure allows for employing heavy computation on a server machine and transferring results to the client.
- Disk-space demanding image data remains on the server and elements can be transferred on demand. High confidentiality standards of the data are met by a password protecting the access.
- Recent developments in *WebGL* applications running in browsers with near-native performance results in many open source libraries, which are well documented and driven by active communities. We use *WebGL* for rendering shape information.

These advantages do not come without drawbacks. Sophisticated libraries/languages, such as the `Visualization Toolkit`[2] or `R`[3] for statistics, are either not available at all or only accessible through complex client-server systems. Therefore, many standard methods had to be written from scratch. The back-end is realized using `NodeJS`[4],

[2]Developed by Kitware Inc; `vtk.org`
[3]Open Source; `r-project.org`
[4]Developed by Joyent Inc, `nodejs.org`

which is based on the Google V8 Javascript runtime environment. Due to its event-driven non-blocking I/O model it is fast and responding even with heavy workload, such as mesh processing. Non-image data for all subjects including the data dictionary is stored in a `JSON` file on the server. Image data are available as raw `DICOM` files. Segmentation masks of anatomical structures are represented as meshes, suited for comparing subjects. The requested data are transmitted when a client connects. The server performs heavy statistical tasks, such as calculation of *Cramér's V* values for all variable combinations in order to keep the computation time on the client as low as possible.

The front-end is created using `Bootstrap`[5] as foundation for the layout and basic UI elements using `HTML5`, `CSS3` and `Javascript`. Information visualizations such as scatter plots and bar charts are created using the popular `Data-Driven Documents (D3.js)` library [3], which works well for attaching data to visible elements like vector graphics and provides powerful transformation and mapping tools. The pivot table implementation uses `PivotTable.js`.[6] `Three.js`[7] allows GPU-accelerated data rendering using `WebGL`. The WebSockets protocol handles the client-server communication. Since our clustering algorithms are written in `MatLab`[8], we had to access them using the `NodeJS` server. We accomplished this by converting it to a parameterized standalone console application, spawned by `NodeJS` on client request. The result is read from the console output and is returned to the client. All parameter-steered console applications can be incorporated in this context.

6 Application

This section describes how the presented *IVA* workflow is used in the epidemiological application. We conducted a qualitative evaluation with two domain experts on a data set compiled to analyze lower back pain. This is one of the most common diseases in the Western civilization [40]. Epidemiological analysis of lumbar back pain, such as the work of Harreby et al. [15], is largely focused on non-image information. In comparable studies, only a few shape-related variables are included [25]. Determining risk factors in this area can lead to particularly affected risk groups, prognostic variables for diagnosis and treatment of lumbar back pain and a better understanding of effects of preventive measures, such as occupational health and safety regulations [12]. Characterizing the healthy aging process of the spine is a long-term goal for determining age-normalized probabilities for spine-related diseases by incorporating individual risk factors.

6.1 The Lumbar Spine Data Set

There are 127 variables describing diagnosed diseases, lifestyle factors, women-specific factors, pain indicators, laboratory values and somatometric variables for 6,753 subjects (4,420 from `SHIP-Trend-0` and 2,333 from `SHIP-2`). Since data acquisition protocols between these two cohorts are identical, the variables between the two cohorts are comparable. The data contains 30 metric, 7 nominal, 29 ordinal and 62 dichotomous variables. Somatometric variables include measures of the human body, such as body height, weight and body fat percentage as well as gender. These measures are reliable and complete. Other variables, such as pain indicators or lifestyle indicators (e.g. physical activity) are more subjective and less reliable. There are also variables missing for each subject, such as variables building upon each other (e.g. Do you have high blood pressure? Which medication is prescribed against it?). Therefore, some manifestations are sparsely populated, which makes statistical evaluation challenging.

The MRI data was acquired for each subject on a 1.5 Tesla scanner (Magnetom Avanto; Siemens Medical Solutions, Erlangen, Germany) by four trained technicians in a standardized way. The spine protocol consisted of a sagittal T1-weighted turbo-spin-echo sequence ($1.1 \times 1.1 \times 4.0$ *mm* voxels) [17].

[5]Developed by Twitter, `getbootstrap.com`

[6]Developed by N. Kruchten, `nicolas.kruchten.com/pivottable`

[7]Originally developed by R. Cabello, `threejs.org`

[8]Owned by The MathWorks, `mathworks.com`

6.1.1 Data Preprocessing

The data processing follows the description in Section 4.1.

Non-Image Data. To ensure a fast and easy data access outside of statistical processors like `SPSS`, the data was exported to the `JSON` file format. Since it lacks export methods for data dictionaries, we used `SPSS` to export our data to the `SAS v9+` format, which saves the data labels, and exported the data values as non-labeled `CSV`. A short script combined both data sources to a `JSON` file. The data types had to be transferred manually. Each variable is stored as an object containing the data as array, its data type, a detailed description and data dictionary translating value or error IDs to values. Continuous variables are discretized to allow for *Cramér's V* contingency coefficient assessment. Following epidemiological publications, we set the number of groups to five, to allow for contingency assessment.

Image Data. The lumbar spine was detected in the image data using a hierarchical finite element method by Rak et al. [32]. This semi-automatic method requires the user to initialize the tetrahedron-based finite element models (FEM) with a click on the L3 vertebra. Two user-defined landmarks on the top and bottom of the L3 vertebra describe an initial model height estimation. The model uses a weighted sum of T1- and T2-weighted MR images to detect the lumbar spine shape. Once registered, it captures information about the shape of the lumbar spine canal as well as the position of the L1-L5 vertebrae [21]. Due to incorrect initialization, strongly deformed spines, contrast differences and artifacts, the model was not able to detect lumbar spines for all subjects. We obtained and worked with 2,540 tetrahedron models of the lumbar spine. For clustering, we extracted the centerline of the lumbar spine canal, which captures information about lordosis and scoliosis (the medical terms for spine curvature) [21].

6.1.2 Shape Visualization and Clustering

The tetrahedron-based detection model consists of corresponding grid points for each structure instance. This allows to calculate shape distance and similarity. This information is used to calculate mean shapes as described in Section 5. The shape distance between meshes is mapped to color (recall Fig. 3).

Shape-based clustering is carried out via agglomerative hierarchical clustering of the spine canal centerlines (recall Section 6.1.1 and [21]). Since the "correct" number of clusters in a given group is unknown, an estimate is computed by means of the knee/elbow method [33]. The method has proven to produce sound results on a preliminary data set and was able to reproduce textbook knowledge [21].

6.2 Participants, Setup and Procedure

Inspired by Lam et al. [23], we conducted an investigation of *Visual Data Analysis and Reasoning* (*VDAR*). This approach aims to characterize the system's ability to explore data, discover knowledge, generate hypotheses and help formulating decisions. Since it is hard to quantify these outcomes, Lam et al. suggest case studies for the *VDAR* by applying the think-aloud protocol to understand the domain expert's observations, inferences and conclusions when using the system.

Our participants are two epidemiological domain experts who also co-authored this publication. HV and KH are physicians with focus on epidemiological research. HV is a specialist in internal medicine (23 years of experience) and head of the SHIP, KH a radiologist (9 years of experience) and responsible for the SHIP MRI data acquisition.

Setup. Due to the large geographical distance, the evaluation was done completely web-based. The experts accessed the prototype by entering the weblink into their browser. User input was observed using screen sharing. Communication was enabled via webcam-supported voice over ip. The total setup time including installing the screen sharing application was about five minutes. Video recordings of the sessions allowed a detailed evaluation afterwards.

Procedure. At first, we controlled mouse and keyboard of the participants' PC and demonstrated the basic functionalities of the system. As they understood the concepts, we handed over the mouse and

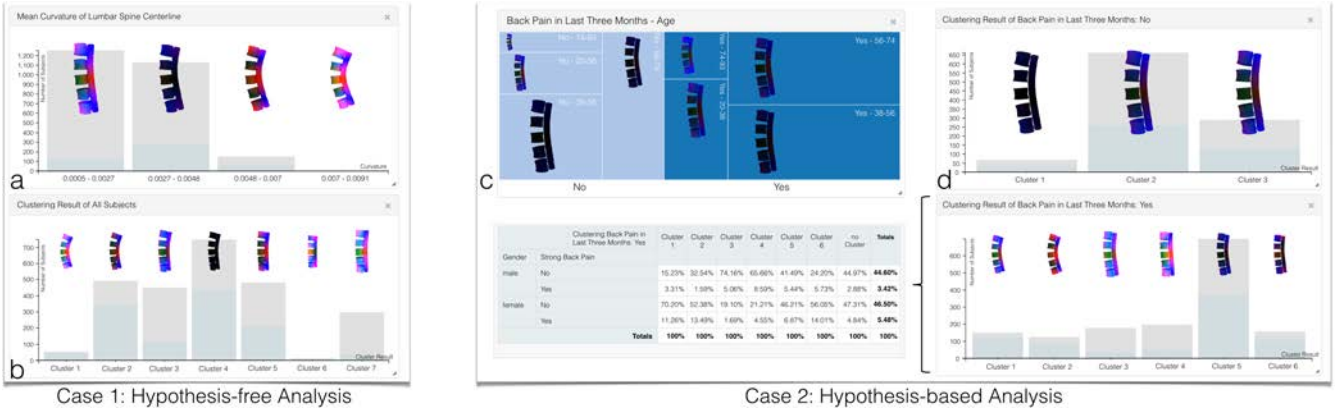


Fig. 6. Various case study results. (a) Mean curvature of lumbar spine canal plotted against the mean shape of 58-74 years old female subjects (light-blue bars). Note the high amount of this subject group relative to the total count in the third group. The last group contains four outliers. (b) Clustering of all subjects yields seven groups, whereas Cluster 4 assembles the mean. The light blue bars indicate the share of females in the group. (c) A mosaic plot mapping age against the dichotomous questionnaire answer to "*Did you experience back pain in the past three months?*". (d) Clustering result of "*Did you experience back pain in the past three months?*" Yes/no with female share in each group. Cluster 1 and 6 for answer "*Yes*" contain mostly women. The pivot table shows how many subjects with strong back pain are in each cluster for answer "*Yes*". Subjects in Cluster 1, 2 and 6 report strong back pain more often than subjects in other clusters.

keyboard control and only observed from this point on. The epidemiologists were given two tasks: one hypothesis-free analysis of the data and one starting with an assumption. For each case we conducted one analysis with each expert.

6.3 Case 1: Hypothesis-free Analysis

Analyzing the data set without prior hypothesis requires a starting point giving an overview over the data [34]. With our tool, there are two ways to achieve this. Performing a *multivariate analysis* by viewing the contingency matrix or a shape grouping step using shape-based clustering. The first was chosen by both experts. Before, they were not able to look at all variables in the context of each other. To cite one expert, the contingency matrix "illumunates the data black box", making it possible to look at the data unbiased of assumptions.

Analysis 1. The radiologist (KH) was looking for correlations with shape-related variables in the data, finding that *spine curvature* correlates with *leg pain*, *age*, *body height* and *hormone replacement therapy status*. Due to the dense mapping of information in the contingency matrix, KH suggested to make this visualization full screen.

After this initial overview, KH performed a *multivariate analysis* by introducing variables, such as *age*, *waist circumference*, *weight* or *lumbar spine canal curvature* as bar chart views into the canvas area and selected subgroups to see how they are distributed and if they could observe unusual behavior in the mean shapes. This pointed out problems with the used categorization method splitting numerical variables into equally-sized ordinal bins. If a variable contains outliers, such as *waist circumference* (e.g. by subjects with morbid obesity), this approach leads to sparse categories, making it hard to calculate associations. The proposed expert solution for this is categorization using quantiles/quartiles and is described in detail in Section 6.5.

A *multivariate analysis* using the *Cramér's V* contingency values for subjects with strong lumbar spine curvature showed, that these subjects are primarily females between 58-74 years who also report pain radiating from their back into other body regions Figure 6(a).

Analysis 2. HV also started with a *multivariate analysis* using the contingency matrix to analyze non-image variables, such as age-associated parameters like *income*, *blood fat values* or *number of born children*, but found no associations of interest. Therefore, he applied the *local investigation* pattern by a shape grouping step using shape-based clustering via dragging *All subjects* from the sidebar into the canvas area, triggering the shape clustering (Fig. 6 b).

Cluster 4 represents subjects with average shape. Other shapes differ with respect to size, such as cluster 2, 3, 7, whereas the last one and cluster 5 also represent a more straight spine, which is usual for subjects with larger body size. Cluster 1 and 6 contain outliers, characterized by their unusual shape and small number. To get an overview of the suggested variables, the user opened the contingency pane (not shown here) to perform a *multivariate analysis* by looking at *Cramér's V* contingency values of all clusters, revealing a strong correlation with *gender* and *body size*. Therefore, another *multivariate analysis* was carried out by dragging the *gender* variable to the canvas and selecting all female subjects (Fig. 6 b). Cluster 1 contained primarily female subjects. Contingency values for this cluster revealed correlations with *leg fatigue*, *physically heavy work*, *body weight*, *dyspnoea* and *headache intensity*. Since it is a pain indicator, headache was of special interest and was further investigated by incorporating a pivot table setting *headache intensity* in relation to cluster affiliation. It was found that cluster 1 subjects report heavy headaches more frequently than other subjects.

The experts emphasized the importance of methods providing an overview over the data for hypothesis generation. With the presented *IVA* approaches they were quickly able to confirm medical knowledge and to elaborate new hypotheses. We observed that the domain experts are more likely interested in variables they are familiar with and have personal clinical experience with.

6.4 Case 2: Hypothesis-driven Analysis

If the user proposes a hypothesis about a relation between a non-image variable and shape, the workflow slightly differs from the hypothesis-free analysis. The starting point follows the *feature localization* pattern, where a variable of interest is selected by dragging it into the canvas area and viewing the subjects' distribution as well as their shape differences.

Analysis 1. Hypothesis: "*Back pain is associated with age and lumbar spine shape*". To validate this hypothesis, a *feature localization* was performed by combining the dichotomous variable "*Did you experience back pain in the last three months?*" with age in a mosaic plot by dropping both variables on the canvas area (Fig. 6 c). HV was not able to observe the expected effect in the visualization. Reasons for this are twofold. Age influences the lumbar spine shape, while the differences between subjects with and without back pain are small. The major differences seen in the visualization are therefore related to the age variable, masking differences related to the back pain parameter. The second explanation is the commonality of back pain in our society. As seen in Figure 6 (c), subjects reporting back pain are the

majority, which makes it difficult to extract parameters that reliably describe back pain. A *multivariate analysis* using the contingency table showed a strong association between *back pain* with both, *gender* and *body height*. *Body height* was explained as a confounder for *gender*, since female subjects on average are smaller than male subjects. The analysis solely based on shape-accentuated body height differences in *gender*, which clouded the differences of *back pain*.

The epidemiologists pointed out that they would like to see a more intuitive and fast way to select subgroups based on different variables to make full use of the analysis capabilities, as discussed in Section 6.5.

Analysis 2. Hypothesis: "*Back pain is related to lumbar spine deformation*". The previously discussed analysis questions the suitability of the lumbar spine segmentation for analyzing back pain, leading to this analysis. Therefore, the dichotomous variable "*Did you experience back pain in the past three months?*" is dropped into the canvas area. Figure 6 (d) shows the results of the automatically triggered shape-based clustering for subjects with and without back pain. The clustering algorithm finds only three homogenous clusters close to the global mean shape for subjects reporting no back pain. The cluster analysis for back pain yields diverse clusters with pathological shape classes. Cluster 5 represents most of the subjects and is similar to the global mean shape. Cluster 1 and 2 present a *hyperlordosis*, a strong curvature of the lumbar spine, while Cluster 3 and 4 present a more straight shape. A *multivariate analysis* using the pivot tables put gender and strong back pain in context to cluster affiliation (Fig. 6 d). It shows that subjects in Cluster 1, 2 and 6 reported strong back pain, while at the same time they also have a considerably higher share of females. To check for unusual correlations, the expert used the *Cramér's V* contingency table. It depicted strong associations with *body fat*, *body weight* and *blood pressure* (Cluster 1) *alcohol consumption* and *attentiveness disorder* (Cluster 2), and *amount of sleep* (Cluster 6). For the experts, these observations are a starting point for a number of new hypotheses about possible relationships, for example the association between overweight and Cluster 1.

In summary, it can be stated that the hypothesis-driven analysis leads to hypothesis generation by design of the framework. It is not suited and intended to statistically validate hypotheses like statistical processors. It rather triggers the analysis of potentially associated variables with a pathology of interest.

6.5 Further Feedback and Lessons Learned

Both domain experts rated the *IVA* approach positively. KH emphasized the way the image data are included into information visualizations, which comes much more natural to her due to her background in radiology. Great potential is also seen in communicating insights efficiently using the presented visualizations.

Multivariate analysis is most important for hypothesis generation. Both experts emphasized the potential of the *multivariate analysis* capabilities of the contingency matrix for gaining insight into a large amount of variables simultaneously. It is also useful to verify established but still controversial risk factors, such as the metabolic syndrome for coronary heart disease and whether the data set provides more suitable risk factors. Creating contingency matrices for subgroups, such as different age bins can help to characterize the aging process by deriving age-specific risk factors. *Multivariate analysis* can be improved by more ways of brushing the data as well as creating subgroups for comparison as a result of the hypothesis-driven analysis case. Too small variable ranges yielding sparse groups could hinder the calculation of statistical resilient measures, since they require a minimum amount of subjects exhibiting the selected variable ranges.

Segmentation quality is crucial. KH pointed out the unusual strong similarity of the L3 vertebrae throughout the population. The medical explanation is that it represents an angular point of curvature of the lumbar spine. A second explanation is the use of the L3 vertebra as initialization point of the lumbar spine model. The experts also emphasized that associations related to shape strongly depend on the segmentation quality. The lumbar spine model used in this case study captures deformation of the spine canal well, but lacks precise definition in vertebrae height and shape. Since deformation of the spine canal is the last stage of pathological lumbar spine deformation and is preceded by vertebrae deformation, the system would strongly benefit from more precise segmentation results capturing these prior changes. For the visual comparison, KH proposed an abstraction of the representation into landmarks, such as centers of the vertebrae and cardinal points of the lumbar spine canal.

Usage of different categorizations depending on expected outcome. Categorizing numerical variables into equal groups possibly creates sparse categories due to outliers.These outliers are only of high interest for finding pathological subjects. The experts therefore suggested two modes of the tool. The outlier mode still creates categories of equal size, producing sparse categories for outliers. Balanced categories are created in the second mode, which uses quartiles or quintiles to set borders between categories.

Web technologies are well suited for rapid feedback. The web-based approach for both implementing the prototype and getting feedback via voice over ip conference calls worked very well. Since the software does not need to be compiled, small changes can even be made on the fly during a testing session. The large data base associated with image-based epidemiological data remains on the server machine and has not to be moved tediously using external hard disks. This approach is well suited for the *VDAR* approach to assess user thought processes using the think-aloud technique.

7 Summary and Conclusion

We presented an *IVA* framework for the analysis of image-centric epidemiological data. Hence, the framework allows both for hypothesis-driven analysis and hypothesis generation. The visualization of multivariate data using multiple connected views allows to get fast visual feedback about subject groups. Brushing and linking makes the data tangible and adaptable to formulated hypotheses. The use of pivot tables is familiar to epidemiologists while embracing the power of interactive adjustment of the shown variables. The automatic suggestion of correlations using contingency methods, such as *Cramér's V* triggers *hypothesis generation* by highlighting variables potentially overlooked by the experts. Shape-based clustering assesses the variability of an anatomical structure in the context of non-image variables, such as disease indicators or lifestyle factors.

Epidemiologists are for the first time able to assess shape information of the lumbar spine and its influence to diseases. Findings from analyzing lumbar back pain using the *IVA* approach range from deriving shape-based groups of subjects to detailed descriptions of variables potentially associated with the disease, such as waist circumference, alcohol consumption and attentiveness disorder. A number of improvements is left open for future work, such as shape brushing methods to intuitively query subjects using image data or the inclusion of more statistical methods and views that are familiar to the epidemiologists (odds ratios, box plots).

As the number of image-centric cohort studies, participating subjects, gathered variables and imaging modalities rises, and advances towards comparability between cohort studies are made, the gap between data complexity and analyzability increases. Our work focuses on closing this gap, allowing the domain experts to dig deep into the data and potentially obtain unexpected findings. We believe that web technologies pave the way to analyze this data in a convenient way.

Acknowledgments

SHIP is part of the Community Medicine Research net of the University of Greifswald, Germany, which is funded by the Federal Ministry of Education and Research (grant no. 03ZIK012), the Ministry of Cultural Affairs as well as the Social Ministry of the Federal State of Mecklenburg-West Pomerania. Whole-body MR imaging was supported by a joint grant from Siemens Healthcare, Erlangen, Germany and the Federal State of Mecklenburg-Vorpommern. The University of Greifswald is a member of the Centre of Knowledge Interchange program of the Siemens AG. This work was supported by the DFG Priority Program 1335: Scalable Visual Analytics. We thank Marko Rak and Klaus Toennies for providing the image detection data.

REFERENCES

[1] P. Angelelli, S. Oeltze, C. Turkay, J. Haasz, E. Hodneland, A. Lundervold, B. Preim, and H. Hauser. Interactive visual analysis of heterogeneous cohort study data. *IEEE Computer Graphics and Applications*, 2014. in print.

[2] J. Blaas, C. Botha, and F. Post. Interactive visualization of multi-field medical data using linked physical and feature-space views. *Proc. of EuroVis' 07*, pages 123–130, 2007.

[3] M. Bostock, V. Ogievetsky, and J. Heer. D^3 data-driven documents. *IEEE Transactions on Visualization and Computer Graphics*, 17(12):2301–2309, 2011.

[4] S. Busking, C. Botha, and F. Post. Dynamic Multi-View Exploration of Shape Spaces. *Computer Graphics Forum*, 29(3):973–982, 2010.

[5] J. J. Caban, P. Rheingans, and T. Yoo. An Evaluation of Visualization Techniques to Illustrate Statistical Deformation Models. *Computer Graphics Forum*, 30(3):821–830, 2011.

[6] K. K. Chui, J. B. Wenger, S. A. Cohen, and E. N. Naumova. Visual analytics for epidemiologists: understanding the interactions between age, time, and disease with multi-panel graphs. *PloS one*, 6(2), 2011.

[7] W. G. Cochran. The $\chi 2$ test of goodness of fit. *The Annals of Mathematical Statistics*, pages 315–345, 1952.

[8] H. Cramér. *Mathematical methods of statistics*, volume 9. Princeton university press, 1946.

[9] X. Dai and M. Gahegan. Visualization based approach for exploration of health data and risk factors. In *Proc. of the International Conference on GeoComputation. University of Michigan, USA*, volume 31, 2005.

[10] J. W. Emerson, W. A. Green, B. Schloerke, J. Crowley, D. Cook, H. Hofmann, and H. Wickham. The generalized pairs plot. *Journal of Computational and Graphical Statistics*, 22(1):79–91, 2013.

[11] L. Ferrarini, H. Olofsen, W. M. Palm, M. A. Van Buchem, J. H. Reiber, and F. Admiraal-Behloul. Games: growing and adaptive meshes for fully automatic shape modeling and analysis. *Medical image analysis*, 11(3):302–314, 2007.

[12] R. H. Fletcher, S. W. Fletcher, and G. S. Fletcher. *Clinical epidemiology: the essentials*. Lippincott Williams & Wilkins, 2012.

[13] O. Gloger, J. Kühn, A. Stanski, H. Völzke, and R. Puls. A fully automatic three-step liver segmentation method on LDA-based probability maps for multiple contrast MR images. *Magnetic Resonance Imaging*, 28(6):882–897, 2010.

[14] D. L. Gresh, B. E. Rogowitz, R. L. Winslow, D. F. Scollan, and C. K. Yung. WEAVE: a system for visually linking 3-D and statistical visualizations applied to cardiac simulation and measurement data. In *Proc. of IEEE Visualization*, pages 489–492, 2000.

[15] M. Harreby, J. Kjer, G. Hesselsøe, and K. Neergaard. Epidemiological aspects and risk factors for low back pain in 38-year-old men and women: a 25-year prospective cohort study of 640 school children. *European Spine Journal*, 5(5):312–318, 1996.

[16] K. Hegenscheid, J. Kuhn, H. Völzke, R. Biffar, N. Hosten, and R. Puls. Whole-Body Magnetic Resonance Imaging of Healthy Volunteers: Pilot Study Results from the Population-Based SHIP Study. *Proc. of RöFo - Fortschritte auf dem Gebiet der Röntgenstrahlen und der bildgebenden Verfahren*, 181(08):748–759, 2009.

[17] K. Hegenscheid, R. Seipel, C. O. Schmidt, H. Völzke, J.-P. Kühn, R. Biffar, H. K. Kroemer, N. Hosten, and R. Puls. Potentially relevant incidental findings on research whole-body MRI in the general adult population: frequencies and management. *European Radiology*, 23(3):816–826, 2013.

[18] M. Hermann, A. C. Schunke, T. Schultz, and R. Klein. A visual analytics approach to study anatomic covariation. In *IEEE PacificVis 2014*, pages 161–168, 2014.

[19] J.-F. Im, M. J. McGuffin, and R. Leung. Gplom: The generalized plot matrix for visualizing multidimensional multivariate data. *IEEE Transactions on Visualization and Computer Graphics*, 19(12):2606–2614, 2013.

[20] P. Klemm, L. Frauenstein, D. Perlich, K. Hegenscheid, H. Völzke, and B. Preim. Clustering Socio-demographic and Medical Attribute Data in Cohort Studies. In *Bildverarbeitung für die Medizin (BVM)*, pages 180–185, 2014.

[21] P. Klemm, K. Lawonn, M. Rak, B. Preim, K. Tönnies, K. Hegenscheid, H. Völzke, and S. Oeltze. Visualization and Analysis of Lumbar Spine Canal Variability in Cohort Study Data. In *VMV 2013 - Vision, Modeling, Visualization*, pages 121–128, 2013.

[22] Z. Konyha, K. Matkovic, and H. Hauser. Interactive visual analysis in engineering: A survey. *Proc. of Spring Conference on Computer Graphics (SCCG 2009)*, pages 31–38, 2009.

[23] H. Lam, E. Bertini, P. Isenberg, C. Plaisant, and S. Carpendale. Empirical studies in information visualization: Seven scenarios. *Transactions on Visualization and Computer Graphics*, 18(9):1520–1536, 2012.

[24] T. Lammarsch, W. Aigner, A. Bertone, S. Miksch, and A. Rind. Towards a concept how the structure of time can support the visual analytics process. In *Proc. of the Int. Workshop Visual Analytics*, pages 9–12, 2011.

[25] M. Lang-Tapia, V. España-Romero, J. Anelo, and M. J. Castillo. Differences on spinal curvature in standing position by gender, age and weight status using a noninvasive method. *Journal of applied biomechanics*, 27(2), 2011.

[26] U. Niemann, H. Völzke, J.-P. Kühn, and M. Spiliopoulou. Learning and inspecting classification rules from longitudinal epidemiological data to identify predictive features on hepatic steatosis. *Expert Systems with Applications*, 2014.

[27] S. Oeltze, H. Doleisch, H. Hauser, P. Muigg, and B. Preim. Interactive Visual Analysis of Perfusion Data. *IEEE Transactions on Visualization and Computer Graphics (TVCG)*, 13(6):1392–1399, 2007.

[28] S. Oeltze, W. Freiler, R. Hillert, H. Doleisch, B. Preim, and W. Schubert. Interactive, graph-based visual analysis of high-dimensional, multi-parameter fluorescence microscopy data in toponomics. *IEEE Trans. on Visualization and Computer Graphics*, 17(12):1882–1891, 2011.

[29] S. Oeltze, H. Hauser, and J. Kehrer. Interactive visual analysis of scientific data, 2013. Half Day Tutorial at IEEE VIS, Seattle, WA, U.S.

[30] K. Pearson. X. on the criterion that a given system of deviations from the probable in the case of a correlated system of variables is such that it can be reasonably supposed to have arisen from random sampling. *The London, Edinburgh, and Dublin Philosophical Magazine and Journal of Science*, 50(302):157–175, 1900.

[31] B. Preim, P. Klemm, H. Hauser, K. Hegenscheid, S. Oeltze, K. Toennies, and H. Völzke. *Visualization in Medicine and Life Sciences III*, chapter Visual Analytics of Image-Centric Cohort Studies in Epidemiology. Springer, 2014. in print.

[32] M. Rak, K. Engel, and K. Toennies. Closed-form hierarchical finite element models for part-based object detection. In *VMV 2013 - Vision, Modeling, Visualization*, pages 137–144, 2013.

[33] S. Salvador and P. Chan. Determining the Number of Clusters/Segments in Hierarchical Clustering/Segmentation Algorithms. In *Proc. of Tools with Artificial Intelligence. ICTAI*, pages 576 – 584, 2004.

[34] B. Shneiderman. The eyes have it: A task by data type taxonomy for information visualizations. In *Proc. of Visual Languages*, pages 336–343. IEEE, 1996.

[35] M. Steenwijk, J. Milles, M. van Buchem, J. H. C. Reiber, and C. Botha. Integrated Visual Analysis for Heterogeneous Datasets in Cohort Studies. *Proc. of IEEE VisWeek Workshop on Visual Analytics in Health Care*, 2010.

[36] S. Thew, A. Sutcliffe, R. Procter, O. de Bruijn, J. McNaught, C. C. Venters, and I. Buchan. Requirements Engineering for e-Science: Experiences in Epidemiology. *Software, IEEE*, 26(1):80–87, 2009.

[37] J. J. Thomas and K. A. Cook. *Illuminating the path: The research and development agenda for visual analytics*. IEEE Computer Society Press, 2005.

[38] E. Tufte. *The visual display of quantitative information*. CT: Graphics Press, 1983.

[39] C. Turkay, A. Lundervold, A. J. Lundervold, and H. Hauser. Hypothesis generation by interactive visual exploration of heterogeneous medical data. In *Human-Computer Interaction and Knowledge Discovery in Complex, Unstructured, Big Data*, pages 1–12. Springer, 2013.

[40] M. van Tulder, B. Koes, and C. Bombardier. Low back pain. *Best Practice & Research Clinical Rheumatology*, 16(5):761 – 775, 2002.

[41] H. Völzke, D. Alte, C. Schmidt, et al. Cohort Profile: The Study of Health in Pomerania. *International Journal of Epidemiology*, 40(2):294–307, Mar. 2011.

[42] M. Q. Wang Baldonado, A. Woodruff, and A. Kuchinsky. Guidelines for using multiple views in information visualization. In *Proc. of the Working Conference on Advanced Visual Interfaces*, pages 110–119. ACM, 2000.

[43] C. Weaver. Cross-filtered views for multidimensional visual analysis. *IEEE Transactions on Visualization and Computer Graphics*, 16(2):192–204, 2010.

[44] Z. Zhang, D. Gotz, and A. Perer. Interactive visual patient cohort analysis. In *Proc. of IEEE VisWeek Workshop on Visual Analytics in Health Care*, 2012.

Interactive Visual Analysis of Heterogeneous Cohort Study Data

Paolo Angelelli, Steffen Oeltze, Judit Haász, Cagatay Turkay, Erlend Hodneland,
Arvid Lundervold, Astri J. Lundervold, Bernhard Preim and Helwig Hauser

Abstract— Cohort studies in medicine are conducted to enable the study of medical hypotheses in large samples. Often, a large amount of heterogeneous data is acquired from many subjects. The analysis is usually hypothesis-driven, i.e., a specific subset of such data is studied to confirm or reject specific hypotheses. In this paper, we demonstrate how we enable the interactive visual exploration and analysis of such data, helping with the generation of new hypotheses and contributing to the process of validating them. We propose a data-cube based model which handles partially overlapping data subsets during the interactive visualization. This model enables seamless integration of the heterogeneous data, as well as linking spatial and non-spatial views on these data. We implemented this model in an application prototype, and used it to analyze data acquired in the context of a cohort study on cognitive aging. We present case-study analyses of selected aspects of brain connectivity by using the prototype implementation of the presented model, to demonstrate its potential and flexibility. .

Index Terms—heterogeneous data, medical visualization, IVA

1 Introduction

Cohort studies in medicine become increasingly common, partly thanks to the availability and to the recent improvements in medical imaging technologies. Such studies are a type of observational study that follows one or more groups of people (samples), called cohorts, over time. They are used to evaluate medical hypotheses in samples sharing common characteristics, for example being healthy, or presenting specific risk factors, to gain a better understanding of the absolute risks of certain pathologies and of the pathology development. Cohort study data is often acquired over longer time periods, following strictly defined protocols, being therefore not trivial to set up. Because of that, they are often designed to deliver a larger variety of data than the focus of the initial study, which, later on, can be the basis for retrospective analyses, evaluating further sets of hypotheses.

There are means to evaluate specific hypotheses, based on such cohort study data, often involving accordingly designed data extraction, transformation, and fusion approaches. However, there is a lack of technology to support the flexible and open-ended exploration of such data, mostly because of its heterogeneity. This means collections of image and non-image (quantitative, often image-derived) data, which in turn can be categorical and numerical, and defined on domains that only partly overlap. Due to the complexities posed by the data heterogeneity, analysts often have to limit their attention to subsets of the data, making the analysis lose the overall relations within different modalities. Integrating all the available data within one visual analysis tool that allows to seamlessly combine them in an on demand fashion is expected to support the experts in the exploration of heterogeneous cohort study data and in the hypothesis generation and verification, and to accelerate their research workflow.

The exploration and analysis of heterogeneous cohort study data generates specific new challenges for visualization. The contribution of this article is therefore two-fold. First, in Section 2, we characterize these challenges, in relation to the substantial heterogeneity of the data, and in relation to the analysis tasks, goals, and typical analysis workflow in the specific context of a cohort study on cognitive aging. Second, in Section 4, we describe our solution, based on a new, general multi data-cube model to support heterogeneous data, and that can be also adapted to other situations of highly heterogeneous problems. Finally, in Section 5 we describe our prototype implementation of our model, that, in Section 6, we use to exemplify how our novel approach can enable the generation of new hypotheses, as well as the swift analysis of relations between otherwise unconnected data parts, thus improving the analysis and exploration process. In Section 6 we also provide an evaluation of our method by two domain experts from the medical and neuropsychological domain.

- *Paolo Angelelli and Helwig Hauser are with the department of Informatics at the University of Bergen. E-Mail: paolo.angelelli@uib.no .*
- *Cagatay Turkay is with giCentre at City University, London.*
- *Steffen Oeltze and Bernhard Preim are with the department of Informatics at the University of Magdeburg.*
- *Judit Haász, Erlend Hodneland and Arvid Lundervold are with the department of Biomedicine at the University of Bergen.*
- *Astri J. Lundervold is with the department of Biological and Medical Psychology at the University of Bergen.*

2 A Scenario of Heterogeneous Data in a Cohort Study

One major goal of this work is to create a solution to enable the explorative visualization and analysis of data that was acquired as part of a longitudinal study on cognitive aging. During this study, more than 100 healthy individuals (mean age 60.8 (7.8), 65% females at inclusion) were recruited through advertisements in local newspapers. At inclusion, all the subjects who responded were interviewed, to exclude those reporting previous or present neurological or psychiatric disorders, a history of substance abuse, or other significant medical conditions. The neuropsychological evaluation confirmed that the participants showed no symptoms indicating mild cognitive impairment (MCI) or dementia. Each participant was examined every three years, starting in year 2004/2005, and then in 2008. The participants were subjected to neuropsychological testing, genetic analysis (data not available for this work), and multimodal MR imaging. The result of each examination consisted of data on white matter fiber integrity, expressed by anisotropy measures computed from diffusion tensor imaging (DTI), cortical and subcortical gray matter measures, automatically calculated from structural MR images, and a number of neuropsychological tests, including the California Verbal Learning Test–Second Version (CVLT-II), the Color–Word Interference Test (CWIT), the Digit Symbol Substitution Task from WAIS-R, and the Mini Mental State Exam (MMSE). To summarize, each examination (per subject and year) consists of:

- white matter fiber bundles with anisotropy measures. Each individual fiber was divided into 100 segments of equal length for the derivation of associated measures.
- gray matter cortical and subcortical regions with quantitative measures for each region.
- scores from different neuropsychological tests.

For a detailed description of the study protocol and for previous selected analyses of this longitudinal study please refer to Ystad et

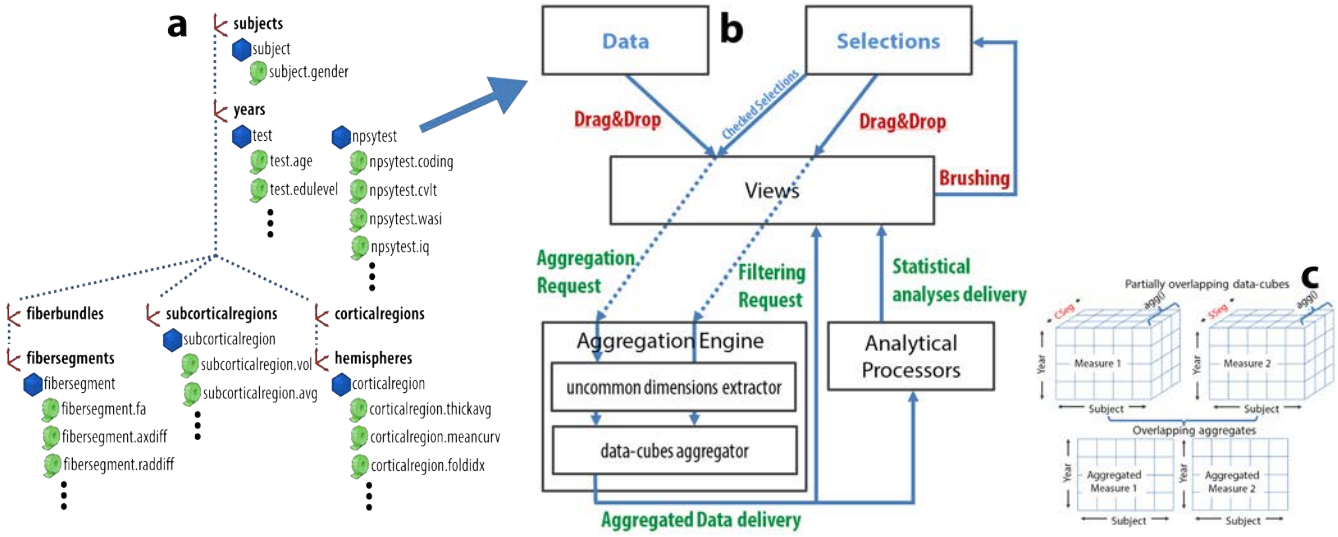


Fig. 1. **a**) Illustration of the dimensions (red), measures (green) and entities (blue) in the dataset of the cohort study on cognitive aging. The hierarchy in the figure is used only for presentation, as the presented model treats the dimensions independently. **b**) Simplified illustration of the proposed model. User interactions are colored in red, automatic operations, transparent to the user, are green, information sources are blue, and in black the components necessary to implement the model. Note that the selections require interaction to be used as filters, but are also automatically re-aggregated upon measure changes in views, or brush changes, and the result is automatically updated in the views. **c**) Illustration of the projection operation. The dimensions which are not common (in red) are processed using a statistical estimator (e.g., average). This operation can be steered by using a selection for each data-cube to filter the elements that are aggregated.

al. [14].

2.1 A heterogeneous dataset

Resulting from this study, a number of measures related to different aspects are available. One specific challenge with respect to the data exploration and analysis is that the measure's domains overlap only partially. Taking a scatterplot as an example, how should two heterogeneous measures be combined? In our case, these measures could be the *fractional white matter fiber anisotropy* (FA), that describes the degree of anisotropy of water diffusion along a fiber, defined for each segment of each fiber bundle, and the *thickness of the cortex*, available for each cortical region in both left and right brain hemisphere. This partial incompatibility of the data domains proved to be one if not *the* key challenge of this work. To overcome this challenge we developed the method presented in this article, able to seamlessly combine heterogeneous measures on the fly.

2.2 Abstract and physical data and their representation

In such studies certain measures, such as white matter FA or gray matter region volume, as well as others, are quantitative abstract measures that relate to physical (anatomical) entities. These, for the example, would be the white matter fibers or the gray matter regions. For these entities additional qualitative data is often also acquired, such as the bundles trajectories, or brain regions meshes or volumes. While analyses are often performed on the quantitative measures, it also becomes necessary to occasionally fetch and inspect the related anatomical data, to explain, for example, data outliers, or to see what effects certain conditions have on the anatomy. For these reasons domain experts would benefit from a system that can link different types of data, and bring up the appropriate sets on demand, e.g. in linked views.

In addition, when dealing with abstract views of measures related to physical entities, domain experts often need to relate groups of entities, such as selections, in abstract views to their physical location. To ease this process we propose to use a view with an illustrative physical model, or atlas, of the entities, which is linked to the other views. Through this atlas, the content of the selections is put in its physical context, to improve the understanding of such data. The definition of this model for the specific case described in this article, and its use, are described in Section 4.5.

3 Related Work

While the majority of visualization research –in particular also medical visualization– was (and still is) focused on the visualization of individual datasets, the visualization of data from population studies has not been a research topic until recently. One recent exception is the work of Bruckner et al. [1], presenting a system to retrieve and visualize anatomical brain data of *Drosophila*, covered in a large database of such flies' brains. This system enables a novel way to perform visual queries, combined with a volume rendering solution called Maximum Intensity Difference Accumulation (MIDA). Still in the biology domain, Jeanquartier and Holzinger presented a visual analytics approach for cell physiology to support the exploration and sensemaking process. [5]. Steenwijk et al. [10] also presented a novel visual analytics framework to query and visualize data from a cohort study , consisting of imaging and non-imaging data for each subject. Their approach was to preprocess and store the imaging and non-imaging data in a searchable relational database, to which a visual interface would perform dynamic queries. Still in the healthcare domain, Simonic et al. [9] presented a visualization system to improve prediction and treatment of patients based on longitudinal data.

More generally, few other visual analysis methods have been proposed for the analysis of higher-dimensional and heterogeneous data. One relevant related solution was presented by North et al. [7], who introduced visualization *schemas* to achieve the concurrent analysis of different sources of information in relational databases. Their system enables building coordinated visualizations in a similar fashion as when constructing relational data schemas. More recently, Weaver uses a method called cross-filtered views [13] to interactively drill down into multidimensional relations between multiple datasets. In his method, different variables are visualized in particular views and brushes in these multiple views are cross-filtered to discover complex relations in the data.

4 A Data-Cube Based Model To Enable Interactive Visual Analysis

The typical workflow approach to analyze the data coming from such studies is to manually extract the pieces of data to analyze from the dataset (e.g., using custom scripts or programs for each analysis), and

then process them using mathematical and statistical packages. Finally, plots of the results are generated either using custom scripts, or by importing the results into applications that can plot the data.

The first, and perhaps the biggest challenge in designing an interactive visualization system targeted at this problem is storing the data acquired with such studies in a way that allows fast and flexible access, retaining the meta-information expressing the relationships between the different pieces of data. Organizing the data in a relational database, similarly to Steenwijk et al. [10], is probably the first solution at hand, and possibly the easiest to design from scratch.

However, organizing data in a relational database is relatively inflexible: the database schema is bound to the specific structure of a particular study, together with the queries associated to it. Using a system designed in such a way to analyze a different dataset would require the redefinition of the database schema, as well as reprogramming the logic for data access. In addition, processing the queried data with mathematical or statistical methods that are not implemented in the database itself would require an additional application layer into which the data should be loaded, thus voiding the benefits of using a relational database. Finally, from a performance point of view, using a relational database to perform complex queries touching all the rows on a large amount of data becomes quickly a performance bottleneck in interactive operations, and this is even more problematic when item selection and measure filtering based on multiple attributes, requiring table joins, are used.

With *Polaris*, Stolte et al. [11] showed how visualization systems can also ground on data organized in a *n*-dimensional, possibly hierarchical, data-cube, which is also known as OLAP cube (for On-Line Analytical Processing) in the field of data warehousing. It has been reported that executing complex queries using OLAP cubes can perform about hundred times faster than doing the same on relational data [4]. A single, hierarchical, data-cube organization however, shows its limitations when the dataset, and its dimensionality, become heterogeneous.

4.1 Data-cubes: dimensions, entities and measures

In our model, data-cubes are constructed using categorical attributes as *dimensions*, while quantitative numerical values are stored as *measures* [11]. The dimensions and measures can be thought of as independent and dependent variables, and dimension coordinates are used to access the measures. Practically, after assigning an order to the dimensions of a cube, a data-cube can be implemented as an in-memory *n*-dimensional array. To make an example taken from the system presented in this paper, a measure for segments of white matter fiber bundles in our dataset, e.g., FA, is represented as a floating point n-dimensional array consisting of $n = 4$ dimensions: *subject*, *year*, *bundle*, and *segment*.

Compared to the model proposed for Polaris, we also introduce a third concept, called *entity*. An entity can be thought of as a row in a database table, and quantitative row fields would be the measures for that entity. In the example above, the measure *fibersegment.fa* (*fa* for fractional anisotropy) would be related to the entity *fibersegment*, being a measure of that entity. When, in our model, a data selection is defined, it also contains selection values for entities, which are then propagated to the measures related to it when it becomes necessary.

4.2 Multiple data-cubes and seamless dimension aggregation

A challenging feature of the data acquired in cohort studies is their heterogeneity. This means that measures are collected for different entities, which do not share the same set of dimensions. In our specific case, when referring to entities, we can talk of white matter fiber segments, grey matter subcortical regions and grey matter cortical regions, as well as neuropsychological tests. As shown in Figure 1a, the dimensions' sets of the measures are only partially overlapping, having all these entities in common only two dimensions, *subject* and *year*. The standard way to organize these data into a single data-cube would be to build a denormalized cube characterized by all the dimensions in the dataset, which would contain all the data. When the data is significantly heterogeneous, however, this strategy may lead to an explosion of the memory requirements caused by the denormalization.

In the model that we present here, the solution to this problem is twofold: on one side we store all the data in multiple, normalized data-cubes, to eliminate any kind of information redundancy and minimize the memory occupancy. Secondly, we propose runtime aggregation of the measures' data-cubes, when data which are held in data-cubes belonging to different entities have to be combined or cross-checked. Such aggregation operation is also referred to as the *projection* of a data-cube [11] (see Fig. 1c). Our model includes an engine to perform aggregation on-the-fly, for reducing the data-cubes' dimensions to their largest common subset, without having any embedded knowledge of the relations between measures. In contrast, this would be necessary when using a relational model for the data, as the system would need to incorporate knowledge about each specific database schema, together with logic for performing the operations.

In our model, when multiple measures are combined in a visualization (e.g., in a scatterplot, a parallel coordinate view, curve view, etc.), each measure is aggregated across those dimensions not belonging to the intersection. For the moment we can consider the mean as measure aggregator, but there are several other options, such as different statistic estimators which can be selected by the user.

In certain cases, it is also useful to change the level of detail. To allow this, we enable toggling which common dimension to keep during the aggregation. This is similar to a *roll-up* operation, with the difference that the dimensions' structure is treated as hierarchy-less.

Finally, even if some of the dimensions may embed a hierarchy, others are independent from each other. For example, it is easy to imagine that *subject* is independent from other dimensions, while *bundle* and *segment* are logically nested, as segments are part of a bundle. However, an imposed dimension hierarchy for all the dimensions would be useful to represent the data in a tree-like visualization, and let the user navigate the dataset (as shown in Fig. 1a). To compute such a hierarchy, we group entities recursively by the number of common dimensions, with each group reflecting dimensions occurring in the same number of entities. By letting the dimensions that occur in more entities floating higher in the tree hierarchy, and then proceeding recursively on subgroups, we can generate a complete hierarchy. Having defined such a hierarchy, it is possible to represent the measures in our cohort study data like in Fig. 1a.

4.3 Selections and selection-based filtering

In section 4.2 we explained how to create projections of a measure by aggregating it over entire dimensions. Obtaining an aggregate of a measure over a whole brain, however, may not always produce specific enough data to answer questions of interest. To enable a more focused analysis, selection techniques can be used in order to restrict the processed or visualized data to specific subsets under investigation. An example is the *Polaris specifications* [11], introduced for defining selections. Interactive visual analysis has introduced the related concept of brushing, a visual method to select items with certain characteristics (e.g., fitting certain ranges on specific measures), by defining a visual brush over a view on the data. These brushes normally contain a value for each data item, either binary or a percentage value, to express if or how much the data item is selected. Our model makes use of brushing to let the user define data selections. Using data-cubes, this brush should be transformed into a data-cube itself, where each item contains the tag information for the related entity. In our case, having several entities in the dataset generates an additional challenge: when tagging one entity, we must also propagate the selection to all those other entities in the dataset sharing at least one dimension with the tagged one. As a clarifying example, let us consider a selection of only those white matter fiber segments above a certain FA threshold. Such selection does not necessarily involve all the examinations, or even all the subjects. Let us say the user wants to cross only items in this selection with the cortical thickness. Then this selection has to be propagated to the entity *cortical region*, knowing that the shared dimensions between the entities *cortical region* and *fiber segment* are *subject* and *year*. This has to be done in an appropriate manner, so

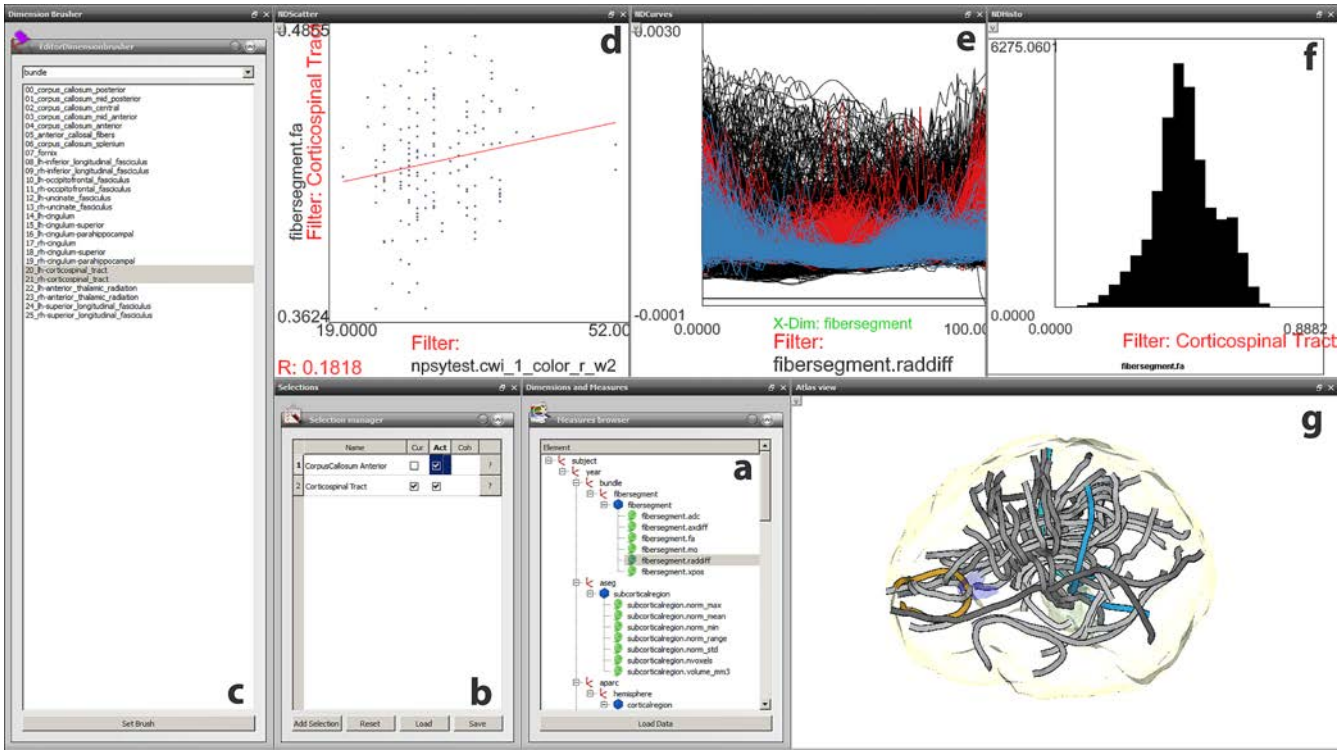


Fig. 2. Screen-shot of the prototype of the proposed model. The Measure Browser **(a)** lets the user drag desired measures into a view, and the Selection Manager **(b)** allows to add new selections, activate them, enable one of them for editing, and drag them into views, to be used as filters. The Dimension Brusher **(c)** enables slicing the data-cubes in the data collection, while the other views **(d,e,f)**, in this setup a scatterplot, a curve view and a histogram view, can be seen as projections of the data, and allow a more advanced definition of the selections, by means of brushing ranges of measures. In each view a drop down menu lets the user adjust the aggregation dimensions as well as the additional analyses to perform. Finally, the Atlas view **(g)** represents the selections in their anatomical context using a brain model. The two selections visualized contain, the first, both the fibers and the brain region of the Corpus Callosum anterior, and the second both the fibers of the corticospinal tract and the brainstem region (colors representing different bundles).

that only those (*subject*, *year*) pairs selected in one entity are selected also in the other one. In our model we propose a propagation scheme where a brush on one entity is propagated to all the other entities in the dataset that share dimensions with the brushed one. The propagation is done by first computing a projection of the brushed entity onto the common dimensions with all the other entities. Such projections of the brush are generated using the *max* operator, which produces, for each set of items being aggregated along one aggregation coordinate, the equivalent of a Boolean value indicating whether or not at least one item was selected. This scheme also allows multiple selections to be combined using Boolean logic, giving the user the necessary flexibility in building up expressive item selections.

Once a selection has been defined, it can be used in two manners. First, selections can be visually highlighted in the views, and thus compared with the whole dataset or with other selections. Second, since most of the views are built upon aggregated data-cubes, this aggregation can be steered, or *filtered*, using a selection (Fig. 2d). By setting a selection as aggregation filter, the aggregation is performed only using those items that are tagged in the selection. In this way, carefully selected information from the dataset can be cross-checked with other aspects, enabling the user to analyze virtually any aspect of the dataset.

4.4 Unrolling dimensions: a first step toward iterated visual analysis

Using a system implementing our model interactively is a flexible way to cross-analyze a wide variety of information in such heterogeneous datasets. In some cases, however, the analysis can benefit from automating certain steps, like repeating selected tests or analyses using a scheme defined by the user on different data, or with varying parameters or methods. This could be seen as extending a purely interactive visual analysis metaphor by using it as a analysis-setup tool for defining what type of actions to automate. The results of this extension could be thought as an *iterated visual analysis*. A clarifying example could be correlating age with subcortical region volume. The user could first define a selection, for example by filtering specific ages, or other parameters such as the IQ. This selection could then be used to filter the aggregation, which could conclude the interactive analysis step. Since it is also interesting to have details of how the volume of each specific subcortical region correlates with age, the user might want to combine his interactively specified selection with another one, selecting only a specific subcortical region, and repeat the process for every subcortical region. To ease this process, enabling at the same time to produce comparable results, we propose a method to automatically dissect and process the measures present in a specific view, by iteratively *slabbing* each measure's data-cube along those dimension that are specific to the data-cube (e.g., not common). In the example above, the only non-common dimension in a view containing only age and subcortical region volume is the *subcortical region*, as both the *year* and *subject* measures are common to both the entities (see Fig. 1a). The expression *unrolling a dimension* here means automatically generating a sequence of selections for an entity having such dimension, each selection containing only data items along one specific coordinate of that dimension at a time. The user can choose one or more of the non common dimensions in the view to unroll, and the automatically generated selection is combined with a user specified one, if present, before aggregation and further analysis take place. When performing dimension unrolling, however, a large amount of data is being generated, and we currently deal with it by outputting only the

analyses results, such as regression or correlation values. We make use of this technique in the case-studies illustrated in Section 6, and the results of the unrolling are shown in Fig. 4.

4.5 Visualizing aggregates in physical space

Sometimes it is of interest to link abstract information of physical entities to these entities in a spatial visualization of the data. Examples could be various, ranging from the analysis of mechanical components to various kinds of simulations. In the case of cohort studies, it can be useful to visualize the content of selections, as well as other parameters in the context of the brain's anatomy. A practical example would be visualizing where the parts of the white matter fiber bundles within a certain range of anisotropy, or having certain properties (e.g., sensitivity to aging) are located. To represent statistical information for a selection in physical space, we propose to use a physical atlas of what the data refer to, which in our case is a brain atlas (Fig. 2g). The selection aggregation is then performed on the dimensions present in the atlas.

4.6 Performance and limitations of the data model

We compared the performance between our implementation of the data model described above and a relational database (SQLite) on simple queries involving aggregation. We found that, using the dataset introduced in Section 2 (approximately 50MB in SQLite form, including only the quantitative measures), operations on our data model were more than ten times faster than the corresponding operations on the relational database. For example, such operations on the largest table/cube in our database, consisting of approximately 500,000 rows, lasted 260ms using our model, compared to the 2700ms using SQLite. Data-cubes are, however, able to provide such performance only when the data fits the system memory, and our model is, at the present, not supporting out-of-core data. In case of datasets not fitting the system memory, a standard database is necessary, but with the awareness that new technology will be necessary to allow systems like this to perform interactively.

5 Prototype realization

The prototype implementation of our model has been specifically realized to explore and analyze selected aspects of the brain aging dataset produced by the cohort study described in Section 2. The prototype consists of a coordinated multiple view application implementing linking and brushing on top of the proposed aggregation engine (2). Measures can be visualized and cross-checked on demand and in different views by using drag-and-drop interaction from the measure browser window into the view of choice. Selections can be initialized and modified by means of brushes on the views. Using selections as filters is implemented via drag-and-drop: dragging a selection into a view opens a selection dialog for the measure to filter. Choosing the measure re-triggers the filtered aggregation process.

To present selections and statistical information in physical space, we employ a brain atlas onto which aggregated statistics can be mapped. For simplification purposes, we treat the brain of a representative subject S and the fine-granular parcellation of its cortical and sub-cortical white and gray matter as the atlas. A more sophisticated approach would require the averaging of brain regions across all subjects and the computation of average fiber tracts. Instead of displaying all fibers of S ($>$20000), we compute a representative fiber for each fiber bundle (Fig. 2g). This reduces visual clutter and facilitates the mapping of statistics, aggregated across all fibers of a bundle and all subjects. Previous work suggests choosing the longest fiber traveling though the densest parts of the bundle as representative fiber [8]. We apply this approach directly to homogenous bundles, i.e. all fibers following a similar course. In heterogeneous bundles, we first subdivide the bundle by grouping similar fibers, and then compute the representative of each group. For the grouping, we employ a spectral clustering technique [2]. The white matter measures (such as *FA*) in our data were extracted after subdividing each individual fiber into 100 segments of equal length, to allow tract analysis. Therefore, we also divide each representative fiber into 100 segments, allowing the system to map the measures to each segment individually. We assign a unique color and add halos to each of them to enhance the visual separation of the representatives. The aggregated values are then encoded, upon normalization, by modifying the color saturation of each fiber segment (high values resulting in high saturation). Segmentations and related measures for brain regions are also included in the study, and were extracted with Freesurfer (http://surfer.nmr.mgh.harvard.edu/). For displaying the measures, an isosurface is constructed per segmented region. The visual separation of brain regions is enhanced by assigning unique colors according to the Freesurfer's color look-up table. Mapping a measured value is then performed upon normalization by modifying the surface transparency (high values resulting in high opacity). Finally, a highly transparent outer surface of the brain is superimposed, to augment the overall atlas visualization (Fig. 3g).

6 Case-studies and Evaluation

We conducted a two-phase study with domain experts: a design requirement phase to understand the analysts' needs, and an evaluation phase to evaluate our method after the suggestions of the experts were included. In the first phase, we have been able to gather initial impressions from two neuropsychologists and one neurologist, as well as some feature requests. The prototype has been received with strong interest. Thanks to the flexibility and simplicity of performing data selections and cross-analyses, it has been seen as a practical alternative to the current way of analyzing data, consisting of extracting the values by various means into separate tables, and loading them into commercial statistical packages or tools. The experts also explicitly requested to be able to get a detailed description, save, and load the selections, and to be able to export filtered data.

After this first cycle, we included these additional functionalities, and performed the second evaluation phase with a neuropsychologist and with a neurologist. These two evaluation sessions were subdivided in three parts, organized as follows: first, a thorough explanation of the application of the underlying model was given. The explanation was followed by few questions about the model, in order to ensure the understanding of the model. This first part of the session was successful, and the domain experts could explain well the difference between our model and the table-based data model present in all the statistical analysis tools used in a standard analysis workflow, where the observations are the rows, and the variates are the columns. Neither of them was previously familiar with relational databases or OLAP cubes. During this first part of the sessions, the demonstration of the application functionalities was also well understood. The second part of the sessions had a dual aim: to verify that our model is capable of producing the same results obtainable with a standard analysis workflow, but in a faster way, and to prove that our model is capable of helping the generation of new hypotheses. For this second part two case-studies, one for each domain, were set up, and are described below. The third part of the sessions was used to gather an assessment of the proposed method by asking the domain experts specific questions, and details are given in Section 6.3.

6.1 Neurologic case-study

Jointly with a neurologist, we attempt to confirm or reject three hypotheses which were already statistically evaluated in previous work [12]:

- The increased age-related anisotropy decline in the anterior callosal fiber (CC-Anterior), as compared to the posterior portion of the corpus callosum, called splenium (CC-Splenium).
- The higher sensitivity to age-related anisotropy decline of superior fibers (Superior-LF), as compared to inferior fibers (Inferior-LF).
- The resistance of the cortico-spinal tract to age-related anisotropy decline.

To confirm the first hypothesis, we begin with selecting the fibers under investigation. Then, we use these selections as filters in scatterplots opposing the age of the subjects and the FA of the fiber segments

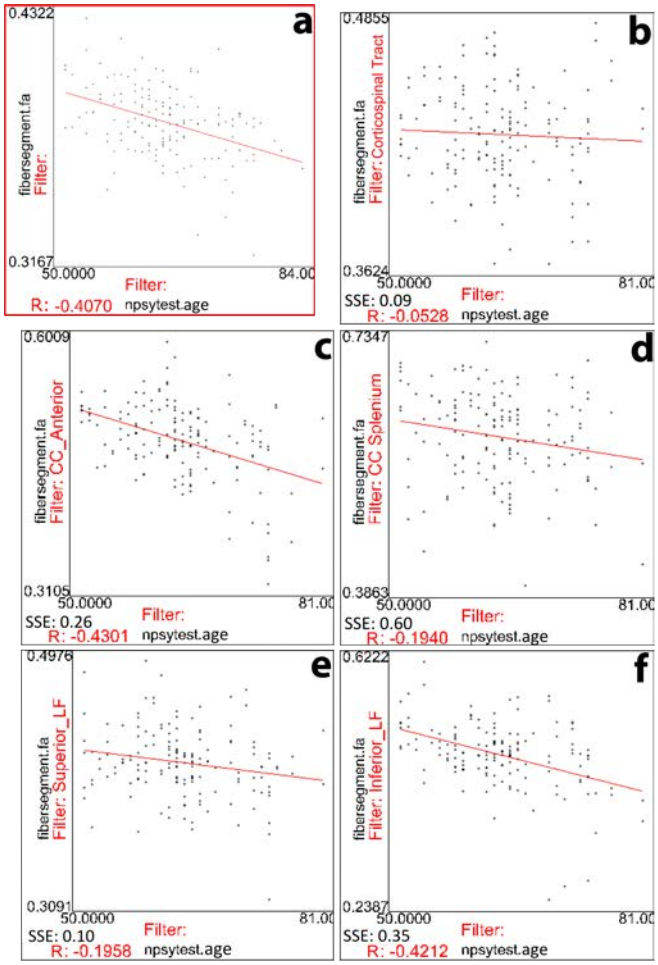


Fig. 3. **a**) Age opposed to FA for each examination (subject, year). Visualizing the linear regressor depicts the general declining trend, also summarized by the negative correlation r-value. **b**,**c**,**d**,**e**,**f**) Age opposed to FA aggregated (across segments and bundles) using a different filter in each plot, as labeled in the views. **c** and **f** show a stronger negative correlation, while **d**) and **e**) show a weaker negative correlation. **b**) shows almost no correlation between FA and age for the corticospinal fiber tracts, which confirms previously published studies, and can be used as control. In each plot: R is the correlation coefficient, SSE is the sum of squared residuals of the regression analysis.

in the subjects' brains. In these scatterplots, shown in Fig. 3, each point represents a single subject examination (*subject*, *year*), while the other dimensions are aggregated for each of the measures. In the case of FA, this aggregation is filtered using the selections above. The system automatically computes the Pearson's r value of the two measures (one aggregated using the filter), the p-value, which, in our case, is below 0.05 except for the corticospinal tract (that, therefore, does not show a correlation that is statistically significant) and the regression line. The regression analysis also provides the regression coefficient and the sum of squared residuals (SSE) as a metric for the goodness of fit. These plots confirm that the corticospinal tract is relatively insensitive to the age effect. They also show that the posterior portion of the corpus callosum is less prone to age effect compared to the frontal portion. But, in contrast to our expectations, superior fibers are less prone to age effect than inferior fibers. This could suggest the new hypothesis that language functions stay normal while the visual integration might decline. Such hypothesis, however, requires further investigations.

In the second part of this case study we decide to perform an explorative investigation of the relation between the anisotropy decline in the white matter fiber tracts and age. We do this by looking at the correlation coefficient, as well as the regression coefficient, between FA and age. The result of this analysis is shown in Fig. 3a, and we discover that there is a significant negative correlation (-0.406) between these two aspects. The regression coefficient (-0.001, not normalized) is small since the data has not been normalized, but the regressor (the red line) provides a better picture of the trend than the value alone. Once we discovered that these aspects are worth investigating, we use the unrolling mechanism described in Section 4.4 to evaluate this relation selectively for each fiber bundle. The system estimates these statistics for the chosen measure by iterating over a user specified dimension, in our case *fiberbundle*. These estimates are presented in two bar charts shown in Fig. 4a and 4b. It is easy to spot one fiber (fornix) that goes against the general declining trend, also showing a bad fitting (sum of squared residual, SSE). We decide to bring this fiber up for inspection in a scatterplot (Fig. 4c), by using the filtering capability of the system. So we manually create a specific selection, defined by *slabbing* the data-cube along the fornix coordinate of the fiberbundle dimension and use it to filter the aggregation. In the scatterplot we detect several zero values (Fig. 4c), probably due to missing data, which tells that the information for this fiber should be discarded or the missing data should be removed. We opt for cleaning the data, by performing a selection with a brush on the scatterplot that excludes the incorrect values. This leads to opposite results (Fig. 4d), in line with the overall declining trend (these results are sketched with a dashed line in the barcharts of Fig. 4a and 4b). We also notice that the corticospinal fiber tracts seem to be particularly insensitive to age decline, while other tracts have very strong decline (anterior callosal fibers and inferior longitudinal fasciculi). Finally, we notice two corresponding tracts, left and right occipitofrontal fasciculi, which are not homogeneous, with the right one showing a more pronounced anisotropy decline, even though they are anatomically symmetric to each other. This finding should be investigated further, to verify the fibers' geometrical path along which the measures have been sampled, in order to possibly formulate a new hypothesis on this phenomenon.

6.2 Neuropsychologic case-study

Jointly with a neuropsychologist, we attempt to verify the relation between the volume of the frontal regions of the cortex and the performance in the Stroop task. We focus on this task since a functional correlation between these brain regions and the Stroop effect has been discovered using functional imaging techniques [6]. Therefore we would expect that subjects with smaller frontal cortices would perform worse at the Stroop task.

We begin by opposing the cortical gray matter volume measure to the scores of the Stroop task (where higher is worse) in a scatterplot (Fig. 4f). A general declining trend (smaller cortical regions causing worse performance) is therefore expected, and it is also what we get. Then we create a selection with the frontal pole cortical area, and filter the aggregation accordingly. The resulting plot shows a minor increase in correlation and regressor slope (these values are outlined in red in Fig. 4e). Such result confirms a mild correlation between the volume of frontal cortical area and Stroop task performance, but the aggregated result tells us that there must be other cortical regions whose decline is even more correlated to the Stroop task than the frontal cortex. Therefore we decide to use the unrolling option to get an overview of how each single cortical region correlates with the Stroop task scores. The resulting statistics (Pearson's r and p-value, non-normalized regression coefficient and sum of the squared residuals) are displayed using a table widget, allowing us to sort the rows by a chosen column (Fig. 4e, sorted by r-value). Sorting these data by r value unveils an area, the parahippocampal cortical region, for which no know hypotheses have been formulated, but which shows the strongest correlation between its volume and the Stroop task scores. According to the neuropsychologist, this is a new finding, but something that requires further investigation because this analysis is done on the raw data, not corrected for basic skills.

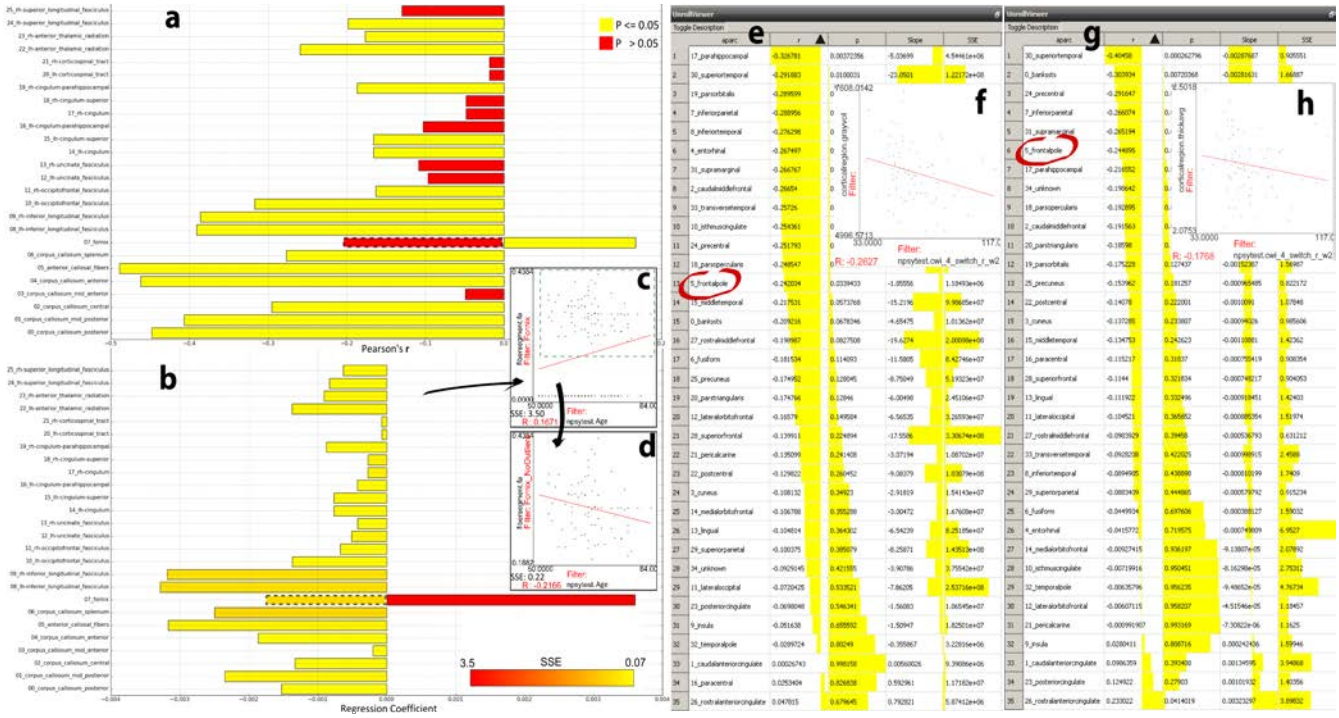


Fig. 4. **a**) The correlation coefficient between age and the FA of the fibers. The dimension *fiberbundle* for the measure FA is unrolled, meaning that FA is filtered by automatically iterating over a chosen dimension, in this case *fiberbundle*. Each bar represents the correlation for a specific coordinate in the *fiberbundle* dimension. **b**) The same type of visualization for the regression coefficient. **c**,**d**) Scatterplots related to the *fornix* fiber bundle, before and after excluding wrong values. **f**) Stroop task scores related to cortical region volume in a scatterplot. The cortical region volume is aggregated along the cortical region dimension (called *aparc*). **e**) The same analysis after unrolling the cortical region dimension. Each row in the table reports the correlation and regression results for the data filtered for a single cortical region, reported in the first column. **d,c**) The same analyses, but for Stroop task scores and cortical region thickness.

At this point we also wonder whether any relation between the Stroop task scores and the cortical thickness is present in the data, as thickness is another measure that has been shown to correlate with level of cognitive functions [3] We proceed as before, opposing the cortical thickness measure to the scores of the Stroop task in a scatterplot (Fig. 4h). A general declining trend is visible also in this case, but less strong than with the cortical volume. We then use the unrolling option to get an overview of how each single cortical region correlates with the Stroop task scores (Fig. 4g). In this case, the frontal pole cortical area shows a stronger correlation as compared to the overall cortical thickness. However, the task performance seems to be even more affected by other areas, most notably the superiortemporal cortical region. This is also a new finding, which however requires, as in the previous case, further investigation in order to formulate a new hypothesis on this phenomenon.

6.3 Assessment of the model

In the third part of the sessions we asked the domain experts the following questions: a) whether or not our prototype system was useful for data exploration tasks, b) whether or not such system was capable of answering specific questions, c) whether or not such system was useful to generate new hypotheses and d) whether or not such system could potentially replace their current tools.

Both scientists answered positively to question a), stating that such a tool, able to load and combine in a quick yet flexible way all the measures from such large studies, would be certainly helpful in data exploration tasks. This answer was supported by the fact that, in the current analysis workflow based on data in tabular form and commercial statistical analysis packages, all the work of data combination and selection has to be done manually for each question to analyze, which makes data exploration tasks especially cumbersome.

The scientists were also particularly positive regarding question c). The key aspects that were regarded as most useful in generating new hypotheses are: having the whole data at hand in one tool, the ease of use, and being able to fire queries in the tool. Moreover, and what impressed them the most, is to be able to automatically generate relevant selections in an iterated way while processing the data with a specific statistical method.

Concerning question b), which is also related to question d), our system proved to be effective in performing basic multivariate statistical analyses on the data. However, the domain experts stated to rarely use only basic multivariate statistics, but rather adding advanced techniques to assess relations between two or more measures. In addition, the neuropsychologist that was interviewed stated to rarely use the raw data alone, but often combine multiple measures into more advanced descriptors (e.g., correcting test results for the basic subject skills). However, the fast and flexible selection and filtering capabilities that the presented model offers were highly appreciated, since both the scientists stated to perform selections on the subjects to include in each analysis based on different parameters that vary from case to case. The conclusion for question b), and also for question d), was that an ideal tool would combine the presented model with more advanced data derivation and statistical analysis tools. This is a good lesson learnt, and a direction that, in some way, was already taken by having the R software environment embedded into our prototype system, even if not all of the requested methods are bound to the prototype yet.

7 Conclusion and Future Work

Medical cohort studies are an excellent starting point for exploratory data analysis, since most of the data acquisitions are standardized before specific hypotheses are formulated. Often, such studies are designed to provide enough data, of very heterogeneous character, such that a large set of possible hypotheses can be tested on them. Accordingly, hypothesis generation becomes an own challenge, when asso-

ciated with populations studies. In this work, we have demonstrated that an exploratory interface, which is capable of flexibly linking up different aspects of the data even if they are not given with respect to exactly the same domain, can help to swiftly identify new and possibly promising research hypotheses. We also showed, that the same approach is also capable of enabling a first quick analysis of the identified hypotheses, leading to an accelerated analysis methodology with respect to such highly rich and versatile data. The prototype system presented here, however, is still relatively limited in features, but such an application could potentially benefit from a broad spectrum of functionalities. In future, we plan to continue in this research direction, and extend the capabilities of this tool.

As future work we also plan to import genotype data for the subjects, that at the time being was not readily available, and to integrate 2D/3D graph views for representing the brain connectivity information. We are also trying to obtain a more thorough evaluation of the system in terms of new requirements, in particular from a statistical and data-mining perspective. Finally, we plan to include the retrieval and visualization of patient-specific image data, to assess whether outliers originate from the image data, or whether they are the result of an erroneous derivation process.

ACKNOWLEDGMENTS

We wish to thank Stefan Bruckner for developing the VolumeShop framework, used to create the application prototype. The cohort study used in this work was financially supported by the Research Council of Norway, University of Bergen, MedViz, and Western Norway Health Authority (grant #911593 to A.L., and grants #911397 and #911683 to A.J.L.).

REFERENCES

[1] S. Bruckner, V. Solteszova, M. Groller, J. Hladuvka, K. Buhler, J. Yu, and B. Dickson. Braingazer-visual queries for neurobiology research. *Visualization and Computer Graphics, IEEE Transactions on*, 15(6):1497–1504, 2009.

[2] A. Brun, H. Knutsson, H.-J. Park, M. Shenton, and C.-F. Westin. Clustering fiber traces using normalized cuts. *Medical Image Computing and Computer-Assisted Intervention (MICCAI'04)*, pages 368–375, 2004.

[3] A. Z. Burzynska, I. E. Nagel, C. Preuschhof, S. Gluth, L. Bäckman, S.-C. Li, U. Lindenberger, and H. R. Heekeren. Cortical thickness is linked to executive functioning in adulthood and aging. *Human brain mapping*, 33(7):1607–1620, 2012.

[4] S. Chaudhuri and U. Dayal. An overview of data warehousing and olap technology. *ACM Sigmod record*, 26(1):65–74, 1997.

[5] F. Jeanquartier and A. Holzinger. On visual analytics and evaluation in cell physiology: A case study. In *Availability, Reliability, and Security in Information Systems and HCI*, pages 495–502. Springer, 2013.

[6] M. Milham, M. Banich, E. Claus, and N. Cohen. Practice-related effects demonstrate complementary roles of anterior cingulate and prefrontal cortices in attentional control. *Neuroimage*, 18(2):483–493, 2003.

[7] C. North, N. Conklin, K. Indukuri, and V. Saini. Visualization schemas and a web-based architecture for custom multiple-view visualization of multiple-table databases. *Information Visualization*, 1(3-4):211, 2002.

[8] L. O'Donnell, C. Westin, and A. Golby. Tract-based morphometry for white matter group analysis. *NeuroImage*, 45(3):832–844, 2009.

[9] K.-M. Simonic, A. Holzinger, M. Bloice, and J. Hermann. Optimizing long-term treatment of rheumatoid arthritis with systematic documentation. In *Pervasive Computing Technologies for Healthcare, 5th International Conference on*, pages 550–554. IEEE, 2011.

[10] M. D. Steenwijk, J. Milles, M. A. Buchem, J. H. Reiber, and C. P. Botha. Integrated Visual Analysis for Heterogeneous Datasets in Cohort Studies. In *IEEE VisWeek Workshop on Visual Analytics in Health Care*, 2010.

[11] C. Stolte, D. Tang, and P. Hanrahan. Polaris: A system for query, analysis, and visualization of multidimensional relational databases. *Visualization and Computer Graphics, IEEE Transactions on*, 8(1):52–65, 2002.

[12] A. Voineskos, T. Rajji, N. Lobaugh, D. Miranda, M. Shenton, J. Kennedy, B. Pollock, and B. Mulsant. Age-related decline in white matter tract integrity and cognitive performance: a dti tractography and structural equation modeling study. *Neurobiology of aging*, 2010.

[13] C. Weaver. Cross-filtered views for multidimensional visual analysis. *Visualization and Computer Graphics, IEEE Transactions on*, 16(2):192–204, 2010.

[14] M. Ystad, A. Lundervold, E. Wehling, T. Espeseth, H. Rootwelt, L. Westlye, M. Andersson, S. Adolfsdottir, J. Geitung, A. Fjell, et al. Hippocampal volumes are important predictors for memory function in elderly women. *BMC medical imaging*, 9(1):17, 2009.

Paolo Angelelli received his Ph.D. in Computer Science from the University of Bergen, Norway. He is currently a Post-Doctoral researcher at the University of Bergen. His research interests are in fields of medical visualization, heterogeneous data visualization and interactive visual analysis.

Steffen Oeltze received his Ph.D. in Computer Science from the University of Magdeburg, Germany. He is currently a Post-Doctoral researcher at the University of Magedburg. His research interests are the visual analysis of perfusion and functional data, the visual exploration of vasculature and biological data.

Judit Haász received her M.D. from the Semmelweis University, Budapest, Hungary. She is currently a Ph.D. candidate at the University of Bergen, Norway. Her research interests are in fields of functional and structural brain changes after strokes and correlation of cognition, brain function and genetic biomarkers.

Cagatay Turkay received his Ph.D. from University of Bergen, Norway. He is a faculty member and lecturer at the gi-Centre at City University, London, UK. His research interests are the integration of interactive visualizations, data analysis techniques and supporting exploratory knowledge of experts. with special focus on bioinformatics and biomolecular modeling.

Erlend Hodneland received his Ph.D. in Physics from the University of Bergen, Norway. He is currently a Post-Doctoral researcher at the University of Bergen. His research interests are in the fields of image processing and pattern recognition, functional imaging, image registration and quantification in medicine and biology.

Arvid Lundervold received his M.D. from the University of Oslo, Norway, and his Ph.D. in physics from the University of Bergen, Norway. He his currently a professor in medical information technology at the University of Bergen, and head of the Neuroinformatics and Image Analysis Laboratory in the Neural Networks Research Group. His research interests are in the fields of image processing and pattern recognition, functional imaging, image registration, quantification and visualization, and mathematical modeling.

Astri J. Lundervold received her Ph.D. in neuropsychology from the University of Oslo, Norway. She is currently a professor in neuropsychology at the University of Bergen, Norway. She leads the Clinical Cognitive Neuroscience group. Her main research interest is to characterize behavior associated with normal function and neuropsychiatric disorders across the life span.

Bernhard Preim received his Ph.D. in Computer Science from the University of Magdeburg, Germany. He is currently a professor for visualization at the University of Magdeburg. His research interests are in medical visualization and applications in diagnosis, surgical education and surgical planning.

Helwig Hauser received his Ph.D. in Computer Science from the Vienna University of Technology, Austria. He is currently a professor for visualization at the University of Bergen, Norway. His research interests are interactive visual analysis, illustrative visualization, the combination of scientific and information visualization and the application of visualization to various domains.

Measurement of the Stratum Radiatum/Lacunosum-Moleculare (SRLM)

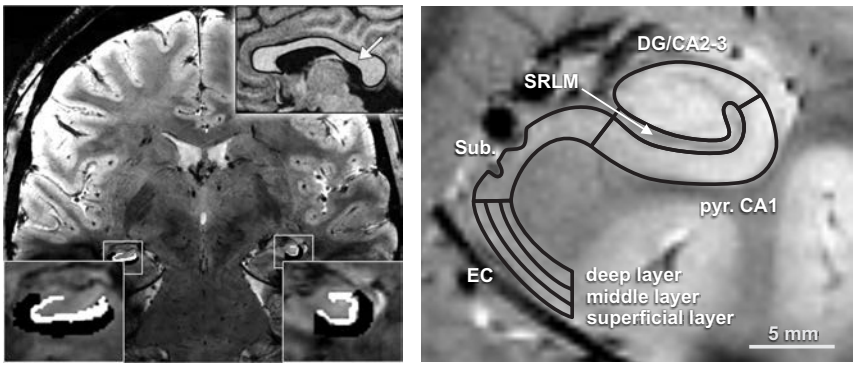
Steffen Oeltze[1], Hartmut Schütze[2], Anne Maaß[2], Emrah Düzel[2,3,4], Bernhard Preim[1]

[1]Department of Simulation and Graphics, University of Magdeburg
[2]Institute of Cognitive Neurology and Dementia Research, University of Magdeburg
[3]German Centre for Neurodegenerative Diseases (DZNE), Standort Magdeburg
[4]Institute of Cognitive Neuroscience, University College London, UK
stoeltze@isg.cs.uni-magdeburg.de

Abstract. Alzheimer disease (AD) at an early stage is characterized by a synaptic loss and atrophy in the apical layer of the CA1 part of the hippocampus, the stratum radiatum and stratum lacunosum-moleculare (SRLM). It was shown in vivo that patients with mild AD exhibit a reduced thickness of the SRLM.
We propose a new approach to measure SRLM thickness in coronal brain sections. It is based on the interpolated contour of the manually segmented SRLM and its medial axis. We automatically compute the axis by combining Voronoi diagrams and methods from graph analysis. While existing measurement approaches require a mental segmentation of the SRLM and a repeated local estimate of its center, we obviate the latter. We evaluate our approach based on coronal $T2^*$-weighted 7-Tesla MR images of 27 subjects.

1 Introduction

At an early stage, Alzheimer disease (AD) is characterized by episodic memory dysfunction. The hippocampus – a brain structure existing in both hemispheres and being part of the limbic system – plays a crucial role in consolidating episodic memory [1]. Postmortem studies found that synaptic loss and atrophy in the apical layer of the CA1 part of the hippocampus, the stratum radiatum and stratum lacunosum-moleculare (SRLM, Fig. 1(a)), coincide with earliest cognitive symptoms [2]. Kerchner et al. [3] showed that patients with mild AD exhibit a reduced SRLM thickness. Their analysis was based on coronal $T2^*$-weighted images from ultra-high field 7-Tesla (7-T) MRI. Only ultra-high field imaging provides an in-plane resolution ($< \approx 0.5\,\text{mm}$) high enough to identify hippocampal subfields (Fig. 1(b)). Since normal SRLM thickness is about 1 mm, its width covers only a few pixels. Within each transection of the hippocampus, the SRLM is heavily quantized and borders to adjacent structures form a step-wise pattern. Since the CA1 apical neuropil is affected in early stages of AD, a correct measurement of SRLM thickness could serve as an objective imaging biomarker for AD pathology and moreover, contribute to judging the (e.g. protective) effects of physical or cognitive training in early AD patients as well as healthy older people.



(a) (b)

Fig. 1. *(a)* Coronal $T2^*$-weighted Magnetic Resonance image with overlaid ROIs of the stratum radiatum and stratum lacunosum-moleculare (SRLM, white) and pyramidal CA1 (black). The upper inset shows the corpus callosum (arrow). Its shape is similar to SRLM shape. *(b)* Schematic of the subfields in the hippocampal body.

Two approaches to measuring SRLM thickness have been published. A manual approach was presented by Kerchner et al. [3]. For each hemisphere in two consecutive slices, the user draws three lines extending over the local width of the SRLM. The approach is subjective, it is restricted to a subset of the slices that show the SRLM, and it poorly acknowledges the variance in thickness along the SRLM. Recently, Kerchner et al. [4] proposed a semi-automatic measurement. The user first draws in the medial axis of the SRLM in all slices. Orthogonal signal intensity profiles are then determined at equidistantly sampled points along a spline that is fit to the user-defined line. A single thickness value is computed per slice based on the average signal intensity change from the medical axis of the SRLM to the surrounding subfields. The approach is less subjective but sensitive to the user's definition of the medial axis. The definition requires a mental segmentation of the SRLM and a concurrent, repeated local estimate of its center. The latter is particularly tedious in regions of very small SRLM thickness (1-2 pixels). Furthermore, the measured local thickness depends on the signal intensity distribution of the surrounding subfields. SRLM sites being equally thick may result in different measurement values. This effect is mitigated by averaging the local intensity profiles but hampers a real local analysis.

The shape of the SRLM in coronal slices is similar to the shape of the *corpus callosum* (cc) in mid-sagittal slices (Fig. 1(a), cc in upper inset). An overview of approaches to measuring callosal thickness is part of [5]. Most approaches rely on the medial axis of the cc and determine thickness orthogonal to it. We propose a related SRLM measurement approach which is based on the interpolated contour of the manually segmented SRLM and the medial axis of this contour, i.e., of the SRLM. In contrast to [4], we automatically compute the axis. The contour increases the range of possible measurements, e.g., by area or curvature. Another advantage is the coherent local computation of SRLM thickness independent of the signal intensities of surrounding structures. We evaluate our approach based on two studies comprising $T2^*$-weighted 7-T MR images of 27 subjects.

2 Materials and Methods

The focus of this section is on the measurement of the SRLM. Before, we describe our study data and briefly our SRLM segmentation.

2.1 Study Data and SRLM Segmentation

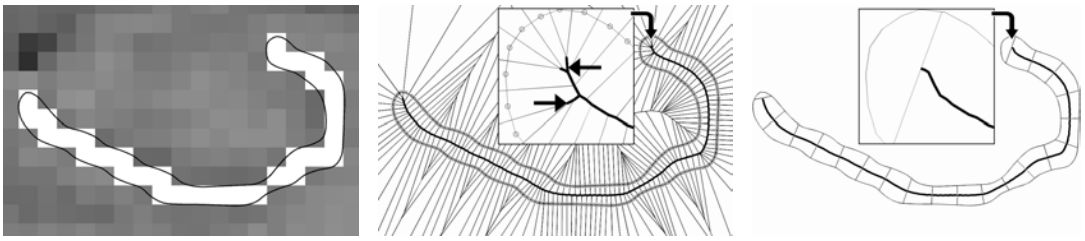
The data were collected in an ultra-high field 7-T MRI study at the Institute of Cognitive Neurology and Dementia Research of the University of Magdeburg, Germany. The study combined a visual learning paradigm with high resolution functional measurements and very high resolution structural images for 14 subjects. A pre-study with 13 subjects was conducted. For both studies, healthy young people were recruited (age 25±2, 14 male). MRI data were acquired using a 7-T MRI system (Siemens, Erlangen, Germany) with a 32-channel head coil. The high resolution partial structural volume was acquired (T2*-weighted imaging, $TE = 18.5\,\text{ms}$, $TR = 680\,\text{ms}$, in-plane resolution $0.33\,\text{mm}\times0.33\,\text{mm}$, slice thickness $1.5\,\text{mm} + 25\%$ gap, 45 slices, FOV $212\,\text{mm}\times179\,\text{mm}$, matrix 640×540), with a slice alignment orthogonal to the hippocampal main axis. The pre-study was conducted with the same sequence and similar parameters.

The segmentation of hippocampal subfields was performed for each hemisphere using MRIcron (Chris Rorden, Version 4, April 2011). First, subfields in the hippocampal body were traced according to [6]. Next, the parahippocampal regions were delineated. Then, the hippocampal head was segmented into subregions according to [7]. The hippocampal tail was not delineated. Overall, the hippocampus was segmented into subiculum (Sub), CA1-stratum pyramidale (pyr. CA1), CA1-stratum radiatum/stratum lacunosum-moleculare (SRLM, Fig. 1(b)) and the remaining portion comprising CA2, CA3, and DG (DG/CA2-3). Only the SRLM part in the hippocampal body is used in the thickness evaluation.

2.2 SRLM Measurement

Our measurement of the SRLM is based on its medial axis. A good survey of medial axis computation algorithms is part of [8]. The pixel-based medial axis generated by topological thinning or distance transform algorithms is too coarse since the width of the SRLM in some regions covers only 1-2 pixels. Surface sampling methods allow for a more fine-grained determination but require a representation of the objects boundary by a dense cloud of sample points. Our measurement algorithm starts by computing this point cloud.

Given the binary mask of the SRLM resulting from segmentation, we process this mask slice-by-slice (Fig. 2(a)). A 3D measurement is not feasible due to the considerable slice thickness ($1.5\,\text{mm}+25\%$ gap). We start by computing a smooth contour of the quantized binary mask. A marching squares algorithm with linear interpolation provides an initial sharp-edged contour. The contour is enhanced via B-spline interpolation followed by a Laplacian smoothing with displacement adjustment to avoid shrinkage [9]. The smoothing parameters have been determined empirically: 10 smoothing passes with a factor of 0.1 and a window of 3



(a) (b) (c)

Fig. 2. Measurement approach. *(a)* Computation of the contour (black) of the SRLM mask (white). *(b)* Voronoi diagram (net-like structure) of the contour points (circles). The medial axis (thick polyline) is part of the diagram. The inset shows erroneous side branches. *(c)* Ideal medial axis and thickness measurements (thin, orthogonal lines).

points. The final contour is resampled equidistantly with a sample point density that fulfills the requirements for an accurate medial axis computation [10].

The medial axis is derived from the Voronoi diagram of the sample points [10] (Fig. 2(b)). A Voronoi diagram divides the space into regions such that each seed (contour sample point) is contained in a separate region which comprises all points that are closer to this seed than to any other. The edges of the Voronoi diagram, which are completely contained within the SRLM contour, constitute its medial axis. They are determined based on point-in-polygon tests. The Voronoi approach is sensitive to noise in the contour. Slight deviations from a perfectly smooth curve cause short side branches originating from the medial axis (inset in Fig. 2(b)). Hence, pruning is often carried out as a post-processing step [10].

We suggest an inverse strategy that separates the ideal medial axis MA_{ideal} from the noisy one MA_{noisy} (Fig. 2(c)). Due to the normally non-branching, tubular shape of the SRLM within a coronal slice, MA_{ideal} is a simple polyline extending from one end to the other. Its separation is based on the observation that MA_{ideal} corresponds to the longest of the shortest paths between any pair of terminal vertices of MA_{noisy}. We treat MA_{noisy} as an undirected, unweighted, acyclic graph. Each of the n vertices is a node and an edge exists between two nodes if they are connected by a line segment in MA_{noisy}. We describe the graph by its $n \times n$ symmetric adjacency matrix A, whose entries equal 1 if the two corresponding nodes are connected by an edge and 0 otherwise. Terminal vertices of MA_{noisy} are characterized by a single 1 in their corresponding row or column of A. For each pair of terminal vertices, we find the shortest in-between path, i.e., along MA_{noisy}, by Breadth-First Search on A. The longest of these shortest paths represents MA_{ideal}.

To measure SRLM thickness, we equidistantly resample MA_{ideal} according to a user-defined number of thickness measurements (Fig. 2(c)). At each sample point, we erect an orthogonal line. Its intersections with the lower and upper part of the SRLM's contour delimit the local thickness. If a line intersects the contour more than twice, the two intersections which are on either side of the line and closest to MA_{ideal} are chosen. The Euclidean distance between the intersection points corresponds to the local thickness of the SRLM.

3 Results

We have applied our measurement approach to data of 27 subjects. For each subject, the SRLM of the two hemispheres has been segmented in ≈10 slices resulting in 594 SRLM contours. The algorithm was set to equidistantly sample SRLM thickness at 20 locations along the medial axis. All medial axes and orthogonal lines defining local SRLM thickness were visually verified. Local thickness was correctly represented in $> 95\%$ of the orthogonal lines. Figure 3 shows typical examples for successful and failed representations, and illustrates the shape variety of the SRLM. In $(a - c)$, common shape variants and their reasonable measurements are displayed. The contours in $(d - f)$ represent increasing deviations from the typical SRLM shape leading to incorrect thickness measurements (thick lines). It can be seen that these errors occur mostly at sites of high bending or where one part of the contour bends significantly different than the opposite part. The branchings seen in $(g - h)$ result from uncertainty during segmentation which is due to similar signal intensities of blood vessels or surrounding structures. Although a reasonable medial axis can be computed, thickness measurements are disturbed by the second branch and it remains unclear which branch represents the SRLM. Instead of neglecting individual thickness measurements, we completely removed cases similar to $(d - h)$ from our analysis (22% of the contours were removed).

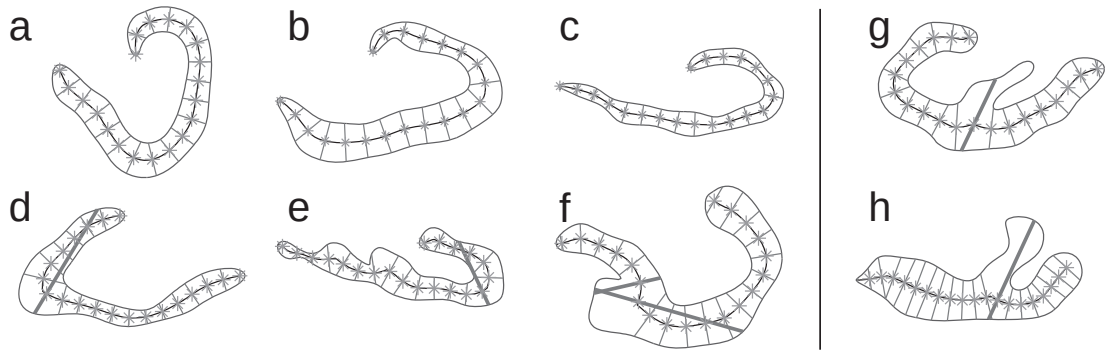


Fig. 3. Successful *(a-c)* and failed *(d-h)* evaluations of SRLM thickness. Thick lines represent unreasonable measurements due to strong local differences in the bending of the lower and upper SRLM contour parts *(d-f)* or due to a branching contour *(g-h)*.

The SRLM thicknesses of all subjects had a mean of 0.95 mm ($\sigma = 0.17$ mm) and showed a very high correlation between both hemispheres ($r = 0.93, p < 0.01$), which suggests that an individual property of the subjects was indeed obtained. Kerchner et al. reported thickness values in the range $0.4 - 0.6$ mm [4]. The differences to our values are most likely due to their conservative estimate of where SRLM ends and where surrounding structures begin based on the signal intensities. While they choose the approximate middle of the unsure transition zone, we include the entire zone during segmentation.

4 Discussion

Our method is largely dependent on the individual rater bias during pixel-wise delineation of the SRLM. However, if either a conservative or a slightly relaxed segmentation strategy is consistently chosen for all subjects of a study, the bias should be minimized. Hence, comparing a group of subjects with mild Alzheimer disease or Mild Cognitive Impairment and a control group is feasible. We aim at correlating thickness and performance measures of recognition memory tests. Hereby, differences in thickness between groups and reproducible measurements are rather important than absolute real thickness values.

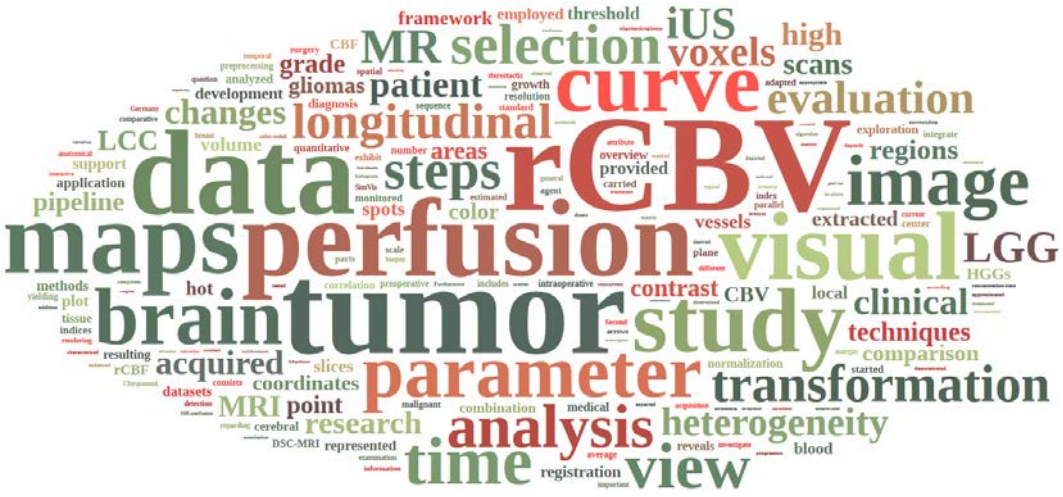
The causes of erroneous thickness measurements as illustrated in Figure 3 ($d-f$), have been also acknowledged by Herron et al. in measuring the corpus callosum (cc) [5]. However, their proposed solution involves a strict anatomically based definition of the cc's center. The computation of a similar center for the SRLM is hampered by its higher shape variability (Figure 3). A promising solution in regions of high SRLM bending is based on electric field lines and was presented for measuring the cerebral cortex in 2D histological sections [11].

References

1. Milner B. The Medial Temporal-Lobe Amnesic Syndrome. Psychiatr Clin North Am. 2015;28(3):599–611.
2. Thal DR, Holzer M, Rüb U, Waldmann G, Günzel S, Zedlick D, et al. Alzheimer-Related τ-Pathology in the Perforant Path Target Zone and in the Hippocampal Stratum Oriens and Radiatum Correlates with Onset and Degree of Dementia. Experimental Neurology. 2000;163(1):98–110.
3. Kerchner GA, Hess CP, Hammond-Rosenbluth KE, Xu D, Rabinovici GD, Kelley DA, et al. Hippocampal CA1 Apical Neuropil Atrophy in Mild Alzheimer Disease Visualized with 7-T MRI. Neurology. 2010;75(15):1381–1387.
4. Kerchner GA, Deutsch GK, Zeineh M, Dougherty RF, Saranathan M, Rutt BK. Hippocampal CA1 Apical Neuropil Atrophy and Memory Performance in Alzheimer's Disease. NeuroImage. 2012;63(1):194–202.
5. Herron TJ, Kang X, Woods DL. Automated Measurement of the Human Corpus Callosum Using MRI. Frontiers in Neuroinformatics. 2012;6(25):1–15.
6. Mueller SG, Stables L, Du AT, Schuff N, Truran D, Cashdollar N, et al. Measurement of Hippocampal Subfields and Age-Related Changes with High Resolution MRI at 4T. Neurobiology of Aging. 2007;28(5):719 – 726.
7. Wisse LEM, Gerritsen L, Zwanenburg JJM, Kuijf HJ, Luijten PR, Biessels GJ, et al. Subfields of the Hippocampal Formation at 7T MRI: In Vivo Volumetric Assessment. NeuroImage. 2012;61(4):1043 – 1049.
8. Foskey M, Lin MC, Manocha D. Efficient Computation of a Simplified Medial Axis. In: ACM Symposium on Solid Modeling and Applications; 2003. p. 96–107.
9. Vollmer J, Mencl R, Müller H. Improved Laplacian Smoothing of Noisy Surface Meshes. Computer Graphics Forum. 1999;18(3):131–138.
10. Brandt JW, Algazi VR. Continuous Skeleton Computation by Voronoi Diagram. CVGIP: Image Understanding. 1992;55(3):329–338.
11. Schmitt O, Böhme M. A Robust Transcortical Profile Scanner for Generating 2-D Traverses in Histological Sections of Richly Curved Cortical Courses. NeuroImage. 2002;16(4):1103 – 1119.

Part IV

Perfusion Data



This part of the postdoctoral thesis cumulates the following publications:

Chapter 12 [63] S. Glaßer, **S. Oeltze**, U. Preim, A. Bjørnerud, H. Hauser, and B. Preim, "Visual Analysis of Longitudinal Brain Tumor Perfusion", *Proc. of the SPIE Medical Imaging*, pp. 86700Z, 2013.

Chapter 13 [28] C. Chalopin, **S. Oeltze**, B. Preim, A. Müns, J. Meixensberger, and D. Lindner, "Method for the Evaluation of US Perfusion for Brain Tumor Surgery", *Proc. of Jahrestagung der Deutschen Gesellschaft für Computer- und Roboter Assistierte Chirurgie (CURAC)*, pp. 198-202, 2013.

Visual Analysis of Longitudinal Brain Tumor Perfusion

Sylvia Glaßer[a], Steffen Oeltze[a], Uta Preim[b], Atle Bjørnerud[c], Helwig Hauser[d] and Bernhard Preim[a]

[a]Department of Informatics, University of Magdeburg, Germany
[b]Municipal Hospital of Magdeburg, Germany
[c]Department of Medical Physics, Rikshospitalet-Radiumhospitalet Medical Centre; Department of Physics, University of Oslo, Norway
[d]Department of Informatics, University of Bergen, Norway

ABSTRACT

In clinical research on diagnosis and evaluation of brain tumors, longitudinal perfusion MRI studies are acquired for tumor grading as well as to monitor and assess treatment response and patient prognosis. Within this work, we demonstrate how visual analysis techniques can be adapted to multidimensional datasets from such studies within a framework to support the computer-aided diagnosis of brain tumors. Our solution builds on two innovations: First, we introduce a pipeline yielding comparative, co-registered quantitative perfusion parameter maps over all time steps of the longitudinal study. Second, based on these time-dependent parameter maps, visual analysis methods were developed and adapted to reveal valuable insight into tumor progression, especially regarding the clinical research area of low grade glioma transformation into high grade gliomas. Our examination of four longitudinal brain studies demonstrates the suitability of the presented visual analysis methods and comprises new possibilities for the clinical researcher to characterize the development of low grade gliomas.

Keywords: Brain Tumor Perfusion, Longitudinal Data, Multiple Coordinated View Systems

1. INTRODUCTION

Magnetic resonance imaging (MRI) is used for evaluating brain tumors due to its high soft-tissue contrast. In addition to morphologic aspects represented by conventional MRI, dynamic susceptibility contrast (DSC) perfusion imaging enables the characterization of dynamic aspects, in particular the cerebral microvasculature that is represented by the quantitative perfusion parameter relative cerebral blood volume (rCBV), see Figure 1. DSC-MRI, in combination with conventional MRI, is a good presurgical indicator for glioma grade and may identify the most malignant parts of a tumor for guiding stereotactic biopsy as well as to monitor and assess treatment response and patient prognosis. Gliomas – tumors with a glial cell origin – are the most common primary brain tumors, varying histopathologically from low grade gliomas (LGGs) to high grade gliomas (HGGs).

Grading of gliomas and thus the differentiation between LGGs and HGGs plays an important role for treatment planning and patient outcome.[1] Further, LGGs may transform into HGGs at some point in time, and an early detection of such a transformation is of significant clinical importance. If a surgical removal or radiation treatment is not possible, e.g., due to the tumor's location or patient's request, patients with LGGs are commonly subject to a life-long MRI monitoring. Here, the clinical research focus lies on the detection of LGG transformation in longitudinal brain data acquired over several years.

Since tumor growth depends on angiogenesis, i.e., the formation of new vessels and/or the sprouting of existing vessels, the increased cerebral microvasculature yields elevated rCBV values. Evaluation of rCBV is a clinical research focus. It is an important indicator for a patient's survival and gliomas with high rCBV values have a significantly faster progression time.[2] Since HGGs have in general foci of higher rCBV values, and rCBV correlates with the tumor grade, rCBV is also employed for differentiation between LGGs and HGGs.[2–4]

For longitudinal studies, rCBV maps of different acquisitions have to be compared, which is a complex and exploratory analysis task, due to the absence of standardized intensity values and the high variability of MRI

Further author information: glasser@isg.cs.uni-magdeburg.de, Telephone: +49 (0)3 91 67 52 527

scanners and patient data. As one challenge, intervoxel correspondence has to be established between scans of different time steps. This is further complicated by tumor growth and increasing tumor heterogeneity. At last, MRI scans from longitudinal studies are acquired over a period of several years and thus employ different scanning protocols, image resolutions, and sequence parameters. This is in contrast to cohort studies, e.g., the SHIP study,[5] where many patients are examined in a short time period and extra care is taken to avoid image parameter changes. Currently, longitudinal studies are assessed with software intended for the diagnosis of a patient at one point in time. Hence, no dedicated support for longitudinal studies, i.e., normalization of image data and visual comparison of data from different points in time, is provided.

In this paper, we present a visual analysis approach for the investigation of longitudinal brain tumor perfusion studies within a clinical research scenario. The approach comprises a preprocessing pipeline and the visual analysis framework for the computer-aided evaluation of longitudinal brain tumors. The framework also includes a new 3D parallel coordinates plot that provides a complete overview of all rCBV values at all time steps and allows for fast detection of trends.

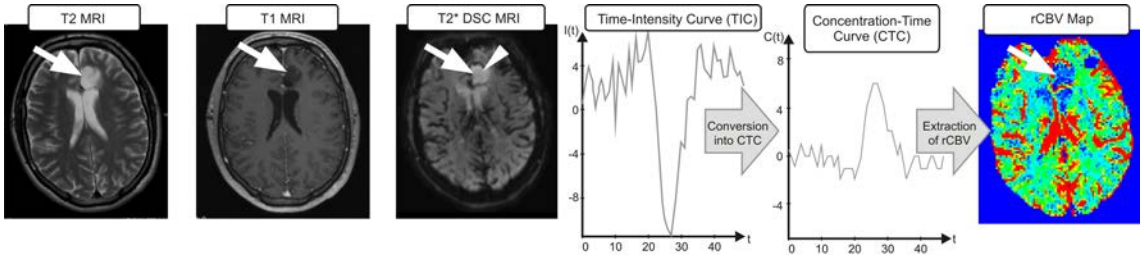


Figure 1. MRI slices with a tumor (see arrow). For a single voxel (see arrowhead) of the DSC-MRI data, the time-intensity curve is extracted and converted into the concentration-time curve (see diagrams). From the concentration-time curve, the parameter CBV is extracted, yielding a CBV map (low values are mapped to blue, high values to red). Note the varying data size, orientation, and image resolution.

2. RELATED WORK

Basic visualization techniques for exploring tumor perfusion data were presented by Behrens et al.,[6] including color-coded parameter maps. Covarrubias et al.[1] employ rCBV parameter maps to specify ROIs with high rCBV values (also called hot spots), indicating tumor growth and malignancy. Wetzel et al.[7] analyze different methods for creating a representative ROI regarding intra- and interobserver reproducibility, but a gold standard for a cut-off value between LGGs and HGGs of a ROI's average rCBV value is still missing.[8] Covarrubias et al.[1] suggest an rCBV value of 1.5 ml / 100 g and Law et al.[4] 1.75 ml / 100 g as cut-off value. These cut-off values are reported to have high sensitivity but low specificity by Emblem et al.[3] They suggest cut-off values between $3.75 - 5.58$ ml / 100 g after minimizing the number of glioma grade misclassifications and maximizing the average sensitivity and specificity. Furthermore, Emblem et al. present a histogram analysis of rCBV heterogeneity leading to increased diagnostic accuracy and interobserver agreement. Lupo et al.[9] extract the spatial heterogeneity of a brain tumor's concentration-time curves to further distinguish microvasculature characteristics.

For scientific visualizations, a 3D parallel coordinates plot has been presented by Wegenkittl et al.[10] The WEAVE system[11] links scientific visualizations with multidimensional statistical representations by brushing facilities applied to scatter plots. The SimVis framework[12] extends this work, and was adapted to the visual exploration of perfusion data by Oeltze et al.[13] Oeltze et al.[14] also analyzed longitudinal ischemic stroke cerebral perfusion data. Kohle et al.[15] introduced an adapted volume rendering of breast MRI data that emphasizes suspicious regions with an appropriate color scheme. Coto et al.[16] present a multidimensional view system with scatter plot-based selections for visual exploration of breast perfusion data, where cutaway views reveal the suspicious breast tumor in its surrounding. Botha et al.[17] present a system for analyzing and exploring medical multi-field data and integrate techniques from pattern analysis to enhance the data exploration process. For visual analysis of breast perfusion data regarding perfusion characteristics and tumor heterogeneity, Glaßer et al.[18] and Preim et al.[19] applied glyphs and region merging.

In clinical research regarding neurological impairments, longitudinal studies are becoming a standard element.[20] So far, visual analysis has seen limited use in medical research, although in many applications the magnitude and heterogeneity of the data would strongly benefit from these methods. Steenwijk et al.[21] apply visual analysis techniques to cohort study image data to explore parameters across patients. In general, there is a wide variety of visual analysis techniques for evaluating *single* perfusion studies. However, specific support for longitudinal studies is needed for the comparison of different points in time – taking into account that imaging parameters may have changed. We want to investigate the applicability of visual analysis techniques to longitudinal perfusion brain studies and investigate rCBV development as well as local tumor heterogeneity approximated with the correlation of rCBV and relative cerebral blood flow (rCBF).[22]

3. APPLICATION SCENARIO AND IMAGE DATA

This section introduces clinical research questions regarding LGGs, their development, and the image data for which the visual analysis techniques were developed.

3.1 Clinical Research Questions

The comparison of rCBV maps is the most important question of our clinical partners. As a prerequisite, preprocessing has to be carried out to incorporate rCBV maps from different perfusion scans. Since LGG transformation relates to increased tumor heterogeneity, the evaluation of the tumor's heterogeneity based on quantitative perfusion parameters is another task. Tumor transformation is also related to tumor growth, which should be visualized. In summary, there are four clinical research question and a last question covering our application that consists of the problem-solving components:

1. How can a comparative rCBV evaluation be achieved for a longitudinal brain perfusion study?
2. When does the LGG transformation start based on the evaluation of rCBV maps?
3. How is the tumor's heterogeneity characterized in terms of local rCBV and rCBF correlation?
4. How does the tumor grow and develop during the longitudinal study?
5. How can visual exploration and analysis of longitudinal studies be carried out?

3.2 Image Data

Our case study consists of a selection of four patients fulfilling two special conditions: First, each patient had a confirmed diagnosis of an inoperable grade II glioma, i.e., an LGG, and was thus monitored over several years. Second, during this time period, a transformation into an HGG took place. We evaluated the brain MRI data of these four patients resulting in four longitudinal studies L_1 – L_4. For each study, up to five MRI protocols including perfusion DSC-MRI sequences, T1 pre- and post-contrast, T2, and Fluid Attenuated Inversion Recovery (FLAIR) MRI were acquired. Typical sequence parameters for the DSC-MRI perfusion studies are gradient-echo echo planar imaging (GRE-EPI) with a temporal resolution TR $= 1.4-1.72$ s, echo time TE $= 30-52$ ms, image matrix $= 128\times128$, slice thickness $= 6.5$ mm, in-plane resolution $= 1.8$ mm $\times 1.8$ mm, number of slices $= 12-19$, number of acquisitions $= 50-75$, and a total acquisition time ranging from 73 s to 119 s.

For each study, the point in time of the LGG transformation was estimated by an experienced radiologist. The estimation is based on *all* MRI protocols – instead of the single perfusion scan – of all time steps for each study. L_1 was acquired over almost three years at four time steps $t_1 - t_4$ for monitoring of an LGG. The transformation into an HGG started between t_1 and t_2. The patient of study L_2 underwent surgical intervention and the remaining LGG was monitored for four years. MRI data was acquired at five time steps and the LGG transformation started between t_3 and t_4. L_3 contains an oligodendroglioma, a glioma type exhibiting foci of high CBV values irrespective of the tumor grade.[23] MRI scans have been acquired at four time steps during two and a half years. The transformation started between t_3 and t_4. L_4 has been acquired at five time steps to supervise an LGG over a time period of three and a half years. The transformation into an HGG started after t_3.

4. VISUAL ANALYSIS OF LONGITUDINAL BRAIN TUMORS

Our computer-aided diagnosis of longitudinal brain tumor studies includes two innovations: a *preprocessing pipeline* for the MRI perfusion scans and the *framework* containing the adapted visual analysis techniques.

4.1 Preprocessing Pipeline

Our preprocessing pipeline facilitates the comparison of a longitudinal study's different perfusion datasets (with possibly different image parameters) and comprises five steps, which are described in the following.

4.1.1 Co-registration of Perfusion Datasets

Motion artifacts in brain perfusion imaging typically result from patient movement. With the skull as static reference object, rigid registration algorithms allow for co-registration of brain perfusion data. We co-registered each study's perfusion DSC-MRI datasets to the study's DSC-MRI dataset acquired last in time with *Rview* (rview.colin-studholme.net), employing a rigid registration algorithm.[24] Thus, a concurrent analysis of all perfusion scans is supported.

4.1.2 Extraction of CBV and CBF maps

To assess CBV and cerebral blood flow (CBF), the distribution of the contrast agent is analyzed. The contrast enhancement results in time-intensity curves for each voxel (see Fig. 1). We employ the software package nordicICE (NordicNeuroLab, www.nordicneurolab.com) to transform these curves into concentration-time curves, applying the regularized singular value decomposition for deconvolution.[25] The arterial input function was extracted from the arteria cerebri media.[26] Contrast agent leakage correction was carried out due to possible contrast agent extravasation in regions of blood-brain-barrier disruption,[27] caused by the tumor. CBV is approximated as the area under the concentration-time curve and defined as the total volume of blood traversing a given region of the brain, measured in ml of blood per 100 grams of brain tissue. CBF is defined as the volume of blood traversing a given region of brain per unit time, measured in milliliters of blood per 100 grams of brain tissue per minute. Although in ischemic stroke diagnosis, CBF is thoroughly analyzed, the role of this parameter in brain tumor diagnosis has not been as extensively studied as CBV.[1] We investigate the correlation between CBF and CBV for tumor heterogeneity evaluation. While a small decrease in CBF is expected as a consequence of normal aging in a longitudinal study, the transformation from LGG to HGG is expected to involve significant stronger changes. However, a gold standard for CBF and CBV evaluation seems hard to establish, due to different imaging modalities, age, and gender.[8]

4.1.3 Normalization of CBV and CBF maps

For comparison of CBV and CBF values from DSC-MRI scans, normalization has to be carried out. We employ the general approach, where CBV and CBF maps are normalized with the averaged white matter's values of the contralateral side,[7] yielding relative CBV (rCBV) and relative CBF (rCBF) values. Visual inspection and ROI placement are realized with *MeVisLab* (www.mevislab.de), a platform for medical image processing and visualization.

4.1.4 Exclusion of Vessels and Adapted Smoothing

Next, vessels were excluded from the rCBV and rCBF maps, since brain vessels exhibit higher values than the surrounding tissue. They can be identified and removed based on the earlier and stronger contrast enhancement in the DSC-MRI data. We extracted a vessel mask with nordicICE based on the cluster analysis of the estimated perfusion-related parameters to separate vessels (both arteries and veins) from other tissue. For each rCBV and rCBF map the corresponding vessel mask is applied. Afterwards, a 3×3 modified average filter is applied to smooth the data as well as to reduce holes caused by the vessel mask. This filter empirically accounts best for rCBV changes due to noise or subtle artifacts. The modified filter only averages over non-vessel voxels in the 3×3 neighborhood of the current filter kernel. Hence, outliers, i.e., isolated voxels with rCBV values higher than a certain threshold, are removed. We employed the .995 quantile of each rCBV map as cut-off value.

4.1.5 Extraction of Tumor Masks

In a last step, a tumor mask for each of the perfusion datasets was created. Hence, the T2 and FLAIR images were co-registered to the perfusion datasets acquired at the same point in time of the studies with Rview. Next, binary tumor masks based on hyperintense areas in T2 and FLAIR data for each study's time step were extracted and applied to the rCBV maps. Tumor mask creation was manually carried out and validated by an experienced radiologist.

4.2 A Framework for Visual Analysis of Longitudinal Brain Tumors

Based on the preprocessed longitudinal scans, we developed a framework for the longitudinal evaluation of brain perfusion data. The framework contains color-coded 2D and 3D visualizations, a local heterogeneity map, and a parallel coordinates plot view.

4.2.1 Direct 2D and 3D Visualizations

The 2D and 3D visualizations directly map each voxel's rCBV value to color and are explained in more detail below.

2D rCBV Maps. For direct rCBV extraction, a standard 2D slice view of the brain tumor with voxelwise rCBV is presented. All tumor slices of a study's scans are provided. We apply a modified rainbow color scale from blue to red, see Figure 2, based on the color scale suggested by Wetzel et al.[7] Red highlights critical rCBV values (values > 5), and suspicious rCBV values greater than 1 are mapped to cyan. Although rainbow color scales do in general lack an intuitive visual interpretation of the data's order, they support visual clustering. Hence, regions with mostly red colors (and thus high rCBV values) or blue regions can be observed. Based on the rCBV maps a selection of voxels, – *the rCBV threshold selection* – can be defined. This selection consists of all voxels with an rCBV value greater than a user-defined threshold and can be combined with all other views.

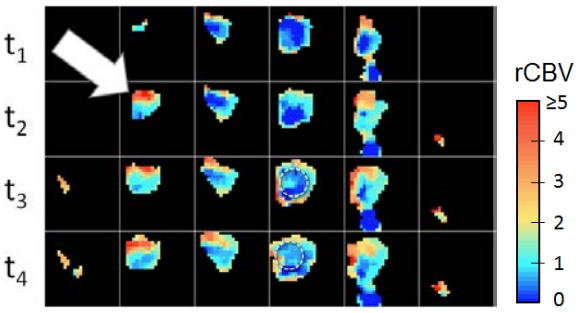


Figure 2. 2D rCBV maps of all slices (horizontally aligned) of study L_1 for all four time steps (vertically aligned) of the study. At t_1 and t_4, the tumor is covered by four and six slices, respectively. Note the hot spot with high rCBV values (arrow). The encircled regions exhibit lower rCBV values.

3D Overview. The 3D overview provides a direct volume rendering applying the same color scale (see Fig. 2) for highlighting tumor voxels. A linear opacity transfer function assigns α-values of 0.2 to low rCBV values ≤ 1 and $\alpha = 1$ to rCBV values ≥ 5. The brain (extracted from the perfusion dataset with thresholding) is displayed as isosurface context object and the tumor is revealed by cutting the isosurface, see Figure 3. The 3D overview visualizes the tumor's progression as well as its spatial localization in the brain. However, only limited information about a tumor's malignancy can be provided, since no exact quantification of rCBV hot spots is possible. The 3D overview reveals the spatial variability of rCBV hot spots. They may occur at slightly different positions across all scans of a longitudinal study due to tumor growth (see Fig. 3).

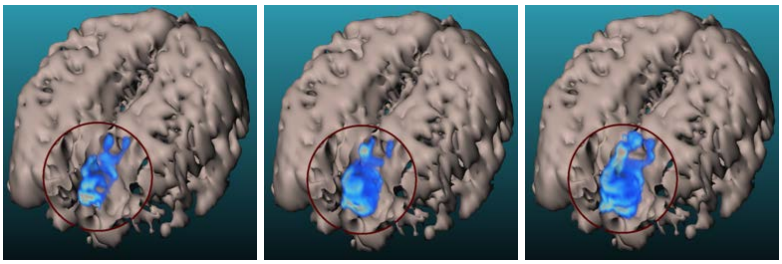


Figure 3. 3D view of tumor growth of L_1 at t_1 (left), t_2 (center), and t_3 (right). The brain is depicted as context object and the tumor is emphasized with a cut-out technique.

4.2.2 Visual Analysis of Local Tumor Heterogeneity

In clinical research, tumor growth and malignancy are associated with increased tumor heterogeneity due to necrosis mainly in the tumor center. Also, some neoangiogenetic tumor parts are expected to exhibit increased heterogeneity. For an assessment of local heterogeneity, approximated as correlation of rCBV and rCBF, we extract the local correlation coefficient (LCC) measure.[28] An important property of LCC is the independence of scaling of the data value range. Thus, the LCC maps are independent of rCBV and rCBF normalization factors, which had been used to normalize CBV and CBF maps. We apply LCC to the preprocessed rCBV and rCBF parameter maps, yielding a color-coded 2D LCC parameter map, see Figure 4. LCC values of 1 indicate an increasing linear relationship. Values ≈ 0 illustrate a missing linear dependency. We apply a heat color map from dark red to bright orange to map increasing values of LCC.

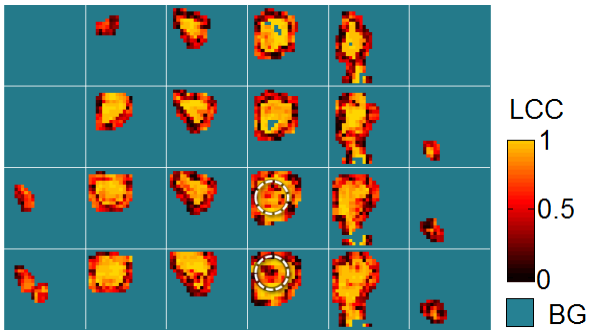


Figure 4. LCC parameter maps of L_1. Heterogeneous areas are mapped to dark red. The encircled regions at time steps t_3 and t_4 exhibit low rCBV values, see Figure 2.

4.2.3 The rCBV Profile Parallel Coordinates Plot

A parallel coordinates plot allows for the exploration of multivariate data, where each axis presents a data dimension. We adapt the parallel coordinates plot as follows: For each voxel, the rCBV values over time (i.e., the scans at different time steps of a study) are extracted, yielding rCBV curves. In Figure 5, the rCBV profile view of L_1 is presented (the same color scale is applied, recall Fig. 2). According to discussions with our clinical partners, the rCBV changes, i.e., the voxel's rCBV differences between two subsequent time steps were of great interest resulting in the rCBV change profile view, see Fig. 5. Hence, shadows support depth perception, and the rainbow color scale's hue changes support the differentiation between rCBV values. The data's order can be easily inferred from the curves' height. In addition, height lines support direct quantitative rCBV or rCBV change extraction, whereas curve differentiation is supported by contour lines and Fresnel shading.

To obtain groups of similar curves, we tested different strategies to sort the curves. We empirically determined the best result (i.e., the lowest amount of occlusions of curves in the background and the easiest detection of trends) with a sorting based on the squared differences of each rCBV change curve's integral. The order is applied to both rCBV profile views. Tumors with larger extents and thus a larger amount of voxels yield a higher number of rCBV change curves, which may lead to visual clutter. To reduce the curve number, the user

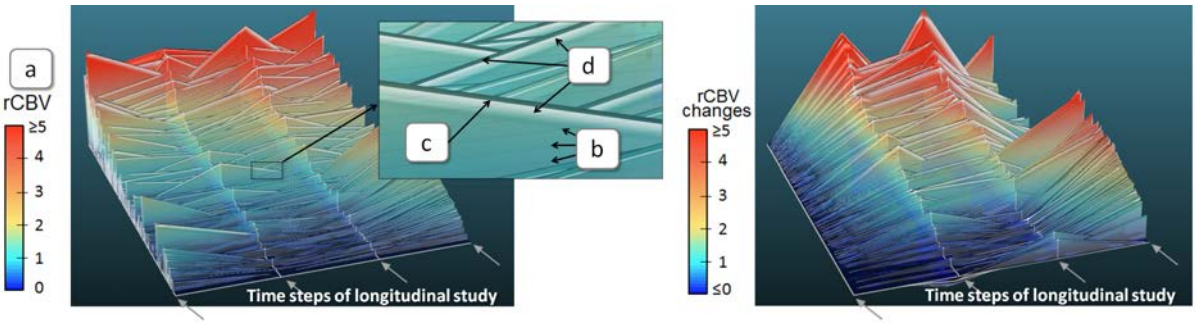


Figure 5. The rCBV profile view (left) with all voxels' rCBV curves. The rCBV values are color-coded (a) and height lines support direct measurement (b). Differentiation of curves is supported by dark contour lines (c), with Fresnel shading attached to (d). The rCBV change profile view presents the rCBV parameter curves' differences (right). Thus, all differences are mapped to zero for t_1.

can apply the rCBV threshold selection, and only curves that have at least one rCBV value greater than this threshold are included. In addition, the rCBV profile views allow for the *rCBV curve index selection*. Hence, all voxels with a corresponding curve index inside the selected range of curve indices are combined. The curve indices are sorted depending on the rCBV profile view. Although the rCBV profile views allow for simultaneous display of all voxels' rCBV changes, the spatial connectivity of the voxels is lost, as the voxels' order purely depends on the rCBV curves, and not on their spatial position.

4.3 Combination of Visual Analysis Techniques

The presented techniques can be combined via the rCBV threshold selection and the rCBV curve index selection. Then, a tumor voxel is only mapped to color or represented as curve when it is part of all chosen selections. Furthermore, the 2D views can be combined with lenses to allow for simultaneous evaluation of rCBV and LCC values.

5. EVALUATION

In this section, the studies $L_1 - L_4$ were discussed with two physicians regarding the clinical research questions (recall Sec. 3).

5.1 How can a comparative rCBV evaluation be achieved?

With the presented pipeline, a concurrent evaluation of each study's perfusion scans can be achieved. User interaction is only necessary during ROI placement in the contralateral healthy brain tissue. The pipeline can be applied to all brain DSC-MRI datasets independent of MRI scanner or specific scanning protocol parameters. Also, distortion of a ROI's rCBV values due to vessel voxels is prevented.

5.2 When does the LGG transformation start based on the evaluation of rCBV maps?

In general, rCBV maps are evaluated with the hot spot method. If the hot spot's average rCBV value is greater than a certain threshold, e.g., > 1.5,[1] the tumor is expected to be an HGG. The transformation of the LGG of L_1 started between t_1 and t_2, since a hot spot occurs at t_2 with an average rCBV value > 1.5. The patient of L_2 was monitored after surgical intervention. In Figure 6(a), the development of the recurrent LGG can be observed: First, there is a lesion dominated by radiation necrosis and low rCBV values. Next, the rCBV view at t_4 reveals some rCBV hot spots, whereas at t_5 increased extents of hot spots with higher rCBV are visible. The reported rCBV differences of oligodendrogliomas in comparison to other brain gliomas is reflected by L_3, since foci with high rCBV values do occur at all time steps (see Fig. 6(e)). Study L_4 comprises an LGG, whereas between t_3 and t_4 more hot spots with increased rCBV values occur, (see Fig. 6(i)). However, in clinical research, it is not always clear at which extent and number of hot spots and which threshold of rCBV a glioma is graded as HGG.[1,3,4] Furthermore, the normalization of rCBV with a possibly too low or too high normalization factor (extracted on the contralateral brain side) could strongly hamper this analysis and ROI placement suffers from interobserver variability.[7] These limitations also hold for our parameter maps.

5.3 How is the tumor's heterogeneity characterized in terms of local rCBV and rCBF correlation?

To support stereotactic biopsy, the most malignant tumor parts have to be determined. These parts are characterized by areas with increased angiogenesis and therefore with highest rCBV values. Furthermore, our clinical research partners presumed increased heterogeneity and thus uncorrelated rCBV and rCBF values for these parts. The LCC maps of L_1 are presented in Figure 4 and of L_2-L_4 in Figure 6. Hence, also the combination of LCC maps and rCBV maps via lenses is demonstrated. Two facts can be observed: First, with increased tumor growth, heterogeneous areas in the tumor center become visible, which may be caused by necrotic areas. Second, at the tumor's boundary, uncorrelated areas of rCBV and rCBF exist. During tumor growth, necrotic areas in the tumor center show up with typical low rCBV values, see the encircled regions in Figures 2 and 4. Necrotic tumor parts typically exhibit low rCBV values. For stereotactic biopsy, these parts should be explicitly spared. In contrast, areas with high rCBV and high heterogeneity in terms of LCC should be aimed at. These areas can be defined by applying an rCBV selection first. Next, a ROI is set in the remaining rCBV map (see Fig. 6(c)). Another finding of the LCC maps is the lack of relationship between LGG transformation and the amount of LCC approximated heterogeneity.

5.4 How does the tumor grow and develop?

For a qualitative overview of tumor growth, the 3D view is provided (see Fig. 3, Fig. 6(d)(h)(l)). Hence, no quantitative evaluation is possible. Quantitative information is provided in the 2D parameter maps and the rCBV profile views. Furthermore, the rCBV profile views provide a fast overview of significantly changing rCBV values and thus a possible LGG transformation. This relationship is illustrated in Figure 6(m)-(o), where the rCBV values and the rCBV changes of L_4 suddenly increase after t_3, matching the estimated transformation point in time. Due to the large tumor extent of L_3, a reduction of the number of curves is necessary, see Figure 6(p)-(r). Still, no rapid rCBV value increase can be observed due to the oligodendroglioma type.

5.5 How can visual exploration and analysis of longitudinal studies be carried out?

A possible user scenario is carried out in the following way: First, the preprocessing pipeline provides co-registered rCBV parameter maps. Now, the clinical researcher starts with the spatial 3D cut-out view for a first overview. Next, the datasets are analyzed in the rCBV profile curve view and the 2D rCBV maps. Hence, the transformation time step can be estimated by examining the rCBV curves, the rCBV change curves, or scalar values of rCBV. The analysis involves combinations of the rCBV threshold selection and the rCBV curve index selection (see Figures 6((c),(g),(k), and (o)) to determine the most malignant tumor part for tumor grading or stereotactic biopsy.

In summary, for all tumors but the oligodendroglioma of L_3, the transformation time step (which initially was estimated also based on structural MRI sequences) could be approximated with the application employing only the perfusion datasets (recall Sec. 3.2).

6. CONCLUSION

We presented an application for the evaluation of longitudinal brain perfusion studies focusing on the development of LGGs within the context of clinical research. A preprocessing pipeline enables the simultaneous evaluation of tumor localization, size and rCBV values of all perfusion scans. The visual analysis techniques include linked 2D parameter maps, a 3D overview and the new 3D rCBV profile view that represents all rCBV values of all voxels at all time steps. Once the user examined the LGGs temporal development and the point in time of a starting LGG transformation, the analysis can be restricted to foci with high rCBV values and local tumor heterogeneity. Combination of the different views as well as selecting a set of voxels which may represent the most malignant tumor part is carried out with lenses in the 2D views, the rCBV threshold selection, and the rCBV curve index selection. The resulting voxel set can be used for ROI placement for tumor grading or stereotactic biopsy.

Without our application, a clinician would have to mentally integrate scans of the same study acquired at earlier time steps to evaluate the LGG progression. In addition, the application prevents vessel voxels to be integrated in ROIs and thus avoids distortions of rCBV averaging. Our clinical partners do especially like the rCBV

change profile view depicting each voxel's rCBV changes, due to their ability to visualize the temporal development. Hence, the tumor changes are analyzed for a retrospective evaluation. However, with the heterogeneity analysis based on the LCC, additional information is provided to guide a stereotactic biopsy.

Our framework is a prototype application for clinical research and may be adapted to longitudinal studies of breast and prostate tumors, where perfusion data is acquired for diagnosis and treatment monitoring. Due to the growing number of longitudinal medical image data and the general need of comparison of sequenced image data, e.g., cohort study data, we expect an increasing demand for visual exploration and analysis of this kind of medical data.

Acknowledgements

This work was partially supported by the German Research Foundation project SPP 1335 "Scalable Visual Analytics" and the "Yggdrasil" mobility programme of the Research Council of Norway.

REFERENCES

[1] Covarrubias, D. J., Rosen, B. R., and Lev, M. H., "Dynamic Magnetic Resonance Perfusion Imaging of Brain Tumors," *Oncologist* **9**(5), 528–37 (2004).

[2] Law, M., Young, R. J., Babb, J. S., and et al., "Gliomas: Predicting Time to Progression or Survival with Cerebral Blood Volume Measurements at Dynamic Susceptibility-weighted Contrast-enhanced Perfusion MR Imaging," *Radiology* **247**(2), 490–8 (2008).

[3] Emblem, K. E., Nedregaard, B., Nome, T., and et al., "Glioma Grading by Using Histogram Analysis of Blood Volume Heterogeneity from MR-derived CBV Maps," *Radiology* **247**(3), 808–17 (2008).

[4] Law, M., Yang, S., Wang, H., and et al., "Glioma Grading: Sensitivity, Specificity, and Predictive Values of Perfusion MR Imaging and Proton MR Spectroscopic Imaging Compared with Conventional MR Imaging," *AJNR* **24**(10), 1989–98 (2003).

[5] Völzke, H., Alte, D., Schmidt, C. O., and et al., "Cohort Profile: The Study of Health in Pomerania," *International Journal of Epidemiology* **40**(2), 294–307 (2011).

[6] Behrens, U., Teubner, J., Evertsz, C. J., Walz, M., Jürgens, H., and Peitgen, H.-O., "Computer-Assisted Dynamic Evaluation of Contrast-Enhanced-MRI," in [*Proc. of CARS*], 362–7 (1996).

[7] Wetzel, S. G., Cha, S., Johnson, G., Lee, P., Law, M., Kasow, D. L., Pierce, S. D., and Xue, X., "Relative Cerebral Blood Volume Measurements in Intracranial Mass Lesions: Interobserver and Intraobserver Reproducibility Study," *Radiology* **224**(3), 797–803 (2002).

[8] Bjørnerud, A. and Emblem, K. E., "A fully automated method for quantitative cerebral hemodynamic analysis using DSC-MRI," *J Cereb Blood Flow Metab* **30**, 1066–1078 (2010).

[9] Lupo, J., Cha, S., Chang, S. M., and Nelson, S. J., "Dynamic Susceptibility-Weighted Perfusion Imaging of High-Grade Gliomas: Characterization of Spatial Heterogeneity," *AJNR* **26**(6), 1446–54 (2005).

[10] Wegenkittl, R., Löffelmann, H., and Gröller, E., "Visualizing the behaviour of higher dimensional dynamical systems," in [*Proc. of IEEE Visualization*], 119–25 (1997).

[11] Gresh, D. L., Rogowitz, B. E., Winslow, R. L., Scollan, D. F., and Yung., C., "Weave: A system for visually linking 3-d and statistical visualizations, applied to cardiac simulation and measurement data.," in [*Proc. of IEEE Visualization*], 489–92 (2000).

[12] Doleisch, H., Gasser, M., and Hauser, H., "Interactive Feature Specification for Focus+Context Visualization of Complex Simulation Data," in [*Proc. of the Symposium on Data Visualisation*], 239–48 (2003).

[13] Oeltze, S., Doleisch, H., Hauser, H., Muigg, P., and Preim, B., "Interactive Visual Analysis of Perfusion Data," *IEEE TVCG* **13**(6), 1392–9 (2007).

[14] Oeltze, S., Hauser, H., Rørvik, J., Lundervold, A., and Preim, B., "Visual Analysis of Cerebral Perfusion Data – Four Interactive Approaches and a Comparison," in [*Proc. of ISPA*], 588–95 (2009).

[15] Kohle, S., Preim, B., Wiener, J., and Peitgen, H.-O., "Exploration of Time-varying Data for Medical Diagnosis," in [*Proc. of VMV*], 31–38 (2002).

[16] Coto, E., Grimm, S., Bruckner, S., Gröller, E., Kanitsar, A., and Rodriguez, O., "Mammo Explorer: An Advanced CAD Application for Breast DCE-MRI," in [*Proc. of VMV*], 91–98 (2005).

[17] Blaas, J., Botha, C. P., and Post, F. H., "Interactive Visualization of Multi-Field Medical Data Using Linked Physical and Feature-Space Views ," in [*IEEE EuroVis*], 123–30 (2007).

[18] Glaßer, S., Preim, U., Tönnies, K., and Preim, B., "A Visual Analytics Approach to Diagnosis of Breast DCE-MRI Data," *Computers & Graphics* **34**(5), 602–11 (2010).

[19] Preim, U., Glaßer, S., Preim, B., Fischbach, F., and Ricke, J., "Computer-Aided Diagnosis in Breast DCE-MRI - Quantification of the Heterogeneity of Breast Lesions," *European Journal of Radiology* **81**(7), 1532 – 1538 (2012).

[20] Ashburner, J., Csernansk, J. G., Davatzikos, C., Fox, N. C., Frisoni, G. B., and Thompson, P. M., "Computer-assisted imaging to assess brain structure in healthy and diseased brains," *Lancet Neurol* **2**(2), 79–88 (2003).

[21] Steenwijk, M., Milles, J., Buchem, M., Reiber, J., and Botha, C., "Integrated visual analysis for heterogeneous datasets in cohort studies," in [*IEEE VisWeek Workshop on Visual Analytics in Health Care.*], (2010).

[22] Netti, P. A., Roberge, S., Boucher, Y., Baxter, L. T., and Jain, R. K., "Effect of transvascular fluid exchange on pressure-flow relationship in tumors: A proposed mechanism for tumor blood flow heterogeneity," *Microvasc Res* **52**(1), 27–46 (1996).

[23] Lev, M. H., Ozsunar, Y., Henson, J. W., and et al., "Glial Tumor Grading and Outcome Prediction Using Dynamic Spin-Echo MR Susceptibility Mapping Compared with Conventional Contrast-Enhanced MR: Confounding Effect of Elevated rCBV of Oligodendroglimoas," *AJNR* **25**(2), 214–21 (2004).

[24] Studholme, C., Hill, D. L. G., and Hawkes, D. J., "An overlap invariant entropy measure of 3d medical image alignment," *Pattern Recognition* **32**(1), 71–86 (1999).

[25] Rosen, B. R., Belliveau, J. W., Vevea, J. M., and Brady, T. J., "Perfusion imaging with NMR contrast agents," *Magn Reson Med* **14**(2), 249–65 (1990).

[26] Østergaard, L., Weisskoff, R. M., Chesler, D. A., Gyldensted, C., and Rosen, B. R., "High resolution measurement of cerebral blood flow using intravascular tracer bolus passages. Part I: mathematical approach and statistical analysis," *Magnetic Resonance in Medicine* **36**(5), 715–25 (1996).

[27] Boxerman, J. L., Schmainda, K. M., and Weisskoff, R. M., "Relative cerebral blood volume maps corrected for contrast agent extravasation significantly correlate with glioma tumor grade, whereas uncorrected maps do not.," *AJNR* **27**(4), 859–67 (2006).

[28] Cachier, P. and Pennec, X., "Non-rigid registration by gradient descent on a gaussian-windowed similarity measure using convolutions.," in [*Proc. of MMBIA*], 182–9 (2000).

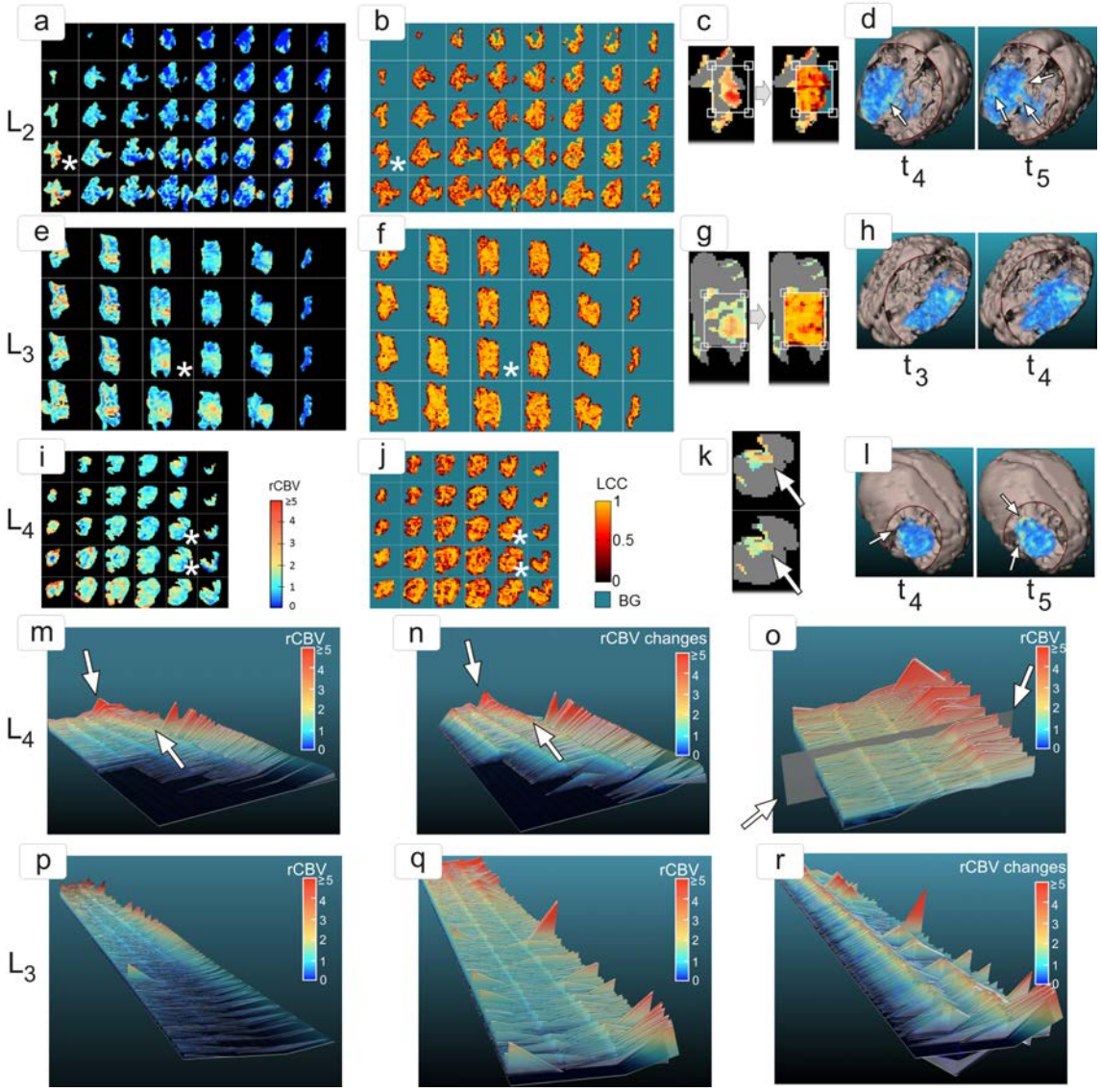


Figure 6. Presentation of the visual analysis techniques applied to L_2-L_4. To improve readability, the first and the last 2 tumor slices of L_2 and L_3 are not presented. In (a), (e) and (i) the rCBV maps with corresponding LCC maps in (b), (f), and (j) are depicted. In (d), (h), and (l), the 3D overviews are presented.
The combination of (a) and (b) via a lens for a selected slice (marked with *) is depicted in (c), revealing a hot spot with increased local heterogeneity at the tumor's boundary. In (a) and (d), increased rCBV values and larger hot spots (marked with arrows) at t_4 and t_5 are revealed. In (g), the combination of (e) and (f) via a lens for a selected slice (marked with *) is depicted, revealing a hot spot with increased local heterogeneity in the tumor's center. In (c) and (g), the rCBV threshold selection (rCBV threshold = 1.5) is applied, mapping lower rCBV values to gray. In (k), the rCBV curve index selection is applied to the slices marked with * in (i). Only the voxels, which were selected in the rCBV profile view in (o), are mapped to color. The region marked with an arrow exhibits only moderate rCBV, but these voxels belong to the selected curve indices. Thus, this tumor part may be one of the most malignant parts, since these voxels correspond to rCBV curves with the highest curve indices, i.e., the highest integrals. The rCBV profile view (m) and the rCBV change profile view (n) reveal increased rCBV values and changes, marked with arrows, after t_3 – matching the estimated transformation starting time. In (o), an rCBV curve index selection is carried out by applying the selection plane (see arrows) such that from the 600 rCBV curves with highest integral 300 curves are selected. The selection is then applied to all other views, see the rCBV map in (k).

Method for the evaluation of US perfusion for brain tumor surgery

C. Chalopin[1], S. Oeltze[2], B. Preim[2], A. Müns[3], J. Meixensberger[3], D. Lindner[3]

[1] Universität Leipzig, Innovation Center Computer Assisted Surgery (ICCAS), Leipzig, Germany
[2] Otto-von-Guericke-Universität, Institut für Simulation und Graphik (ISG), Magdeburg, Germany
[3] Universitätsklinik Leipzig, Klinik und Poliklinik für Neurochirurgie, Leipzig, Germany

Kontakt: claire.chalopin@iccas.de

Abstract:

This paper presents a method to evaluate intraoperative ultrasound (iUS) perfusion imaging of brain tumors acquired during resection surgeries. It consists in comparing the iUS perfusion with the standard preoperative MR perfusion. In a first step the iUS and MR perfusion data are represented in a common frame with the same pixel size. This is performed using image registration methods to achieve a pixel-wise correspondence between both data sets. In a second step the perfusion data are analyzed and visualized with the SimVis framework. It is possible to select region of interests of the tumor, such as the margins or the center, and to perform a region-based comparison between the iUS and MR perfusion data. The pipeline is demonstrated for one representative surgical case.

Keywords: US perfusion, MR perfusion, brain tumor

1 Problem

Ultrasound (US) perfusion is an imaging modality used to analyze the perfusion of human tissue. This technique is in comparison to MR perfusion less heavy and can be easily used in the operating room. An ultrasound contrast agent constituted of gas micro-bubbles is injected and its absorption by the tissue is qualitatively (visually) or quantitatively analyzed in the temporal sequence of the US images. The quantitative analysis consists in plotting the image intensities measured at a given pixel over time and in computing different perfusion parameters from the obtained time-intensity curve. The most common parameters are the peak intensity, the time to peak, and the area under the curve corresponding to the cerebral blood volume. So far there is no standard for the choice of the most relevant parameters and their visualization [8].
An important medical application of perfusion imaging is its ability to differentiate lesions from healthy tissue. Ultrasound perfusion imaging is nowadays routinely performed for the detection and operation of hepatic lesions [10, 13]. Cerebral US perfusion is also beneficial for the transcranial examination of brain tumors [5, 6, 11, 12], and was moreover tested intraoperatively during tumor surgeries [4]. In comparison to B-mode ultrasound imaging, intraoperative US (iUS) perfusion enables to depict more accurately the tumor margins and is therefore a promising control tool for the detection of possible remnants of tumor tissue. This imaging modality needs however to be still evaluated in the context of brain tumor surgery. In this paper, we present a method for the evaluation of intraoperative US perfusion by comparison with a gold standard, namely preoperative MR perfusion.

2 Material and Methods

Perfusion data acquisition
One day before the brain tumor surgery, preoperative MR data of the patient are acquired. The examination includes 3D contrasted T1-weighted MR anatomical data (mostly isotropic) and 3D+t T2*-weighted MR perfusion data. The MR perfusion data consists of several volumes acquired at different points in time rendering the contrast agent accumulation visible. The in-plane resolution is 1.75 mm x 1.75 mm, the slice thickness is 5mm, and t is 40.
The tumor surgery is guided using a sononavigation system (Sononavigator, Localite, Sankt Augustin, Germany) and a conventional US device (AplioXG, Toshiba, Medical Systems Europe, Zoetermeer, Netherland). At the beginning of the intervention, the anatomical MR volume is registered with the patient based on anatomical landmarks and the result is improved using a head surface registration technique. At any moment the surgeon is able to acquire an iUS volume by scanning the region of interest using a tracked 2D US transducer. The acquired volume is superimposed on the preoperative MR data on the monitor of the sononavigation system.
During the surgery, right after the skull opening (craniotomy), a bolus of 1.5 ml of US contrast agent (SonoVue, Bracco s.p.a., Milano, Italy) is injected into the patient. Based on US B-mode images of the tumor, the surgeon localizes with the US transducer the cross-section plane situated at the tumor middle. The position of the probe in the patient coordinate system is identified through the navigation system. A temporal sequence of 2D iUS perfusion images is then acquired with at a rate of 19.0 frames per second over about one minute. The in-plane resolution is 0.35 mm x 0.35 mm.

iUS-perfusion and MR-perfusion data registration
Three main difficulties hamper the comparison of the iUS-perfusion data with the preoperative MR-perfusion data:

- The different data dimension: 2D+t data and 3D+t data;
- The differences in image size and in-plane resolution;
- The orientation of the data in different coordinate systems.

We propose a registration pipeline establishing a pixel-to-pixel correspondence between the iUS-perfusion data and a reformatted slice through the MR-perfusion data. All steps have been implemented in a prototypical application within the MeVisLab framework (MeVis Medical Solutions AG).

Step 1: preregistration. The cross-section plane corresponding to the 2D iUS perfusion data is registered with the preoperative anatomical 3D MR data using the transform matrix M_{sono} provided by the sononavigation system. However, the tumors in both data are offset due to the brain shift after craniotomy.
Step 2: brain shift correction. The tumor margin is manually delineated in the 2D iUS perfusion data (Figure 1, left) and in the anatomical 3D MR data resulting in two sets of contour points: P_{iUS} (2D) and P_{MR} (3D). An Iterative Closest Point (ICP) algorithm is used to register P_{iUS} to P_{MR} [1] using a rigid transformations and an isotropic scaling. It provides a transformation matrix $M_{shift,}$ which together with M_{sono} enables us to find the plane in the anatomical 3D MR data corresponding to the acquisition plane of the iUS perfusion data.
Step 3: plane correspondance. Then the anatomical 3D MR data are transformed into the MR perfusion frame. The MR data are acquired within the same protocol and are therefore represented in the same coordinate system but with differ-

ent in-plane resolution. A rigid registration is employed using a Newton-type optimizer and normalized mutual information to cope with the non-linear image intensity relations [2] and provides the transformation matrix $M_{\text{perf.}}$ The matrix multiplication $M_{\text{perf}} \; x \; M_{\text{shift}} \; x \; M_{\text{sono}}$ yields the final transformation matrix M_{result}, which transforms the original iUS perfusion data into the corresponding plane in the MR perfusion data.

Step 4: pixel-wise correspondence. The MR perfusion data are resampled along the transformed iUS image plane leading to 2D+t MR perfusion data with the same x-y dimensions and the same in-plane resolution as the iUS data (Figure 1). A Trilinear interpolation is employed. A valid pixel-to-pixel correspondence in the iUS and MR perfusion images is obtained. Only the number of points in time still differs.

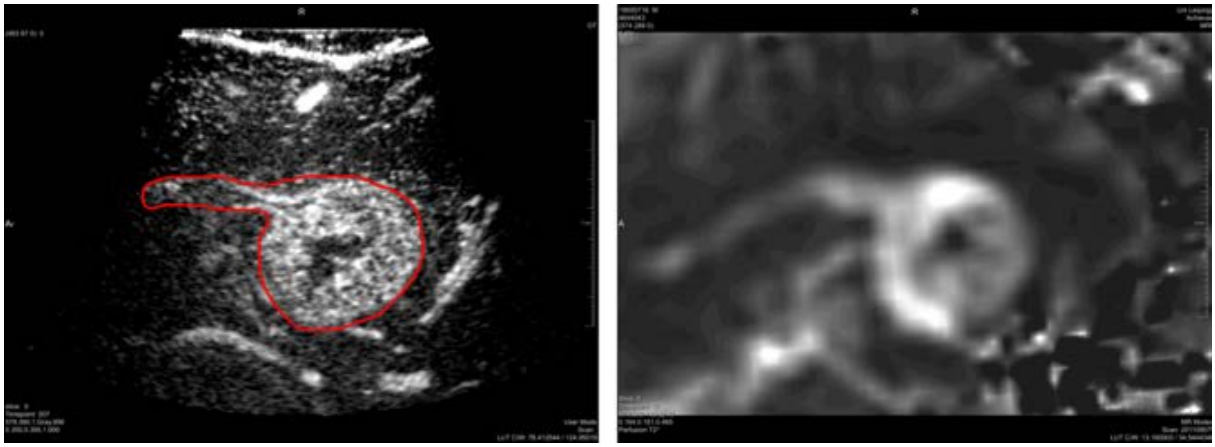


Figure 1: Registered intraoperative, contrast-enhanced Ultrasound (iUS) data (*left*) and preoperative MR perfusion data (*right*). The MR perfusion data has been resampled along the registered iUS image plane. The tumor margin is manually delineated in the iUS data and employed for registration as well as visualization purposes.

Intraoperative US-perfusion and MR-perfusion comparison

We compare iUS-perfusion and MR-perfusion based on perfusion parameters derived from the corresponding time-intensity curves [8]. Two pre-processing steps are carried out before parameter computation. First, the signal intensities of the T2*-weighted MR perfusion data are converted to contrast agent concentration [9]. This is a prerequisite for the determination of perfusion parameters, which are at least proportional to real quantitative hemodynamic parameters, such as cerebral blood volume and cerebral blood flow. Second, the iUS data is smoothed in the spatial as well as in the temporal domain to reduce the effects of strong speckle noise. A mean filter kernel of size 5 x 5 x 19, where 19 is the number of points in time, has been empirically determined as appropriate.

After the pre-processing, a set of seven perfusion parameters is computed pixel-wise from the iUS and the MR perfusion data [8]. The pre-processing and the parameter derivation have been implemented in MeVisLab. The fourteen 2D perfusion parameter images are aggregated for a concurrent interactive visual analysis and a comparison within the SimVis framework [3]. SimVis is a multiple coordinated view framework where each view is equipped with interactive drill-down operations for focusing on data features. Two classes of views are integrated: physical views, such as direct volume rendering, show information in the context of the spatiotemporal observation space while attribute views, such as scatter plots and histograms, show relationships between multiple data attributes. The user may drilldown the data by selecting interesting regions of the observation space or attribute ranges leading to a consistent highlighting of this selection in all other views (*brushing-and-linking*). SimVis has been applied to perfusion data in breast cancer diagnosis and in the diagnosis of ischemic stroke and Coronary Heart Disease [7].

3 Results

This pipeline was tested on data of a patient with a glioblastoma multiform located in the left frontal area of the brain. The iUS perfusion data reveals the enhanced tumor margins as well as a necrosis at the tumor center (Figure 1, left). For this patient T2*-weighted MR perfusion data are available.

A typical analysis session with SimVis is shown in Figure 2. The upper view shows the tumor in its spatial context. A gradient image of the point in time employed for tumor delineation in the iUS data serves as background (Figure 1, left). Blood vessels as well as the tumor itself appear elevated due to their high contrast agent accumulation at this point in time leading to a strong separation from the surrounding tissue, i.e. a high gradient. The whole tumor extent is indicated by a highly transparent layer in front of the gradient image. The current attribute selection is colored according to the area under the curve derived from the iUS data. The selection has been defined in the histogram view (lower middle). The histogram shows the results of an Euclidean distance transform, which determines for each pixel of the tumor

the distance to the surrounding tissue. High distances (dark bars) are brushed by means of a rectangular attribute selection thereby restricting the analysis to the center of the tumor. Moving the brush inside the histogram triggers an instant update of all other views and allows for a seamless inspection of tumor zones from the center to the margin.
The lower left scatter plot opposes the area under the curve of the iUS (x) and the MR (y) data. Each image pixel is represented by a colored dot. The transparency of each dot is modified with respect to the frequency of the underlying attribute pair. The current selection is rendered dark. Its shape illustrates the correlation of the chosen parameter between the two imaging modalities. The lower right parallel coordinates plot opposes three more pairs of perfusion parameters. The vertical axes are alternately associated with the parameter derived from the iUS and from the MR perfusion data, respectively. Each image pixel is represented by polyline connecting the axes. The current selection is rendered dark. The parallel coordinates facilitate a visualization of the entire high-dimensional space of perfusion parameters. From the course of the polylines, correlations as well as data clusters may be inferred. The current selection may be further refined by brushing the scatter plot and/or the parallel coordinates.

4 Discussion

The brain tumor tested here has an irregular shape, which is of benefit to the registration step. The appropriateness of the ICP algorithm for nearly spherical tumors, such as metastases, has to be investigated. Moreover, the success of the registration method is dependent on the manually performed tumor segmentations. It would be interesting to test the influence of slight changes to the tumor delineation. Further, it should be investigated whether perfusion parameters are best derived after resampling the MR perfusion data or before. In the latter case, they would be computed for each voxel

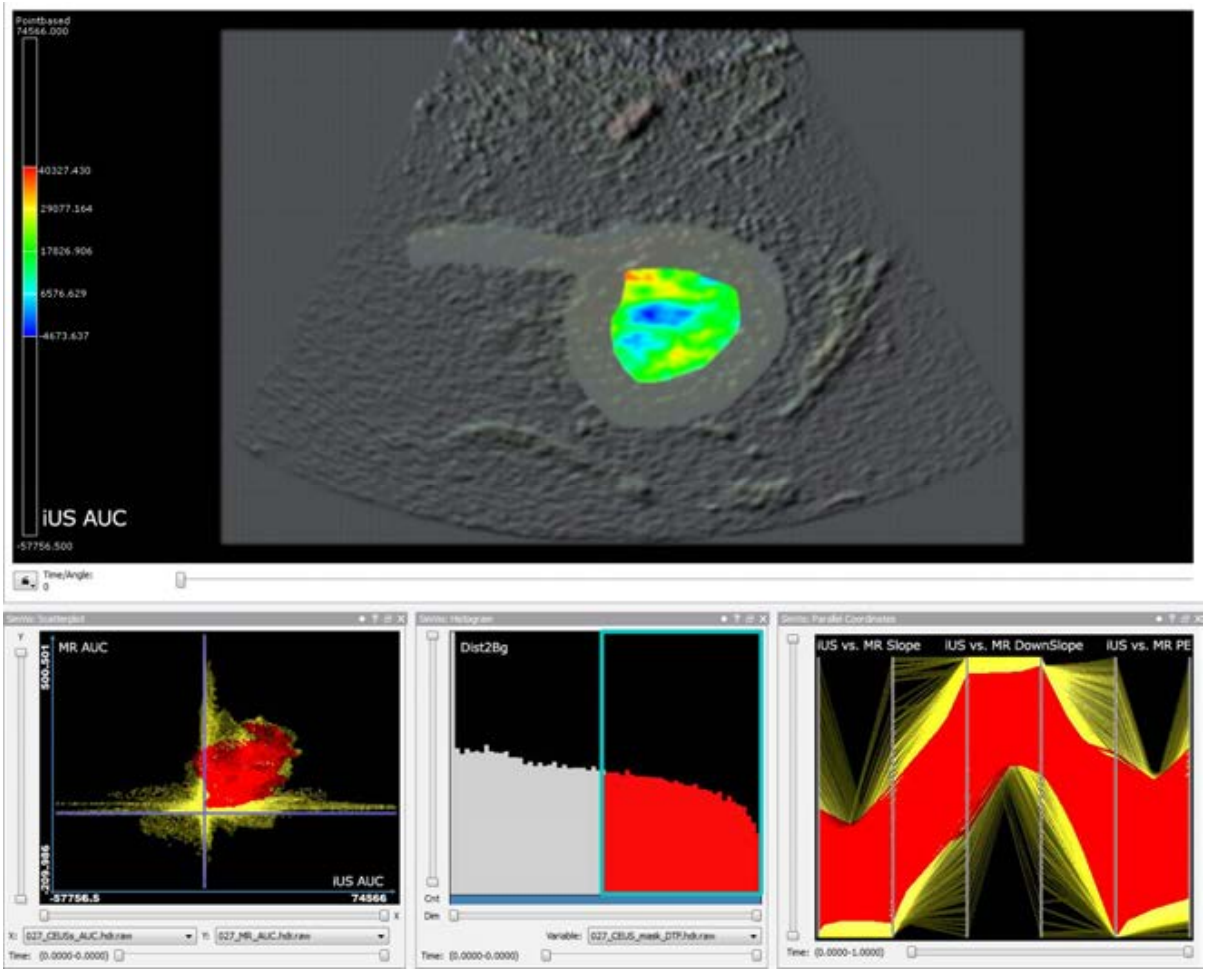


Figure 2: Concurrent, interactive visual analysis of perfusion parameters derived from intraoperative contrast-enhanced Ultrasound data (iUS) and preoperative MR perfusion data. The user may select and interesting range of parameter values within one or more of the attribute views (lower row) causing a colored emphasis of the associated tumor part in the spatial view (upper row).

and the resulting parameter volumes would then be resampled along the iUS image plane. So far, a mean filter was used to smooth the iUS perfusion data. Alternatively, fitting the perfusion curves with a gamma variate function could be more appropriate. This could attenuate as well the motion artifacts due to hand jittering during the intraoperative acquisition. Finally, the possible selection of further regions of interest in the image data such as blood vessels should be added.

5 Conclusion

In this paper, we presented a pipeline for the evaluation of intraoperative US perfusion data of patients with brain tumors. The pipeline consists of registration steps and an interactive visual analysis step. It facilitates a comparison of the perfusion parameters of the iUS perfusion data with those of the MR perfusion data based on a pixel-wise correspondence between the data. The implementation of the visual analysis step enables the user to select regions of interest of the tumor according to its anatomy, for example the margin or a possible necrosis, and according to its features, i.e. the perfusion parameters. Testing the pipeline on more patient data is required to improve the registration steps and the visualization and analysis of perfusion parameters according to clinical purposes.

6 References

[1] PJ Besl, ND McKay, *A Method for Registration of 3-D Shapes*, IEEE T Pattern Anal, **14(2)** 239–256 (1992)

[2] T Boehler, D van Straaten, S Wirtz, H-O Peitgen, *A robust and extendible framework for medical image registration focused on rapid clinical application deployment*, Comput Biol Med, **41(6)** 340-349 (2011).

[3] H Doleisch, M Gasser, H Hauser, *Interactive feature specification for focus+context Visualization of complex simulation data*, In: Joint Eurographics / IEEE TCVG Symposium on Visualization (VisSym), Grenoble, France, May 26-28 (2003).

[4] C Hansen, W Wilkening, H Ermert, *Intraoperative contrast enhanced perfusion imaging of cerebral tumors*, In: IEEE Ultrasonics Symposium, Rotterdam, The Netherlands, September 18-21 (2005).

[5] JU Harrer, L Mayfrank, M Mull, C Klötzsch, *Second harmonic imaging: a new ultrasound technique to assess human brain tumor perfusion*, J Neurol Neurosurg Psychiatry, **74** 333-338 (2003).

[6] JU Harrer, W Möller-Hartmann, MF Oerterl, C Klötzsch, *Perfusion imaging of high-grade gliomas: a comparison between contrast harmonic and magnetic resonance imaging*, J Neurosurg, **101** 700-703 (2004).

[7] S Oeltze, H Doleisch, H Hauser, P Muigg, B Preim, *Interactive Visual Analysis of Perfusion Data*, IEEE T Vis Comput Gr, **13(6)** 1392–1399 (2007).

[8] B Preim, S Oeltze, M Mlejnek, E Gröller, A Hennemuth, S Behrens, *Survey of the Visual Exploration and Analysis of Perfusion Data*, IEEE T Vis Comput Gr, **15(2)** 205-220 (2009).

[9] BR Rosen, JW Belliveau, JM Vevea, TJ Brady, *Perfusion imaging with NMR contrast agents*, Magnet Reson Med, **14(2)** 249-65 (1990).

[10] AK Siösteen, A Elvin, *Intra-operative uses of contrast-enhanced ultrasound*, Eur Radiol Suppl **14(8)** P87-P95 (2004).

[11] K van Leyen, C Klötzsch, JU Harrer, *Brain tumor imaging with transcranial sonography: state of the art and review of the literature*, Ultraschall in Med, **32** 572-581 (2011).

[12] E Vicenzini, R Delfini, F Magri, F Puccinelli, M Altieri, A Santoro, MF Giannoni, L Bozzao, V Di Piero, GL Lenzi, *Semiquantitative human cerebral perfusion assessment with ultrasound in brain space-occupying lesions*, J Ultrasound Med, **27** 685-692 (2008).

[13] J Zacherl, C Scheuba, M Imhof, M Zacherl, F Längle, P Pokieser, F Wrba, E Wenzl, F Mühlbacher, R Jakesz, R Steininger, *Current value of intraoperative sonography during surgery of hepatic neoplasms*, World J Surg, **26** 550-554 (2002).

Part V

Surveys

This part of the postdoctoral thesis cumulates the following publications:

Chapter 14 [191] T. Ropinski, **S. Oeltze**, and B. Preim, "Survey of Glyph-Based Visualization Techniques for Spatial Multivariate Medical Data", *Computers and Graphics (C&G)*, vol. 35, no. 2, pp. 392-401, 2011.

Chapter 15 [173] **S. Oeltze-Jafra** and B. Preim, "Survey of Labeling Techniques in Medical Visualizations", *Eurographics Workshop on Visual Computing for Biology and Medicine (EG VCBM)*, pp. 199-208, 2014.

Survey of Glyph-based Visualization Techniques for Spatial Multivariate Medical Data

Timo Ropinski, Steffen Oeltze and Bernhard Preim

Abstract

In this survey article, we review glyph-based visualization techniques which have been exploited when visualizing spatial multivariate medical data. To classify these techniques, we derive a taxonomy of glyph properties that is based on classification concepts established in information visualization. By considering both the glyph visualization as well as the interaction techniques that are employed to generate or explore the glyph visualization, we are able to classify glyph techniques into two main groups: those supporting pre-attentive and those supporting *attentive* processing. With respect to this classification, we review glyph-based techniques described in the medical visualization literature. Based on the outcome of the literature review, we propose design guidelines for glyph visualizations in the medical domain.

Keywords: Glyphs, medical visualization, multivariate data

1. Introduction

Data sets acquired in the medical domain, contain a multitude of information that provides a huge potential for diagnosis and individualized therapy planning. However, to exploit this potential, the data needs to be interpreted efficiently. In the past, mainly the increasing resolution of the scalar volume data sets posed a challenge for medical visualization. Algorithms had to be developed in order to extract and emphasize structures of interest. Today, also the multiple variables which can be derived from different modalities or time steps, pose a challenge when interpreting visualizations. When dealing with 3D data sets, as done in this article, additional problems arise, as for instance occlusion handling or choosing an appropriate projection.

One essential example, where multivariate data sets accrue is the 4D ultrasound acquisition of the human heart, which recently became a routine proceeding. With this method it is possible to derive multiple variables during one examination, i. e., information regarding the structure of the heart as well as direction and amount of blood flow. Another example are data sets acquired with multimodal medical scanners. When, for instance combining positron emission tomography (PET) with computed tomography (CT) it becomes possible to obtain an integrated visualization of metabolism activity within a high resolution structural context. To address the visualization challenges posed by these multivariate medical data sets, glyph-based techniques are a viable option [1, 2].

Glyphs are considered as symbolic or iconic representations of one or more variables of a data set. They are usually geometric objects, whose visual representation can be altered through changing the glyph properties. By using a *parameter mapping function* (PMF), the variables that have to be represented can be associated with one or more properties of a glyph, e. g., shape, size or color. The parameter mapping can either be continuous or discrete. When interpreting glyph-based visualizations, the visual representation of individual glyphs as well as the overall structure given by the arrangement of all glyphs can be exploited [1, 2]. Bürger and Hauser describe glyphs as just one visualization technique for multi-variate data [2]. They mention the benefit that a large number of data dimensions can be incorporated and the mapping can be used to reflect semantics. Another advantage is the ability to combine glyphs with other visualization techniques. On the other hand continuity is not given, and thus the data set depicted by the glyphs is only visualized in a discrete manner. Therefore, the visualization designer should carefully review the visualization requirements before using glyphs.

In this article, we describe and classify the glyph-based techniques exploited when visualizing spatial multivariate medical data. Ward classifies different glyph placing strategies and proposes rules for their usage in the context of information visualization [1]. In this article, we restrict ourselves to the review of glyph-based visualization techniques for medical data sets. In contrast to most applications in information visualization, in medical visualization the spatial embedding of the glyphs is of major interest. In almost every case, the data is regularly sampled, whereby for each sampling point either one or a set of scalar variables are present. We propose a taxonomy of glyph properties specifically designed for medical visualization that is based on findings from the perception literature. Rather than classifying glyph techniques with respect to their technical properties, we address the way they communicate the information to be visualized. This is essential since in medical diagnosis time is usually the limiting factor, and thus an efficient communication is essential. The described taxonomy extends the taxonomy proposed by Ropinski and Preim [3]. In particular, we now also consider the findings from higher-order tensor glyphs which play an important role in the literature. Further-

more, we introduce a more strict differentiation between *pre-attentive* and *attentive* glyph properties. Thus, we introduce the appearance property as a pre-attentive glyph property. These new considerations lead to a concise set of usage guidelines for glyph-based medical visualizations, which we believe has the potential to improve future glyph applications in the medical domain.

After discussing our glyph taxonomy in the next section, we will describe the usage of glyphs in selected medical application areas in Section 3. There, we will cover visualization of cardiac MRI data, and diffusion weighted MRI (DW-MRI) data in general. Based on the best practice we have identified in our literature review, we propose some usage guidelines in Section 4, before concluding the paper in Section 5.

2. Glyph-Based Medical Visualization Taxonomy

Efficiency as well as accuracy of information processing play a crucial role in clinical practice, and therefore an intuitive association between the visualization of data and their meaning is important. Glyph representations are not only used in medical research. It has also been shown in user studies that their use can be beneficial in the medical diagnosis [4]. However, due to the time constraints during a diagnosis, it is of great interest that the most important of these associations can be done in a very short time frame. Semiotic theory concentrates on signs and how they convey meaning. According to semiotic theory, stimuli are processed in two phases [5]. The first phase is the pre-attentive one, where impulses are perceived *in parallel* and instantly (within 250 ms) as one entity. Within this phase facts that can be easily perceived are extracted, for example the overall structure of a visualization, strong differences in shape, and strong differences in color [6]. In the second more goal-driven phase of the perception process the visualization is analyzed *sequentially*, i. e., parts of the visualization are identified and observed more detailed, one after another. Due to the conscious nature of this phase, we refer to it as *attentive* phase in the remainder of this article. According to Treisman, the initial pre-attentive phase is the major step towards improved comprehension [7]. This theory is supported by recent findings in fMRI-based brain research [8]. Thus, visualization designers should choose the pre-attentive stimuli wisely in order to communicate the desired information effectively. In fact, most glyph-based visualizations exploit pre-attentive stimuli, such that the distribution of glyph size, shape and color as well as the glyph aggregation aid visual comprehension. Attentive stimuli are mainly considered in glyph-based visualizations by exploiting interaction metaphors. Thus, after the pre-attentively perceivable information is extracted, the user can interactively explore the glyph visualization. This explorative process is also reflected in Shneidermans *overview, zoom, filter out, details-on-demand* concept that describes a general explorative task to be performed through these four steps [9].

Therefore, our taxonomy that is based on the preliminary work presented in [3], allows to distinguish between pre-attentive stimuli (Section 2.1) and attentive stimuli processing through appropriate interaction metaphors (Section 2.2). To be able to incorporate also future glyph-based techniques, we do not consider entire glyph visualizations, but regard different aspects as glyph shapes, parameter mapping or placement, which are integrated into the taxonomy.

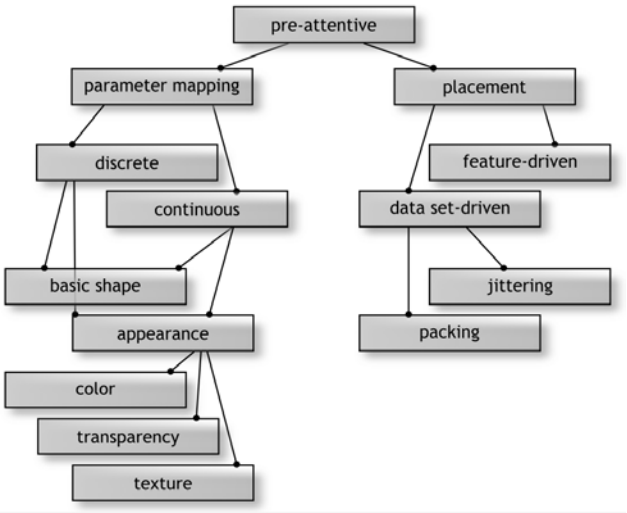


Figure 1: We consider parameter mapping and glyph placement as the two main groups of glyph properties, which are perceived pre-attentively.

2.1. Pre-Attentive Stimuli

Pre-attentive stimuli, relevant for glyph-based visualization, can be classified with respect to glyph shapes as well as glyph appearance, such as color, transparency and texture (see Fig. 1). Furthermore, the glyph placement strategies have an influence, since the spatial distribution can also be perceived pre-attentively. These pre-attentive stimuli can be exploited only in order to extract basic features of objects in the display which include colors, closure, line ends, contrast, tilt, curvature, and size [7]. A quantitative analysis occurs in the attentive processing phase which we address in Section 2.2.

2.1.1. Glyph Shape

The shape of a glyph is the main characteristic and it is important that it can be perceived easily and unambiguously [10]. Since the perception of shapes as well as spatial relationships is more accurate than the perception of quantities depicted by colors [11], the glyph shapes, and also glyph placements (see Section 2.1.3), are primarily used to convey information. We distinguish two main groups of glyph shapes:

1. *Basic* glyph shapes are geometric objects which can be modified by changing their geometric properties, such as size or orientation. Wide-spread examples are spheres, cuboids, and ellipsoids.
2. *Composite* glyph shapes are composed of the basic glyph shapes. Composite glyph shapes require a more specialized mapping function, i. e., parameters usually cannot be mapped to geometric properties as radius or length. They are often used to display multivariate data, whereby each variable can be communicated using another property of the composed building blocks.

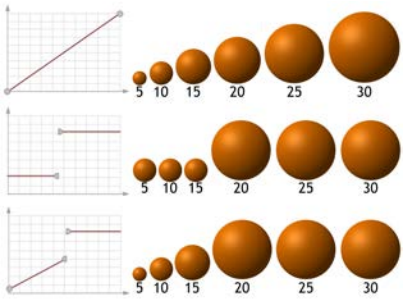
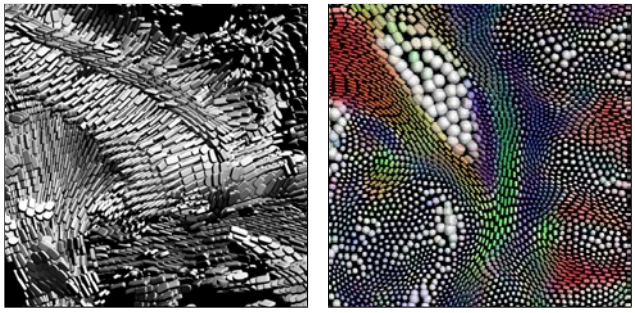


Figure 2: The mapping function (left) maps a parameter to the size of a glyph (right). Depending on the continuity of the mapping function, its reconstruction can be rather challenging without having a legend.

Regarding the complexity of composite glyph shapes, we assume that mainly basic glyph shapes benefit from the improved perception of shapes and spatial relationships in the pre-attentive phase, and consequently discuss composite glyph shapes within Section 2.2.

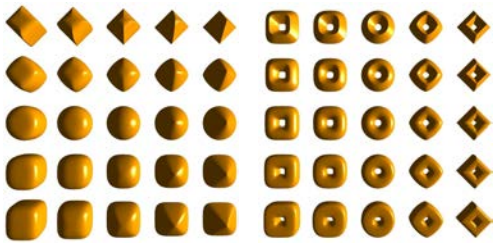
Continuous and discrete mapping. For both basic and composite glyph shapes either a *continuous* or a *discrete* mapping can be used in order to communicate information based on the glyph's properties. Ropinski et al. described variations of mapping functions [12]. They point out that also a *step mapping function* can be used in order to allow a better differentiation of the values to be visualized. Such a mapping is shown in Fig. 2, where glyphs representing different data values are shown next to the respective PMF. The mapping type should always be application-driven. For instance, to support the decision process in therapy and diagnosis, a discrete mapping might be appropriate, where color could be used to classify tissue states, e. g., benign suspicious, malign. In contrast, when quantifying perfusion, the whole range of scalar values might be of interest. When using a spherical glyph, such a scalar data value can be continuously mapped to its radius (recall Fig. 2). Discrete mappings often better support a quantitative analysis. Since a perspective projection is used in many cases, the size of a glyph cannot be *measured* in image space, without considering the perspective distortion. While color scales are sufficient to enable the interpretation of color-coded information, a similar concept is not sufficient to show the meaning of glyph sizes. Only a tendency can be expressed by using such a scale.

As mentioned above, it must be also ensured that the shapes are distinguishable independent of the viewing direction. Superquadrics do not only satisfy the criterion of unambiguous perception [10], they can also be used to map a multitude of variables (see Fig. 3(a)). According to Barr's definition of superquadrics [14], they are specified besides their size based on two parameters α and β, which influence their roundness. In Fig. 4, an ellipsoid and a toroid superquadric shape are shown with varying α and β parameters. Thus, when considering their size as another glyph property, ellipsoid superquadrics provide three degrees of freedom (DoF) for parameter mapping and toroid superquadrics provide four DoF for parameter mapping, since the radius of the tube can be additionally exploited for parameter mapping. For both cases, the glyph orientation can be considered as additional DoF. In contrast, a spherical shape only provides a mapping with one DoF, given by the radius.



(a) (b)

Figure 3: Superquadrics are unambiguously perceivable [10] (a). Glyphs showing diffusion directions are packed in order to remove undue visual emphasis of the regular sampling grid of the data and to illustrate larger-scale continuous features [13] (b). (Images courtesy of G. Kindlmann.)



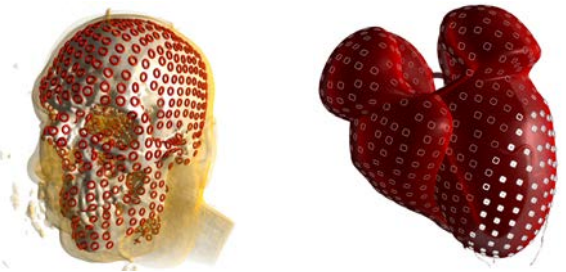
(a) (b)

Figure 4: Superquadric glyph shapes can convey multiple variables by changing their α and β parameter. This variation influences the roundness of the ellipsoid (a) and the toroid superquadric (b) along different principle directions.

2.1.2. Glyph Appearance

Besides a glyph's shape, its appearance is the property most commonly used to convey information. In the context of this article, we refer to the combination of color, transparency and texture as *glyph appearance*. Similar to the shape-based parameter mapping, the mapping of values to the glyph appearance can also be either continuous or discrete. When using a continuous color mapping, an absolute quantification is difficult to achieve because differences in color are harder to perceive than for instance spatial distances [11]. Thus color perception can only be used to get an overview during the pre-attentive phase, i. e., local maxima and minima as well as gradients may be identified. In the following attentive phase, color scales can assist the user when interpreting the visualization. Due to the relatively high occurrence of color blindness, a color scale is preferable, which not only varies in hue, but also in luminance.

A common extension of color-coding is to employ transparency. Transparency, however, makes the perception of occlusion relations more difficult. Since occlusion is one of the strongest depth cues, and psychophysical experiments indicate that spatial perception is improved when using multiple depth cues [15], the use of transparency is expected to hamper the visual perception of objects [16]. Therefore, transparency may not be used for quantification either. In contrast it may be used



(a) (b)

Figure 5: Feature-driven placement of glyphs within two data sets, where the glyphs are oriented according to the surface normal. Toroidal glyphs aligned to the surface of the skull (a). Glyphs depicting PET intensities as positioned on the surface of the human heart extracted from simulated CT data [12] (b). To emphasize potentially interesting regions an inverse parameter mapping is used.

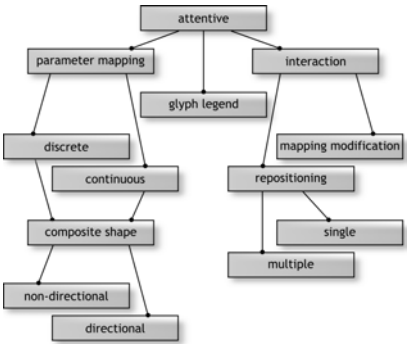


Figure 6: We consider composite glyph shapes, glyph legends and interactive techniques as the three main groups of glyph properties, which are relevant during the attentive processing.

to de-emphasize less important glyphs or to depict uncertainty. To communicate more complex information through the glyph appearance, as for instance directional information, also textures can be used.

2.1.3. Glyph Placement

To achieve a beneficial glyph visualization, also the glyph placement is crucial. Ward has already proposed a taxonomy for glyph placement strategies in the context of information visualization [1]. Some concepts of his taxonomy can be transferred to medical visualization, whereas some concepts should be omitted. In our opinion, neither a data-driven nor a structure-driven approach for glyph placement should be exploited in medical visualization (data-driven placement is based on the data dimensions and structure-driven placement on the relationship between data points), and thus both can be omitted. In contrast, in most medical applications, a reproduction strategy as described in [5] is exploited. Thus, glyphs are either placed based on the underlying regular grid, or based on the location of features present in the data set. Therefore, we distinguish between *data set-driven* and *feature-driven* placement. Placement on the regular grid is a data set-driven placement strategy, while the isosurface placement [12] is a feature-driven placement strategy (see Fig. 5). Both placement strategies are usually combined with a spatial context and may contain overlapping or non-overlapping glyphs.

Avoiding image space clustering. When choosing a data set-driven placement, the underlying structure of the regular grid usually has a major influence on the visualization. Thus it can unintentionally emphasize or even feign a non-existent glyph aggregation. However, there is a variety of techniques, which help to avoid this inadvertent effect of the underlying grid structure. Laidlaw et al. [17] proposed a jittered placement in order to reduce the aliasing introduced by the regular grid. Similarly, Meyer-Spradow et al. proposed a random distribution with relaxation, in order to get a uniform glyph distribution, when using feature-based glyph placement [4]. Bokinsky presented data-driven spots, which are used to display multiple scalar fields [18]. These spots are colored Gaussian splats, which are placed on a jittered grid. A more sophisticated approach has been proposed by Kindlmann and Westin [13]. They exploited a particle system in order to generate a packed glyph placement, which combines the continuous character of a texture with the used glyph technique (see Fig. 3(b)).

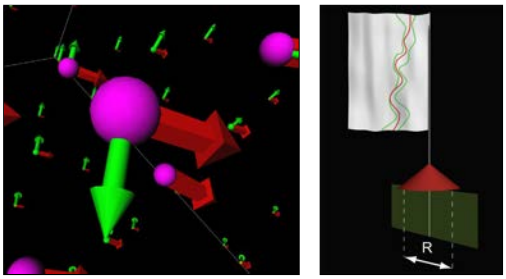
2.2. Attentive Stimuli

While the techniques described in the previous subsection primarily provide a first impression, the techniques described in this subsection allow to get into detail and possibly derive quantitative results. Thus, besides the composite glyph shapes mentioned above, we mainly focus on glyph legends and glyph interaction, which are integrated into our taxonomy as shown in Figure 6.

2.2.1. Composite Glyph Shape

In comparison to basic glyph shapes, composite glyph shapes (recall Section 2.1.1) may communicate more complex information in terms of dimensionality and arbitrary mappings, whereby the DoF for a parameter mapping is highly dependent on the type of glyph.

We consider *directional glyphs* as a subset of composite glyphs. When using directional glyphs, the semantic of the visualized data values is taken into account and is expressed by means of the glyph shape, e. g., an arrow is used to communicate a direction of movement. In medical visualization,



(a) Customizable glyph (b) Profile flag glyph

Figure 7: Composite glyphs can be customized in order to depict multiple data values [19] (a). More complex information can be visualized by integrating projection surfaces into composite glyphs [20] (b). (Images courtesy of M. Kraus and M. Mlejnek.)

directional glyph shapes may be used to visualize blood flow or tissue movement.

Composite glyph shapes are frequently used in information visualization [21, 22]. In scientific visualization, Kraus and Ertl have proposed a system that allows non-programmers to intuitively customize composite glyphs [19]. With their system, the user can generate a composite glyph by selecting shapes from a provided set of basic shapes, and configure the orientation and scaling of each shape to be dependent on the data values. An example glyph generated using their system is shown in Fig. 7(a). In medical visualization, the *profile flag* glyph has been proposed as a combination of basic geometric primitives combined with a surface used to project more complex information (see Fig. 7(b)) [20]. The glyph consists of several basic primitives and allows an efficient exploration of knee MRI scans. It has a projection surface for displaying a profile through the scan, while the base cone shows which part of the scan is considered (see Fig. 7(b)). Emphasizing their attentive nature, profile flags have been proposed together with probing-like interaction metaphors: One or two glyphs are manually positioned to visualize the data value at a desired position. Further extensions of this glyph technique are described in [23].

2.2.2. Glyph Legend

To aid the interpretation of visually displayed information, often graphical legends are exploited. In many cases these legends are restricted to represent a color scale annotated with the according range of values. Alternatively, often symbolic maps are used as graphical legends, where a symbol is defined as a certain feature shown in the visualization. In glyph visualization, legends have to be considered as a hybrid between graphical and symbolic legends. On the one hand, the legend should allow the user to mentally reconstruct the parameter mapping, on the other hand, it might be also helpful to display glyphs for certain values which define an important threshold. Thus, with such a legend, the user is able to visually compare the glyphs shown in the current visualization with characteristic glyphs included in the legend [12] (see Fig. 8). However, since glyphs need to be visually matched with the glyph legend, absolute quantification is still difficult. Additionally, when certain parameters are mapped to a glyph's size, this might be influenced by the perspective distortion, making the visual matching of equally sized glyphs contained in the legend even more difficult. A typical glyph legend is given by multiple rows, each depicting the range of values for one glyph property. Thus, a glyph legend implicitly depicts the DoF of the used PMF.

2.2.3. Glyph Interaction

We consider the probing-like repositioning of glyphs as well as the interactive modification of the parameter mapping as the most important interactions influencing the attentive processing of glyph visualizations.

Repositioning glyphs. Many glyph visualization techniques exist, where the initial glyph placement can be interactively modified by moving one or two proxy glyphs through the data set [20, 24]. These glyphs, which are also referred to as probing tools [25], adapt their visual appearance based on the PMF to the new location they are moved to. Thus, repositioning glyphs can be compared to using a color picker tool.

Sigfridsson et al. have proposed such glyphs for tensor field visualization, where the glyphs can be positioned within a continuous field representation, in order to get quantitative values at the desired position [26]. A more complex probing glyph, is the previously mentioned profile flag glyph described in [20] (recall Fig. 7(b)). To allow a semi-quantitative analysis, two profile flags can be visualized and repositioned simultaneously such that their visualization is comparable. Thus, it has to be distinguished between single and multiple probing glyphs, which can be repositioned.

Parameter mapping modification. In comparison to reposition individual glyphs, interactively modifying the PMF is much less frequently considered in the literature. Only a few efforts have been undertaken in order to support interactive parameter mapping [27]. Instead, most researchers have focussed on how glyphs can be visualized when a certain parameter mapping is present [19, 10, 28, 12]. However, generating such a parameter mapping is crucial and should also be interactive. Especially because different glyph representations may be developed with the goal to emphasize certain features. With glyph filtering, it is possible to display only glyphs, which satisfy a certain selection criterion. For instance, in the context of diffusion tensor imaging (DTI) visualization, it would be possible to visualize only glyphs exceeding a certain level of fractional anisotropy. Another example would be blood flow visualization, where it might be desirable to exclude glyphs representing values above a certain speed level. In all cases, filtering has a major influence on the overall glyph distribution. Modifying this glyph distribution can be used to direct the user's attention, or to emphasize critical data values. Thus, glyph filtering could alternatively be also considered as a glyph placement strategy, since less meaningful glyphs are omitted from display. While most filtering techniques are performed automatically by the system, the user should also have the possibility to filter glyphs by changing the PMF.

3. Applications

In this section, we describe medical application areas, in which glyph-based visualization techniques are exploited. The goal is not to give a comprehensive overview of the described glyph techniques. Instead, we pick out DW-MRI and cardiac visualization as prominent examples. An overview on stress and strain tensor glyphs in non-medical domains can be found in [29].

Table 1 gives an overview of the applications described in this section. It relates the techniques to our taxonomy, by listing their most important properties, as the modality of the visualized data, the used glyph shapes, the glyph placement and the used parameter mapping.

3.1. Diffusion Weighted MRI Visualization

In medical visualization, DW-MRI is probably the domain where the usage of glyphs has been most intensively investigated. DW-MRI exploits that the diffusion rate of water

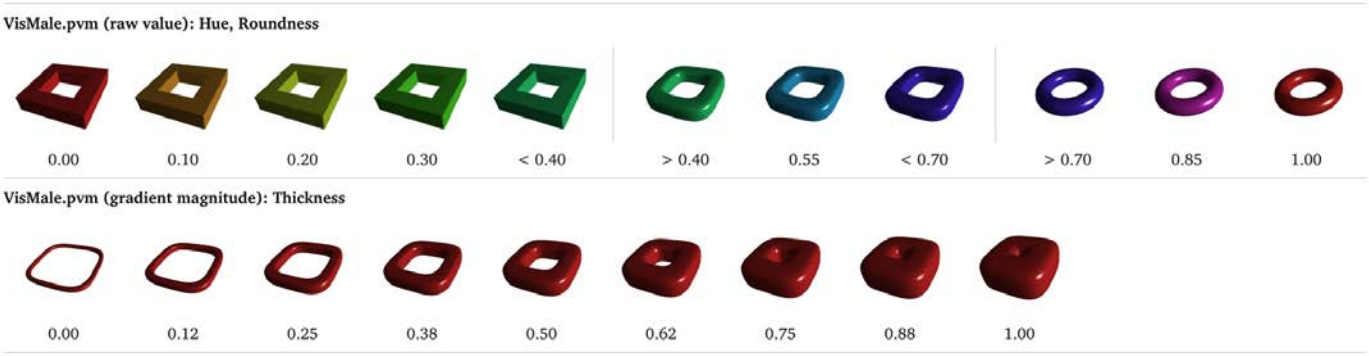


Figure 8: Glyph legends help with the interpretation of a glyph visualization [12]. Legends can be especially helpful when dealing with discontinuous mapping functions as shown in the two bottom rows of Figure 2.

Table 1: A listing of the most important glyph properties used in the applications described in Section 3: the modality of the visualized data, the used glyph shapes, their placement and the used parameter mapping.

Modality	Glyph Shape	Glyph Placement	Parameter Mapping	Reference
DTI	spherical harmonics	grid	higher order tensor to shape	[30]
DTI	HOME	-	higher order tensor to shape	[31]
DTI	ellipsoids	grid	one tensor value to glyph size	[17]
DTI	stroke layers	random	all tensor value to shape and color	[17]
DTI	superquadrics	grid	eigenvector and eigenvalue to shape and orientation	[10]
DW-MRI	deformed spheres	grid	deformation depicts diffusion direction	[28]
DTI	ellipsoids	*interactive*	eigenvector and eigenvalue to shape	[26]
DTI	lines	*interactive*	eigenvector and eigenvalue to length and direction	[26]
DTI	superquadrics	packing	eigenvector to color	[13]
DTI	Gabor patches	grid	eigenvector to color, anisotropy to transparency	[32]
DW-MRI	ellipsoids	grid	deformation to color	[33]
PET/CT	superquadrics	feature-driven	*interactive*	[12]
MRI	cuboids & custom	data set-driven	peak enhancement to color, Up Slope to size	[34]

molecules allows to derive information about the structures of the underlying tissue. However, glyphs are only suitable for getting a first impression immediately after the acquisition. Glyph visualizations cannot be exploited during a neurosurgical intervention since glyphs do not explicitly show the underlying fibers which must not be affected during an intervention.

Laidlaw et al. have proposed glyph techniques for the representation of DTI data derived from the mouse spinal cord [17]. They exploited an array of ellipsoids, where the shape of the ellipsoids present one tensor value, whereas their size is equal due to an introduced normalization. Their second technique used multiple layers of varying brush strokes, to represent all tensor values. The authors state that the ellipsoids are easier to interpret, while the brush stroke visualization is more quantitative. Integrated approaches of stream- and glyph-based techniques have been proposed by Hlawitschka and Scheuermann [30] as well as Chen et al. [35].

Kindlmann proposed superquadric glyph shapes to convey the principal eigenvectors of a diffusion tensor in order to depict the microstructure of white-matter tissue of the human brain [10]. The distinct glyphs are placed at a regular grid and controlled by a fractional anisotropy threshold in order to minimize visual clutter. Jankun-Kelly et al. [36] have evaluated the use of four different glyph visualizations for depicting traceless tensor data. They could show, that among the tested techniques, superquadric glyphs led to lower total error and lower response times. Their approach, that has been originally developed for nematic liquid crystal alignment tensors, is not based on the offsets of the eigenvalues, but on physically-linked metrics.

Domin et al. criticize that most glyphs proposed for DTI visualization are not sufficient [28]. The major drawback is that the used glyphs cannot convey the possibly arbitrary diffusion directions, and the data is usually reduced to 6 DoF in the modeling stage. Several approaches avoid this limitation by scaling vectors on the sphere based on the diffusion coefficients [37, 38, 28]. Fig. 9(a) shows the approach by Domin et al. [28]. However, it should be investigated how far the deformed sphere geometry can be perceived without introducing a cognitive overload.

Sigfridsson et al. have presented a hybrid approach for visualizing tensor fields [26]. Their approach integrated an overview of the field, which is generated through adaptive filtering (see Fig. 9(b)). While this provides the context, glyphs can be used in the attentive phase in order to get more detailed

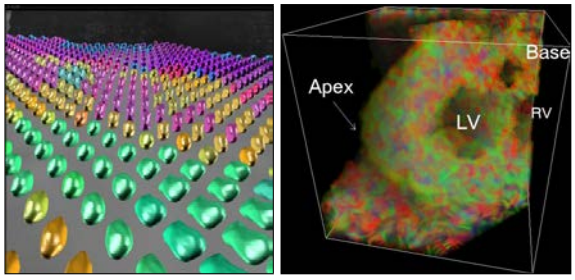
information. Therefore, the glyphs can be positioned freely.

To avoid the perception of false glyph aggregation in DTI data, Kindlmann and Westin have proposed a glyph packing algorithm [13] that exploits a modified particle-system (recall Fig. 3(b)). Hlawitschka et al. have also proposed a packing algorithm for the same application case [24] that achieves fast clustering. However, they describe the packing on a single slice only, which can be moved through the volume.

Since DTI data are primarily analyzed with respect to the direction of principal diffusion, it is reasonable to use transparency to convey the amount of anisotropy [32]. With this strategy, directional information is only pronounced if it is reliable. The use of transparency requires some background information, either a constant background color or anatomic information, such as a T2-weighted image (see Fig. 10(a)). Benger et al. also employ the human shape perception capabilities in particular with respect to pattern discrimination [32]. As a general strategy, they suggest to map tensorial quantities to texture patterns. More specifically, they employ the Gabor filter (see Fig. 10(b) and [39]), which conveys directional information well. A Gabor patch is mapped to the plane formed by the principal and median eigenvector (the y-direction of the Gabor patch is aligned with the principal eigenvector). Mapping tensor information to Gabor textures is optionally combined with the previously defined mapping to color and transparency. Thus, anisotropy characteristics are mapped to color, transparency, and texture (see Fig. 10(b)), while directional information is only visualized if it is assessed as reliable based on the relation between the eigenvalues.

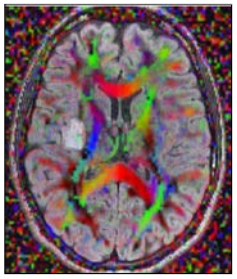
3.2. Higher-Order Tensor Glyphs

While most DTI visualization approaches focus on second-order tensor fields derived from MRI data, some techniques have been also developed for higher-order tensor data [40, 41, 42, 43]. Hlawitschka and Scheuermann presented a technique for higher order tensor fields [30]. They exploit the analogy of higher order tensors and spherical harmonics, and propose a glyph visualization inspired by this analogy. Thus, the authors generate glyphs of symmetric fourth order tensors by deforming a subdivided icosahedron based on the pre-computed spherical harmonic representation. Schultz and Kindlmann have presented another glyph approach for higher-order tensors [31].



(a) Deformed spheres (b) Probing glyphs

Figure 9: Deformed spheres are used to visualize diffusion directions [28] (a). Glyphs can be positioned within a continuous tensor field representation [26] (b). (Images courtesy of L. Linsen and A. Sigfridsson.)



(a) The components of the principal eigenvector are mapped to the red-, green- and blue-component of color. Transparency indicates linear anisotropy.

(b) Semitransparent colored Gabor patches are used to depict anisotropy from tensor data. If the tensor is isotropic (top corner of the triangle), transparency is 100%. Red indicates that planar anisotropy dominates, whereas green indicates predominantly linear anisotropy.

Figure 10: Usage of color and transparency for DTI visualization. (Images courtesy of W. Benger.)

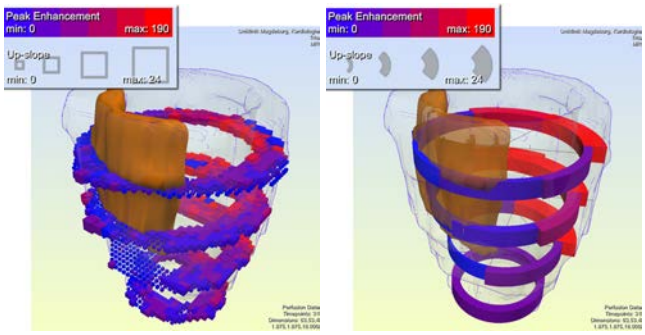
The proposed glyphs are colored based on the parameters to be visualized and allow to emphasize the depicted maxima values by introducing sharp edges, at the cost of smoother shapes around the minima. Therefore, the proposed glyphs are referred to as higher-order maximum enhanced (HOME) glyphs. A composite DTI glyph has been also proposed by Westin et al. [44]. For each glyph, they combine a sphere, a disc and a rod. The radius of the disc is determined based on the largest eigenvalue, while the radius of the disk is determined by the second largest eigenvalue. Finally, the length of the rod is set to twice the largest eigenvalue.

3.3. Cardiac Visualization

In this section, we address glyph-based visualizations for assessing functional and structural parameters of the heart. Choi et al. present such glyphs specifically designed for cardiac visualization [45]. They propose a technique for accurately measuring ventricular volume, mass, wall thickness, and wall motion, based on a 3D shape reconstruction through fitting a deformable model. The data extractable from the model can be visualized by using glyph techniques also in a quantitative manner. With their technique, different variables can be mapped to properties of the glyph allowing a comprehensible visualization of these multivariate data sets.

Wünsche and Lobb introduced a glyph-based technique to visualize wall motion of the heart [33]. Based on a generated finite element model, they position ellipsoid glyphs to show the movement of the myocardium. Each glyph is divided into six regions, whereas the region color encodes whether a dilation or a contraction is present.

Ropinski et al. have proposed easily modifiable superquadric glyphs which they also apply to cardiac visualization [12]. With their surface-based placement strategy, they are able to position the glyphs directly on the myocardium. Thus, the high resolution CT data provides the morphological context, while the low resolution PET data is depicted by using the glyphs (see Fig. 5). By choosing appropriate PMFs, it is possible to guide the user's



(a) Grid placement (b) 17-segment model placement

Figure 11: Glyph visualizations developed to support the analysis of cardiac MR data. Perfusion parameters are encoded voxel-wise by cuboid glyphs (a) or segment-wise by a more sophisticated glyph shape (b) [34].

attention to regions of interest, e. g., areas of the myocardium with low PET activity.

Oeltze et al. exploited glyphs to visualize cardiac MRI data in order to also allow an exploration of the structure and the function of the myocardium [34]. Cuboid-shaped glyphs map two parameters describing the perfusion of the myocardium [46]: peak enhancement to color and Up-Slope to size. The glyphs are placed voxel-wise for each slice of the perfusion scan and are integrated in a visualization of the left ventricle and scar tissue derived from MRI late enhancement data (see Fig. 11(a)). Glyphs may also be aggregated according to the 17-segment model of the American Heart Association (AHA). A more sophisticated glyph shape represents a segment of the myocardium (see Fig. 11(b)). Another approach has been presented by Ennis et al. [47]. They have described the use of superquadric glyphs, which allow a better differentiation of the fiber direction as compared to ellipsoidal glyphs.

4. Usage Guidelines

In this section we propose guidelines for the usage of glyph techniques. It should be mentioned that apart from the work by Spradow et al. [4] and Jankun-Kelly et al. [36] to our knowledge no other formal evaluation of glyph techniques has been conducted so far. Therefore, these guidelines should rather be considered as conjectures derived from the best practice described in the literature. We do not state that these guidelines are complete or describe a theory, but they distill the observations, we have made during our literature review. We believe they can be a helpful condensation for the interested reader. Therefore, we have derived the following six usage guidelines for the integration of glyph-based techniques in medical visualization, which in our opinion reflect the current state of the art:

1. Parameter mapping functions should
   - visually emphasize important variables.
   - incorporate the range of values.
   - guide the user's focus of attention to encode relevance.
   - incorporate semantics of the data.
   - be *mentally reconstructable* based on the visualization.
2. Glyph placement should be well-balanced and avoid unwanted glyph aggregations in image space, e. g., by applying jittering or relaxation procedures.
3. Glyph shapes should be unambiguously perceivable independent of the viewing direction.
4. Glyph visualizations should support quantitative analysis in the attentive phase.
5. Hybrid visualization should be exploited to provide the anatomical context.
6. When using the glyph size to convey information, perspective projections should be avoided.

Most of these guidelines are focused on choosing an appropriate PMF, which is crucial to allow comprehensive glyph visualizations. According to the **first guideline**, the most important variables should be more prominent in the final visualization. When, for instance, using a torus-like glyph, the color is better perceivable than the roundness of the glyph. Since the range of values present in a data set has a major impact on the right choice of a mapping function, this range should be considered carefully when specifying the function. Furthermore, an appropriate parameter mapping can also be used to guide the user's focus of attention. In Fig. 5, an inverse parameter mapping is used, i. e., low PET activity is mapped to thick glyphs, while high PET activity is mapped to thin glyphs. By using this inverse mapping, the region of interest, namely the region with reduced PET activity, is visually emphasized. Additionally, the parameter mapping should be intuitive, i. e., in cases where a glyph property fits semantically to a parameter to be mapped, it should be assigned to this parameter. For instance, in cases where a parameter represents the dimensions of a feature, it should be mapped to the glyph's size. Intuitive mappings are presented in [48].

Also colors should be chosen wisely when specifying a color PMF. In some applications, widely accepted color mappings are present. These should be considered when exploiting glyph techniques. For instance, PET data sets are often visualized exploiting a heat map, i. e., a yellow-to-red gradient. In cases, where no widely accepted color mapping is available, the mapping should be chosen by considering color perception, and eventually the semantics of the variables to be visualized. When taking into account the opponent color model [49], two-colored gradients can be generated, whereas the two colors are perceived as lying along the opposite directions of a coordinate axis. According to the opponent color theory, the three color axis are specified by red and green, blue, and yellow as well as black and white. However, it should also be considered that the displayed colors influence spatial comprehension. The chromadepth technique [50] supports depth perception based on the fact that the lens of the eye refracts colored light with different wavelengths at different angles. Although this effect can be supported by diffraction grating glasses, watching images without instrumentation can also result in a depth effect.

Furthermore, for all parameter mappings it is important that the mapping supports a *mental interpolation*, i. e., the user is able to mentally reconstruct the PMF when viewing distinct glyphs. This can be supported when equal perceptual distances match distances in the range of values which is true when specifying colors in the CIE color model.

As expressed by the **second guideline**, glyph placement is also important for the comprehension of glyph-based visualizations. Since a regular placement on the grid may convey non-existent aggregations unintendedly [12, 13], this has to be avoided. Therefore, when using a data set-driven placement in combination with regular grids often occurring in medical imaging data, at least a jittering has to be applied [17]. Moreover, when a texture-like appearance is desired, glyph packing strategies are sufficient in order to avoid misleading aggregations [13]. In general, the placement should be chosen in a way that the observer can perceive existent aggregations easily. Thus, also the glyph size and the question whether glyphs are overlapping or non-overlapping have to be taken into account. Since the glyph size and thus the spacing between adjacent glyphs is dependent on the resolution of the data set, general guidelines cannot be proposed.

Independently of parameter mapping and placement, the used glyph shapes should also satisfy certain criteria to allow a comprehensible visualization. First of all, glyph shapes should be unambiguously perceivable independent of the viewing direction as stated by the **third guideline**. Another important criterion possibly resulting in improved perception is the usage of intuitive glyph shapes. Similar to choosing an intuitive parameter mapping, a glyph shape can be chosen which represents the semantics of the variables to be shown. For instance, when considering tissue motion or blood flow direction, directional glyphs such as arrows are sufficient.

While the previous guidelines are focussed on the pre-attentive phase, in some cases a quantitative analysis in the attentive phase is desired. For these application cases, interaction metaphors, e. g., probing tools [26], and glyph legends [12] should be integrated, as expressed by the **fourth guideline**.

Furthermore, especially when choosing a rather large glyph spacing and thus a lot of context would become visible, the **fifth guideline** should be considered, i. e., the visualization should be enhanced by integrating a spatial context, for instance, by visualizing morphological structures through rendering selected structures from a CT data set. Since such a hybrid visualization is often described in literature, dealing with glyph-based medical visualization [12, 26], it can be assumed as helpful in many cases.

Finally, in order to support the quantification in the attentive phase, according to the **sixth guideline**, perspective projections should be avoided, when mapping parameters to glyph size. Instead an orthographic projection ensures that glyphs are quantifiable and the user can visually compare glyphs to glyphs at different locations or in the legend.

As mentioned above, all these guidelines have been derived from the best practice. Thus, they just reflect the current state of the art, and have to be subject to a systematic evaluation in the near future. However, although these guidelines have been derived from applications within the area of medical visualization, they may also be transferable to a certain extent to other domains.

5. Conclusions and Future Work

In this article we have described a taxonomy for glyph properties to be used in medical visualizations. In our taxonomy we classify these properties by considering aspects from the area of perception. Thus, we were able to identify glyph properties supporting pre-attentive and attentive visual processing. This distinction is important, since medical diagnoses are often performed under time pressure, and thus it is essential that the most relevant information is conveyed in a very direct way. Based on our literature review, we have proposed six guidelines with the goal to support improved glyph-based visualizations in the future. Since these guidelines are only derived from observations, a systematic evaluation is necessary in order to prove their usefulness. However, our taxonomy and to some extent also the guidelines have been motivated by the feature integration theory of attention. The pre-attentive processing is known to be partially dependent on the experiment setup that would be in our case equivalent to the glyph design. For instance, it is known that the expectation of the viewer as well as the similarity of the distracters have an influence on the pre-attentive phase [51, 52]. Furthermore, in some cases the existence of a certain feature is better perceivable within a group of objects not having this feature, than its absence within a group having this feature [53]. The latter is not an issue when only two variables are depicted by the shown glyphs. Nevertheless, when multiple variables are depicted, the distracters might not look similar enough to fully support pre-attentive processing. To be able to conduct the previously mentioned evaluation, all these aspects need to be taken into account.

Acknowledgments

We thank Christian Döring, Arvid Malyczyk, Jennis Meyer-Spradow, Lydia Paasche and Michael Specht for contributing some of the techniques described in this paper. Furthermore, we thank the reviewers for the very fruitful comments, which added great value to this paper. Finally, we thank all authors for permitting to use their images. This work was supported by the DFG Priority Program 1335: Scalable Visual Analytics.

References

[1] M. O. Ward, A taxonomy of glyph placement strategies for multidimensional data visualization, Information Visualization 1 (3) (2002) 194–210.

[2] R. Bürger, H. Hauser, Visualization of multi-variate scientific data, in: EuroGraphics 2007 State of the Art Reports (STARs), 2007, pp. 117–134.

[3] T. Ropinski, B. Preim, Taxonomy and usage guidelines for glyph-based medical visualization, in: Proc of Simulation and Visualization, 2008, pp. 121–138.

[4] J. Meyer-Spradow, L. Stegger, C. Döring, T. Ropinski, K. H. Hinrichs, Glyph based spect visualization for the diagnosis of coronary artery disease, IEEE Transactions on Visualization and Computer Graphics (TVCG) (Vis Conference Issue) (2008) 1499–1506.

[5] J. F. J. Rodrigues, A. J. M. Traina, M. C. F. de Oliveira, C. J. Traina, Reviewing data visualization: an analytical taxonomical study, in: Information Visualization, 2006, pp. 713–720.
[6] C. Ware, Information visualization: perception for design, Morgan Kaufmann Publishers Inc., 2000.
[7] A. Treisman, Preattentive processing in vision, Comput. Vision Graph. Image Process. 31 (2) (1985) 156–177.
[8] J. R. Anderson, Cognitive Psychology and its Implications, 6th Edition, Worth Publishers, 2004.
[9] B. Shneiderman, The eyes have it: A task by data type taxonomy for information visualizations, in: IEEE Symposium on Visual Languages, 1996, p. 336.
[10] G. L. Kindlmann, Superquadric tensor glyphs., in: Joint Eurographics - IEEE TCVG Symposium on Visualization, 2004, pp. 147–154.
[11] W. S. Cleveland, R. McGill, Graphical Perception: Theory, Experimentation, and Application to the Development of Graphical Methods, Journal of the American Statistical Association 79 (387) (1984) 531–554.
[12] T. Ropinski, M. Specht, J. Meyer-Spradow, K. H. Hinrichs, B. Preim, Surface glyphs for visualizing multimodal volume data, in: International Fall Workshop on Vision, Modeling, and Visualization (VMV07), 2007, pp. 3–12.
[13] G. Kindlmann, C.-F. Westin, Diffusion tensor visualization with glyph packing, IEEE Transactions on Visualization and Computer Graphics 12 (5) (2006) 1129–1135.
[14] A. H. Barr, Superquadrics and angle-preserving transformations, IEEE Computer Graphics and Applications 1 (1) (1981) 11–23.
[15] N. Bruno, J. Cutting, Minimodularity and the perception of layout, Journal of Experimental Psychology 117 (2) (1988) 161–170.
[16] V. Interrante, H. Fuchs, S. M. Pizer, Conveying the 3d shape of smoothly curving transparent surfaces via texture, IEEE Transactions on Visualization and Computer Graphics 3 (2) (1997) 98–117.
[17] D. H. Laidlaw, E. T. Ahrens, D. Kremers, M. J. Avalos, R. E. Jacobs, C. Readhead, Visualizing diffusion tensor images of the mouse spinal cord, in: IEEE Visualization, 1998, pp. 127–134.
[18] A. Bokinsky, Multivariate data visualization with data-driven spots, Ph.D. thesis, University of North Carolina - Chapel Hill (2003).
[19] M. Kraus, T. Ertl, Interactive data exploration with customized glyphs, in: WSCG, 2001, pp. 20–23.
[20] M. Mlejnek, P. Ermes, A. Vilanova, R. van der Rijt, H. van den Bosch, F. Gerritsen, E. Gröller, Profile flags: a novel metaphor for probing of T2 maps, in: IEEE Visualization, 2005, pp. 599–606.
[21] D. S. Ebert, R. M. Rohrer, C. D. Shaw, P. Panda, J. M. Kukla, D. A. Roberts, Procedural shape generation for multi-dimensional data visualization, in: Data Visualization, 1999, pp. 3–12.
[22] M. D. Lee, R. E. Reilly, M. E. Butavicius, An empirical evaluation of chernoff faces, star glyphs, and spatial visualizations for binary data, in: Asia-Pacific Symposium on Information visualisation, 2003, pp. 1–10.
[23] M. Mlejnek, P. Ermes, A. Vilanova, R. van der Rijt, H. van den Bosch, E. Gröller, F. Gerritsen, Application-oriented extensions of profile flags, in: EG/IEEE VGTC Symposium on Visualization, 2006, pp. 339–346.
[24] M. Hlawitschka, G. Scheuermann, B. Hamann, Interactive glyph placement for tensor fields, in: Advances in Visual Computing, Third International Symposium, ISVC, 2007, pp. 331–340.
[25] L. Hong, X. Mao, A. Kaufman, Interactive visualization of mixed scalar and vector fields, in: IEEE Visualization, 1995, pp. 240–247.
[26] A. Sigfridsson, T. Ebbers, E. Heiberg, L. Wigström, Tensor field visualization using adaptive filtering of noise fields combined with glyph rendering, in: IEEE Visualization, 2002, pp. 371–378.
[27] M. van Almsick, T. H. Peeters, V. Prckovska, A. Villanova, B. ter Haar Romeny, Gpu-based ray-casting of spherical functions applied to high angular resolution diffusion imaging, IEEE Transactions on Visualization and Computer Graphics 15 (1).
[28] M. Domin, S. Langner, N. Hosten, L. Linsen, Direct glyph-based visualization of diffusion MR data using deformed spheres, in: Visualization in Medicine and Life Sciences, 2007.
[29] Y. Hashash, J. Yao, D. Wotring, Glyph and hyperstreamline representation of stress and strain tensors and material constitutive response, International Journal of Numerical and Analytical Methods in Geomechanics 27 (2003) 603–626.
[30] M. Hlawitschka, G. Scheuermann, Hot- lines: Tracking lines in higher order tensor fields, in: IEEE Visualization, 2005, pp. 27–34.
[31] T. Schultz, G. Kindlmann, A maximum enhancing higher-order tensor glyph, Computer Graphics Forum 29 (3) (2010) 1143–1152.
[32] W. Benger, H. Bartsch, H.-C. Hege, H. Kitzler, A. Shumilina, A. Werner, Visualizing Neuronal Structures in the Human Brain via Diffusion Tensor MRI, International Journal of Neuroscience (2006) 461–514.
[33] B. Wünsche, R. Lobb, A. Young, The visualization of myocardial strain for the improved analysis of cardiac mechanics, in: GRAPHITE, 2004, pp. 90–99.
[34] S. Oeltze, A. Hennemuth, S. Glaer, C. Khnel, B. Preim, Glyph-based visualization of myocardial perfusion data and enhancement with contractility and viability information, in: EG Workshop on Visual Computing for Biology and Medicine, 2008, pp. 11–20.
[35] W. Chen, S. Zhang, S. Correia, D. F. Tate, Visualizing diffusion tensor imaging data with merging ellipsoids, in: IEEE Pacific Visualization Symposium, 2009, pp. 145–151.
[36] T. J. Jankun-Kelly, Y. S. Lanka, J. E. Swan II, An evaluation of glyph perception for real symmetric traceless tensor properties, Computer Graphics Forum 29 (3) (2010) 1133–1142.
[37] A. L. Alexander, K. M. Hasan, M. Lazar, J. S. Tsuruda, D. L. Parker, Analysis of partial volume effects in diffusion-tensor MRI, Magnetic Resonance in Medicine 45 (2001) 770–780.
[38] L. R. Frank, Anisotropy in high angular resolution diffusion-weighted MRI, Magnetic Resonance in Medicine 45 (2001) 935–939.
[39] M. Porat, Y. Y. Zeevi, The generalized gabor scheme of image representation in biological and machine vision, IEEE Trans. Pattern Anal. Mach. Intell. 10 (4) (1988) 452–468.
[40] L. R. Frank, Characterization of anisotropy in high angular resolution diffusion-weighted MRI, Magnetic Resonance in Medicine 47 (2002) 1083–1099.
[41] D. C. Alexander, G. J. Barker, S. R. Arridge, Detection and modeling of non-gaussian apparent diffusion coefficient profiles in human brain data, Magnetic Resonance in Medicine 48 (2002) 331–340.
[42] E. Özarslan, T. Mareci, Generalized diffusion tensor imaging and analytical relationships between diffusion tensor imaging and high angular resolution diffusion imaging, Magnetic Resonance in Medicine 50 (2003) 955–965.
[43] J.-D. Tournier, F. Calamante, D. G. Gadian, A. Connelly, Direct estimation of the fiber orientation density function from diffusion-weighted MRI data using spherical deconvolution, NeuroImage 23 (2004) 1176–1185.
[44] C.-F. Westin, S. Maier, H. Mamata, A. Nabavi, F. Jolesz, R. Kikinis, Processing and visualization for diffusion tensor MRI, Medical Image Analysis 6 (2002) 93–108.
[45] S. Choi, D. Lee, S.-J. Yoo, M.-H. Kim, Interactive visualization of diagnostic data from cardiac images using 3D glyphs., in: ISMDA, 2003, pp. 83–90.
[46] B. Preim, S. Oeltze, M. Mlejnek, E. Gröller, A. Hennemuth, S. Behrens, Survey of the visual exploration and analysis of perfusion data, IEEE Transactions on Visualization and Computer Graphics 15 (2) (2009) 205–220.
[47] D. B. Ennis, G. Kindlmann, I. Rodriguez, P. A. Helm, E. R. McVeigh, Visualization of tensor fields using superquadric glyphs, Magnetic Resonance in Medicine 53 (1) (2005) 169–176.
[48] S. Oeltze, A. Malyszczyk, B. Preim, Intuitive mapping of perfusion parameters to glyph shape, in: Bildverarbeitung für die Medizin (BVM 2008), 2008, pp. 262–266.
[49] M. Bollmann, B. Mertsching, Opponent color processing based on neural models, in: International Workshop on Advances in Structural and Syntactical Pattern Recognition, 1996, pp. 198–207.
[50] R. Steenblik, The chromostereoscopic process: A novel single image stereoscopic process, in: Proc of SPIE - True 3D Imaging Techniques and Display Technologies, 1987.
[51] J. Duncan, Boundary conditions on parallel search in human vision, Perception 18 (1989) 457–469.
[52] J. Duncan, G. Humphreys, Visual search and stimulus similarity, Psychological Review 96 (3) (1989) 433–458.
[53] A. Treisman, Properties, parts and objects, Wiley, 1986.

Survey of Labeling Techniques in Medical Visualizations

S. Oeltze-Jafra[1] and B. Preim[1]

[1]Department of Simulation and Graphics, University of Magdeburg, Germany

Abstract

Annotations of relevant structures and regions are crucial in diagnostics, treatment planning, medical team meetings as well as in medical education. They serve to focus discussions, present results of collaborative decision making, record and forward diagnostic findings, support orientation in complex or unfamiliar views on the data, and study anatomy. Different techniques have been presented for labeling the original data in 2D slice views, surface representations of structures extracted from the data, e.g., organs and vasculature, and 3D volume rendered representations of the data. All aim at a clear visual association of labels and structures, visible and legible labels, and a fast and aesthetic labeling while considering individual properties of the data and its representation and tackling various issues, e.g., occlusion of structures by labels, overlapping labels, and crossings of lines connecting labels with structures.
We survey the medical labeling work and propose a classification with respect to the employed labeling technique. We give guidelines for choosing a technique dependent on the data representation, e.g., surface rendering or slice view, the type of structures to be labeled, and the individual requirements on an effective label layout.

Categories and Subject Descriptors (according to ACM CCS): I.3.m [Computer Graphics]: Miscellaneous—J.3 [Computer Applications]: Life and Medical Sciences—

1. Introduction

The term *labeling* is used ambiguously in medical computing. On the one hand, it refers to automatically identifying anatomical structures in medical data such as specific segments of a vascular tree [BPC*13] or individual vertebrae of the spine [ACC11]. On the other hand, it is used for the process of annotating structures in medical visualizations by textual labels. This survey is dedicated to the latter.

Labeling medical visualizations has a centuries-long tradition in medical textbooks and anatomy atlases. Here, its main function is to communicate anatomical structures for education. In the age of modern medical imaging devices and computerized medicine, the range of possible applications has increased. Labeling plays an important role in diagnostics, treatment planning, medical team meetings, and in the education of medical students and patients. It is an everyday task for radiologists in diagnosing image data. Annotations serve to:

- record and forward diagnostic findings, e.g., to the transferring doctor or a medical specialist,
- focus discussions in team meetings, e.g., a tumor board discussing the therapeutic strategy,
- present results of collaborative decision making,
- support orientation in complex or unfamiliar views on the data, e.g., 3D views of highly branched vascular trees or (virtual) endoscopic views in sinus surgery,
- study anatomy in computer learning systems, e.g., the *VOXEL-MAN-series* [SFP*00],
- explain an intervention in patient education, and
- practice an intervention in a surgery training system, such as the *LiverSurgeryTrainer* [MMH*13].

Various labeling techniques which mimic hand-crafted visualizations of medical illustrators and techniques tailored to the particularities of computer support, e.g., a third dimension and interactivity, have been developed. Numerous approaches were proposed to labeling the original image data in 2D slice views, 3D surface representations of structures extracted from the data, e.g., organs and vasculature, and 3D volume rendered representations of the data.

In Section 2, we provide the foundation of labeling medical visualizations by discussing general and medicine specific requirements on an effective labeling and by introducing the employed labeling techniques. The latter serve as a classification scheme for our overview of medical labeling work in Section 3. We conclude in Section 4 with guidelines for choosing a labeling technique.

2. Foundation of Labeling Medical Visualizations

Labeling may be accomplished automatically or interactively. Interactive labeling is performed, e.g., by radiologists when annotating their findings in a slice view. For instance, a label depicting measurement values and a comment shall be added to a pathologic structure that has been segmented before. Modern radiological workstations support this process by providing lines, arrowheads, and textual labels. While these features are useful, automatic labeling where the system (re)arranges the labels, relieves the user from taking care of, e.g., crossing lines, labels overlapping each other or other important findings, and proximity of a label to the related structure. This is gaining importance with an increasing number of labels.

In labeling scenarios where the structures and label texts are predefined, e.g., in a 3D view of an anatomy learning system showing the human body, interactive labeling is inappropriate. Instead, automatic labeling is accomplished and adapted to user-interactions, such as pan and zoom.

2.1. Requirements on an Effective Labeling

Ali, Hartmann, and colleagues pose some general requirements on an effective label layout for interactive 3D illustrations [AHS05, HGAS05]. While they do not specifically aim at medical visualizations, their working examples are mostly borrowed from that domain. The requirements are:

- **Readability** Labels must not overlap,
- **Unambiguity** Labels clearly refer to their objects,
- **Pleasing** Prevent visual clutter,
- **Real-Time** Compute layouts at interactive rates,
- **Frame-Coherency** Prevent visual discontinuities,
- **Compaction** Reduce the layout area.

Besides theses general requirements, medical visualizations pose specific ones. **(1)** The labels must neither occlude diagnostically relevant information, such as potential pathologies, nor patient information, e.g., from the DICOM header, which is superimposed on the visualization. **(2)** The slice-based investigation of image data is a specialty of the medical domain. Here, the interactive labeling of findings is crucial. Once all findings have been annotated, an automatic post-processing, rearranging the labels for an effective layout, is desirable. **(2a)** An important aspect of the layout is the slice-coherency of annotations. If a structure covers multiple slices, it should be labeled in each of them and annotations should not abruptly change their position while browsing the slices. **(2b)** Branching structures and elongated structures, e.g., vessels, appear disconnected in slice views. As long as they are located close together, they should be summarized by a single label. **(3)** It may be necessary to label also hidden structures to remind the viewer of their existence, e.g., in the course of evaluating all lymph nodes in 3D views of the neck anatomy [MP09]. **(4)** Transparent surfaces often serve as context information in medical visualizations, e.g., the transparent liver or brain surface with the respective inner vascular system rendered opaque. Visibility tests for checking whether a structure must be labeled need to take transparency into account. **(5)** In volume-rendered views, object visibility is also dependent on the transparency transfer function. If physicians adjust this function dependent on the anatomy of interest, e.g., bony structures or soft tissue, visibility computations and the labeling must be updated accordingly.

The fulfillment of all requirements is hard to achieve and may conflict, e.g., with the desire to label as many visible objects as possible [AHS05]. Sometimes, objective criteria are missing to evaluate whether a requirement has been met, e.g., for readability and unambiguity. In 3D visualizations, interactivity aggravates the compliance with each requirement since the labeling has to be updated once the object is rotated or zoomed in. Structures which were visible become hidden and vice versa, empty screen space used for a label may now be occupied by a structure, and graphics objects relating labels to structures may start to cross.

2.2. Labeling Techniques

In a review of medical labeling work, we identified five different labeling techniques. None of them has been specifically invented for or is restricted to medical visualizations. However, all have been extended and tailored to a specific type of medical data, e.g., containing tubular structures such as vessels, a particular use case, e.g., an anatomy learning system or a surgery trainer, or a certain type of medical visualization, e.g., surface-based, volume-rendered, or slice-based. Adhering to the identified techniques, we classify the labeling work into the categories (Fig. 1):

- internal labels
- external labels
- boundary labeling
- excentric labeling
- necklace maps

Internal labels are superimposed on the structure of interest and should fit its screen representation (Fig. 1a). A good legibility is achieved if enough screen space is available for a sufficient font size, a good contrast between label text and background structure is provided, and the label text is aligned horizontally. If horizontal text extends beyond the structure, it should rather be aligned along the centerline of the structure's screen representation [GAHS05]. For strongly bended centerlines, smoothing is advisable. Ropinski and colleagues argue that in 3D visualizations, internal labels should also convey an objects 3D shape and hence, be projected onto it [RPRH07].

External labels are positioned on empty screen space and connected to their structure by a line (Fig. 1b). This so-called *leader* connects an *anchor point* on the structure and

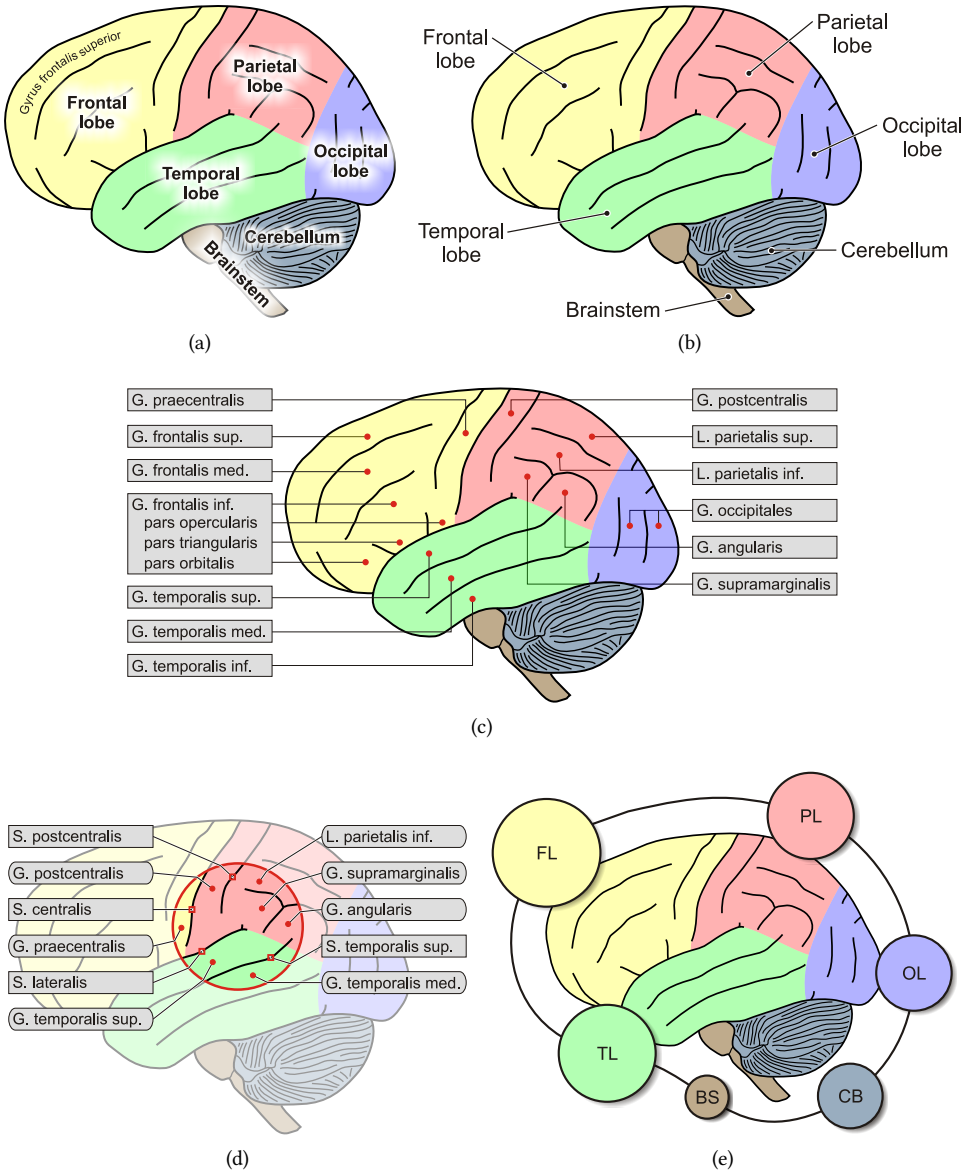


Figure 1: *Overview of labeling techniques in medical visualizations. (a) Internal labels are superimposed on the structures of interest. (b) External labels are positioned on empty screen space and connected to an anchor point on the structure of interest by a line. (c) Boundary labeling organizes the labels along a virtual rectangle enclosing all structures. (d) Excentric labeling annotates structures located inside a draggable, flexible focus region. Labels are stacked to the left and/or right of the region. (e) Necklace maps abandon connection lines and instead relate labels to structures by matching colors and spatial proximity.*

a point on the label box holding the label's textual representation. The definition of an anchor point is crucial. While the center of mass of an objects screen projection is suitable for convex objects, thinning algorithms shrinking the projection to a single pixel [HAS04] or algorithms computing the skeleton of a mesh in 3D [PR98] are generally applicable. Multiple anchor points may exist if an object is partially occluded and it must be decided which parts are to be labeled. External labels are often aligned along the silhouette of all objects in the scene [HAS04]. Close-up views form an exception since the entire screen space may be covered by objects. Mogalle and colleagues formulate requirements on external labels in 2D slice views, which can be generalized to 3D visualizations [MTSP12]. In summary, labels must not

overlap with other labels and structures, they should identify a structure unambiguously, and visual clutter must be avoided. To meet the latter two requirements,

- the number of leader crossing must be minimum,
- labels must be placed in close proximity to the structure, i.e. the total leader length must be minimum, and
- leader shapes should be simple, e.g., horizontal or vertical lines instead of zigzagging polylines.

Internal and external labels may be combined in a dynamic labeling. If a structure covers more and more screen space while being zoomed in, its external label can be replaced by an internal one at some point [GAHS05]. Further aspects of dynamic labeling, e.g., level-of-detail dependent labeling and interactive labeling speed, were discussed in the context of street maps [BDY06].

Boundary labeling generates a very tidy layout by organizing all labels along a virtual rectangle enclosing the entire scene (Fig. 1c). While the term was coined by Bekos and colleagues [BKSW05], Preim and colleagues already used this technique for the exploration of anatomical models [PRS97] and Ali and colleagues referred to it as "flush layout" [AHS05]. Each label is connected by a leader to an anchor point on its associated structure. Optimization approaches for minimizing the number of leader crossings, the total leader length, and the number of leader bends have been proposed in the context of static 2D maps [BKSW05,BHKN09] and 3D interactive visualizations [AHS05]. A circular boundary shape has been employed in [BSF*11]. Note that the tidiness of a boundary layout comes at the expense of a restricted freedom in label positioning which must be accounted for in optimization.

For structures being partially occluded or structures of the same type spread over multiple locations in the scene, it might be desirable to connect a single label to multiple anchor points (*many-to-one labeling problem*, see the label "G. occipitales" in Fig. 1c). Solutions to this problem tailored to boundary labeling have been presented in [BCF*13,Lin10]).

Excentric labeling is dedicated to annotating subsets of dense data and was presented by Fekete and Plaisant [FP99] (Fig. 1d). Labeling a subset of the scene is in contrast to the previous techniques, which often aim at labeling large parts. It is accomplished by means of a moveable, flexible focus region which can be dragged by the user. The labels of the focused structures are displayed in stacks to one or both sides of the focus region and connected to the structures by leaders. Fink and colleagues extended excentric labeling by techniques for creating a visually pleasing annotation, e.g., the use Bézier curves instead of zigzagging polylines and the optimization of total leader length [FHS*12].

Necklace maps abandon leaders and instead relate labels to structures by matching colors and spatial proximity in order to generate an uncluttered visualization (Fig. 1e). They were proposed by Speckmann and Verbeek for visualizing statistical data on geographical maps [SV10]. In the necklace map approach, labels are referred to as *symbols*. They are organized on a one-dimensional curve (the necklace) that surrounds the map or a subregion. Circles and bars have been implemented as symbol shapes. A data attribute is mapped to the area of the circular symbols or to the length of the bar-shaped symbols, respectively. Optimizing symbol sizes and positions is NP-hard. Speckmann and Verbeek contribute an algorithm that is exact up to a certain symbol density.

Boundary labeling, excentric labeling, and necklace maps may be seen as variants of external labeling since all position labels outside the structures of interest. We think however that they exhibit sufficient unique characteristics to be treated as unique labeling techniques.

3. Overview of Medical Labeling Work

Preim and Botha dedicate a section of their book to labeling medical visualizations [PB13]. We extend their set of reviewed techniques, update and extend their classification, and we provide guidelines for choosing an appropriate labeling technique. We collected labeling work from the IEEE and ACM electronic libraries and a Google search and categorize it according to the employed labeling techniques. We dedicate an extra category to labeling slice-based visualizations since they are most prevalent in clinical routine.

3.1. Internal Labels

Mori and colleagues describe a method for the automatic extraction of the bronchial tree from Computed Tomography (CT) images and for the automatic identification and naming of the bronchial branches [MHST00]. The surface of the extracted tree serves as the input for a virtual bronchoscopy system facilitating flights through the bronchus. Interpreting the images rendered from a viewpoint inside the tree is hampered by a lack of spatial orientation. To improve this situation, the name of the current branch is superimposed and outgoing child branches are annotated.

Petrovic and colleagues present a GPU-based approach for efficiently rendering very large sets of fiber tracts derived from whole-brain Diffusion Tensor Imaging (DTI) data [PFK07]. They propose a Level-of-Detail management system and a streamtube imposter construct for fast rendering and the reduction of overdraw. Curvature-correct text labels are employed for annotating the simulated tubes. The labeling is integrated in the fragment processing leading to the impression of text being attached to the imposter's surface. Special care is taken to orient the text right-side-up and to draw labels only if their corresponding geometry is close enough for the text to be legible. Fading labels in and out by alpha blending prevents label popping.

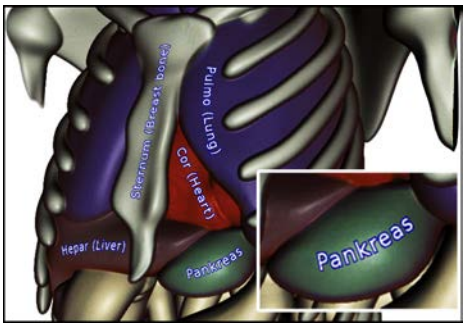


Figure 2: *Internal labels are projected onto the surface in order to convey its 3D shape. Image adapted from [RPRH07].*

Ropinski and colleagues argue that in surface-based 3D medical illustrations, internal labels should not only match the screen representation of an object but also convey its 3D shape, i.e. its varying depth structure [RPRH07]. Hence, they project a label onto the surface (Fig. 2). Special care must be taken to maintain legibility in case of noisy, strongly bended surfaces, and highly occluded regions, e.g., the sulci of the brain's surface. As a solution, the label is projected onto a smooth intermediate surface, a bezier patch in [RPRH07] and a *text scaffold* in [CG08], whose adherence to the original surface can be adjusted. Furthermore, the intermediate surface is oriented along the medial axis of its object as it is defined in image space and such that perspective distortion of labels is minimized [RPRH07].

Jiang and colleagues propose a method for annotating vascular structures in volume rendered views integrated in computer-assisted surgery systems [JNH*13]. While investigating highly-branching structures, surgeons strongly benefit from guidance by labeling. First, a surface model of the vasculature is constructed based on centerline and radius information. Then, in a two-pass rendering process, labels are projected from the current viewpoint onto the surface model which is after that rendered into a depth buffer image (first pass) followed by a ray-casting of the original data volume considering the depth buffer (second pass). Since vessels are often partially occluded by other vessels or organs, they are assigned multiple identical labels at intervals along their run. The legibility of labels may be hampered along surface parts generated from jagged centerlines. The problem is mitigated by centerline smoothing. The impact of transfer function adjustment on the visibility of vessels and the legibility of labels is not discussed.

Major an colleagues present the automated landmarking and labeling of spinal columns in CT images [MHSB13]. Disks are first superimposed on the automatically detected intervertebral spaces. The mean of the disks' positions is then employed for placing the corresponding vertebral text labels. The labeling approach is integrated in 2D slice views as well as in 3D volume rendered views.

3.2. External Labels

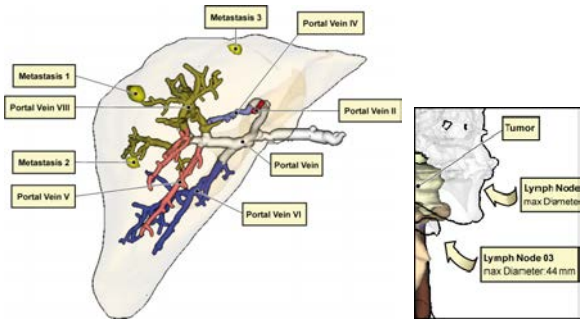
Hartmann, Ali, and Strothotte employ dynamic potential fields to generate effective and appealing label layouts for complex-shaped anatomical 3D models [HAS04]. Requirements on a layout, such as proximity of object and label and prevention of overlapping labels, are formalized in terms of attractive and repulsive forces steering the label placement. Anchor points are computed via thinning an object's screen projection to a single pixel. Limiting to the labeling approach is its inability to prevent leader crossings and visual discontinuities during interaction (frame-coherency).

Ali, Hartmann, and Strothotte extend their work by a variety of real-time label layout algorithms eliminating these limitations [AHS05]. Each algorithm is designed for a combination of a particular layout and leader style and demonstrated by an anatomical model. The proposed *flush* layout corresponds to boundary labeling (Fig. 1c) while the circular layout aligns the labels along the silhouette of the 3D model. Straight and orthogonal leaders are supported. The latter represent axis-aligned lines with their bends made at orthogonal angles. Anchor points are computed by applying a distance transform to an object's screen space projection. The pixel with the largest distance is chosen.

Sonnet and colleagues augment interactive explosion diagrams of complex 3D models by dynamic, scrollable annotations [SCS04]. They demonstrate their approach amongst others by anatomical models. The user may move the pointer over an object causing its textual description to be displayed. The closer the pointer gets to the object's centroid, the larger the label box becomes revealing more and more of the text. At entering the object, only a small box is displayed to account for the possibly unintentional or temporary hovering on the way to another object. The box is connected to the object's centroid (anchor point) by a transparent triangle emphasizing togetherness.

Bruckner and Gröller integrate external labels in the *VolumeShop* system to simplify orientation in an interactive environment, e.g., when exploring anatomical models [BG05]. They propose a simple algorithm aligning labels along the convex hull of the projected bounding volumes of all visible objects. This resembles the silhouette-based circular layout by Ali and colleagues [AHS05]. Special care is taken to resolve leader crossings and overlapping labels, which however causes visual discontinuities in animated views due to the extra computational effort. Details on the computation of anchor points are not given.

Mühler and Preim present techniques for annotating 3D structures reconstructed from medical image data in surgery planning [MP09]. They extend the work of Ali and colleagues [AHS05] by tackling the labeling of structures located inside or behind semi-transparent objects, e.g., the portal vein and metastases inside the liver parenchyma (Fig. 3a). Standard visibility tests by means of depth buffering return no objects to be labeled in such situations.



(a) (b)

Figure 3: *(a) Labeling of the portal vein and metastases located inside the semi-transparent liver parenchyma. Standard visibility tests by means of depth buffering would return no objects to be labeled here. Image from [MP09]. (b) Hidden lymph nodes labeled by bended arrows indicating presence and location. Image adapted from [MP09].*

Hence, a multi-buffering approach is proposed treating all objects as visible and computing a set of anchor point candidates for each of them. The candidates are derived by distance transforms applied on the buffers. Next, rays are cast from each anchor point to the viewer and the opacity of intersected objects is accumulated. If it is above a threshold for each anchor point, the object is not labeled. Otherwise, the point with the "smallest" occlusion is chosen.

A further contribution is the labeling of currently hidden objects to recall their existence. For instance, no lymph node must be overlooked in planning neck dissections (Fig. 3b). Bended, arrow-shaped leaders indicate the presence and location of currently invisible lymph nodes.

3.3. Internal and External Labels

Götzelmann and colleagues propose a hybrid label layout comprising internal and external labels [GAHS05]. The label type is chosen depending on the zoom level. For instance, if an object gets closer to the camera and occupies more screen space, an external label is replaced by an internal one to exploit the gained space. During interaction, the entire scene is continuously projected to screen space and the skeleton of each object's projection is determined. It is then tested, whether an internal label would be given sufficient space to be placed along the skeleton while guaranteeing minimal readability. If this is not the case, an external label is drawn whose anchor point is computed according to [AHS05]. The labeling approach has been integrated in a framework for anatomical education [VGHN08].

Ropinski and colleagues also propose a hybrid layout [RPRH07]. If the screen coverage of a projected object is sufficient to place a label along the pojection's medial axis while guaranteeing a minimum label size, an internal label is drawn. Otherwise, external labels are employed (Fig. 4).

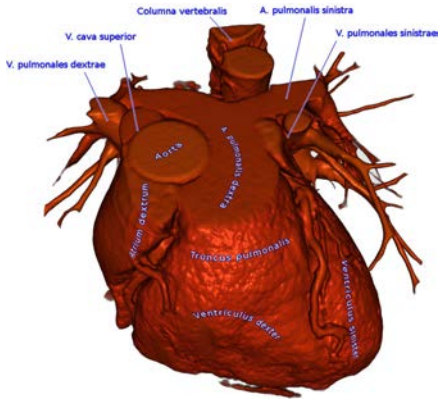


Figure 4: *Hybrid layout comprising internal and external labels. If the screen coverage of a projected object is sufficiently large, an internal label is drawn. Image from [RPRH07].*

3.4. Boundary Labeling

Preim and colleagues present a system for the exploration of anatomical models which combines zooming techniques, fisheye views, and interactive external labels [PRS97]. The labels are aligned on the left and right boundary of a rectangle enclosing the model (Fig. 5). They are connected via straight lines to anchor points on the model parts. The computation of the anchor points is not described. The focus is on the interaction with the model and the labels. For instance, selecting a label causes (1) an enlargement of the corresponding model part and a simultaneous shrinking of the other parts (fisheye technique) as well as (2) an enlargement of the label box gaining space for a more detailed description. In (2), neighboring labels are automatically relocated and minimized if necessary. Assigning a single label to multiple model parts (many-to-one labeling) is supported, e.g., to facial muscles in both sides of the face. However, optimization with respect to leader crossings and total leader length seems to be missing (Fig. 5).

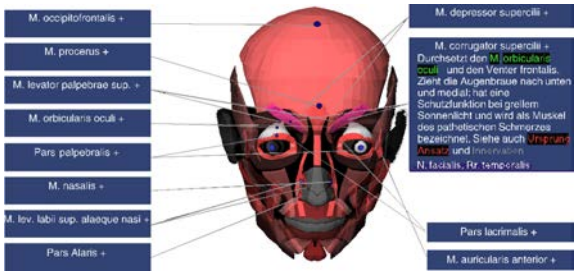


Figure 5: *Boundary labeling of facial muscles. A label of muscles above the eye has been selected causing a description to be displayed (right), the muscles to be enlarged, and two labels to be pushed downward. Image from [PRS97].*

Eichelbaum and colleagues visualize human brain connectivity derived from Diffusion-weighted magnetic resonance imaging (DW-MRI) data [EWH*10]. They employ fiber tracking and clustering to generate fiber bundles which illustrate the connection of brain regions. The bundles are displayed together with the regions inside a semi-transparent surface of the brain. The regions are labeled according to [BKSW05] for improving spatial orientation.

Battersby and colleagues [BSF*11] employ *ring maps* for visualizing multivariate epidemiological data [BSF*11]. A ring map shows a 2D geographical map enclosed by a virtual circular boundary shape along which glyphs are aligned. The glyphs are composed of n parts for encoding n variates. The parts are located at an uniformly increasing distance to the boundary. Each set of parts with equal distance to the boundary represents a ring. The glyphs and county names are connected to their respective map region via straight leaders. Label and anchor point positions are chosen such that glyphs are uniformly distributed, located close to their region, and leaders do not cross.

3.5. Excentric Labeling

Fekete and Plaisant introduce excentric labeling for the annotation of dense, point-based data representations, e.g., scatter plots [FP99]. A circular focus region is dragged across the representation and labels of the objects in focus are displayed in stacks to the left and/or right of the region. Multiple labeling variants are proposed. In the basic variant, straight lines connect points – coincident with the anchor points here – and labels. In the radial variant, leader crossings are prevented by first connecting the labeled point with a point on the boundary of the focus region and then, bending towards the sorted stack of labels. In further variants, the label order and justification reflect the y- and x-position of the labeled points, respectively. However, this is at the expense of crossing-free leaders. Plaisant and colleagues integrate excentric labeling in *LifeLines* – a system for visualizing personal histories [PMR*96]. They demonstrate how the investigation of patient records benefits from labeling health-related events (Fig. 6).

Luboschik and colleagues present a point-feature labeling approach, which is fast and avoids overlapping labels as well as the occlusion of other visual representatives such as leaders and icons [LSC08]. In contrast to the work of Fekete and Plaisant [FP99], each point is initially labeled. In the first step of an iterative, particle-based approach, all points with sufficient empty space in their direct neighborhood are annotated by an adjacent label (no leader). Then, the remaining points are annotated by positioning the label as close as possible and connecting it to the point by a straight line. The labeling approach is coupled with a movable label lens. Labels of focused points are relocated along the outside of the lens such that they do not overlap other labels. Straight leaders are drawn to convey correspondence. The

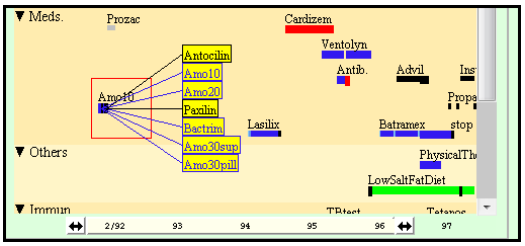


Figure 6: *Excentric labeling in LifeLines [PMR*96]. A rectangular focus region is dragged across events in a patient record. Drugs administered in a narrow time frame become readable.*

approach has been demonstrated amongst others for the labeling of symbol maps encoding health data.

3.6. Necklace Maps

Glaßer and colleagues apply necklace maps to labeling clusters of breast tumor tissue with cluster-specific perfusion information [GLP14]. The necklace surrounds an abstract representation of the tumor (Fig. 7). Each set of equally-colored, spindle-shaped extensions represents a cluster of voxels exhibiting similar perfusion characteristics. The extensions originate at the cluster's center and are directed towards the subregions of the tumor. For each cluster, a label is strung on the necklace and colored according to the cluster's color. Proximity of labels and clusters is not optimized. Each label shows an iconic representation of the wash-in and wash-out of a contrast agent.

Oeltze-Jafra and colleagues combine dynamic excentric labeling and static necklace maps for the interactive visual exploration of multi-channel fluorescence microscopy data [OPH*14]. Nested necklaces show information aggregated over all channels as well as individual channels.

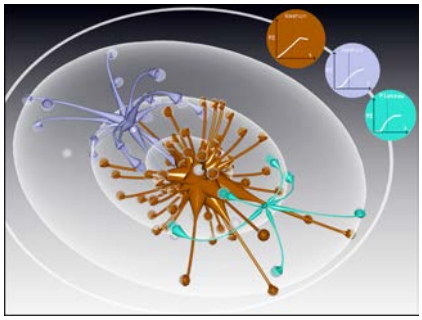


Figure 7: *Abstract visualization of three regions with distinct perfusion in a breast tumor. Labels showing plots of contrast agent accumulation are strung on a surrounding necklace. Correspondence is conveyed by color. Image based on [GLP14].*

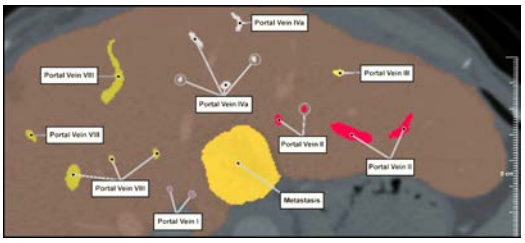


Figure 8: *In a slice view, branches of the portal vein and metastases inside the liver parenchyma (brown, large region) are annotated. Disconnected, but close parts of the same branch are summarized by one label. Image from [MP09].*

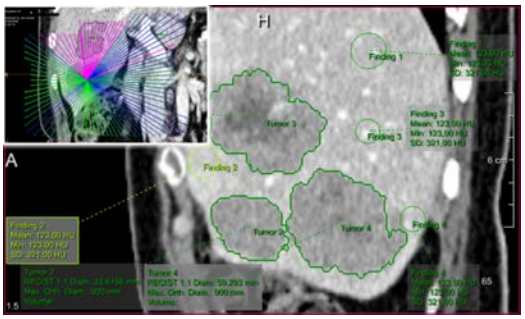


Figure 9: *Labeling radiological findings in 2D slice data. Label positions for each finding are searched in discrete directions starting at the finding's anchor point (inset). Green rays represent directions complying to a set of constraints, e.g., no occlusion of other findings. Image adapted from [MTSP12].*

3.7. Labeling Slice-Based Visualizations

The manual annotation of digital images is crucial in clinical routine. Efforts were made to advance the generation, management, and dissemination of annotations. Cai and colleagues present a web-based system for the collaborative generation and editing of labels supporting collaborative decision making [CFF01]. Goede and colleagues propose a methodology and implementation for annotating digital images [GLC*04]. They define a set of rules to standardize the annotation process.

Mühler and Preim discuss important aspects of automatically labeling slice views [MP09]. If empty space exists in the image, e.g., around the head in images of the brain, external labels should be placed there. Otherwise, they should be positioned on less important structures, e.g., on the liver parenchyma in an examination of inner metastases and vessels (Fig. 8). Elongated structures with a small diameter, e.g., vessels, often appear as disconnected components. They should be summarized by one label if they are located close together (Fig. 8). Slice coherency of labels must be guaranteed to support their visual tracking and to avoid flickering artifacts. Mühler and Preim lock the position of a label across multiple slices until it overlaps with a crucial image region. They also employ many-to-one labeling and achieve crossing-free leaders.

Mogalle and colleagues present an optimal placement of external labels representing radiological findings in 2D slice data [MTSP12]. They focus on avoiding leader crossings, mutually overlapping labels and labels occluding findings, and on minimizing total leader length. A local optimization algorithm achieves a trade-off between speed and labeling quality. It samples directions for label placement starting at the anchor point of a structure and assesses a direction's compliance with each of the requirements (inset of Fig. 9). This results in a set of weighted candidate directions for each object. The final layout is derived from these sets either by a greedy optimization or a label shifting approach (Fig. 9). The labeling is limited to ≈ 10 annotations, which is however realistic for radiological data.

4. Guidelines for Labeling Medical Visualizations

The search for a suitable labeling technique is first guided by the visual representation of the data, second, the type of structures to be labeled, and third, the individual requirements on an effective label layout (Sec. 2.1). This order is reflected by the decision diagram in Figure 10. The decision process leads to publications describing a suitable technique. For instance, if surfaces of arbitrary shape shall be labeled and readability of the labels, unambiguity of the association between label and structure, and the tidiness of the label layout are the may concerns, boundary labeling according to [PRS97, EWH*10] is suitable.

In general, *internal labels* facilitate an easy visual association with a structure and lead to a compact label layout. *External labels*, do not occlude their associated structure, are easier to read, and better suited for small structures and dense data. However, they demand extra effort to establish the visual association with a structure. Leaders, proximity, and color are employed requiring optimization steps, e.g., to reduce leader crossings, achieve sufficient proximity for all labels, and avoid overlapping labels. A dynamic application of both *internal and external labels* dependent on a structure's screen coverage is appropriate in interactive 3D views where zooming is frequently used. The main strength of *boundary labeling* is the tidiness of the label layout providing a fast overview of all labeled structures. *Excentric labeling* is particularly suited for the piece-wise exploration of very dense data. So far, it has only been demonstrated for representations of abstract data, such as patient records [FP99]. Another potential application area could be the exploration of large, annotated microscopic images, e.g., in histology. *Necklace maps* avoid the visual clutter caused by leaders and show little occlusion of the data at the expense of additionally required screen space and weaker visual association of labels and structures. They could prove beneficial in *volume renderings* where separate

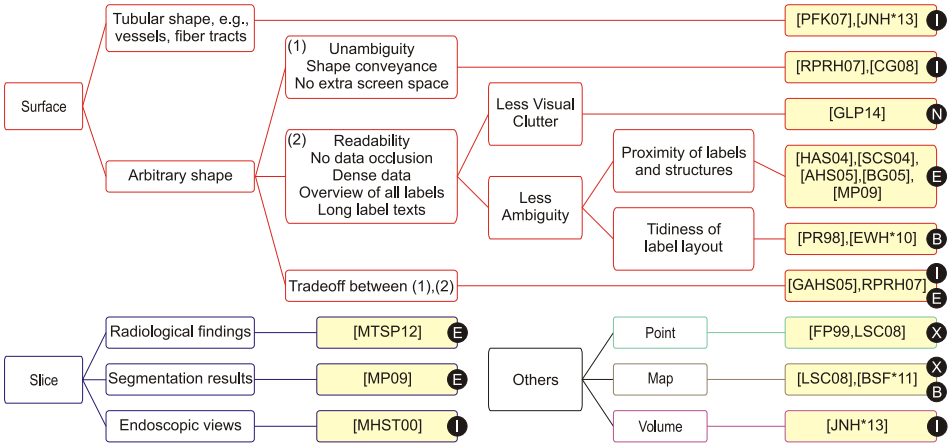


Figure 10: *Decision diagram for choosing a labeling technique dependent on the data representation (surface rendering, slice view, point set, map, or volume rendering), the type of structures to be labeled, and the individual requirements on an effective label layout. The yellowish boxes show the related work. The circled letters encode the labeling technique: I=internal labels (Sec. 3.1), E=external labels (Sec. 3.2), B=boundary labeling (Sec. 3.4), X=excentric labeling (Sec. 3.5), and N=necklace maps (Sec. 3.6).*

structures are discernible due to different colors and opacities but not processable, e.g., for anchor point computation.

In *medical education* systems, 3D surface models should be annotated by a combination of internal and external labels according to customs in hand-drawn medical illustrations. External labels and boundary labeling are better suited for displaying long label texts, e.g., descriptions or links to other structures, and for generating foldout groups of labels, e.g., all bones of the foot. In *intervention planning* and *medical training* systems, external labels should be used since they are more legible than internal ones, provide a clearer overview of all relevant structures, may indicate currently hidden objects in 3D scenes [MP09], and above all, they do not occlude structures. Occlusion is critical since, e.g., the irregular and complex shape of an object may influence the interventional strategy. In 3D views of *vasculature* and *fiber tracts*, internal labels should be employed. Thin, elongated structures are often only visible at intervals due to mutual occlusion and occlusion by other structures. Identical internal labels applied to these intervals avoid the visual clutter that would be induced by an external many-to-one labeling [JNH*13]. An exception are 2D and 3D views for *vascular diagnosis*. Here, external labels are more appropriate since internal ones may interfere with the perception of pathologic shape variances, e.g., stenoses, and the evaluation of vascular cross-sections and compositions of the vessel wall, e.g., in plaque detection. External labels are generally recommended for slice views in *radiology* since they do not occlude the associated findings. Boundary labeling treating the image border as boundary would be best here in terms of occlusion but would also require shrinking the image to make space for the labels. Optimization with respect to placing labels on less important image parts is more promising [MTSP12]. A similar situation exists in *virtual endocsopy* where the endoscopic view should occupy maximum screen space. However, internal labels are more appropriate here, e.g, for annotating branches, since their in-site position causes less distraction while navigating the endoscope. Pathologies, e.g., polyps should again be annotated by external labels.

5. Concluding Remarks

We provided an overview of the existing medical labeling work and proposed a classification with respect to the employed labeling techniques. Furthermore, we gave guidelines for choosing a suitable technique. The labeling of 3D surfaces is the most extensively researched subfield. Labeling medical volume renderings and slice views are underrepresented measured against their wide-spread use and may pose interesting directions for future work. Labeling slice views may benefit from transferring more knowledge in cartography, where labeling is widely studied.

References

[ACC11] Alomari R., Corso J., Chaudhary V.: Labeling of Lumbar Discs Using Both Pixel- and Object-Level Features With a Two-Level Probabilistic Model. *IEEE Trans. Med. Imag. 30*, 1 (2011), 1–10.

[AHS05] Ali K., Hartmann K., Strothotte T.: Label Layout for Interactive 3D Illustrations. *Journal of WSCG 13*, 1 (2005), 1–8.

[BCF*13] Bekos M., Cornelsen S., Fink M., Hong S.-H., Kaufmann M., Nöllenburg M., Rutter I., Symvonis A.: Many-to-One Boundary Labeling with Backbones. In *Graph Drawing*, vol. 8242 of *LNCS*. Springer, 2013, pp. 244–255.

[BDY06] Been K., Daiches E., Yap C.: Dynamic Map Labeling. *IEEE Trans. Vis. Comput. Graphics 12*, 5 (2006), 773–780.

[BG05] Bruckner S., Gröller E.: VolumeShop: An Interactive System for Direct Volume Illustration. In *IEEE Visualization* (2005), pp. 671–678.

[BHKN09] Benkert M., Haverkort H. J., Kroll M., Nöllenburg M.: Algorithms for Multi-Criteria Boundary Labeling. *J. Graph Algorithms Appl. 13*, 3 (2009), 289–317.

[BKSW05] Bekos M. A., Kaufmann M., Symvonis A., Wolff A.: Boundary Labeling: Models and Efficient Algorithms for Rectangular Maps. In *Graph Drawing*, vol. 3383 of *LNCS*. Springer, 2005, pp. 49–59.

[BPC*13] Bogunovic H., Pozo J., Cardenes R., San Roman L., Frangi A.: Anatomical Labeling of the Circle of Willis Using Maximum A Posteriori Probability Estimation. *IEEE Trans. Med. Imag. 32*, 9 (2013), 1587–1599.

[BSF*11] Battersby S. E., Stewart J. E., Fede A. L.-D., Remington K. C., Mayfield-Smith K.: Ring Maps for Spatial Visualization of Multivariate Epidemiological Data. *Journal of Maps 7*, 1 (2011), 564–572.

[CFF01] Cai W., Feng D., Fulton R.: Web-Based Digital Medical Images. *IEEE Comput. Graph. Appl. 21*, 1 (Jan. 2001), 44–47.

[CG08] Cipriano G., Gleicher M.: Text Scaffolds for Effective Surface Labeling. *IEEE Trans. Vis. Comput. Graphics 14*, 6 (2008), 1675–1682.

[EWH*10] Eichelbaum S., Wiebel A., Hlawitschka M., Anwander A., Knösche T., Scheuermann G.: Visualization of Effective Connectivity of the Brain. In *Vision, Modeling, and Visualization (VMV)* (2010), pp. 155–162.

[FHS*12] Fink M., Haunert J.-H., Schulz A., Spoerhase J., Wolff A.: Algorithms for Labeling Focus Regions. *IEEE Trans. Vis. Comput. Graphics 18*, 12 (2012), 2583–2592.

[FP99] Fekete J.-D., Plaisant C.: Excentric Labeling: Dynamic Neighborhood Labeling for Data Visualization. In *SIGCHI Conference on Human Factors in Computing Systems* (1999), pp. 512–519.

[GAHS05] Götzelmann T., Ali K., Hartmann K., Strothotte T.: Form Follows Function: Aesthetic Interactive Labels. In *Computational Aesthetics in Graphics, Visualization and Imaging* (2005), Computational Aesthetics'05, pp. 193–200.

[GLC*04] Goede P. A., Lauman J. R., Cochella C., Katzman G. L., Morton D. A., Albertine K. H.: A Methodology and Implementation for Annotating Digital Images for Context-appropriate Use in an Academic Health Care Environment. *J Am Med Inform Assoc 11*, 1 (2004), 29–41.

[GLP14] Glaßer S., Lawonn K., Preim B.: Visualization of 3D Cluster Results for Medical Tomographic Image Data. In *Computer Graphics Theory and Applications (GRAPP)* (2014), pp. 169–176.

[HAS04] Hartmann K., Ali K., Strothotte T.: Floating Labels: Applying Dynamic Potential Fields for Label Layout. In *Symposium on Smart Graphics* (2004), pp. 101–113.

[HGAS05] Hartmann K., Götzelmann T., Ali K., Strothotte T.: Metrics for Functional and Aesthetic Label Layouts. In *Smart Graphics* (2005), pp. 115–126.

[JNH*13] Jiang Z., Nimura Y., Hayashi Y., Kitasaka T., Misawa K., Fujiwara M., Kajita Y., Wakabayashi T., Mori K.: Anatomical Annotation on Vascular Structure in Volume Rendered Images. *Comput Med Imaging Graph 37*, 2 (2013), 131 – 141.

[Lin10] Lin C.-C.: Crossing-Free Many-to-One Boundary Labeling With Hyperleaders. In *IEEE Pacific Visualization Symposium (PacificVis)* (2010), pp. 185–192.

[LSC08] Luboschik M., Schumann H., Cords H.: Particle-Based Labeling: Fast Point-Feature Labeling Without Obscuring Other Visual Features. *IEEE Trans. Vis. Comput. Graphics 14*, 6 (2008), 1237–1244.

[MHSB13] Major D., Hladůvka J., Schulze F., Bühler K.: Automated landmarking and labeling of fully and partially scanned spinal columns in ct images. *Medical Image Analysis 17*, 8 (2013), 1151 – 1163.

[MHST00] Mori K., Hasegawa J., Suenaga Y., Toriwaki J.: Automated Anatomical Labeling of the Bronchial Branch and its Application to the Virtual Bronchoscopy System. *IEEE Trans. Med. Imag. 19*, 2 (2000), 103–114.

[MMH*13] Mönch J., Mühler K., Hansen C., Oldhafer K.-J., Stavrou G., Hillert C., Logge C., Preim B.: The LiverSurgeryTrainer: training of computer-based planning in liver resection surgery. *Int J Comput Assist Radiol Surg 8*, 5 (2013), 809–818.

[MP09] Mühler K., Preim B.: Automatic Textual Annotation for Surgical Planning. In *Vision, Modeling, and Visualization (VMV)* (2009), pp. 277–284.

[MTSP12] Mogalle K., Tietjen C., Soza G., Preim B.: Constrained Labeling of 2D Slice Data for Reading Images in Radiology. In *Eurographics Workshop on Visual Computing for Biology and Medicine (VCBM)* (2012), pp. 131–138.

[OPH*14] Oeltze S., Pieper F., Hillert R., Preim B., Schubert W.: Interactive Labeling of Toponome Data. In *Eurographics Workshop on Visual Computing for Biology and Medicine (VCBM)* (2014), p. this volume.

[PB13] Preim B., Botha C.: *Visual Computing for Medicine: Theory, Algorithms, and Applications.* Morgan Kaufmann, San Francisco, 2013. Section 10.7.

[PFK07] Petrovic V., Fallon J., Kuester F.: Visualizing Whole-Brain DTI Tractography with GPU-based Tuboids and LoD Management. *IEEE Trans. Vis. Comput. Graphics 13*, 6 (2007), 1488–1495.

[PMR*96] Plaisant C., Milash B., Rose A., Widoff S., Shneiderman B.: LifeLines: Visualizing Personal Histories. In *SIGCHI Conference on Human Factors in Computing Systems* (1996), pp. 221–227.

[PR98] Preim B., Raab A.: Annotation von Topographisch Komplizierten 3D-Modellen. In *SimVis* (1998), pp. 128–140.

[PRS97] Preim B., Raab A., Strothotte T.: Coherent Zooming of Illustrations with 3D-Graphics and Text. In *Graphics Interface* (1997), pp. 105–113.

[RPRH07] Ropinski T., Praßni J.-S., Roters J., Hinrichs K.: Internal Labels as Shape Cues for Medical Illustration. In *Vision, Modeling, and Visualization (VMV)* (2007), pp. 203–212.

[SCS04] Sonnet H., Carpendale S., Strothotte T.: Integrating Expanding Annotations with a 3D Explosion Probe. In *Conference on Advanced Visual Interfaces* (2004), AVI '04, pp. 63–70.

[SFP*00] Schiemann T., Freudenberg J., Pflesser B., Pommert A., Priesmeyer K., Riemer M., Schubert R., Tiede U., Höhne K.: Exploring the Visible Human using the VOXEL-MAN framework. *Comput Med Imaging Graph 24*, 3 (2000), 127 – 132.

[SV10] Speckmann B., Verbeek K.: Necklace Maps. *IEEE Trans. Vis. Comput. Graphics 16*, 6 (2010), 881–889.

[VGHN08] Vázquez P.-P., Götzelmann T., Hartmann K., Nürnberger A.: An Interactive 3D Framework for Anatomical Education. *Int J Comput Assist Radiol Surg 3*, 6 (2008), 511–524.

Bibliography

[1] A. K. Al-Awami, J. Beyer, H. Strobelt, N. Kasthuri, J. W. Lichtman, H. Pfister, and M. Hadwiger. NeuroLines: a subway map metaphor for visualizing nanoscale neuronal connectivity. *IEEE Transactions on Visualization and Computer Graphics*, 20(12):2369–2378, 2014.

[2] N. Allen, C. Sudlow, P. Downey, T. Peakman, J. Danesh, P. Elliott, J. Gallacher, J. Green, P. Matthews, J. Pell, et al. UK Biobank: Current status and what it means for epidemiology. *Health Policy and Technology*, 1(3):123–126, 2012.

[3] P. Alliez, G. Ucelli, C. Gotsman, and M. Attene. Recent advances in remeshing of surfaces. In *Shape Analysis and Structuring*, pages 53–82. Springer, 2008.

[4] P. Angelelli, S. Oeltze, C. Turkay, J. Haász, E. Hodneland, A. Lundervold, A. J. Lundervold, B. Preim, and H. Hauser. Interactive visual analysis of heterogeneous cohort study data. *IEEE Computer Graphics and Applications*, 34(5):70–82, 2014.

[5] L. Antiga, M. Piccinelli, L. Botti, B. Ene-Iordache, A. Remuzzi, and D. A. Steinman. An image-based modeling framework for patient-specific computational hemodynamics. *Medical & Biological Engineering & Computing*, 46(11):1097–1112, 2008.

[6] R. Arias-Hernández, J. Dill, B. Fisher, and T. M. Green. Visual analytics and human-computer interaction. *Interactions*, 18(1):51–55, 2011.

[7] R. Arias-Hernández, L. Kaastra, T. Green, and B. Fisher. Pair analytics: Capturing reasoning processes in collaborative visual analytics. In *Hawaii International Conference on System Sciences (HICSS)*, pages 1–10, 2011.

[8] D. Barh, V. Zambare, and V. Azevedo. *Omics: Applications in Biomedical, Agricultural, and Environmental Sciences*. CRC Press, 2013.

[9] A. Barysenka, A. W. Dress, and W. Schubert. An information theoretic thresholding method for detecting protein colocalizations in stacks of fluorescence images. *Journal of Biotechnology*, 149(3):127–131, 2010.

[10] J. Bernard, D. Sessler, T. May, T. Schlomm, D. Pehrke, and J. Kohlhammer. A visual-interactive system for prostate cancer cohort analysis. *IEEE Computer Graphics and Applications*, 35(3):44–55, 2015.

[11] J. Beyer, A. Al-Awami, N. Kasthuri, J. W. Lichtman, H. Pfister, and M. Hadwiger. Connectomeexplorer: Query-guided visual analysis of large volumetric neuroscience data. *IEEE Transactions on Visualization and Computer Graphics*, 19(12):2868–2877, 2013.

[12] S. Bhattacharya, G. Mathew, E. Ruban, D. B. Epstein, A. Krusche, R. Hillert, W. Schubert, and M. Khan. Toponome imaging system: in situ protein network mapping in normal and cancerous colon from the same patient reveals more than five-thousand cancer specific protein clusters and their subcellular annotation by using a three symbol code. *Journal of Proteome Research*, 9(12):6112–6125, 2010.

[13] A. Biswas, D. Thompson, W. He, Q. Deng, C.-M. Chen, H.-W. Shenk, R. Machiraju, and A. Rangarajan. An uncertainty-driven approach to vortex analysis using oracle consensus and spatial proximity. In *IEEE Pacific Visualization Symposium (PacificVis)*, pages 223–230, 2015.

[14] B. M. Boesiger and J. R. Shiber. Subarachnoid hemorrhage diagnosis by computed tomography and lumbar puncture: Are fifth generation CT scanners better at identifying subarachnoid hemorrhage? *The Journal of Emergency Medicine*, 29(1):23–27, 2005.

[15] G.-P. Bonneau, H.-C. Hege, C. R. Johnson, M. M. Oliveira, K. Potter, P. Rheingans, and T. Schultz. Overview and state-of-the-art of uncertainty visualization. In *Scientific Visualization*, pages 3–27. Springer, 2014.

[16] S. Born, M. Pfeifle, M. Markl, M. Gutberlet, and G. Scheuermann. Visual analysis of cardiac 4D MRI blood flow using line predicates. *IEEE Transactions on Visualization and Computer Graphics*, 19(6):900–912, 2013.

[17] A. Brun, H. Knutsson, H.-J. Park, M. E. Shenton, and C.-F. Westin. Clustering fiber traces using normalized cuts. In *Medical Image Computing and Computer-Assisted Intervention (MICCAI)*, pages 368–375. Springer, 2004.

[18] R. Bürger, P. Muigg, M. Ilcík, H. Doleisch, and H. Hauser. Integrating local feature detectors in the interactive visual analysis of flow simulation data. In *Eurographics Conference on Visualization (EuroVis)*, pages 171–178, 2007.

[19] G. Byrne and J. R. Cebral. Vortex dynamics in cerebral aneurysms. *arXiv:1309.7875v1*, 2013.

[20] G. Byrne, F. Mut, and J. R. Cebral. Quantifying the large-scale hemodynamics of intracranial aneurysms. *American Journal of Neuroradiology*, 35(2):333–338, 2014.

[21] J. J. Caban and D. Gotz. Visual analytics in healthcare – opportunities and research challenges. *Journal of the American Medical Informatics Association*, 22(2):260–262, 2015.

[22] B. Cabral and L. C. Leedom. Imaging vector fields using line integral convolution. In *SIGGRAPH*, pages 263–270. ACM, 1993.

[23] M. A. Castro, C. M. Putman, M. Sheridan, and J. Cebral. Hemodynamic patterns of anterior communicating artery aneurysms: a possible association with rupture. *American Journal of Neuroradiology*, 30(2):297–302, 2009.

[24] J. Cebral, F. Mut, M. Raschi, E. Scrivano, R. Ceratto, P. Lylyk, and C. Putman. Aneurysm rupture following treatment with flow-diverting stents: computational hemodynamics analysis of treatment. *American Journal of Neuroradiology*, 32(1):27–33, 2011.

[25] J. R. Cebral, M. Castro, S. Appanaboyina, C. M. Putman, D. Millan, A. F. Frangi, et al. Efficient pipeline for image-based patient-specific analysis of cerebral aneurysm hemodynamics: Technique and sensitivity. *IEEE Transactions on Medical Imaging*, 24(4):457–467, 2005.

[26] J. R. Cebral, F. Mut, J. Weir, and C. Putman. Quantitative characterization of the hemodynamic environment in ruptured and unruptured brain aneurysms. *American Journal of Neuroradiology*, 32(1):145–151, 2011.

[27] J. R. Cebral, F. Mut, J. Weir, and C. M. Putman. Association of hemodynamic characteristics and cerebral aneurysm rupture. *American Journal of Neuroradiology*, 32(2):264–270, 2011.

[28] C. Chalopin, S. Oeltze, B. Preim, A. Müns, J. Meixensberger, and D. Lindner. Method for the evaluation of US perfusion for brain tumor surgery. *Jahrestagung der Deutschen Gesellschaft für Computer- und Roboter Assistierte Chirurgie*, pages 198–202, 2013.

[29] C.-K. Chen, S. Yan, H. Yu, N. Max, and K.-L. Ma. An illustrative visualization framework for 3D vector fields. *Computer Graphics Forum*, 30(7):1941–1951, 2011.

[30] E. S. Connolly, A. A. Rabinstein, J. R. Carhuapoma, C. P. Derdeyn, J. Dion, R. T. Higashida, B. L. Hoh, C. J. Kirkness, A. M. Naidech, C. S. Ogilvy, et al. Guidelines for the management of aneurysmal subarachnoid hemorrhage: A guideline for healthcare professionals from the American heart association/American stroke association. *Stroke*, 43(6):1711–1737, 2012.

[31] G. Corrao, V. Bagnardi, A. Zambon, and C. La Vecchia. A meta-analysis of alcohol consumption and the risk of 15 diseases. *Preventive Medicine*, 38(5):613–619, 2004.

[32] C. D. Correa, Y.-H. Chan, and K.-L. Ma. A framework for uncertainty-aware visual analytics. In *IEEE Symposium on Visual Analytics Science and Technology (VAST)*, pages 51–58, 2009.

[33] D. R. Cox and E. J. Snell. *Analysis of binary data*, volume 32 of *Chapman & Hall/CRC Monographs on Statistics & Applied Probability*. Chapman and Hall/CRC, 2nd edition, 1989.

[34] T. N. Dang, P. Murray, and A. G. Forbes. PathwayMatrix: Visualizing binary relationships between proteins in biological pathways. *BMC Proceedings*, 9(Suppl 6):S3, 2015.

[35] N. de Hoon, R. van Pelt, A. Jalba, and A. Vilanova. 4D MRI flow coupled to physics-based fluid simulation for blood-flow visualization. *Computer Graphics Forum*, 33(3):121–130, 2014.

[36] J. Dietzsch, J. Heinrich, K. Nieselt, and D. Bartz. Spray: a visual analytics approach for gene expression data. In *IEEE Symposium on Visual Analytics Science and Technology (VAST)*, pages 179–186, 2009.

[37] J. Dill, R. Earnshaw, D. Kasik, J. Vince, and P. C. Wong, editors. *Expanding the Frontiers of Visual Analytics and Visualization*. Springer, 2012.

[38] H. Ding, C. Wang, K. Huang, and R. Machiraju. GRAPHIE: Graph based histology image explorer. *BMC Bioinformatics*, 16(Suppl 11):S10, 2015.

[39] H. Doleisch, M. Gasser, and H. Hauser. Interactive feature specification for focus+context visualization of complex simulation data. In *IEEE TCVG - EUROGRAPHICS Symposium on Visualization*, pages 239–248, 2003.

[40] H. Doleisch and H. Hauser. Smooth brushing for focus+context visualization of simulation data in 3D. *Journal of WSCG*, 10(1):147–154, 2002.

[41] R. Doll and R. Peto. Cigarette smoking and bronchial carcinoma: Dose and time relationships among regular smokers and lifelong non-smokers. *Journal of Epidemiology and Community Health*, 32(4):303–313, 1978.

[42] A. R. T. Donders, G. J. van der Heijden, T. Stijnen, and K. G. Moons. Review: a gentle introduction to imputation of missing values. *Journal of Clinical Epidemiology*, 59(10):1087–1091, 2006.

[43] W. Dubitzky, O. Wolkenhauer, H. Yokota, and K.-H. Cho. *Encyclopedia of Systems Biology*. Springer Publishing Company, Incorporated, 2013.

[44] S. G. Eick and G. J. Wills. High interaction graphics. *European Journal of Operational Research*, 81(3):445–459, 1995.

[45] J. W. Emerson, W. A. Green, B. Schloerke, J. Crowley, D. Cook, H. Hofmann, and H. Wickham. The generalized pairs plot. *Journal of Computational and Graphical Statistics*, 22(1):79–91, 2013.

[46] A. Endert, M. S. Hossain, N. Ramakrishnan, C. North, P. Fiaux, and C. Andrews. The human is the loop: New directions for visual analytics. *Journal of Intelligent Information Systems*, 43(3):411–435, 2014.

[47] M. Ester, H.-P. Kriegel, J. Sander, and X. Xu. A density-based algorithm for discovering clusters in large spatial databases with noise. In *International Conference on Knowledge Discovery and Data Mining*, pages 226–231, 1996.

[48] European Population Imaging Infrastructure. `http://populationimaging.eu/` [2016-14-02].

[49] Facebook company info. `http://newsroom.fb.com/company-info/` [2015-11-26].

[50] J.-D. Fekete. Visual analytics infrastructures: From data management to exploration. *Computer*, 46(7):22–29, 2013.

[51] W. Freiler, K. Matković, and H. Hauser. Interactive visual analysis of set-typed data. *IEEE Transactions on Visualization and Computer Graphics*, 14(6):1340–1347, 2008.

[52] M. Friedenberger, M. Bode, A. Krusche, and W. Schubert. Fluorescence detection of protein clusters in individual cells and tissue sections by using toponome imaging system: Sample preparation and measuring procedures. *Nature Protocols*, 2(9):2285–2294, 2007.

[53] J. Frösen, R. Tulamo, A. Paetau, E. Laaksamo, M. Korja, A. Laakso, M. Niemelä, and J. Hernesniemi. Saccular intracranial aneurysm: Pathology and mechanisms. *Acta Neuropathologica*, 123(6):773–786, 2012.

[54] R. Gasparotti and R. Liserre. Intracranial aneurysms. *European Radiology*, 15(3):441–447, 2005.

[55] R. Gasteiger. *Visual Exploration of Cardiovascular Hemodynamics*. PhD thesis, Dept. of Simulation and Graphics, Univ. of Magdeburg, Germany, 2014.

[56] R. Gasteiger, D. J. Lehmann, R. van Pelt, G. Janiga, O. Beuing, A. Vilanova, H. Theisel, and B. Preim. Automatic detection and visualization of qualitative hemodynamic characteristics in cerebral aneurysms. *IEEE Transactions on Visualization and Computer Graphics*, 18(12):2178–2187, 2012.

[57] N. Gehlenborg, S. I. O'Donoghue, N. S. Baliga, A. Goesmann, M. A. Hibbs, H. Kitano, O. Kohlbacher, H. Neuweger, R. Schneider, D. Tenenbaum, et al. Visualization of omics data for systems biology. *Nature Methods*, 7:S56–S68, 2010.

[58] R. P. Gelber, L. J. Launer, and L. R. White. The Honolulu-Asia Aging Study: Epidemiologic and neuropathologic research on cognitive impairment. *Current Alzheimer Research*, 9(6):664–672, 2012.

[59] A. Gerasch, D. Faber, J. Küntzer, P. Niermann, O. Kohlbacher, H.-P. Lenhof, and M. Kaufmann. Bina: a visual analytics tool for biological network data. *PloS one*, 9(2):e87397, 2014.

[60] German National Cohort Consortium. The German National Cohort: Aims, study design and organization. *European Journal of Epidemiology*, 29(5):371–382, 2014.

[61] E. S. Gilbert. Ionising radiation and cancer risks: What have we learned from epidemiology? *International Journal of Radiation Biology*, 85(6):467–482, 2009.

[62] S. Glaßer, K. Lawonn, T. Hoffmann, M. Skalej, and B. Preim. Combined visualization of wall thickness and wall shear stress for the evaluation of aneurysms. *IEEE Transactions on Visualization and Computer Graphics*, 20(12):2506–2515, 2014.

[63] S. Glaßer, S. Oeltze, U. Preim, A. Bjørnerud, H. Hauser, and B. Preim. Visual Analysis of Longitudinal Brain Tumor Perfusion. In *Proc. of the SPIE Medical Imaging*, page 86700Z, 2013.

[64] S. Glaßer, U. Preim, K. Tönnies, and B. Preim. A visual analytics approach to diagnosis of breast DCE-MRI data. *Computers & Graphics*, 34(5):602–611, 2010.

[65] D. Gotz, A. Perer, and Z. Zhang. Iterative refinement of cohorts using visual exploration and data analytics, Nov. 8 2012. US Patent App. 13/672,000.

[66] L. Goubergrits, B. Thamsen, A. Berthe, J. Poethke, U. Kertzscher, K. Affeld, C. Petz, H.-C. Hege, H. Hoch, and A. Spuler. In vitro study of near-wall flow in a cerebral aneurysm model with and without coils. *American Journal of Neuroradiology*, 31(8):1521–1528, 2010.

[67] D. L. Gresh, B. E. Rogowitz, R. L. Winslow, D. F. Scollan, and C. K. Yung. WEAVE: a system for visually linking 3-D and statistical visualizations, applied to cardiac simulation and measurement data. In *Conference on Visualization*, pages 489–492, 2000.

[68] T. Groenfeldt. At NYSE, the data deluge overwhelms traditional databases, 2013. Forbes Technology `http://www.forbes.com/sites/tomgroenfeldt/2013/02/14/at-nyse-the-data-deluge-overwhelms-traditional-databases/` [2015-11-26].

[69] A. Grzesik, J. Bernarding, J. Braun, H.-C. Koennecke, K. J. Wolf, and T. Tolxdorff. Characterization of stroke lesions using a histogram-based data analysis including diffusion-and perfusion-weighted imaging. In *SPIE Medical Imaging: Physiology and Function from Multidimensional Images*, volume 3978, pages 23–31, 2000.

[70] M. Hadwiger, J. Beyer, W.-K. Jeong, and H. Pfister. Interactive volume exploration of petascale microscopy data streams using a visualization-driven virtual memory approach. *IEEE Transactions on Visualization and Computer Graphics*, 18(12):2285–2294, 2012.

[71] J. Han, M. Kamber, and J. Pei. *Data mining: Concepts and Techniques*. Morgan Kaufmann Publishers, 3rd edition, 2011.

[72] C. D. Hansen and C. R. Johnson, editors. *Visualization Handbook*. Academic Press, 2011.

[73] J. A. Hartigan and B. Kleiner. Mosaics for contingency tables. In *Computer Science and Statistics: Proceedings of the Symposium on the Interface*, pages 268–273, 1981.

[74] T. Hassan, E. V. Timofeev, M. Ezura, T. Saito, A. Takahashi, K. Takayama, and T. Yoshimoto. Hemodynamic analysis of an adult vein of Galen aneurysm malformation by use of 3D image-based computational fluid dynamics. *American Journal of Neuroradiology*, 24(6):1075–1082, 2003.

[75] H. Hauser. Generalizing focus+context visualization. In *Scientific Visualization: The Visual Extraction of Knowledge from Data*, pages 305–327. Springer, 2006.

[76] H. Hauser. Basics of interactive visual analysis, 2013. Notes of the IEEE Visualization conference tutorial "Interactive Visual Analysis of Scientific Data" `http://www.vismd.de/lib/exe/fetch.php?media=teaching_tutorials:iva_basics_hauser.pdf`.

[77] H. Hauser, F. Ledermann, and H. Doleisch. Angular brushing of extended parallel coordinates. In *IEEE Symposium on Information Visualization*, pages 127–130, 2002.

[78] K. Hegenscheid, J. Kühn, H. Völzke, R. Biffar, N. Hosten, and R. Puls. Whole-body magnetic resonance imaging of healthy volunteers: Pilot study results from the population-based SHIP study. *RoFo: Fortschritte auf dem Gebiete der Rontgenstrahlen und der Nuklearmedizin*, 181(8):748–759, 2009.

[79] E. Heiberg, T. Ebbers, L. Wigström, and M. Karlsson. Three-dimensional flow characterization using vector pattern matching. *IEEE Transactions on Visualization and Computer Graphics*, 9(3):313–319, 2003.

[80] N. Henry and J.-D. Fekete. Matrixexplorer: a dual-representation system to explore social networks. *IEEE Transactions on Visualization and Computer Graphics*, 12(5):677–684, 2006.

[81] C. M. Hentschke, K. D. Tönnies, O. Beuing, and R. Nickl. A new feature for automatic aneurysm detection. In *IEEE International Symposium on Biomedical Imaging (ISBI)*, pages 800–803, 2012.

[82] R. Hillert. Toponome analysis. *Encyclopedia of Systems Biology*, pages 2188–2191, 2013.

[83] R. Hillert, A. Gieseler, A. Krusche, D. Humme, H.-J. Röwert-Huber, W. Sterry, P. Walden, and W. Schubert. Large molecular systems landscape uncovers T cell trapping in human skin cancer. *Scientific Reports*, 6(19012), 2016.

[84] A. Hofman, G. G. Brusselle, S. D. Murad, C. M. van Duijn, O. H. Franco, A. Goedegebure, M. A. Ikram, C. C. Klaver, T. E. Nijsten, R. P. Peeters, et al. The Rotterdam Study: 2016 objectives and design update. *European Journal of Epidemiology*, 30(8):661–708, 2015.

[85] Homepage of the "Symposium on Biological Data Visualization". `http://biovis.net/2016/index.html` [2016-07-02].

[86] Homepage of the "Virtual Intracranial Stenting Challenge". http://www.ics-meeting.net/ics_visc_past_.html [2016-19-01].

[87] Homepage of the "Workshop on Visual Analytics in Healthcare". http://visualanalyticshealthcare.org/ [2016-06-02].

[88] R. P. Horgan and L. C. Kenny. 'Omic' technologies: genomics, transcriptomics, proteomics and metabolomics. *The Obstetrician & Gynaecologist*, 13(3):189–195, 2011.

[89] A. Humayun, C. W. Shan-e Ahmed Raza, S. Abouna, M. Khan, and N. M. Rajpoot. A novel framework for molecular co-expression pattern analysis in multi-channel toponome fluorescence images. In *Microscopic Image Analysis with Applications in Biology (MIAAB)*, pages 109–112, 2011.

[90] IEEE Xplore Digital Library. http://ieeexplore.ieee.org/Xplore/home.jsp [2015-11-26].

[91] J. E. Iglesias, J. C. Augustinack, K. Nguyen, C. M. Player, A. Player, M. Wright, N. Roy, M. P. Frosch, A. C. McKee, L. L. Wald, et al. A computational atlas of the hippocampal formation using ex vivo, ultra-high resolution MRI: Application to adaptive segmentation of in vivo MRI. *NeuroImage*, 115:117–137, 2015.

[92] M. A. Ikram, A. van der Lugt, W. J. Niessen, P. J. Koudstaal, G. P. Krestin, A. Hofman, D. Bos, and M. W. Vernooij. The Rotterdam Scan Study: Design update 2016 and main findings. *European Journal of Epidemiology*, pages 1–17, 2015.

[93] M. Inaba, L. White, C. Bell, R. Chen, H. Petrovitch, L. Launer, R. D. Abbott, G. W. Ross, and K. Masaki. White matter lesions on brain magnetic resonance imaging scan and 5-year cognitive decline: the Honolulu-Asia Aging Study. *Journal of the American Geriatrics Society*, 59(8):1484–1489, 2011.

[94] A. Inselberg. The plane with parallel coordinates. *The Visual Computer*, 1(2):69–91, 1985.

[95] J. A. Jacko, editor. *Human Computer Interaction Handbook: Fundamentals, Evolving Technologies, and Emerging Applications*. CRC press, 2012.

[96] G. Janiga, L. s. Daróczy, P. Berg, D. Thévenin, M. Skalej, and O. Beuing. An automatic CFD-based flow diverter optimization principle for patient-specific intracranial aneurysms. *Journal of Biomechanics*, 48(14):3846–3852, 2015.

[97] G. Janiga, C. Rössl, M. Skalej, and D. Thévenin. Realistic virtual intracranial stenting and computational fluid dynamics for treatment analysis. *Journal of Biomechanics*, 46(1):7–12, 2013.

[98] M. Jarema, I. Demir, J. Kehrer, and R. Westermann. Comparative visual analysis of vector field ensembles. In *IEEE Conference on Visual Analytics Science and Technology (VAST)*, pages 81–88, 2015.

[99] F. Jeanquartier and A. Holzinger. On visual analytics and evaluation in cell physiology: a case study. In *Availability, Reliability, and Security in Information Systems and HCI*, pages 495–502. Springer, 2013.

[100] L.-D. Jou, C. M. Quick, W. L. Young, M. T. Lawton, R. Higashida, A. Martin, and D. Saloner. Computational approach to quantifying hemodynamic forces in giant cerebral aneurysms. *American Journal of Neuroradiology*, 24(9):1804–1810, 2003.

[101] L. T. Kaastra and B. Fisher. Field experiment methodology for pair analytics. In *Workshop on Beyond Time and Errors: Novel Evaluation Methods for Visualization*, pages 152–159, 2014.

[102] J. Kehrer and H. Hauser. Visualization and visual analysis of multifaceted scientific data: A survey. *IEEE Transactions on Visualization and Computer Graphics*, 19(3):495–513, 2013.

[103] J. Kehrer, F. Ladstadter, P. Muigg, H. Doleisch, A. Steiner, and H. Hauser. Hypothesis generation in climate research with interactive visual data exploration. *IEEE Transactions on Visualization and Computer Graphics*, 14(6):1579–1586, 2008.

[104] D. Keim, F. Mansmann, J. Schneidewind, and H. Ziegler. Challenges in visual data analysis. In *International Conference on Information Visualization*, pages 9–16, 2006.

[105] D. A. Keim, J. Kohlhammer, F. Mansmann, T. May, and F. Wanner. Visual analytics. In D. A. Keim, J. Kohlhammer, G. Ellis, and F. Mansmann, editors, *Mastering the Information Age - Solving Problems with Visual Analytics*, pages 7–18. Eurographics Association, 2010.

[106] G. A. Kerchner, G. K. Deutsch, M. Zeineh, R. F. Dougherty, M. Saranathan, and B. K. Rutt. Hippocampal CA1 Apical Neuropil Atrophy and Memory Performance in Alzheimer's Disease. *NeuroImage*, 63(1):194–202, 2012.

[107] M. Khan and C. Waddington. Clinical aspects of the Toponome Imaging System (TIS). *Encyclopedia of Systems Biology*, pages 412–414, 2013.

[108] P. Klemm. Interaktive 3D-Visualisierung von Protein-Clustern in Toponomdaten. Master's thesis, Dept. of Computer Science, Univ. of Magdeburg, Germany, 2012.

[109] P. Klemm, L. Frauenstein, D. Perlich, K. Hegenscheid, H. Völzke, and B. Preim. Clustering socio-demographic and medical attribute data in cohort studies. In *Bildverarbeitung für die Medizin (BVM)*, pages 180–185, 2014.

[110] P. Klemm, S. Glaßer, K. Lawonn, M. Rak, H. Völzke, K. Hegenscheid, and B. Preim. Interactive visual analysis of lumbar back pain. In *International Conference on Information Visualization Theory and Applications (IVAPP)*, pages 85–92, 2015.

[111] P. Klemm, K. Lawonn, S. Glaßer, U. Niemann, K. Hegenscheid, H. Völzke, and B. Preim. 3D Regression heat map analysis of population study data. *IEEE Transactions on Visualization and Computer Graphics*, 22(1):81–90, 2016.

[112] P. Klemm, K. Lawonn, M. Rak, B. Preim, K. D. Tönnies, K. Hegenscheid, H. Völzke, and S. Oeltze. Visualization and analysis of lumbar spine canal variability in cohort study data. In M. Bronstein, J. Favre, and K. Hormann, editors, *International Workshop on Vision, Modeling and Visualization (VMV)*, pages 121–128, 2013.

[113] P. Klemm, S. Oeltze-Jafra, K. Lawonn, K. Hegenscheid, H. Völzke, and B. Preim. Interactive visual analysis of image-centric cohort study data. *IEEE Transactions on Visualization and Computer Graphics*, 20(12):1673–1682, 2014.

[114] B. Köhler, S. Born, R. van Pelt, U. Preim, and B. Preim. A survey of cardiac 4D PC-MRI data processing. *Computer Graphics Forum*, page to appear, 2016.

[115] B. Köhler, R. Gasteiger, U. Preim, H. Theisel, M. Gutberlet, and B. Preim. Semi-automatic vortex extraction in 4D PC-MRI cardiac blood flow data using line predicates. *IEEE Transactions on Visualization and Computer Graphics*, 19(12):2773–2782, 2013.

[116] J. Kölling, D. Langenkämper, S. Abouna, M. Khan, and T. W. Nattkemper. Whide - a web tool for visual data mining colocation patterns in multivariate bioimages. *Bioinformatics*, 28(8):1143–1150, 2012.

[117] Z. Konyha, K. Matković, D. Gračanin, M. Jelović, and H. Hauser. Interactive visual analysis of families of function graphs. *IEEE Transactions on Visualization and Computer Graphics*, 12(6):1373–1385, 2006.

[118] B. Kovalerchuk, F. Delizy, L. Riggs, and E. Vityaev. Visual discovery in multivariate binary data. In *IS&T/SPIE Electronic Imaging*, pages 75300B–75300B. International Society for Optics and Photonics, 2010.

[119] J. Krause, A. Perer, and H. Stavropoulos. Supporting iterative cohort construction with visual temporal queries. *IEEE Transactions on Visualization and Computer Graphics*, 22(1):91–100, 2016.

[120] J. Kretschmer, C. Godenschwager, B. Preim, and M. Stamminger. Interactive patient-specific vascular modeling with sweep surfaces. *IEEE Transactions on Visualization and Computer Graphics*, 19(12):2828–2837, 2013.

[121] T. Krings, D. M. Mandell, T.-R. Kiehl, S. Geibprasert, M. Tymianski, H. Alvarez, F.-J. Hans, et al. Intracranial aneurysms: from vessel wall pathology to therapeutic approach. *Nature Reviews Neurology*, 7(10):547–559, 2011.

[122] A. Kuhn, D. J. Lehmann, R. Gasteiger, M. Neugebauer, B. Preim, and H. Theisel. A clustering-based visualization technique to emphasize meaningful regions of vector fields. In *International Workshop on Vision, Modeling and Visualization (VMV)*, pages 191–198, 2011.

[123] R. S. Laramee, C. Garth, H. Doleisch, J. Schneider, H. Hauser, and H. Hagen. Visual analysis and exploration of fluid flow in a cooling jacket. In *IEEE Visualization*, pages 623–630, 2005.

[124] K. Lawonn, S. Glaßer, A. Vilanova, B. Preim, and T. Isenberg. Occlusion-free blood flow animation with wall thickness visualization. *IEEE Transactions on Visualization and Computer Graphics*, 22(1):728–737, 2015.

[125] M. D. Lee, M. A. Butavicius, and R. E. Reilly. Visualizations of binary data: a comparative evaluation. *International Journal of Human-Computer Studies*, 59(5):569–602, 2003.

[126] D. Lesage, E. D. Angelini, I. Bloch, and G. Funka-Lea. A review of 3D vessel lumen segmentation techniques: Models, features and extraction schemes. *Medical Image Analysis*, 13(6):819–845, 2009.

[127] Y. Li, C. Wang, and C.-K. Shene. Streamline similarity analysis using bag-of-features. In *IS&T/SPIE Electronic Imaging*, pages 90170N–90170N. International Society for Optics and Photonics, 2013.

[128] J. W. Lichtman and J.-A. Conchello. Fluorescence microscopy. *Nature Methods*, 2(12):910–919, 2005.

[129] B. Linke, S. Pierre, O. Coste, C. Angioni, W. Becker, T. J. r. Maier, D. Steinhilber, C. Wittpoth, G. Geisslinger, and K. Scholich. Toponomics analysis of drug-induced changes in arachidonic acid-dependent signaling pathways during spinal nociceptive processing. *Journal of Proteome Research*, 8(10):4851–4859, 2009.

[130] W. E. Lorensen and H. E. Cline. Marching cubes: a high resolution 3D surface construction algorithm. *SIGGRAPH Computer Graphics*, 21(4):163–169, 1987.

[131] K. Lu, A. Chaudhuri, T.-Y. Lee, H.-W. Shen, and P. C. Wong. Exploring vector fields with distribution-based streamline analysis. In *IEEE Pacific Visualization Symposium (PacificVis)*, pages 257–264, 2013.

[132] J. Ma, C. Wang, C.-K. Shene, and J. Jiang. A graph-based interface for visual analytics of 3D streamlines and pathlines. *IEEE Transactions on Visualization and Computer Graphics*, 20(8):1127–1140, 2014.

[133] A. Maaß, H. Schütze, O. Speck, A. Yonelinas, C. Tempelmann, H.-J. Heinze, D. Berron, A. Cardenas-Blanco, K. H. Brodersen, K. E. Stephan, et al. Laminar activity in the hippocampus and entorhinal cortex related to novelty and episodic encoding. *Nature Communications*, 5(5547), 2014.

[134] L. MacDonald, G. Baldini, and B. Storrie. Does super-resolution fluorescence microscopy obsolete previous microscopic approaches to protein co-localization? *Methods in Molecular Biology*, 1270:255–275, 2015.

[135] A. M. MacEachren. Visual analytics and uncertainty: Its not about the data. In E. Bertini and J. C. Roberts, editors, *EuroVis Workshop on Visual Analytics (EuroVA)*. The Eurographics Association, 2015.

[136] R. Maciejewski, R. Hafen, S. Rudolph, S. G. Larew, M. A. Mitchell, W. S. Cleveland, and D. S. Ebert. Forecasting hotspots—a predictive analytics approach. *IEEE Transactions on Visualization and Computer Graphics*, 17(4):440–453, 2011.

[137] S. Malik, F. Du, M. Monroe, E. Onukwugha, C. Plaisant, and B. Shneiderman. Cohort comparison of event sequences with balanced integration of visual analytics and statistics. In *International Conference on Intelligent User Interfaces*, pages 38–49, 2015.

[138] K. Matković, W. Freiler, D. Gračanin, and H. Hauser. ComVis: a coordinated multiple views system for prototyping new visualization technology. In *International Conference Information Visualisation*, pages 215–220, 2008.

[139] K. Matković, D. Gračanin, B. Klarin, and H. Hauser. Interactive visual analysis of complex scientific data as families of data surfaces. *IEEE Transactions on Visualization and Computer Graphics*, 15(6):1351–1358, 2009.

[140] K. Matkovic, D. Gracanin, R. Splechtna, M. Jelovic, B. Stehno, H. Hauser, and W. Purgathofer. Visual analytics for complex engineering systems: Hybrid visual steering of simulation ensembles. *IEEE Transactions on Visualization and Computer Graphics*, 20(12):1803–1812, 2014.

[141] T. McLoughlin, M. W. Jones, R. S. Laramee, R. Malki, I. Masters, and C. D. Hansen. Similarity measures for enhancing interactive streamline seeding. *IEEE Transactions on Visualization and Computer Graphics*, 19(8):1342–53, 2013.

[142] H. Meng, Y. Feng, S. H. Woodward, B. R. Bendok, R. A. Hanel, L. R. Guterman, and L. N. Hopkins. Mathematical model of the rupture mechanism of intracranial saccular aneurysms through daughter aneurysm formation and growth. *Neurological Research*, 27(5):459–465, 2005.

[143] H. Meng, V. Tutino, J. Xiang, and A. Siddiqui. High WSS or low WSS? Complex interactions of hemodynamics with intracranial aneurysm initiation, growth, and rupture: Toward a unifying hypothesis. *American Journal of Neuroradiology*, 35(7):1254–1262, 2014.

[144] R. M. Merrill. *Introduction To Epidemiology*. Jones & Bartlett Learning, 6th edition, 2012.

[145] M. Meuschke, K. Lawonn, B. Köhler, U. Preim, and B. Preim. Clustering of aortic vortex flow in cardiac 4D PC-MRI data. In *Bildverarbeitung für die Medizin (BVM)*, page to appear, 2016.

[146] J. Meyer, J. Thomas, S. Diehl, B. D. Fisher, D. A. Keim, D. H. Laidlaw, S. Miksch, K. Mueller, W. Ribarsky, B. Preim, and A. Ynnerman. From visualization to visually enabled reasoning. In H. Hagen, editor, *Scientific Visualization: Advanced Concepts*, volume 1, pages 227–245. Dagstuhl Follow-Ups, 2010.

[147] M. Meyer, T. Munzner, and H. Pfister. MizBee: a multiscale synteny browser. *IEEE Transactions on Visualization and Computer Graphics*, 15(6):897–904, 2009.

[148] T. Mönch, R. Gasteiger, G. Janiga, H. Theisel, and B. Preim. Context-aware mesh smoothing for biomedical applications. *Computers & Graphics*, 35(4):755–767, 2011.

[149] T. Mönch, M. Neugebauer, and B. Preim. Optimization of vascular surface models for computational fluid dynamics and rapid prototyping. In *International Workshop on Digital Engineering (IWDE)*, pages 16–23, 2011.

[150] M. Monroe, R. Lan, H. Lee, C. Plaisant, and B. Shneiderman. Temporal event sequence simplification. *IEEE Transactions on Visualization and Computer Graphics*, 19(12):2227–2236, 2013.

[151] H. G. Morales, I. Larrabide, A. J. Geers, L. San Roman, J. Blasco, J. M. Macho, and A. F. Frangi. A virtual coiling technique for image-based aneurysm models by dynamic path planning. *IEEE Transactions on Medical Imaging*, 32(1):119–129, 2013.

[152] P. D. Morris, A. Narracott, H. von Tengg-Kobligk, D. A. S. Soto, S. Hsiao, A. Lungu, P. Evans, N. W. Bressloff, P. V. Lawford, D. R. Hose, et al. Computational fluid dynamics modelling in cardiovascular medicine. *Heart*, 102(1):18–28, 2016.

[153] P. Muigg, M. Hadwiger, H. Doleisch, and H. Hauser. Scalable hybrid unstructured and structured grid raycasting. *IEEE Transactions on Visualization and Computer Graphics*, 13(6):1592–1599, 2007.

[154] P. Muigg, J. Kehrer, S. Oeltze, H. Piringer, H. Doleisch, B. Preim, and H. Hauser. A four-level focus+context approach to interactive visual analysis of temporal features in large scientific data. *Computer Graphics Forum*, 27(3):775–782, 2008.

[155] R. F. Murphy. Putting proteins on the map. *Nature biotechnology*, 24(10):1223–1224, 2006.

[156] T. W. Nattkemper, H. Wersing, W. Schubert, and H. Ritter. A neural network architecture for automatic segmentation of fluorescence micrographs. *Neurocomputing*, 48(1):357–367, 2002.

[157] M. Neugebauer. *Computergestützte Exploration von Blutfluss in zerebralen Aneurysmen - geometrische Verarbeitung und interaktive Visualisierung*. PhD thesis, Dept. of Simulation and Graphics, Univ. of Magdeburg, Germany, 2014.

[158] M. Neugebauer, V. Diehl, M. Skalej, and B. Preim. Geometric reconstruction of the ostium of cerebral aneurysms. In *International Workshop on Vision, Modeling and Visualization (VMV)*, pages 307–314, 2010.

[159] M. Neugebauer, K. Lawonn, O. Beuing, P. Berg, G. Janiga, and B. Preim. AmniVis–a system for qualitative exploration of near-wall hemodynamics in cerebral aneurysms. *Computer Graphics Forum*, 32(3pt3):251–260, 2013.

[160] M. Neugebauer, K. Lawonn, O. Beuing, and B. Preim. Automatic generation of anatomic characteristics from cerebral aneurysm surface models. *International Journal of Computer Assisted Radiology and Surgery*, 8(2):279–289, 2013.

[161] C. B. Nielsen, M. Cantor, I. Dubchak, D. Gordon, and T. Wang. Visualizing genomes: Techniques and challenges. *Nature Methods*, 7:S5–S15, 2010.

[162] U. Niemann, H. Völzke, J.-P. Kühn, and M. Spiliopoulou. Learning and inspecting classification rules from longitudinal epidemiological data to identify predictive features on hepatic steatosis. *Expert Systems with Applications*, 2014.

[163] S. I. O'Donoghue, A.-C. Gavin, N. Gehlenborg, D. S. Goodsell, J.-K. Hériché, C. B. Nielsen, C. North, A. J. Olson, J. B. Procter, D. W. Shattuck, et al. Visualizing biological data—now and in the future. *Nature Methods*, 7:S2–S4, 2010.

[164] S. Oeltze. *Visual Exploration and Analysis of Perfusion Data*. PhD thesis, Dept. of Simulation and Graphics, Univ. of Magdeburg, Germany, 2010.

[165] S. Oeltze, H. Doleisch, H. Hauser, P. Muigg, and B. Preim. Interactive visual analysis of perfusion data. *IEEE Transactions on Visualization and Computer Graphics*, 13(6):1392–1399, 2007.

[166] S. Oeltze, W. Freiler, R. Hillert, H. Doleisch, B. Preim, and W. Schubert. Interactive, graph-based visual analysis of high-dimensional, multi-parameter fluorescence microscopy data in toponomics. *IEEE Transactions on Visualization and Computer Graphics*, 17(12):1882–1891, 2011.

[167] S. Oeltze, P. Klemm, R. Hillert, B. Preim, and W. Schubert. Visualization and exploration of 3D toponome data. In *Eurographics Workshop on Visual Computing for Biology and Medicine (VCBM)*, pages 115–122, 2012.

[168] S. Oeltze, D. J. Lehmann, A. Kuhn, G. Janiga, H. Theisel, and B. Preim. Blood flow clustering and applications in virtual stenting of intracranial aneurysms. *IEEE Transactions on Visualization and Computer Graphics*, 20(5):686–701, 2014.

[169] S. Oeltze, D. J. Lehmann, H. Theisel, and B. Preim. Evaluation of streamline clustering techniques for blood flow data. Technical Report FIN-06-2012, Otto-von-Guericke-Universität Magdeburg, Dezember 2012.

[170] S. Oeltze, H. Schütze, A. Maaß, E. Düzel, and B. Preim. Measurement of the Stratum Radiatum/Lacunosum-Moleculare (SRLM). In *Bildverarbeitung für die Medizin (BVM)*, pages 264–269, 2014.

[171] S. Oeltze-Jafra, J. R. Cebral, G. Janiga, and B. Preim. Cluster analysis of vortical flow in simulations of cerebral aneurysm hemodynamics. *IEEE Transactions on Visualization and Computer Graphics*, 22(1):757–766, 2016.

[172] S. Oeltze-Jafra, F. Pieper, R. Hillert, B. Preim, and W. Schubert. Interactive labeling of toponome data. In *Eurographics Workshop on Visual Computing for Biology and Medicine (EG VCBM)*, pages 79–88, 2014.

[173] S. Oeltze-Jafra and B. Preim. Survey of labeling techniques in medical visualizations. In *Eurographics Workshop on Visual Computing for Biology and Medicine (EG VCBM)*, pages 199–208, 2014.

[174] A. Perer, F. Wang, and J. Hu. Mining and exploring care pathways from electronic medical records with visual analytics. *Journal of Biomedical Informatics*, 56:369–378, 2015.

[175] M. V. Peters. Cutting the "gordian knot" in early breast cancer. *Annals of the Royal College of Physicians and Surgeons of Canada*, 8:186–192, 1975.

[176] S. E. Petersen, P. M. Matthews, F. Bamberg, D. A. Bluemke, J. M. Francis, M. G. Friedrich, P. Leeson, E. Nagel, S. Plein, F. E. Rademakers, et al. Imaging in population science: Cardiovascular magnetic resonance in 100,000 participants of UK Biobank-rationale, challenges and approaches. *Journal of Cardiovascular Magnetic Resonance*, 15(1):46, 2013.

[177] F. Pieper. Interaktives Labeling von Toponomdaten, 2013. Bachelor's thesis, Dept. of Computer Science, Univ. of Magdeburg, Germany.

[178] C. Plaisant, B. Milash, A. Rose, S. Widoff, and B. Shneiderman. LifeLines: visualizing personal histories. In *SIGCHI Conference on Human Factors in Computing Systems*, pages 221–227, 1996.

[179] M. S. Porta, S. Greenland, M. Hernán, I. dos Santos Silva, and J. M. Last, editors. *A Dictionary of Epidemiology*. Oxford University Press, 6th edition, 2014.

[180] B. Preim and C. P. Botha. *Visual Computing for Medicine*, chapter Visual Exploration of Simulated and Measured Flow Data, pages 665–714. Morgan Kaufmann, 2nd edition, 2013.

[181] B. Preim, P. Klemm, H. Hauser, K. Hegenscheid, S. Oeltze, K. Tönnies, and H. Völzke. *Visual Analytics of Image-Centric Cohort Studies in Epidemiology*, volume 3 of *Mathematics and Visualization*, chapter Visualization in Medicine and Life Sciences, page in print. Springer Verlag, 2016.

[182] B. Preim, S. Oeltze, M. Mlejnek, E. Gröller, A. Hennemuth, and S. Behrens. Survey of the visual exploration and analysis of perfusion data. *IEEE Transactions on Visualization and Computer Graphics*, 15(2):205–220, 2009.

[183] U. Preim, S. Glaßer, B. Preim, F. Fischbach, and J. Ricke. Computer-aided diagnosis in breast DCE-MRI-quantification of the heterogeneity of breast lesions. *European Journal of Radiology*, 81(7):1532–1538, 2012.

[184] C. J. Prestigiacomo, editor. *Endovascular Surgical Neuroradiology: Theory and Clinical Practice*. Thieme, 2014.

[185] R. G. Raidou, M. Breeuwer, A. Vilanova, U. A. van der Heide, and P. van Houdt. The iCoCooN: Integration of cobweb charts with parallel coordinates for visual analysis of DCE-MRI modeling variations. In *Eurographics Workshop on Visual Computing for Biology and Medicine (VCBM)*, pages 11–20, 2014.

[186] R. G. Raidou, U. Van Der Heide, C. V. Dinh, G. Ghobadi, J. F. Kallehauge, M. Breeuwer, and A. Vilanova. Visual analytics for the exploration of tumor tissue characterization. *Computer Graphics Forum*, 34(3):11–20, 2015.

[187] M. Rak, K. Engel, and K. D. Tönnies. Closed-form hierarchical finite element models for part-based object detection. In *International Workshop on Vision, Modeling and Visualization (VMV)*, pages 137–144, 2013.

[188] A. Rind, T. D. Wang, W. Aigner, S. Miksch, K. Wongsuphasawat, C. Plaisant, and B. Shneiderman. Interactive information visualization to explore and query electronic health records. *Foundations and Trends in Human-Computer Interaction*, 5(3):207–298, 2013.

[189] J. C. Roberts. State of the art: Coordinated & multiple views in exploratory visualization. In *International Conference on Coordinated and Multiple Views in Exploratory Visualization*, pages 61–71, 2007.

[190] S. Rogers. Big data is scaling BI and analytics. *Information Management*, 21(5):14, 2011. `http://www.information-management.com/issues/21_5/big-data-is-scaling-bi-and-analytics-10021093-1.html`.

[191] T. Ropinski, S. Oeltze, and B. Preim. Survey of glyph-based visualization techniques for spatial multivariate medical data. *Computers and Graphics*, 35(2):392–401, 2011.

[192] C. Rössl and H. Theisel. Streamline embedding for 3D vector field exploration. *IEEE Transactions on Visualization and Computer Graphics*, 18(3):407–420, 2012.

[193] A. Ruitenberg, J. C. van Swieten, J. C. Witteman, K. M. Mehta, C. M. van Duijn, A. Hofman, and M. M. Breteler. Alcohol consumption and risk of dementia: the Rotterdam Study. *The Lancet*, 359(9303):281–286, 2002.

[194] C. Russ, R. Hopf, S. Hirsch, S. Sundermann, V. Falk, G. Székely, and M. Gessat. Simulation of transcatheter aortic valve implantation under consideration of leaflet calcification. In *Conference of the IEEE Engineering in Medicine and Biology Society (EMBC)*, pages 711–714, 2013.

[195] D. Sacha, H. Senaratne, B. C. Kwon, G. Ellis, and D. Keim. The role of uncertainty, awareness, and trust in visual analytics. *IEEE Transactions on Visualization and Computer Graphics*, 22(1):240–249, 2016.

[196] D. Sacha, A. Stoffel, F. Stoffel, B. C. Kwon, G. Ellis, and D. A. Keim. Knowledge generation model for visual analytics. *IEEE Transactions on Visualization and Computer Graphics*, 20(12):1604–1613, 2014.

[197] R. Sadana, T. Major, A. Dove, and J. Stasko. Onset: a visualization technique for large-scale binary set data. *IEEE Transactions on Visualization and Computer Graphics*, 20(12):1993–2002, 2014.

[198] O. Sahni, K. E. Jansen, M. S. Shephard, C. A. Taylor, and M. W. Beall. Adaptive boundary layer meshing for viscous flow simulations. *Engineering with Computers*, 24(3):267–285, 2008.

[199] O. Sahni, J. Müller, K. E. Jansen, M. S. Shephard, and C. A. Taylor. Efficient anisotropic adaptive discretization of the cardiovascular system. *Computer Methods in Applied Mechanics and Engineering*, 195(41):5634–5655, 2006.

[200] S. Salvador and P. Chan. Determining the number of clusters/segments in hierarchical clustering/segmentation algorithms. In *International Conference on Tools with Artificial Intelligence (ICTAI)*, pages 576 – 584, 2004.

[201] T. Salzbrunn, H. Jänicke, T. Wischgoll, and G. Scheuermann. The state of the art in flow visualization: Partition-based techniques. In *Simulation and Visualization (SimVis)*, pages 75–92, 2008.

[202] T. Salzbrunn and G. Scheuermann. Streamline predicates. *IEEE Transactions on Visualization and Computer Graphics*, 12(6):1061–1612, 2006.

[203] G. Scheuermann, W. Kollmann, X. Tricoche, and T. Wischgoll. *Numerical Simulations of Incompressible Flows*, chapter Evolution of topology in axi-symmetric and 3-d viscous flows, pages 622–643. World Scientific Publishing, 2003.

[204] J. Schmid, C. Chênes, S. Chagué, P. Hoffmeyer, P. Christofilopoulos, M. Bernardoni, and C. Charbonnier. MyHip: supporting planning and surgical guidance for a better total hip arthroplasty. *International Journal of Computer Assisted Radiology and Surgery*, 10(10):1–10, 2015.

[205] W. Schubert. Topological proteomics, toponomics, melk-technology. *Advances in Biochemical Engineering/Biotechnology*, 83:189–209, 2003.

[206] W. Schubert. Exploring molecular networks directly in the cell. *Cytometry Part A*, 69(3):109–112, 2006.

[207] W. Schubert. A three-symbol code for organized proteomes based on cyclical imaging of protein locations. *Cytometry Part A*, 71(6):352–360, 2007.

[208] W. Schubert. On the origin of cell functions encoded in the toponome. *Journal of Biotechnology*, 149(4):252–259, 2010.

[209] W. Schubert. Toponomics. In *Encyclopedia of Systems Biology*, pages 2191–2212. Springer, 2013.

[210] W. Schubert, M. Bode, R. Hillert, A. Krusche, and M. Friedenberger. Toponomics and neurotoponomics: a new way to medical systems biology. *Expert Review of Proteomics*, 5(2):361–369, 2008.

[211] W. Schubert, B. Bonnekoh, A. J. Pommer, L. Philipsen, R. Böckelmann, Y. Malykh, H. Gollnick, M. Friedenberger, M. Bode, and A. W. M. Dress. Analyzing proteome topology and function by automated multidimensional fluorescence microscopy. *Nature Biotechnology*, 24(10):1270–1278, 2006.

[212] W. Schubert, M. Friedenberger, M. Bode, A. Krusche, and R. Hillert. Functional architecture of the cell nucleus: Towards comprehensive toponome reference maps of apoptosis. *Biochimica et Biophysica Acta*, 1783(11):2080–2088, 2008.

[213] W. Schubert, A. Gieseler, A. Krusche, and R. Hillert. Toponome mapping in prostate cancer: Detection of 2000 cell surface protein clusters in a single tissue section and cell type specific annotation by using a three symbol code. *Journal of Proteome Research*, 8(6):2696–2707, 2009.

[214] W. Schubert, A. Gieseler, A. Krusche, P. Serocka, and R. Hillert. Next-generation biomarkers based on 100-parameter functional super-resolution microscopy tis. *New Biotechnology*, 29(5):599–610, 2012.

[215] T. Schultz and G. L. Kindlmann. Open-box spectral clustering: Applications to medical image analysis. *IEEE Transactions on Visualization and Computer Graphics*, 19(12):2100–2108, 2013.

[216] C. Schumann, M. Neugebauer, R. Bade, B. Preim, and H.-O. Peitgen. Implicit vessel surface reconstruction for visualization and CFD simulation. *International Journal of Computer Assisted Radiology and Surgery*, 2(5):275–286, 2008.

[217] C. Schumann, S. Oeltze, R. Bade, B. Preim, and H.-O. Peitgen. Model-free surface visualization of vascular trees. In *Eurographics Conference on Visualization (EuroVis)*, pages 283–290. Citeseer, 2007.

[218] D. M. Sforza, C. M. Putman, and J. R. Cebral. Hemodynamics of cerebral aneurysms. *Annual Review of Fluid Mechanics*, 41:91–107, 2009.

[219] H.-W. Shen, T.-Y. Lee, A. Chaudhuri, and B. Nouanesengsey. Visual analytics for enabling extreme-scale scientific discovery. In *Scientific Discovery through Advanced Computing Program (SciDac)*, 2011.

[220] J. Shi and J. Malik. Normalized cuts and image segmentation. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 22(8):888–905, 2000.

[221] K. Shi, H. Theisel, H. Hauser, T. Weinkauf, K. Matkovic, H.-C. Hege, and H.-P. Seidel. *Path line attributes-an information visualization approach to analyzing the dynamic behavior of 3D time-dependent flow fields*, chapter Topology-Based Methods in Visualization II, pages 75–88. Mathematics and Visualization. Springer, 2009.

[222] B. Shneiderman. The eyes have it: A task by data type taxonomy for information visualizations. In *IEEE Symposium on Visual Languages*, pages 336–343, 1996.

[223] M. Shojima, M. Oshima, K. Takagi, R. Torii, M. Hayakawa, K. Katada, A. Morita, and T. Kirino. Magnitude and role of wall shear stress on cerebral aneurysm computational fluid dynamic study of 20 middle cerebral artery aneurysms. *Stroke*, 35(11):2500–2505, 2004.

[224] SIEMENS SOMATOM Force Technical Specifications. `http://www.healthcare.siemens.com/computed-tomography/dual-source-ct/somatom-force/technical-specifications` [2015-11-26].

[225] R. Snowden, R. J. Snowden, P. Thompson, and T. Troscianko. *Basic Vision: An Introduction to Visual Perception*. Oxford University Press, 2012.

[226] B. Speckmann and K. Verbeek. Necklace maps. *IEEE Transactions on Visualization and Computer Graphics*, 16(6):881–889, 2010.

[227] R. Spence. *Information Visualization: An Introduction*. Springer, 3rd edition, 2014.

[228] D. A. Steinman, J. S. Milner, C. J. Norley, S. P. Lownie, and D. W. Holdsworth. Image-based computational simulation of flow dynamics in a giant intracranial aneurysm. *American Journal of Neuroradiology*, 24(4):559–566, 2003.

[229] R. J. Sternberg and K. Sternberg. *Cognitive Psychology*. Wadsworth Inc Fulfillment, 6th edition, 2011.

[230] C. D. Stolper, A. Perer, and D. Gotz. Progressive visual analytics: User-driven visual exploration of in-progress analytics. *IEEE Transactions on Visualization and Computer Graphics*, 20(12):1653–1662, 2014.

[231] W. Sturm, T. Schreck, A. Holzinger, and T. Ullrich. Discovering medical knowledge using visual analytics. In *Eurographics Workshop on Visual Computing for Biology and Medicine (VCBM)*, pages 71–81, 2015.

[232] K. R. Subramanian, J. P. Brockway, and W. B. Carruthers. Interactive detection and visualization of breast lesions from dynamic contrast enhanced MRI volumes. *Computerized Medical Imaging and Graphics*, 28(8):435–444, 2004.

[233] G.-D. Sun, Y.-C. Wu, R.-H. Liang, and S.-X. Liu. A survey of visual analytics techniques and applications: State-of-the-art research and future challenges. *Journal of Computer Science and Technology*, 28(5):852–867, 2013.

[234] T. Tanoue, S. Tateshima, J. Villablanca, F. Vinuela, and K. Tanishita. Wall shear stress distribution inside growing cerebral aneurysm. *American Journal of Neuroradiology*, 32(9):1732–1737, 2011.

[235] J. Tao, C. Wang, C.-K. Shene, and R. Shaw. A vocabulary approach to partial streamline matching and exploratory flow visualization. *IEEE Transactions on Visualization and Computer Graphics*, page to appear, 2016.

[236] S. Thew, A. Sutcliffe, R. Procter, O. De Bruijn, J. McNaught, C. C. Venters, and I. Buchan. Requirements engineering for e-Science: Experiences in epidemiology. *IEEE Software*, 26(1):80–87, 2009.

[237] J. J. Thomas and K. A. Cook, editors. *Illuminating the Path: The Research and Development Agenda for Visual Analytics*. IEEE Computer Society, 2005.

[238] K. D. Tönnies, O. Gloger, M. Rak, C. Winkler, P. Klemm, B. Preim, and H. Völzke. Image analysis in epidemiological applications. *it-Information Technology*, 57(1):22–29, 2015.

[239] T. Torsney-Weir, A. Saad, T. Möller, H.-C. Hege, B. Weber, J.-M. Verbavatz, and S. Bergner. Tuner: Principled parameter finding for image segmentation algorithms using visual response surface exploration. *IEEE Transactions on Visualization and Computer Graphics*, 17(12):1892–1901, 2011.

[240] M. Tukenova, C. Guibout, M. Hawkins, E. Quiniou, A. Mousannif, H. Pacquement, D. Winter, A. Bridier, D. Lefkopoulos, O. Oberlin, et al. Radiation therapy and late mortality from second sarcoma, carcinoma, and hematological malignancies after a solid cancer in childhood. *International Journal of Radiation Oncology* Biology* Physics*, 80(2):339–346, 2011.

[241] C. Turkay, F. Jeanquartier, A. Holzinger, and H. Hauser. On computationally-enhanced visual analysis of heterogeneous data and its application in biomedical informatics. In *Interactive Knowledge Discovery and Data Mining in Biomedical Informatics*, pages 117–140. Springer, 2014.

[242] C. Turkay, A. Lundervold, A. J. Lundervold, and H. Hauser. Hypothesis generation by interactive visual exploration of heterogeneous medical data. In *Human-Computer Interaction and Knowledge Discovery in Complex, Unstructured, Big Data*, pages 1–12. Springer, 2013.

[243] Twitter company facts. `https://about.twitter.com/company` [2015-26-11].

[244] E. Uysal, B. Yanbuloglu, M. Erturk, B. M. Kilinc, and M. Basak. Spiral CT angiography in diagnosis of cerebral aneurysms of cases with acute subarachnoid hemorrhage. *Diagnostic and Interventional Radiology*, 11(2):77–82, 2005.

[245] S. van den Elzen and J. J. van Wijk. Baobabview: Interactive construction and analysis of decision trees. In *IEEE Conference on Visual Analytics Science and Technology (VAST)*, pages 151–160, 2011.

[246] S. van den Elzen and J. J. van Wijk. Small multiples, large singles: a new approach for visual data exploration. *Computer Graphics Forum*, 32(3pt2):191–200, 2013.

[247] I. van der Schaaf, A. Algra, M. Wermer, A. Molyneux, M. J. Clarke, J. van Gijn, and G. J. Rinkel. Endovascular coiling versus neurosurgical clipping for patients with aneurysmal subarachnoid haemorrhage. *The Cochrane Database of Systematic Reviews*, 4, 2005.

[248] J. van Gijn, R. S. Kerr, and G. J. Rinkel. Subarachnoid haemorrhage. *The Lancet*, 369(9558):306–318, 2007.

[249] R. van Pelt, R. Gasteiger, K. Lawonn, M. Meuschke, and B. Preim. Comparative blood flow visualization for cerebral aneurysm treatment assessment. *Computer Graphics Forum*, 33(3):131–140, 2014.

[250] R. van Pelt, S. Jacobs, B. M. ter Haar Romeny, and A. Vilanova. Visualization of 4D blood-flow fields by spatiotemporal hierarchical clustering. *Computer Graphics Forum*, 31(3pt2):1065–1074, 2012.

[251] R. van Pelt and A. Vilanova. Understanding blood-flow dynamics: New challenges for visualization. *Computer*, 46(12):60–67, 2013.

[252] J. J. van Wijk. Evaluation: a challenge for visual analytics. *Computer*, 46(7):56–60, 2013.

[253] C. Vehlow, J. Hasenauer, A. Kramer, J. Heinrich, N. Radde, F. Allgower, and D. Weiskopf. Uncertainty-aware visual analysis of biochemical reaction networks. In *IEEE Symposium on Biological Data Visualization (BioVis)*, pages 91–98, 2012.

[254] A. Vilanova, B. Preim, R. van Pelt, R. Gasteiger, M. Neugebauer, and T. Wischgoll. Visual exploration of simulated and measured blood flow. In *Scientific Visualization*, Mathematics and Visualization, pages 305–324. Springer, 2014.

[255] M. H. Vlak, A. Algra, R. Brandenburg, and G. J. Rinkel. Prevalence of unruptured intracranial aneurysms, with emphasis on sex, age, comorbidity, country, and time period: a systematic review and meta-analysis. *The Lancet Neurology*, 10(7):626–636, 2011.

[256] H. Völzke, D. Alte, C. O. Schmidt, D. Radke, R. Lorbeer, N. Friedrich, N. Aumann, K. Lau, M. Piontek, G. Born, and et al. Cohort Profile: The Study of Health in Pomerania. *International Journal of Epidemiology*, 40(2):294–307, 2011.

[257] E. von Elm, D. G. Altman, M. Egger, S. J. Pocock, P. C. Gøtzsche, and J. P. Vandenbroucke. The Strengthening the Reporting of Observational Studies in Epidemiology (STROBE) statement: Guidelines for reporting observational studies. *The Lancet*, 370(9596):1453 – 1457, 2007.

[258] T. D. Wang, K. Wongsuphasawat, C. Plaisant, and B. Shneiderman. Visual information seeking in multiple electronic health records: Design recommendations and a process model. In *ACM International Health Informatics Symposium*, pages 46–55, 2010.

[259] Z. Wang, H.-P. Seidel, and T. Weinkauf. Multi-field pattern matching based on sparse feature sampling. *IEEE Transactions on Visualization and Computer Graphics*, 22(1):807–816, 2016.

[260] M. Q. Wang Baldonado, A. Woodruff, and A. Kuchinsky. Guidelines for using multiple views in information visualization. In *Working Conference on Advanced Visual Interfaces*, pages 110–119, 2000.

[261] C. Ware. *Information Visualization: Perception for Design*. Morgan Kaufmann, 3rd edition, 2012.

[262] G. H. Weber and H. Hauser. Interactive visual exploration and analysis. In *Scientific Visualization*, pages 161–173. Springer, 2014.

[263] V. L. West, D. Borland, and W. E. Hammond. Innovative information visualization of electronic health record data: a systematic review. *Journal of the American Medical Informatics Association*, 22(2):330–339, 2015.

[264] Wikipedia entry for "Massive parallel sequencing". `https://en.wikipedia.org/wiki/Massive_parallel_sequencing` [2015-11-26].

[265] P. C. Wong, H. Foote, G. Chin Jr, P. Mackey, and K. Perrine. Graph signatures for visual analytics. *IEEE Transactions on Visualization and Computer Graphics*, 12(6):1399–1413, 2006.

[266] P. C. Wong, H.-W. Shen, and C. Chen. Top ten interaction challenges in extreme-scale visual analytics. In J. Dill, R. Earnshaw, D. Kasik, J. Vince, and P. C. Wong, editors, *Expanding the Frontiers of Visual Analytics and Visualization*, pages 197–207. Springer, 2012.

[267] K. Wongsuphasawat, J. A. Guerra Gómez, C. Plaisant, T. D. Wang, M. Taieb-Maimon, and B. Shneiderman. LifeFlow: visualizing an overview of event sequences. In *SIGCHI Conference on Human Factors in Computing Systems*, pages 1747–1756, 2011.

[268] J. Wu, M. Wei, Y. Li, X. Ma, F. Jia, and Q. Hu. Scale-adaptive surface modeling of vascular structures. *Biomedical Engineering Online*, 9(1):75, 2010.

[269] J. Xiang, S. K. Natarajan, M. Tremmel, D. Ma, J. Mocco, L. N. Hopkins, A. Siddiqui, E. Levy, and H. Meng. Hemodynamic-morphologic discriminants for intracranial aneurysm rupture. *Stroke*, 42(1):144–152, 2011.

[270] J. Yang, S. A. Wagner, and P. Beli. Illuminating spatial and temporal organization of protein interaction networks by mass spectrometry-based proteomics. *Frontiers in Genetics*, 6(344):1–7, 2015.

[271] J. S. Yi, Y. ah Kang, J. T. Stasko, and J. A. Jacko. Toward a deeper understanding of the role of interaction in information visualization. *IEEE Transactions on Visualization and Computer Graphics*, 13(6):1224–1231, 2007.

[272] M. A. Ystad, A. J. Lundervold, E. Wehling, T. Espeseth, H. Rootwelt, L. T. Westlye, M. Andersson, S. Adolfsdottir, J. T. Geitung, A. M. Fjell, et al. Hippocampal volumes are important predictors for memory function in elderly women. *BMC Medical Imaging*, 9(1):17, 2009.

[273] H. Yu, C. Wang, C.-K. Shene, and J. H. Chen. Hierarchical streamline bundles. *IEEE Transactions on Visualization and Computer Graphics*, 18(8):1353–67, 2012.

[274] S. Zachow, P. Muigg, T. Hildebrandt, H. Doleisch, and H.-C. Hege. Visual exploration of nasal airflow. *IEEE Transactions on Visualization and Computer Graphics*, 15(6):1407–1414, 2009.