OTTO VON GUERICKE UNIVERSITÄT MAGDEBURG

EIT

FAKULTÄT FÜR
ELEKTROTECHNIK UND
INFORMATIONSTECHNIK

Quality assessment of clinical thorax CT images

Dissertation

zur Erlangung des akademischen Grades

Doktoringenieurin
(Dr.-Ing.)

von Zahra Passand, M.Sc.
geb. am 01.03.1985 in Mashhad, Iran
genehmigt durch die Fakultät für Elektrotechnik und Informationstechnik
der Otto-von-Guericke-Universität Magdeburg

Gutachter: Prof. Dr. Christoph Hoeschen (Otto-von-Guericke-Universität Magdeburg)
Prof. Dr. Magnus Båth (University of Gothenburg, Sweden)

Promotionskolloquium am 21.12.2022

Abstract

Computed tomography (CT) is one of the most important and widely used medical imaging modalities. It is used in clinical diagnostics as well as for minimally invasive therapeutic procedures. As with traditional X-ray imaging, the radiation exposure to the patient associated with CT is a disadvantage of this method. Therefore, one goal when using CT imaging is to reduce the radiation exposure or dose as much as possible. However, it should be noted that a reduction in dose is accompanied by a reduction in image quality. For this reason, careful consideration should be given to the level of image quality required for a particular diagnostic purpose and how the dose required for this purpose can be adjusted accordingly. Such optimizations are usually performed manually based on empirical values and the subjective perception of the image quality by the user. Different parameters, such as the body region and the associated absorption of the X-rays or the resolution or slice thickness of the images, play a key role in determining which dose is required for a specific image quality. Important criteria for assessing the quality of an image or a certain region of the image are the spatial resolution and the noise.

The goal of this work was to develop a method for an automated, objective, application-specific quality assessment of clinical thorax CT images. In this work, the focus was on the area of the major fissure of the left lung as it was characterized by radiologists as a good indicator of the overall diagnostic image quality. The quality of the CT images in the region covering the major fissure was assessed by radiologists. This subjective assessment was used for comparison to the objective image quality obtained from the automated assessment system developed in this thesis. The major fissure and the areas surrounding it were automatically segmented to determine various features for image quality assessment. Based on these features, a classifier was trained which assigned the annotated image data to two classes. It was shown, that a reliable classification of the clinical CT images regarding their quality was possible with classical feature engineering and a support vector machine.

Within the scope of this work, a method for an automated evaluation of image quality was developed. For this purpose, classical features characterizing the spatial resolution or the noise of an image were used in combination with a classifier. The major advantage over other methods is that this procedure can be applied to clinical CT images of patients and is not limited to the use of phantoms.

Zusammenfassung

Die Computertomographie (CT) ist eines der wichtigsten und am häufigsten eingesetzten medizinischen Bildgebungsverfahren. Es wird sowohl in der klinischen Diagnostik als auch für minimalinvasive therapeutische Eingriffe eingesetzt. Wie auch bei der klassischen Röntgenbildgebung ist die mit der CT verbundene Strahlenbelastung des Patienten ein Nachteil dieser Methode. Ein Ziel bei der Anwendung der CT-Bildgebung ist es daher, die Strahlenbelastung bzw. Strahlendosis so weit wie möglich zu reduzieren. Dabei ist jedoch zu beachten, dass mit einer Reduktion der Dosis eine Verminderung der Bildqualität einhergehen kann. Aus diesem Grund sollte sorgfältig abgewogen werden, welches Level an Bildqualität für einen bestimmten diagnostischen Zweck benötigt wird und wie die dafür nötige Dosis entsprechend angepasst werden kann. Solche Optimierungen werden in der Regel manuell auf Basis von Erfahrungswerten und der subjektiven Wahrnehmung der Anwender durchgeführt. Dabei spielen unterschiedliche Parameter wie etwa die Körperregion und die damit verbundene Absorption der Röntgenstrahlung oder auch die Auflösung und Schichtdicke der Bilder eine entscheidende Rolle. Wichtige Kriterien für die Qualität eines Bildes sind die räumliche Auflösung und das Rauschen innerhalb des Bildes bzw. einer bestimmten Region.

Das Ziel dieser Arbeit war es, eine Methode zur automatisieren, applikationsspezifischen Qualitätsbewertung von klinischen CT-Bildern des Thorax zu entwickeln. Dabei wurden zunächst ausgewählte Region innerhalb der Lunge analysiert. Die Qualität der Thorax-CT Aufnahmen wurde von Radiologen anhand unterschiedlicher Kriterien bewertet. Diese subjektive Einschätzung bildete die Grundlage für die Entwicklung eines automatisierten Bewertungssystems. Die große Herausforderung bei einem derartigen Ansatz ist die Anwendung des Verfahrens an realen Patientenbildern. Typischerweise wird die Qualität bildgebender Systeme mit Phantomen ermittelt, welche klar definierte Strukturen aufweisen. Im Rahmen der Arbeit wurde sich bei klinischen CT-Bildern auf den Bereich der Fissura obliqua des linken Lungenflügels konzentriert. In diesem Bereich wurden zunächst unterschiedliche Bereiche automatisch segmentiert, um die Features zur Bewertung der Bildqualität zu bestimmen. Basierend auf diesen Features wurde ein Klassifikator trainiert, welcher die annotierten Bilddaten zwei Klassen zugeordnet hat. Dabei konnte gezeigt werden, dass mit klassischem Feature Engineering und einer Support Vector Maschine eine zuverlässige Klassifikation der klinischen CT-Bilder bzw. deren Qualität möglich war.

Im Rahmen dieser Arbeit wurde ein Verfahren zur automatisierten Bewertung der Bildqualität entwickelt. Dafür wurden unter anderem klassische Features zur Bestimmung der räumlichen Auflösung oder der Charakteristik des Bildrauschens in Verbindung mit einem Klassifikator eingesetzt.

For my wonderful mother, for my father and for my beautiful sister Sahar.

Contents

1

Introduction

Computed tomography (CT) is an important imaging modality which is widely used for various clinical applications. It is well known for its capability of quickly producing high quality, three dimensional (3D) images of a patient. For being effective in patient care, the quality of the diagnostic CT images has to be assured under different conditions [Cheng 19]. Unlike many of the daily images which are taken with digital cameras, medical images are acquired out of medical necessity. The quality of a medical image is related to how well it transports functional information or shows the anatomical structures and probable abnormalities to the physician or radiologist such that an accurate diagnosis can be made. When acquiring diagnostic medical images, several trade-offs have to be made. The quality of an image is affected by the applied radiation dose. The goal is to optimize the quality for a certain diagnostic task for which the CT exam was requested while keeping the applied dose at a minimum level. Hence, a task-based image quality assessment should be performed, which means that the quality of the image is assessed in terms of the diagnostic application of the image.

Various factors and parameters can have an impact on the subjective and objective image quality. There are various parameters which can be adjusted in a CT scanner such as the tube voltage, the applied dose, the slice thickness or the reconstruction algorithm and kernel which all affect the image quality. While the X-ray photons are traveling from the source to the detector through the patient tissue, several physical effects and interactions occur such as scattering or beam hardening. The patient itself can also have an impact on the obtained image quality. Patient motion, especially respiratory motion and cardiac motion, can cause severe distortions in the affected regions. The size and weight of the patient can have a relevant influence on the absorption of the X-ray photons and on the scattering effects. Metallic implants with high attenuation can cause severe artifacts in the corresponding slice. If the quality of a CT image is extremely degraded, it might not be usable by the radiologist for diagnostic purposes. Consequently, such images are discarded and the scan needs to be repeated which causes an additional, unnecessary exposure to the patient.

Considering the different influences mentioned above, quality assessment during image acquisition is an important aspect for reducing the radiation exposure to the patient and to the clinical staff, maintaining a sufficient image quality as well as having an objective quality management of various imaging systems in a clinical setting. The analysis and documentation of the image quality during and after acquisition plays an important role and several advances were made in the recent years. A recently introduced legislation law in Germany is one of the indication of this importance. This law requires a continuous assessment and documentation of the medical image quality [EU 13, StrSchG 17].

The ultimate goal would be a system which enables a task-specific quality assessment of medical images. Based on a retrospective analysis of the images and the estimated image quality, one could optimize image acquisition protocols, clinical work flows or even identify systemic errors made during the acquisition process. Having such long-term information could reduce the number of images with insufficient quality, avoiding repeated scans and consequently reducing unnecessarily applied radiation dose.

Within this thesis, different methods and techniques are explored to perform an objective image quality assessment of clinical CT images. The main goal was to develop several task-specific methods which enable an automatic quality assessment of clinical patient CT images in specific parts of the lung. For this purpose, thorax CT images labeled by radiologists were used and different parameters were estimated in predefined regions of interest (ROIs). The various parameters were analyzed regarding their significance and were used to train a classifier.

Chapter 2 explains the medical background and technical foundation of the CT imaging modality and the different type of artifacts which can affect the quality of the acquired images. Chapter 3 summarizes the state of the art and the research works which have been done in recent years in this field. A detailed description of the different methods which have been developed in this thesis is provided in Chapter 4. Chapter 5 summarizes the different results obtained from the different methods. A discussion of the results can be found in Chapter 6. Chapter 7 summarizes the results and the main findings of this thesis and gives an outlook for the future research and development works.

2

Technical and medical foundations

This chapter gives an overview of the history and the basic principles of X-ray and CT imaging. The most important technical parameters which are affecting the quality of acquired CT images, typical distortions in clinical CT as well as biological effects of ionizing radiation on the body are explained in detail. Furthermore, a brief introduction about the anatomy of the region analyzed in this work is given at the end of this chapter.

2.1 X-ray and computed tomography imaging

2.1.1 A brief history of CT

X-ray imaging is the first non-invasive imaging technique which was introduced more than a century ago. The lack of depth information in the X-ray radiographical images for precise localization of the lesions has led to the invention of a tomographic imaging method: CT. The word tomography is composed of the words tomo (slice) and graph (picture). This cites to the fact that imaging of the body using CT is performed in a slice by slice way. When introduced as a clinical imaging method in the early 1970s, CT imaging was the first medical imaging technology allowing a tomographic or sectional imaging [Hsieh 09]. It combines several X-ray projections from various angles around the patient and uses different reconstruction algorithms to generate a tomographic image. The advantage of CT over conventional X-ray radiography is its ability to display the anatomy as 3D slices. In this way, the superposition of anatomical structures is eliminated, presenting the physician a detailed and unobstructed view of the anatomy.

CT images represent the attenuation coefficient distribution inside the body. The body consists of different materials with different densities and attenuation coefficients. Depending on the distribution of the attenuation coefficients, the computed CT image appears darker or brighter at different locations. Due to its remarkable properties in producing high resolution images, CT is widely used for diagnostic and therapeutic purposes in various clinical applications.

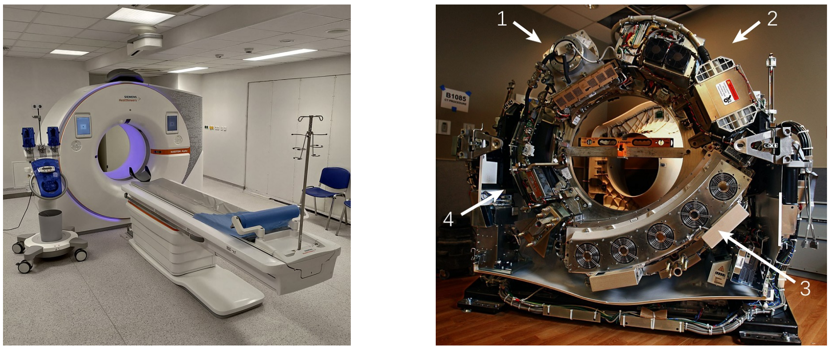
During the last century, remarkable properties in high resolution imaging with CT have been developed. These properties led to a significant increase in CT applications in terms of industrial and medical diagnosis in recent years. With the developments in technology and improvements in science, different generations of CT scanners were developed. The quality of the obtained images has been significantly improved in the different CT scanner generations. For example, due to the longer scan time in the first generation of CT scanners, the quality of the images was mostly affected by patient motion artifacts. To improve the performance in CT scanning, several operations in the acquisition and reconstruction of the data were parallelized.

One important parameter is the time required to complete a scan, either of a single slice or a stack of slices. Complete data acquisition and image production in the first CT scanners was taking several days. Some developments such as using an X-ray tube with higher intensity, modification of the interpolation method and implanting a photo-multiplier in the crystal detector did not only led to scan accuracy improvements but also reduced the time from days to hours [Hsieh 09]. With further developments, this time even has been reduced to a few minutes. Important progress considering the scan time was made by using the slip ring technology which enabled power supply and information transfer to the rotating gantry. In this way, continuous 360° rotations were possible without pausing to rewind cables. The concept of helical or spiral scanning was introduced in the late 1980s allowing to continuously translate the patient table through the gantry without having a start and stop motion. The main limiting factors in the late 1990s were the inertial forces caused by the rotating gantry and the limited amount of X-ray dose that a tube could deliver [Kalender 00]. To overcome these challenges, CT scanners with multi-row detectors were developed, utilizing more of the cone beam by multiple adjacent rows. Recent technologies use dual X-ray sources and multi-row detector banks mounted orthogonal to each other in the gantry which leads to a doubling of the scan speed. Each innovation had a significant implication on the improvement of the whole system. The dual source scanning is on the other hand, as some other developments, increasing the scatter causing image deterioration and artifacts.

The development of the early CT scanner generations and the corresponding reduction of the scanning times can be briefly summarized as follows [Kalender 00]:

- 1st Generation (1970): Single beam and single detector, stepwise translation and rotation, $24\,\mathrm{h}$ per slice
- 2nd Generation (1972): Partial fan beam with multiple detectors, stepwise translation and rotations, $300\,\mathrm{s}$ per slice
- 3rd Generation (1976): Fan beam and detector bank rotating, $5\,\mathrm{s}$ per slice
- 4th Generation (1978): Rotating fan beam and stationary detector bank, $5\,\mathrm{s}$ per slice

Figure 2.1 shows a typical modern CT scanner as well as an overview from the inside of the scanner. Today's most common CT systems are belonging to the third generation scanner type. In order to increase the resolution of the CT images, several versions of third generation CT systems were developed using techniques such as offsetting the rotation center or flying focus X-ray tubes. The functional principle of the third generation scanners is illustrated in Fig. 2.2. In these scanners, several detector cells are installed opposite of the X-ray source. The detector array and the X-ray tube rotate around the patient at the same time ("rotate-rotate" geometry). While the entire system rotates around the patient, the detector and the X-ray tube remain stationary with respect to each other. Using different reconstruction techniques, the collected projection data is used to reconstruct



(a) Typical CT scanner

(b) CT device view from inside

Figure 2.1: View of CT scanners from outside (a) and inside (b). Different parts of the typical CT device are shown in image (b): 1. X-ray tube, 2. Power supply, 3. Detectors including fans, 4. Cooling system for the X-ray tube including a radiator and a fluid pump. The whole system rotates 2 to 3 times per second around the patient. (Image sources: [Vendis 18, RSB 13])

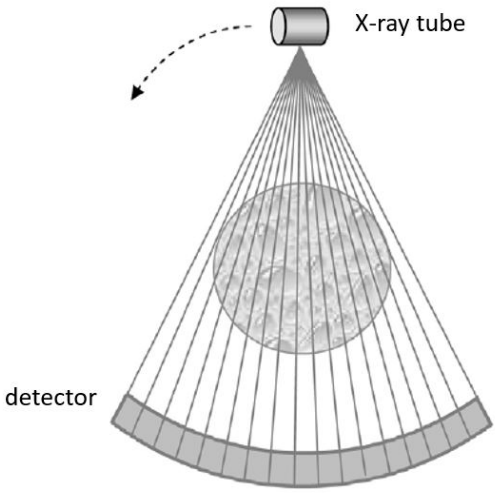


Figure 2.2: Schematic illustration of the third generation CT scan geometry. The arrow indicates the movement during data acquisition. The X-ray paths are shown as slid lines between source and detector. In this schematic, the object is stationary while the source and detector move around it. This technique is used in medical scanners. The contrary approach, i. e. a moving object and a stationary source and detector are often used in industrial scanners. The rotation is performed along a circularly trajectory. (Image source: [Hsieh 09])

the CT image, e. g. by iterative reconstruction (IR) or filtered back projection (FBP) algorithms [Hsieh 09]. These two types of algorithms are briefly explained in Section 2.1.3.

2.1.2 Technical parameters

A CT system can be described by a wide range of parameters. Some of the most important parameters of a CT system, are:

- Tube current
- Tube voltage
- Detector configuration (number and thickness of detector channels)
- Reconstruction techniques (see Section 2.1.3)
- Reconstructed slice thickness

Tube current: When the tube current (typically given in mA) or the product of scan time and tube current (mAs) is increased, image quality is improved, image noise is decreased, and patient dose increases. Tube current and patient dose are linearly related, i. e. with an increase of the current time product, the patient dose increases by a comparable percentage [Saini 04]. In modern CT scanners, the tube current is modulated automatically (automatic exposure control) for most applications. The tube current modulation algorithms automatically increase the current in parts of the body with larger attenuation coefficients whereas the current is decreased in those parts with less attenuation. For imaging of the shoulder and hip areas, which have a relatively high attenuation, higher tube currents are applied. In the thorax with its relatively low attenuation, tube currents are lowered correspondingly [Goldman 08, Raman 13a]. Automatic modulation of the tube's current is used to reduce patient dose and to minimize photon deficiency artifacts.

To realize the tube current modulation, different principles and strategies are applied. The tube current is varied depending on the global assessment of the overall patient size which is obtained from a scout scan. The current is modulated along the head-feet direction of the patient whereas the applied current depends on the attenuation of the patient at each position. In the angular modulation, the tube current is varied for different angles of the X-ray to account for different attenuation coefficients along the different projection directions of the X-ray beam. [Raman 13a, Raman 13b]

Tube voltage: The reduction of the tube voltage, typically given in kV or kVp (kilovoltage peak), can lead to a decreased dose applied to the patient during an examination since, under the same beam conditioning (e. g. applied filtration and mAs), X-ray exposure is approximately proportional to the square of the kVp [Park 17]. This, however, is dependent on the patient size (i. e. on total photon attenuation) and clinical imaging task (i. e. on the disease-specific diagnostic procedure, region or organ of interest, etc.), which strongly affects the CT image quality. For instance, at the same total radiation dose, image noise in pediatric-sized patients does not increase significantly with decreasing tube voltage. However, when compared to adult-sized patients, there is a critical increase in noise level with nonlinear dependence on the tube voltage [McCollough 12].

In vascular imaging or angiography, where iodine is injected to highlight blood vessels, a reduction of the tube voltage increases the attenuation of iodine which results in a higher contrast-to-noise ratio. However, the reduction of the tube voltage causes a nonlinear, exponential increase of image noise. Consequently, the tube current needs to be increased to compensate the noise increase. Nevertheless, it was shown in different studies that a lowered tube voltage can markedly reduce the applied doses while maintaining a certain image quality level [Feuchtner 10, Dong 12, Gnannt 12]. However, this strongly depends on the clinical task and the absorption characteristics of the region to be imaged.

Nowadays, it is widely accepted in clinical routine that protocols with a low tube voltage (e. g. 100 , 80 or $70\,\mathrm{kV}$) are especially useful in non-obese, thin patients when angiographic or vascular protocols are applied where the contrast-to-noise ratio can be increased by an increased attenuation level of the applied iodine [McCollough 12]. The parameterization of the tube voltage is usually performed by the radiologist, who is taking the different conditions into account (acquisition in the arterial or venous phase, portion of the body to be imaged, and the overall anatomy of the patient). Here, the resulting image quality is not only affected by the patient's body mass index but also by the actual distribution of the body fat.

Detector configuration: A detector is defined by its number of data channels or slices. The effective thickness of a detector channel, e. g. $0.5\,\mathrm{mm}$, determines the smallest possible slice thickness. The configuration of the slice thickness is determined based on the study to be performed. For many routine image acquisitions, extremely thin detector thicknesses are often not necessary. For these cases, where $5\,\mathrm{mm}$ images in different view planes (coronal, sagittal or axial) are sufficient for the diagnostic application, the effective thickness of the detector can be increased to $1.25\,\mathrm{mm}$ or $1.5\,\mathrm{mm}$ (instead of $0.5\,\mathrm{mm}$ or $0.75\,\mathrm{mm}$), without a relevant impact on the image quality. However, if multiplanar reformations with high-resolution 3D images or a thin section are required, a smaller effective detector thickness (e. g. $0.5\,\mathrm{mm}$ or $0.75\,\mathrm{mm}$) might be required, although this leads to an increased dose [Raman 13b].

Slice thickness: The slice thickness is determined according to the clinical examination requirements. Slice thicknesses usually range from $1\,\mathrm{mm}$ to $10\,\mathrm{mm}$. The noise in an image is inversely proportional to the slice thickness, i. e. the noise increases with decreasing slices thickness [Alshipli 17]. However, thinner slices improve the visibility of small lesions and provide more diagnostic content, although the image noise increases [Nagel 07]. Hence, slice thickness should be balanced with image noise and required diagnostic content. Several research works investigated the relation between slice thickness, image quality or noise and diagnostic content, e. g. as described in [Yao 16]. In a study using phantoms, which mimiced the mandible and its surrounding tissue, it was found that thinner slices improved the visibility of certain anatomical structures despite the higher noise [Katkar 16].

2.1.3 Reconstruction techniques

The evolution of mathematical methods developed for the reconstruction of CT images is driven by the goal to produce an accurate but efficient image reconstruction while keeping the applied dose at a minimum. The models and techniques used in CT image reconstruction try to compute the attenuation coefficients of various X-ray absorption paths that are obtained as projection data. Since its clinical introduction, huge hardware developments were made in the field of CT (see also Section 2.1.1). In addition to the increased number of detector rows, the data processing methods, i. e. the reconstruction algorithms, were significantly improved during the last decades [Geyer 15]. Various algorithms are used for CT image reconstruction. In the following, the two most widely used methods are briefly introduced:

- Filtered back projection or analytical image reconstruction
- Iterative reconstruction

Filtered back projection

The principle of back projection is that the attenuation profiles acquired at different angle are back projected across the image space. Therefore, the attenuation values, which are represented by the sinogram, are normalized by the length of the projection path, i. e. they are divided by the number of image pixels along this path. Each pixel along this path is assigned by the average value of the attenuation. When adding up the pixel values along the projection direction, the acquired attenuation values (represented by the sinogram) can be estimated. This computation is performed for every acquisition angle. The sum of all attenuation profiles back projected into the image space form the final image. This procedure is shown in Fig. 2.3. [Schofield 20]

The blurred images caused by this simple back projection approach can be improved by

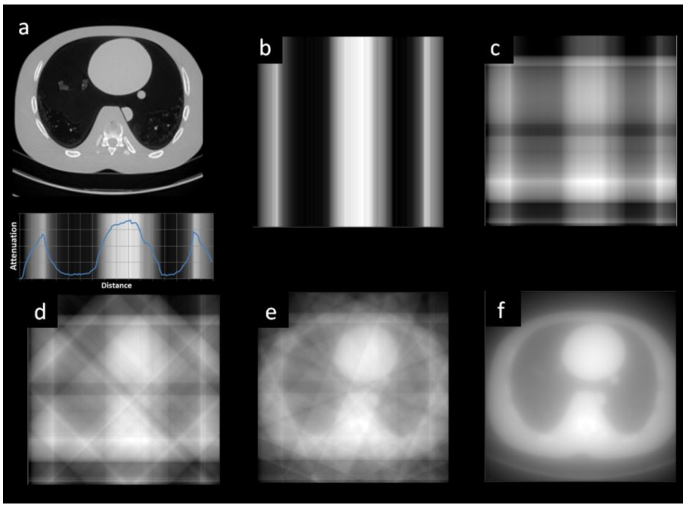


Figure 2.3: Basic principle of the back projection procedure: the slice of a chest phantom and a single attenuation profile acquired in the anterior-posterior direction are shown in (a). The back projection of this profile is shown in (b). An additional perpendicular back projection is shown in (c). The original image is becoming clearer and clearer after back projecting more and more attenuation profiles: four (d), eight (e) and 720 profiles (f). This simple, unfiltered back projection procedure results in blurred images. (Image source: [Schofield 20])

filtering the attenuation profiles before back projection. Simple back projection in the trans axial plane can be understood as a convolution of the object distribution with the point-spread function of the inherent form having a distance dependence on the contributing image points. If the object distribution is a single point (impulse), the back projection of its projection profiles appears as decreasing line density with increasing distance away from the reconstruction point. Furthermore, with sparsely sampled data, star-like artifacts appear in the back projected image. In the frequency domain, the Fourier transform coefficients of the back projection can be seen as Fourier transform coefficients of the object distribution weighted by the magnitude of the spatial frequency. This smearing effect can be reduced by high-pass filtering of the projection profiles which leads to a weighted compensation of the spatial components of the projection profiles. This improved method is referred to as FBP. Simple ramp filters or other filters can be used for this step. The filtering is computationally easy which is one of the reasons why the FBP enables the reconstruction of axial CT images in very short times. Convolution-based filtering techniques are used to filter the projection data in FBP. The resulting image noise and the texture of the image are strongly dependent on the filtering kernels. While a smooth kernel suppresses noise but reduces the spatial resolution, a sharp kernel enhances small structures or sharp edges but increases the noise in the reconstructed image. Besides these, FBP images are also affected by beam hardening and photon starvation effects, since FBP assumes that the effective photon energy of the X-rays are constant along their path through the body. FBP makes several other assumptions which do not match real X-ray physics: it assumes that all X-ray photons have the same energy, that they travel in straight lines, that there is an exponential attenuation inside the body (Beer's law) and that the X-ray source is a point source. [Schofield 20]

Iterative reconstruction

An iterative algorithm is characterized by its stepwise, repetitive calculation. Figure 2.4 shows the principle of the iterative reconstruction algorithm. The result of a previous step is used for the current calculation step, e. g. as an update or correction term. This iterative cycle is stopped if e. g. a convergence criterion is met (when the difference between the results of two subsequent iterative computation steps is sufficiently small) or after a fixed number of iterations. Iterative image reconstructions aim at calculating the image data which truly corresponds to the acquired projection data. In an ideal iterative reconstruction process, multiple forward and back projections are applied with transitions from the raw projection data to the computed image space and vice versa. The goal is to iteratively improve the reconstructed image data. Models of the CT imaging system

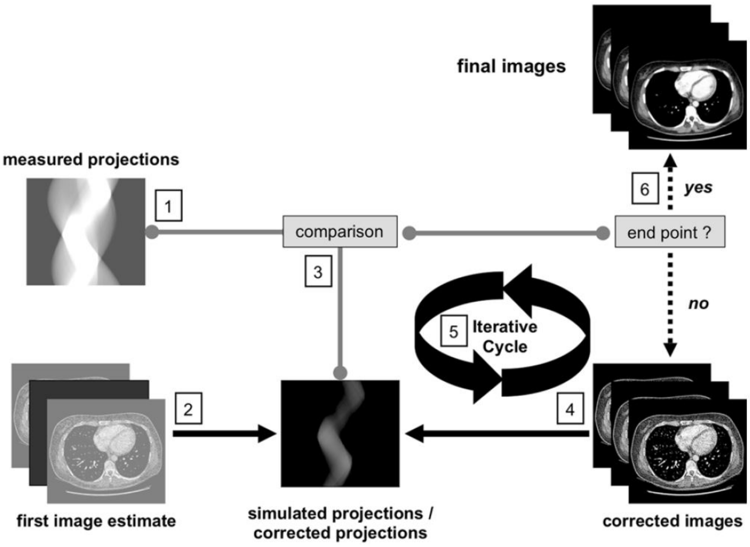


Figure 2.4: The principle steps of iterative image reconstruction algorithms (schematic). Based on the acquired projection (1), an initial image is estimated, e. g. by using FBP reconstruction. Using a simulation of the CT system, synthetic projection data is generated from the image estimate (2). The simulated projections are compared to the measured projections (3). A correction to the image computed in the previous step is applied based on the difference between the two projections (4). This process is repeated (5) until a predefined threshold is reached (6). (Image source: [Geyer 15])

or process can be used for the forward projection step in order to generate synthetic projection data. This synthetic projection data is then compared to the actual acquired projection data. In subsequent back projection step, e. g. using FBP, a correction to the image space is applied which is computed from the difference between the real and the simulated projection data. The steps of an ideal IR process can be summarized by the following [Beister 12, Skornitzke 18, Stiller 18]:

1. Compute an initial estimate of the image, e. g. using an FBP based reconstruction.
2. Simulate synthesized projections by forward projection, i. e. by a switching from image space to projection space.
3. Compare the measured projections to the synthesized projections and calculate a correction term from the difference between these projections.
4. Update the current estimate of the image by a back projecting the correction term, i. e. by a transitioning from the projection to the image space.
5. Repeat this procedure until a predefined stopping criterion is reached, e. g. defined by a threshold.

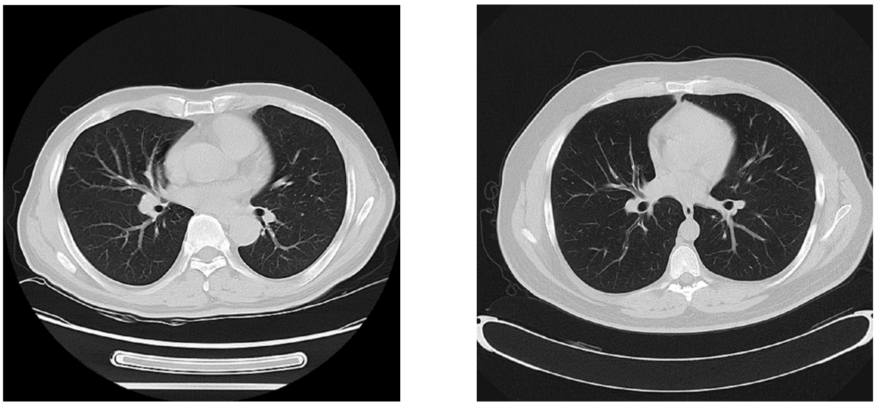
Summary

The main advantage of FBP is its robustness paired with a high image reconstruction speed. The properties of an FBP-reconstructed image can be adjusted by the user, e. g. by the applied reconstruction kernels resulting in images with familiar texture and noise structure. Due to the linear properties of the FBP method, typical metrics used for image quality assessment such as the modulation transfer function (MTF), spatial resolution, the image noise (which can be estimated by the standard deviation inside a given image region), the contrast-to-noise ratio (CNR), the signal-to-noise ratio (SNR), and noise power spectrum (NPS), are clearly defined and are valid for all FBP-reconstructed images. The disadvantages of the FBP technique are the poor contrast, the high noise, and possible artifacts, e. g. due to the starvation of photons. Hence, the dose reduction potential in the FBP method is very limited. [Stiller 18]

The advantages of the IR methods are the reduction of artifacts and image noise, e g. streaking artifacts, while the attenuation values are preserved. This is achieved because IR methods include the physics of the imaging system during reconstruction. The SNR and CNR ratios are increased and image features with low contrast are enhanced. In addition, model based IR can also slightly improve the spatial resolution. Hence, IR based methods can either be used to increase the quality of an image at a constant dose level (compared to an image reconstructed by FBP) or it can be used as a noise reduction technique in low dose images. The improvements in image quality which are important for the usage in clinical routine are generally more pronounced in images reconstructed by a model-based rather than a statistical IR method [Stiller 18]. The potential disadvantages of IR methods are the lower image reconstruction speed when compared to that of FBP. The processing time depends on the available hardware (e. g. using graphics processing units) and the algorithm's implementation. With an increasing complexity of the CT system model, the processing time also increases. However, since the computational power increases, this is not a limitation for IR methods in clinical practice anymore. IR based methods can be subject to over-smoothing which can degrade the representation of very fine details up to a complete loss. Figure 2.5 shows two CT images reconstructed with an iterative and a filtered back projection technique.

2.2 Artifacts and distortions in CT images

During the acquisition of a CT image, various types of artifacts can occur. These artifacts can cause serious degradation in the quality of the achieved images. Recognition of the artifacts is important to not be confused with tissue properties. The knowledge about different artifacts can help to recognize and prevent them, i. e. by properly preparing the patient and adjustment of imaging parameters. Artifacts can caused by physical effects, by



(a) Iterative reconstruction **(b)** Filtered back projection

Figure 2.5: Comparison of two thorax CT images acquired with different reconstruction algorithms: (a) iterative reconstruction (GE Revolution HD, Lung kernel (ASiR-V), $5\,\mathrm{mm}$, $120\,\mathrm{kV}$) and (b) filtered back projection (Siemens SOMATOM Sensation 64, B70f kernel, $5\,\mathrm{mm}$, $120\,\mathrm{kV}$). (Image sources: [Ptrump16 22])

technological system parameters of the CT system, by the parameter configuration applied by the user or by the patient itself due to weight and tissue composition, implants or motion. The careful positioning of the patient and selection of the CT scanner parameters can have a strong influence on the presence of the artifacts in the output image.

For optimizing the quality of images, it is useful to understand what causes artifacts in a CT image and how they can be suppressed or prevented. In general, artifacts in CT images can be divided into three groups:

- Physics based artifacts
- Patient based artifacts
- Scanner based artifacts

Selected physics and patient based artifacts are described in the following sections. To a certain extent, these artifacts can be controlled by the operator of the CT system. Details about the scanner based parameters which mainly depend on the scanner hardware are described in [Barrett 04]. The most important artifacts cause a relevant degradation of the image quality in clinical practice are pathological changes like fibrosis, motion artifacts and scattering.

2.2.1 Physics based artifacts

Scatter artifacts

When radiation or a photon propagates trough a material it is partly scattered and consequently diffracted from the original path [Schulze 11]. The diffraction occurs with random angles. It depends on the object how the scattered photons are distributed on the detector. Scattering is caused by interactions of the photons with the material, e. g. by Compton scattering. When the scattered photons hit the detector in an area which is otherwise hit by only a few photons, the error is highest. The contrast between highly attenuating structures and other parts surrounding it is reduced due to the scattering. Scatter artifacts can be reduced by reference measurements. Collimators which are placed in front of the detector can be used to ensure that only those photons traveling in a straight path hit the detector [Hsieh 09]. Additional methods for scatter reduction are described in [Schörner 12].

Beam hardening artifacts

The photons of an X-ray beam cover a broad spectrum of energies or frequencies. The beam becomes "harder" when passing through an object, which means that its mean energy increases since the photons with lower energies are absorbed earlier (see Fig. 2.6). Beam hardening can cause streaks or dark bands between highly attenuating objects or it can lead to cupping artifacts in the image. [Barrett 04]

Figure 2.7 shows the basic principle of cupping artifacts. The X-rays passing through the center of the cylinder shaped phantom shown in Fig. 2.7a are hardened more than the X-rays passing the outer parts of the phantom. This is because the center beams have to pass trough more material. The attenuation rate of the hardened beam decreases. Hence, the beam has a higher intensity when it hits the detectors. As a consequence, the acquired attenuation profile is different from the profile which is not affected by beam hardening.

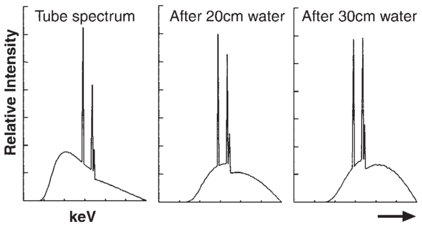
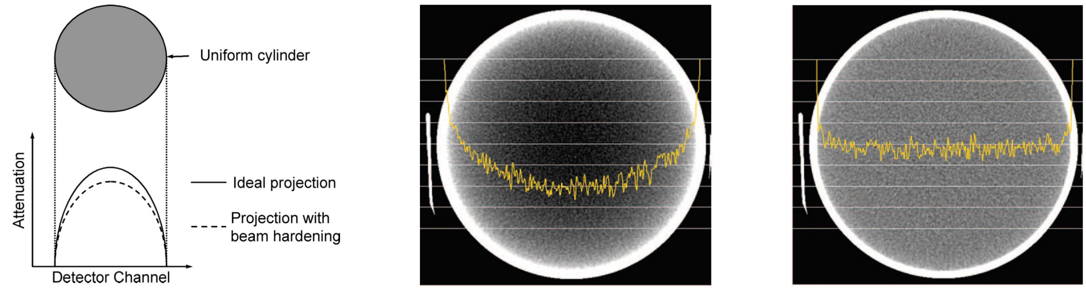


Figure 2.6: Energy spectrum of an X-ray beam passing through water. With increasing depth, the mean energy of the X-rays is increased. For having equivalent amplitudes with unattenuated spectra, the spectra were rescaled. (Image source: [Barrett 04])

Figure 2.7b shows the resulting CT image together with a profile of the CT numbers which has the typical cupped shape. Figure 2.8 shows an example for streaking artifacts inside a thorax CT image, which were caused by a contrast medium. Streaking artifacts can occur in heterogeneous cross sections between two or more relatively dense, highly attenuating objects. The artifacts occur when the beam either passes just through one or through both objects at different tube positions. [Barrett 04]

Partial volume artifacts

Partial volume effects occurring during image acquisition can also cause image artifacts. This artifact can occur when a dense, off-center object partly protrudes the width of the X-ray beam. Figure 2.9 illustrates the effect with a highly exaggerated X-ray beam divergence. In this example, the dense material inside the subject is in the pathway of the X-rays coming from the left side, but it is outside of the beam when coming from the right side (Fig. 2.9a). Shading artifacts are caused by this inconsistency between the



(a) Attenuation profile with and without beam hardening

(b) CT number profiles without calibration

(c) CT number profiles with calibration

Figure 2.7: Cupping artifacts caused by beam hardening. Due to beam hardening, the attenuation profiles differ from an ideal profile which is not affected by beam hardening (a). Profiles of the CT numbers without (b) calibration have characteristic cupped shape. This can be corrected by an additional calibration step (c). (Image sources: [Barrett 04])

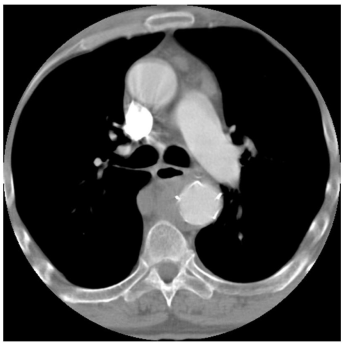
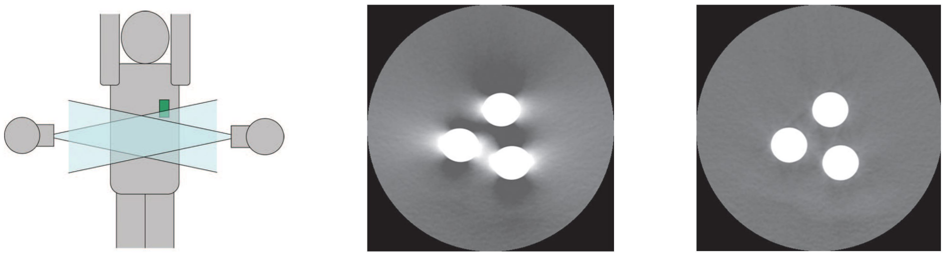


Figure 2.8: Streaking artifacts in a thorax CT image caused by the beam hardening effects as a result of injected contrast medium. (Image source: [Barrett 04])



(a) Principle: dense off-center object attenuate only parts of the X-ray beam.

(b) Phantom rods partially intruded into the scanning section results in artifacts.

(c) Phantom rods fully covered by the scanning region showing no artifacts.

Figure 2.9: Partial volume artifacts can be caused by dense objects which protrude the pathway of an X-ray beam originating from a certain direction (a). Due to the resulting inconsistency in the acquired data, artifacts occur in the CT image which was acquired using a phantom containing three rods (b). The artifacts disappear when the rods are fully intruded into the imaging plane (c). (Image sources: [Barrett 04])

different projection angles or views. The artifacts appear in the reconstructed image as shown in Fig. 2.9b. By using thinner slices, this artifact can be avoided in most of the cases (Fig. 2.9c). This is especially necessary when imaging is performed in parts of the body where the tissue density is rapidly changing along the head-feet direction. If thicker image slices are required, e. g. for noise reduction, several thinner slices can be averaged in a post-processing step. [Barrett 04]

Photon starvation

Photon starvation can occur in areas with highly attenuating structures, e. g. at the shoulders (see Fig. 2.10). In this example, the X-ray beam has to travel a long distance between the left and right shoulder which leads to a large attenuation of the radiation. Hence, an insufficient amount of photons reaches the detector at these angulations. During the reconstruction, this results in horizontal streaks. The problem of photon starvation can be compensated by increasing the tube current for such a scan. The disadvantage of such an approach would be an unnecessarily high patient dose for angulations where the beam passes through less attenuating parts. Hence, different techniques were developed to minimize the effect of photon starvation. Common techniques are tube current modulation or adaptive filtration. [Barrett 04]

Undersampling

CT image quality is also affected by the the number of projections which are used to reconstruct a CT image. In the case of undersampling, i. e. when the interval between projections is too large, a misregistration of information, especially at small objects and sharp edges, can occur. This effect which is known as view aliasing can lead to fine

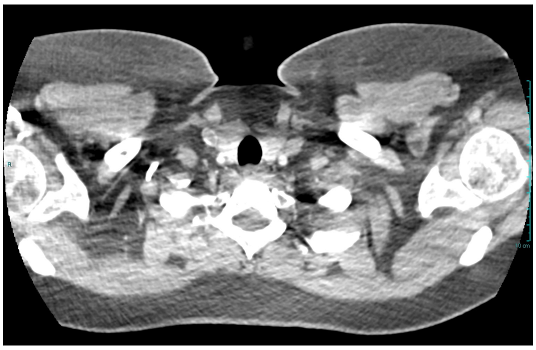


Figure 2.10: Photon starvation artifacts in the center of the image caused by the high attenuation of the clavicles and shoulders. (Image source: [Chaudhary 16])

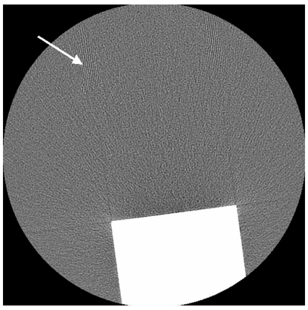


Figure 2.11: Undersampling the edge of the block caused aliasing (arrow). (Image source: [Barrett 04])

stripes from a sharp edge of a dense structure but at a certain distance from the object (see Fig. 2.11). Since the evenly spaced lines caused by the undersampling normally do not mimic anatomic structures, aliasing does not have a serious negative effect on the diagnostic image quality in most cases. However, when one requires a high resolution for analyzing structures of very fine details, aliasing artifacts should be avoided if possible. To minimize aliasing artifacts, a larger number of projections per rotation should be acquired. [Barrett 04]

2.2.2 Patient based artifacts

Patient motion

Patient motion during image data acquisition can cause streaking or shading artifacts in the reconstructed image (see Fig. 2.12). This can be caused by the movement of the organs such as the heart or the lung or by a moving patient. Hence, patient motion should be avoided whenever possible. Patient motion can be reduced in many cases by positioning aids. For some patients, a sedation might be required to minimize the motion. The minimization of scan times in regions which are prone to movements also helps to reduce motion artifacts. If patients hold their breath during a scan, this can reduce respiratory motion. Certain motion artifacts can be also reduced with advanced scanning techniques, e. g. over-scan or under-scan modes, cardiac gating, and software correction [Barrett 04]. For example, in cardiac gating, the acquisition of the CT image data is synchronized with the rhythmic motion of the heart. With this method, data is acquired in phases of the cardiac cycle which have the least cardiac motion.

Incomplete projections

If parts of the patient are outside the scanner's field or view, projections are incomplete for these portions or scan angles leading to streaking or shading artifacts. An example is shown in Fig. 2.13 where the arms of the patient were outside the field of view. During a scan of the thorax, the patient's arms should be raised in order to avoid such artifacts. Such effects can also caused by other objects like tubes filled with contrast agent lying outside the scan range which also absorb a significant amount of radiation at certain scan angles. Hence, to avoid artifacts caused by incomplete projections, no parts of the patient or other objects should lie outside of the scan field [Barrett 04].

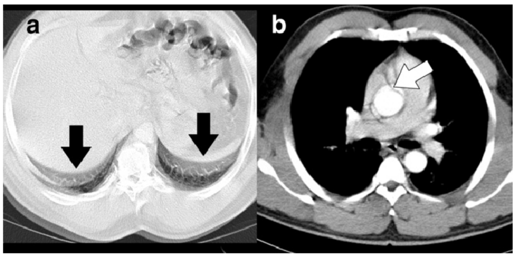
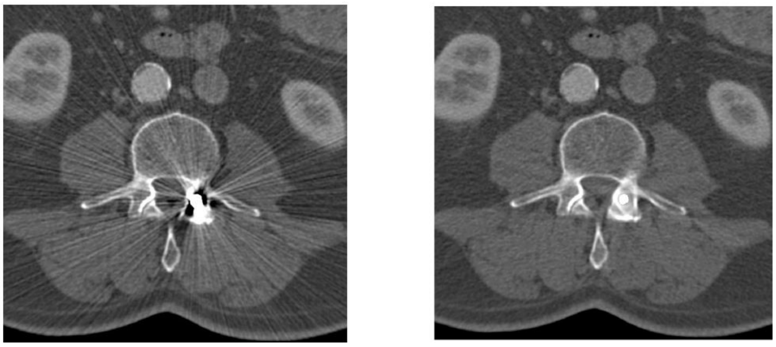


Figure 2.12: Example of motion artifacts in CT. (a) On each side, two diaphragms are visible caused by respiratory motion (arrows). (b) The pulsatile movements of the aorta and the heart simulate a dissection membrane in the ascending aorta (arrow). (Image source: [Christe 10])

Figure 2.13: Streaking artifacts in a CT caused by incomplete projections. During image acquisition, the patient's arms were outside the scanner's field of view. (Image source: [Barrett 04])



(a) Metal artifacts, no reduction. (b) Metal artifacts, with reduction.

Figure 2.14: Metallic spine implants causing artifacts in the CT image. (a) Image reconstructed without artifact correction and (b) with reduction of the metal artifact. (Image sources: [Barrett 04])

Metal artifacts

Metal objects can cause severe streaking artifacts when being present in the scan field. Due to their high absoprtion (because of the high density), metal objects can cause incomplete attenuation profiles. Beam hardening, partial volume effects, and aliasing can cause additional artifacts which are likely to occur when very dense objects are being scanned.

Any removable item such as jewelry need to be taken off before the scan. If the items are not removable (e. g. surgical clips, dental fillings or prosthetic devices), the angulation of the gantry can be modified to exclude these items from the scans. When it is not possible to avoid metal objects within a scan, higher tube voltages can be used to penetrate some objects. Using thinner scanning sections can also help to reduce metal artifacts due to the partial volume artifacts. Figure 2.14 shows an exemplary metal artifact in a CT image caused by a spine implant. This example also shows the general possibilities of software correction methods which can be employed to reduce those artifacts. [Barrett 04]

2.3 Biological effects of ionizing radiation

Radiation can have harmful effects on humans. When X-rays penetrate an object, their energy is partly transferred to the object. This energy transfer can cause changes in the object's material. Cell or tissue damage through ionizing radiation occurs either directly when the radiation hits a target or indirectly by free radicals which are generated by the radiation (see Fig. 2.15).

The direct effect mechanism causes damages in atoms or molecules of the cells, e. g. of the deoxyribonucleic acid (DNA) by breaking or striking molecular bonds, of the ribonucleic

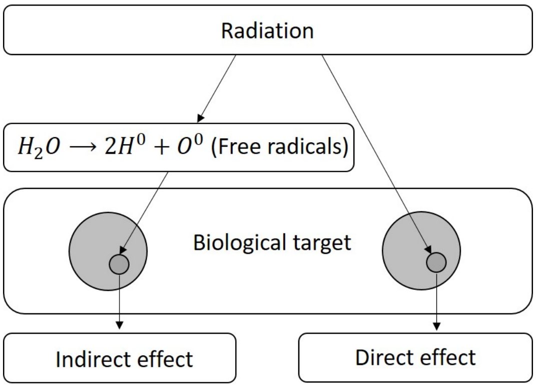


Figure 2.15: Effects of ionizing radiation on biological tissue. The mechanisms can be divided into direct and indirect effects. The indirect mechanism affects the tissues through the generation of free radicals which cause a damage of the tissue.

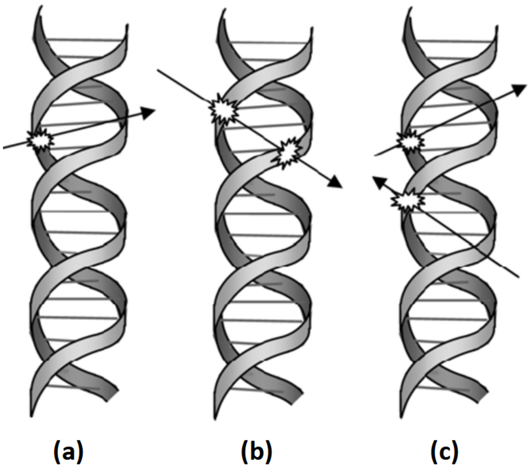


Figure 2.16: Illustration of radiation damage to DNA showing how a single event causes (a) a single-strand break or (b) a double-strand break and how two independent events cause (c) a double-strand break. (Image source: [Hsieh 09])

acid (RNA) and of enzymatic and structural proteins, by direct hits of the ionizing radiation, whereas all atoms and molecules are considered as vulnerable. In DNA, the radiation causes single- or double-stranded chromosomal breaks (see Fig. 2.16).

The indirect radiation effects are caused by the radiolysis of cellular water which results in ion pairs or free radicals. The free radicals generated by the ionizing radiation interact with the cell's atoms and molecules, especially in the DNA, resulting in chemical modifications. Due to the interaction of X-rays and water, the free radicals hydrogen and hydroxy are formed. The formation of additional free radicals is allowed by an excess of oxygen during the radiation of cells. This can lead to the formation of hydroxyperoxy free radicals and the highly unstable hydrogen peroxide. [Elgazzar 06]

Low linear energy transfer radiation causes about 70 % of biological damage due to indirect effects [Prasad 95]. High linear energy transfer is primarily responsible for the direct ionization effects. Delayed effects of radiation have been observed in cells which were not directly irradiated but which were in a spatial or temporal distance from the initial radiation. These observations can be described by the bystander effect and by the genomic instability [Elgazzar 06].

The bystander mechanism can be explained by a cell-to-cell interaction. Directly irradiated cells communicate with adjacent cells [Elgazzar 06]. In this way, they spread the radiation effects to the surroundings cells. Although the mechanism behind this effect is not fully understood, it was proposed that the release of soluble factors, e. g. of cytokines [Ramesh 96] or gap junctional intercellular communication [Azzam 01] can cause the radiation spread. The bystander effect is mainly associated with densely ionizing radiation (e. g. alpha particles) [Iyer 00], but can also be observed with low linear energy transfer radiation such as X-rays or gamma rays.

The maximal genetic damage caused by radiation and resulting in genomic instability occurs shortly (minutes to hours) after the exposure. Delayed effects were observed in the irradiated cells as well as in the descendants. In cells which sustained a nonlethal damage of the DNA, several generations after the exposure increased mutation rates were observed in descendant cells [Morgan 03]. The delayed reproductive death was observed up to six generations after the primary radiation exposure [Suzuki 03, Elgazzar 06].

Depending on their temporal occurrence, one can differentiate between early and delayed radiation effects. Acute radiation sickness, whole-body exposure syndromes, central nervous system syndrome, gastrointestinal and hematopoietic (bone marrow) syndrome, or acute regional effects are considered as early effects [Elgazzar 06]. Delayed radiation effects include genetic effects, cancer, effects on unborn children and other delayed somatic effects.

CT imaging and its associated X-ray dose has received increasing attention in the last decades which is not surprising when considering the large number of CT scans performed every year. Different studies led to the conclusion that even the low dose exposure caused by the state of the art CT devices can lead to the serious side effects such as cancer development [Council 06, Smith-Bindman 12]. Therefore, significant efforts were made to reduce the applied dose during CT exams while maintaining the quality of the acquired images in acceptable levels [Chun 15]. These efforts are mainly driven by government agencies, by the clinical and research community as well as the industry. The clinical community has agreed to adopt to the ALARA (as low as reasonably achievable) principle. The ALARA principle assumes that every radiation dose applied to a patient can have a harmful effect that may increase the risk of cancer or genetic mutations. Nowadays it is a regulatory requirement to reduce the X-ray radiation to the lowest reasonably possible but diagnostic useful level [Hsieh 09].

However, there is a close relation between the level of the applied X-ray dose and the noise in the CT image. When other parameters are kept constant and for a given reconstruction algorithm, the noise in a CT image is approximately inversely proportional to the square root of the tube current. This means that when the current of the X-ray tube (and as a consequence the dose) is reduced, the noise in the image increases. When applying the ALARA principle, the aim is to reduce the radiation dose without creating a negative impact on the image quality or its diagnostic accuracy. [Hsieh 09]

2.4 Anatomy of the region of interest

CT image quality analysis can be performed either over the complete image or in a certain part of it, i. e. inside an ROI. This work focused on the analysis of thorax CT images, more specifically on the analysis of the major fissure of the left lung. Special diseases such as fibrosis can be diagnosed by investigation of this region. Other areas in the lung relevant for diagnostic applications are the pulmonary veins, the apical bronchus of the right upper and right lower lobe (B1 and B6) and the right apical pulmonary artery of the right lower lobe (A6) [Jardin 86]. These latter structures can be recognized with subjective image quality assessment as well but are less dependent on resolution properties and differ quite a lot between patients, so that they are not as suitable for objective image quality assessment. Thus, concentration in this work is on the major fissure of the left lung trying to prove that classification of image quality works based on this structure alone.

Major fissure of the left lung

Fissures are the anatomical structures which divide the lungs into separate lobes. While special diseases can be distinguished based on their location on the lower or upper lobes, knowing the exact place of these areas can be really important for the correct diagnostic of the disease. A human lung has three fissures: one on the left side and two on the right side of the lung. The shape and location of the heart has led to the reduced volume of the left lung. Therefore, the left side has normally only two lobes (upper and lower) [Doel 15]. As it can be seen in the Fig. 2.17, on the left side of the lung, one can find a single fissure known as oblique fissure or major fissure. It divides the left lung into an upper and a lower lobe. The two fissures of the right side of the lung are the oblique and the horizontal fissure. These fissures segment the right lung into upper (superior), lower (inferior) and middle lobes. The superior lobe and the middle are separated by the horizontal fissure. In radiological imaging, the fissures are often used as landmarks in the processes of locating and identifying abnormal structures like lesions. The orientation of these fissures changes from superior to inferior. There are some factors which make finding the fissure in the images challenging. The shape of the fissure varies from one patient to another. Sometimes there are incomplete fissures in a specific person. Special diseases can also affect the position, shape and even the existence of the fissures. Additionally, noise artifacts can challenge the clear visibility of them. The quality of the CT image has an important effect on distinguishing the fissure from other tissues. Different resolutions of the image leads to different appearance of the fissure. Figure 2.18 shows a thorax CT image highlighting a part of the major fissure of the left lung.

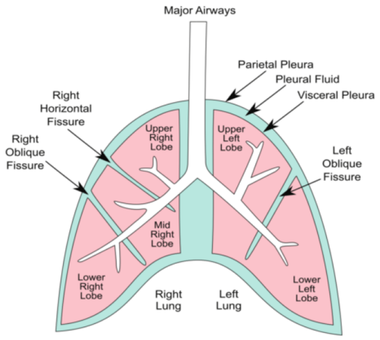


Figure 2.17: Schematic of the lung showing the main airways, the major fissures and the lobes. The fissures are formed by visceral pleura, which together with the parietal pleura forms a double layer surrounding the lungs. (Image source: [Doel 15])

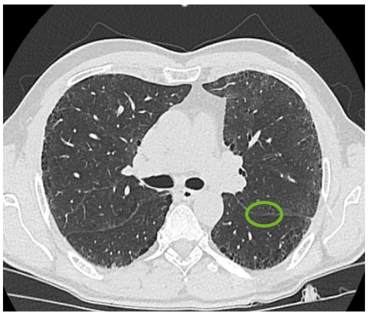


Figure 2.18: Thorax CT image. The marked area shows a part of the major fissure of the left lung.

3

Problem formulation and state of the art

This chapter gives a general overview of several studies which have been performed on image quality assessment (IQA) in CT images. In the recent years, different methods and algorithms have been developed and introduced for this purpose.

3.1 Clinical need for automated image quality analysis

The reduction of ionizing radiation is a crucial aspect in the application of CT or X-ray imaging techniques. However, there is a trade-off between the applied dose and the image quality. Generally speaking, an increased dose leads to an increased image quality. Due to the severe biological effects (see Chapter 2), the dose cannot be increased to an arbitrary level. The applied radiation dose is of special importance in thorax CT imaging because of the exposure of the lung, the breasts and the thyroid which are all radiosensitive organs [Revel 15]. The goal in CT imaging is to minimize the applied dose while maintaining the perceived quality of the image at an acceptable level. The perceived image quality usually depends on the perspective of the radiologist or technician operating the CT system. In addition to the dose, other parameters affect the quality in a CT image (see Chapter 2).

It became clinical practice to track and compare the applied radiation dose across different institutions, vendors and scanner models. Since 2015 it is necessary to record the radiation dose for all CT studies [Commission 14]. While monitoring and recording the radiation dose is an important tool for the identification of overdoses or the evaluation of a consistent application, it does not provide information about the achieved image quality [Christianson 12]. Tracking the radiation dose alone is not sufficient to determine if a certain CT protocol is optimized. This is due to the fact that image quality is affected by various physical, patient-related and technical parameters, e. g. the reconstruction kernel or reconstruction technique [Goldman 07a]. For instance, iterative reconstruction will result in less dependency of the quality of the acquired images on the dose level. Hence, to optimize a CT protocol and monitor its performance in clinical practice, an automated assessment of the obtained CT image quality would be ideal.

IQA can be divided into subjective and objective evaluation. The subjective quality evaluation can be considered as the most important and most reliable method since usually the images are analyzed and evaluated by the human observers, i. e. the radiologists who are the final users of them [Mohammadi 14, Wang 21].

In subjective quality assessment, the quality of an image is measured by different subjects, ideally professionals of their respective field, who rate the quality of images under clearly defined conditions. The ratings of different observers might be different and are summarized by a weighted average of the individual's results. The subjective evaluation is usually considered as the ground truth or gold standard, but it is slow and time consuming in practical applications [Wang 11].

The standard method for physical or objective IQA of CT imaging systems is to use dedicated phantoms. However, this does not necessarily reflect the quality of a diagnostic patient CT image. Current methods for IQA in clinical patient CT images require a lot of manual work which makes them highly time consuming with regards to either the manual, subjective image evaluation or the manual selection of ROIs for an algorithm based evaluation. These ROIs are typically used to compute noise, contrast or other parameters. The main goal of an automated, algorithm based IQA should be to assess the quality of patient CT images during clinical routine. In contrary to a phantom based IQA, the quality of patient CT images depends on various technical but also patient-related parameters. Measuring the image quality in clinical CT images is a complex task due to the variable anatomical structures and larger inter- and intra-patient variability of tissue distributions. The requirements regarding the image quality differ with respect to the anatomical region which shall be analyzed. Various parameters such as the size of the anatomical structures, material composition of the tissue, surrounding structures, motion, and others affect the resulting quality. Hence, it is difficult to define a general definition or threshold while analyzing and evaluating the image quality.

Various reasons exist to develop and establish an automated IQA of clinical patient CT images. The IQA can be implemented to control the quality across different vendors and model platforms, different operators, shifts, and hospitals or institutions. Once IQA has been established for a longer period of time, the information obtained from the acquired data could be used to prospectively optimize acquisition protocols where the goal is to achieve a certain diagnostic accuracy while keeping the dose at a minimum. To achieve this goal, both, image quality and the applied dose, need to be monitored and analyzed. Ideally, future CT exams could be personalized by configuring the scanning parameters if one would be able to predict the achievable image quality for a certain patient.

The following section gives an overview of different IQA-related studies with a focus on medical images and CT images.

3.2 CT image quality assessment - state of the art

In recent years, many different approaches have been developed to cope with the increasing importance and requirements of image quality documentation during clinical routine. However, most of these studies were performed and evaluated on phantom images and not on clinical patient data. Among different features used for image quality, calculation of the NPS is a challenging process. In the following, different works aiming for an automated IQA are briefly summarized.

Båth describes the challenges with deriving a subjective metric for IQA [Båth 10]. It was shown that certain SNR measurements have a limited agreement with the human observers. For example, the detection of an anatomical structure or of a pathology may not be mainly hampered by the noise in the image but by the properties of the anatomical background surrounding the structure or area of interest. This also means that metrics such as the detective quantum efficiency have a small influence on the image quality since the anatomical background is not taken into account. It was also stated that at the time of the publication, there was no model observer which had correlations with the performance and properties of a human observer, especially when considering the complexity of medical images with their different background structures. Although some parameters such as the SNR might have very little in common with a human observer, it was suggested to use such quantitative measures for constancy control. Due to the inherent fluctuations of the ratings performed by human observers, it was concluded that human observer studies generally have a low reliability. Consequently, a large number of observations are required to achieve reliable results. However, since humans are finally using the clinical image, human observers are required for the fundamental rating of images, e. g. the ones used in this work.

In [Båth 07], a general description of performing IQA using a non-parametric, rank-invariant method was described. A visual grading analysis of the images was used as a basis to grade the visibility of important anatomical structures of the images on an ordinal scale. Whereas such ordinal scales are often labeled using numbers, these numbers do not represent the numbers of an interval scale. In such case, a mean value of the ordinal scale values would not have a true meaning. For this purpose, an receiver operating characteristics analysis was suggested.

Objective IQA can be automatized using mathematical models and algorithms which fit the human evaluation. This can save human resources and can be applied at a larger scale to get new insights from existing data. IQA can be classified in three types of approaches: full reference, reduced reference and no reference [Wang 04, Dost 22]. The full and reduced reference methods are often used for multimedia applications where image or video data is compressed and broadcasted.

Full reference IQA: The reference image is compared to a received or processed image to evaluate the visual quality. This is difficult in many practical applications since the original image is often not available at the receiving side. Different metrics for IQA were developed and are in practical use for the full reference method, e. g. feature similarity index [Zhang 11], structural similarity image index [Wang 04], multi-scale structural similarity index [Li 10], and others [Dost 22].

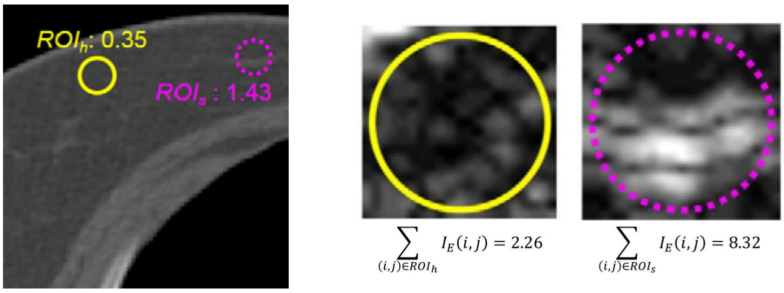
Reduced reference IQA: These methods do not require an original image as a reference. For IQA, information extracted from the images such as transformation coefficients or other relevant features of the original images are used. This approach is practically used in real-time scenarios including modified versions of the structural similarity image index and multi-scale structural similarity index, spatial efficient entropic differencing [Bampis 17] and reduced reference entropic differencing [Soundararajan 11].

No reference IQA: This methods are also referred to as “blind” IQA as there is no reference image available and IQA is solely performed using features extracted from the image to be analyzed. This is a challenging approach because the information contained in the different features can be limited. Different methods exist for the no reference IQA, e. g. blind/referenceless image spatial quality evaluator [Chow 17], distortion identification-based image verity and integrity evaluation [Zhang 14], natural image quality evaluator [Goodall 14], and many others [Dost 22].

In many practical applications such as in clinical CT imaging, a (reduced) reference image is not available. Therefore, a no reference IQA is required for the analysis and evaluation of CT images during clinical routine. There are many dedicated features which are often used for assessing the quality of phantom CT images. In the following, several works using these features are summarized. The goal was to investigate if these feature can be used for an automated IQA of clinical CT images.

Assessment of noise and noise power spectrum

An algorithm for the automatic measurement of noise in patient CT images of the chest and abdomen has been developed by [Chun 15]. To determine a homogeneous ROI, a structure coherence feature was devised. This feature represents the likelihood that a given pixel belongs to an anatomical structure. Structure coherence was represented by two different features: edginess and randomness. Edginess provides the probability of the pixels being located on a structure. The randomness of the pixel orientation was used to determine the presence or absence of anatomical structures. Both features allow a distinction between structural and homogeneous tissue regions to a certain degree. Figure 3.1 shows an example for the structure coherence levels and the edginess in the fat tissue region of a patient CT image. However, the structure coherence feature combing the two features improves the discriminability of the two different types of regions. To estimate the noise level, a three step process was implemented. Firstly, a fat tissue segmentation was done using Gaussian filtering, fat intensity thresholding, and morphological operations.



(a) ROIs in fat tissue **(b)** Edginess values

Figure 3.1: Computation of the structure coherence feature in the fat region of a patient CT image. (a) ROI_h and ROI_s are typical homogeneous and structural ROIs in the fat tissue region. The numerical values are the structure coherence levels. (b) Pixel-wise edginess values of the two regions. (Image source: [Chun 15])

Secondly, the structural coherence features were determined to localize multiple homogeneous ROIs within segmented image. Thirdly, the noise level in the image was obtained by calculating the average pixel standard deviation over selected ROIs. This method proved to be accurate for noise measurements during validation with manual observations. One of the limitations was that noise was only estimated in fat tissue regions. The authors highlighted that there might be relevant differences of the noise in different tissue types due to the different absorption coefficients.

The work of [Malkus 17] measured the noise level based on the air around the patient and proposed it as a substitute for the noise measured inside the patient. A global noise index was estimated as a reference by subtracting two adjacent slices, removing the remaining morphological variations and computing the standard deviations in small ROIs not containing air or bone. In a second step, the standard deviation was computed in the air surrounding the patient. Both signals, i. e. the global noise index and the noise value estimated from the air were correlated. A linear correlation was found between two different noise metrics.

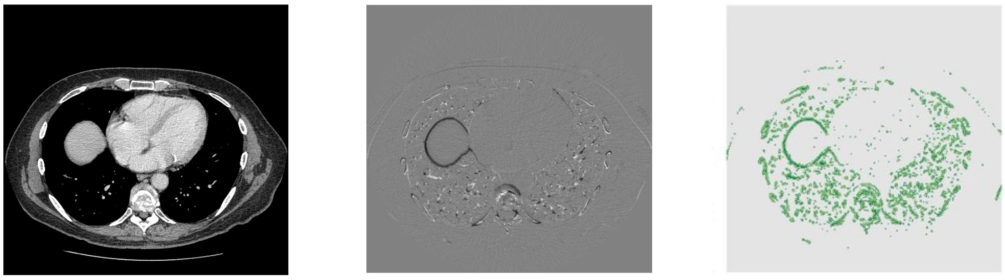
Using the standard deviation for noise measurement does not provide complete information of the spectral characteristics of the noise. In CT images, the detection and classification of low contrast objects also depends on the noise texture. The noise component in an image reflects changes in both, the spatial frequency distribution and the magnitude. The NPS is another metric which is used to describe the noise in spatial correlations and magnitude [Boedeker 07, Greffier 19]. Computing the NPS requires a homogeneous or stationary background. Otherwise, an accurate noise approximation in CT images is not possible. The NPS resolution corresponds to the spatial frequency resolution, which is determined by the size of the ROI, i. e. by the number of available pixels. In different studies, ROIs with different sizes ranging from 16 x 16 pixels to 128 x 128 pixels were chosen for NPS calculations from phantom images [Dolly 16, Passand 20]. Prior to the NPS computation, detrending of the ROI is usually performed to remove underlying signal fluctuations which are not associated with the noise. In [Dolly 16], different methods to remove the effect of artifacts on NPS estimation are summarized. These methods are based on image subtraction, first order or second order polynomial fitting. In the subtraction method, two image data sets containing the same anatomy are acquired under identical operating conditions. These images are then subtracted from each other and only the noise remains. However, such an approach is not feasible when acquiring patient data, since this causes an unnecessary increase of the applied dose. As an alternative, two adjacent slices can be subtracted, although this will result in a differential image containing information of the anatomical structures. In [Passand 20], the gray values were normalized by the mean gray value of the ROI before computation of the NPS.

An approach for noise assessment in clinical CT images was developed by [Christianson 15]. The method was evaluated using anthropomorphic phantom CT images. The authors estimated the noise in various homogeneous ROIs distributed across the whole image. Noise estimation was limited to soft tissue regions (Hounsfield unit between 0 and 100). Different ROI sizes ranging from $4\,\mathrm{mm}^2$ to $900\,\mathrm{mm}^2$ were evaluated. It was assumed that very small ROI sizes do not fully cover the low-frequency characteristic of the noise while too large ROI sizes may be contaminated by anatomical structures. For the global noise measurement, a $36\,\mathrm{mm}^2$ ROI size was identified as ideal. The standard deviation was computed for each ROI and a histogram was generated from the different standard deviations. The histogram was separated into two areas: homogeneous and transitional. The transitional areas, which are the areas containing different anatomical structures or which were at the boundary of different tissue types, constituted only a small portion of the image. Hence, the resulting histogram was dominated by two characteristic components: a higher, narrow peak caused by the homogeneous image areas and a short, broad tail which corresponded to the areas containing structures, i. e. the transitional areas. Finally, the global noise level was estimated based on the identification of the histogram peak corresponding to the homogeneous tissue areas.

Tian *et al.* proposed another technique to predict the noise in clinical CT images [Tian 16]. Clinical images containing chest, abdomen, and pelvic regions were utilized for the noise measurements. An adjacent slice subtraction method was performed on the clinical images to separate noise from anatomical structures (see Fig. 3.2). A Sobel edge filter was applied to the subtraction image to detect the residual structures. Areas of 30 x 30 pixels not containing any structures were used for the noise estimation by computing their standard deviation. This method is practically applicable to clinical CT images. It has to be considered that a reliable edge detection within the differential image is required and that a sufficient number of ROIs remains in the subtraction slice. With thinner slices, better results are to be expected due to the smaller differences in the adjacent slices.

Assessment of the spatial resolution by means of modulation transfer function

The analysis of the spatial resolution of CT images by means of MTF was rarely used with patient CT images. This was due to the fact that the MTF computation usually relies on the analysis of the edge scan function, i. e. the analysis of an edge between two objects (see Section 4.4). It is a challenging task to identify edges suitable for MTF analysis within



(a) Chest CT image **(b)** Subtraction image **(c)** Edges in subtraction image

Figure 3.2: Subtraction technique to eliminate anatomical structures from an image. The CT image shown in (a) was subtracted from its adjacent slice, resulting in an subtraction image (b). This difference image contains information about the anatomical structures. By identifying the edges (c), homogeneous region within the CT image could be located and used for the NPS analysis. (Image source: [Tian 16])

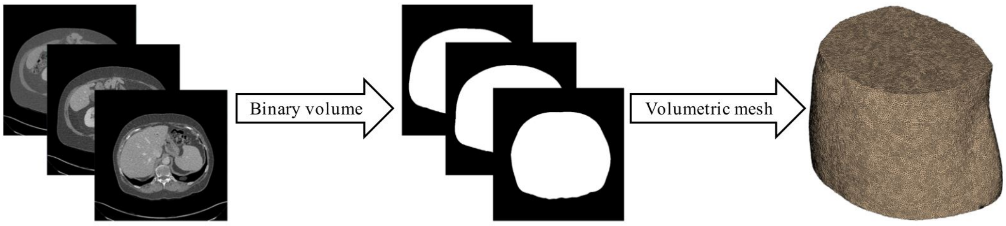


Figure 3.3: Process of generating a tetrahedral mesh from a CT dataset. A binary volume was created in the first step to separate the patient from surrounding objects (e. g. the table, cables or tubes). A mesh was generated based on this volume which was then used to compute the edge scan functions along the air-skin interface perpendicular to the surface of the skin. (Image source: [Sanders 16])

a patient CT image.

Sanders *et al.* developed a method for MTF estimation in patient CT images [Sanders 16]. The MTF was estimated at the air-skin interface. It was shown that such a method was sensitive enough to represent the changes in the spatial resolution caused by different reconstruction techniques. Figure 3.3 shows the computation of a volume mesh model which was used to support the estimation of the edge scan function at the air-skin interface. The main challenges to implement this method were a reliable segmentation of the air-skin boundary and to avoid the inclusion of the patient table or the table cushion into the analysis. Another limitation is that the air-skin interface might not represent the spatial resolution of other image parts which can be treated differently during the reconstruction process. This approach was further developed by [Cheng 19]. Besides other parameters as a normalized NPS, the MTF estimated from the air-skin boundary was used to compute a new “clarity” metric. This metric which was mainly based on the MTF and NPS was used to describe how clear a lesion with a given shape is represented in a CT image whereas the MTF and the NPS represent the signal and noise transfer properties.

There are many other works considering the MTF computation using phantoms. Carton *et al.* developed an MTF measurement for quality control in digital mammography using dedicated phantoms and acquisition techniques [Carton 05]. Anam *et al.* developed an automatic algorithm for MTF calculation using a transparent, thermoplastic phantom [Anam 18]. Starting from the center of the phantom image, lines were constructed towards the edge of the image. At the edge, an ROI was extracted used to derive the edge spread function (ESF) and line spread function (LSF) functions which were then used to estimate the MTF.

Image quality assessment using machine learning or deep learning

Bosse *et al.* proposed a method for IQA using a convolutional neural network [Bosse 16]. The network was trained using the LIVE database containing 981 images which were generated by distorting 29 reference images [Wang 04]. The distortions included different JPEG compression levels, Gaussian blur or white noise. The authors could achieve good results in IQA, but local distortions of the images were not further assessed. Additionally, due to their artificial nature, the images used in this study did not represent the distortions occurring in clinical CT images. Another work using the LIVE database for IQA was published by Bianco *et al.* [Bianco 18]. In this work, different networks were trained and evaluated showing the applicability of neural networks for IQA.

Piccini *et al.* used convolutional neural networks for estimating the image quality of cardiac magnetic resonance images [Piccini 20]. For the training, 424 magnetic resonance images from randomly selected patients were used. The training images were graded by an expert reader on a 9 level quality scale. After training, 100 images were used for testing the algorithm whereas these 100 images were graded a second time by the first expert reader and by an additional reader. The results showed a high overall correlation between the algorithm and the expert's grading. Very high and very low grades were slightly under- and overestimated and the variability increased for the medium grades.

Other approaches

A task-based method for image quality quantification using a model observer was described by [Racine 18]. To assess the quality of CT images, a channelized Hotelling observer with dense difference of Gaussian channels was used. The channelized Hotelling observer is an anthropomorphic numerical observer which can be seen as a generalized likelihood ratio test that uses channels which mimic the response of the human visual system. This observer was developed by [Myers 87, Yao 92] and is widely used for evaluating the quality of medical images. It acts as a surrogate for human observers and is used for the assessment and optimization of image processing algorithms imaging or imaging algorithms, especially in the early research and development stages. In the work of [Racine 18], a score (the decision variable) was computed by the channelized Hotelling observer using phantom CT images. The score was then related to a volume-specific dose index. Using different phantom sizes and doses, interesting relations between the image quality, phantom size and applied dose were shown. The authors recommended to perform such studies for clinical questions with the goal to ensure sufficient image quality while reducing the applied dose.

3.3 Thesis objectives

Often, phantom based image quality descriptors are used to measure the quality of CT images. Typically, images are obtained from dedicated test phantoms to estimate and evaluate parameters such as MTF or NPS. The usage of these Fourier based parameters for IQA was suggested by [ICRU 12]. While it is comparatively straight forward to analyze such phantom images, there is no high correlation between the parameters derived from such measurements and a subjective evaluation of the image quality. This is even more difficult when using iterative reconstruction techniques: when reconstructing an object with a very limited number of structures such as a phantom, these reconstruction techniques can reduce the noise to a very high extent without significantly deteriorating the structures. However, this is not comparable with real patient images which contain a large amount of fine and coarse structures. Hence, evaluating the image quality using only phantoms is not sufficient for the practical clinical application.

Depending on the clinical or diagnostic application, only certain parts of CT image are relevant for the radiologist. Therefore, this specific region should have a sufficient quality to enable a reliable diagnostic analysis. Consequently, the objective of this work was to take further steps towards an automated image quality assessment in a specific, application dependent region within clinical CT images. The features including MTF, NPS and power spectrum (PS) were estimated at the structure of interest (the major fissure of the left lung) or in its direct vicinity. Considering the state of the art, several adoptions and additional developments were required. This can be exemplary explained considering the MTF based method proposed by [Sanders 16]. In this approach, the air-skin interface of the patient was used to derive and evaluated the spatial resolution of the CT image. However, inside the patient's body, image parameters such as noise, contrast or scatter effects are different when compared to the regions close to the skin. Hence, the parameters or features used for IQA should be estimated directly at the structures or in their vicinity.

Based on an statistical analysis, the different features were evaluated with respect to their applicability for distinguishing between different images quality classes, i. e. if there was a significant difference of the values for a certain feature between the different classes. Due to the limited amount of images available for this work, neural networks were not considered for IQA. Clinical CT images of the lung from several patients served as a basis for the development of the algorithms to quantify the image quality in those regions. The CT images were annotated by experienced radiologists enabling the evaluation of the developed method. The evaluation should reveal how this method could be used as a new tool to analyze the quality and consistency of CT images. Such a tool would enable an easier optimization of scanning protocols with reduced patient exposure while maintaining a sufficient, task-related image quality.

4

Material and methods

4.1 Overview

This chapter illustrates the methods used in this work for the development of an automated image quality assessment tool to be used with real patient images. The CT images used in this work were acquired from a phantom as well as from various patients at different clinical sites located in Germany and France. The acquisition process of the phantom images and annotation of the clinical images is described in Sections 4.2 and 4.3. The images were used for the development and evaluation of the different algorithms.

The main concept for image quality assessment was based on the computation and analysis of the MTF and NPS which are explained in Section 4.4 and 4.5. In addition, the ratio between the PS and the MTF were introduced as another metric for this purpose (Section 4.6). These methods were combined to define a new image quality metric for the analysis and evaluation of clinical CT images. Those parameters or features were used to analyze the applicability of a machine learning approach for differentiating between different image quality levels (Section 4.8).

4.2 Clinical thorax CT images – acquisition and annotation

The thorax CT images used in this work were collected from two different hospitals (University Hospital Magdeburg, Magdeburg, Germany and Cochin Hospital, Paris Descartes University, Paris, France) involved in the European "MEDIRAD" project. The acquired images were from the patients with certain lung diseases such as fibrosis. The regions of interest considered on the CT images were predefined by several radiologists from the same clinical centers where the data were collected. CT images from 31 patients were available for the development of the IQA method and its evaluation within this thesis. The images were acquired using two different CT scanner models: a Revolution HD (GE Healthcare, USA) and an Aquilion Prime (Toshiba, Japan).

Different radiologists annotated the CT images regarding their quality levels. This annotation was performed by analyzing specific regions of the image which contained relevant diagnostic information, e. g. about certain lung diseases such as infection (pulmonary tuberculosis and mycobacterial infections), interstitial pathology or pulmonary nodules and metastases. The radiologists considered different aspects during annotation such as the noise level in the image, the spatial resolution and the overall diagnostic acceptability. Due to this region-specific annotation procedure, it was possible that cardiac motion artifacts existed in one part of the image while the other part of it which was used for diagnosis was free from distortions or artifacts.

One criterion during the image quality assessment was the *clear reproduction of the major fissure of the left lung*. The quality was rated by the radiologists on an ordinal scale ranging from 1 to 5 with 1="confident that the criterion is fulfilled"; 2="somewhat confident that the criterion is fulfilled"; 3="I do not know if the criterion is fulfilled or not"; 4="Somewhat confident that the criterion is not fulfilled"; 5="Confident that the criterion is not fulfilled". The images in which the reproduction of the major fissure was rated as 1 or 2 were labeled as "acceptable" (ACC) and the ones rated as 4 or 5 were rated as "not acceptable" (NAC).

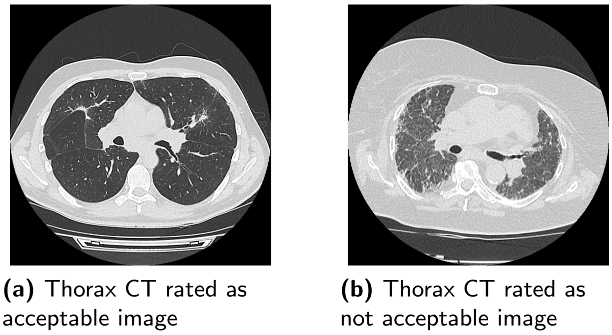
All the images which were used in this work were analyzed in the axial plane of the thorax CT. Table 4.1 lists various acquisition parameters of the CT images used in this thesis. Two exemplary thorax CT images are depicted in Figure 4.1.

Table 4.1: Scan parameters used during the data acquisition. 16 images were annotated as ACC and 15 as NAC. It should be considered that the given tube currents were not constant but modulated.

Dataset	Pixel spacing [mm]	Tube current [mA]	Tube voltage [kV]	Total collimation width [mm]	Slice thickness [mm]	Spiral pitch factor
			Acceptable (ACC) cases			
10	0.891	218.0	120	38.4	1.0	1.2
11	0.844	439.0	120	38.4	1.0	1.2
12	0.740	80.0	120	40.0	1.0	0.8
19	0.727	222.0	100	38.4	1.0	1.2
20	0.654	94.0	100	38.4	1.0	1.2
21	0.797	100.0	120	40.0	1.0	0.8
22	0.654	94.0	100	38.4	1.0	1.2
60	0.723	80.0	120	40.0	1.0	0.8
61	0.797	115.0	120	40.0	1.0	0.8
62	0.746	480.0	100	38.4	1.0	1.2
63	0.740	80.0	120	40.0	1.0	0.8
64	0.844	273.0	120	38.4	1.0	1.2
65	0.891	181.0	120	38.4	1.0	1.2
66	0.717	80.0	120	40.0	1.0	0.8
67	0.654	111.0	100	38.4	1.0	1.2
68	0.727	148.0	100	38.4	1.0	1.2
			Not acceptable (NAC) cases			
13	0.656	10.0	100	40.0	1.2	1.0
16	0.667	124.0	120	40.0	1.0	0.8
18	0.740	80.0	120	40.0	1.0	0.8
48	0.741	86.0	120	40.0	1.0	0.8
49	0.827	103.0	120	40.0	1.0	0.8
50	0.671	80.0	120	40.0	1.0	0.8
51	0.667	107.0	120	40.0	1.0	0.8
52	0.763	139.0	120	40.0	1.0	0.8
53	0.827	101.0	120	40.0	1.0	0.8
54	0.759	121.0	120	40.0	1.0	0.8
55	0.748	247.0	100	38.4	1.0	1.2
56	0.740	80.0	120	40.0	1.0	0.8
57	0.703	10.0	100	40.0	1.2	1.0
58	0.637	10.0	100	40.0	1.2	1.0
59	0.748	231.0	100	38.4	1.0	1.2

4.3 Phantom images

For more objective evaluation purposes, the CIRS 701 Adult male phantom (Cirs Inc. USA) was scanned with different parameters using a CT scanner (Aquilion Prime, Toshiba, Japan). Figure 4.2 shows the phantom used in this work. An iterative reconstruction technique was used to compute the CT images. The tube voltage and the tube current

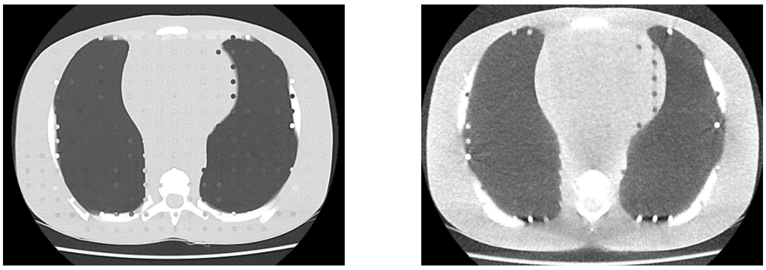


(a) Thorax CT rated as acceptable image

(b) Thorax CT rated as not acceptable image

Figure 4.1: Two thorax CT images which were labeled as acceptable (a) and not acceptable (b). The images were acquired at the University Hospital Magdeburg, Germany.

Figure 4.2: CIRS 701 Adult male phantom (Cirs Inc. USA). This phantom was designed to simulate the anatomy of the patient and for rendering more reliable and accurate CT dose data. (The phantom was provided by the Bundesamt für Strahlenschutz, Munich, Germany.)



(a) 120 kV, 80 mA

(b) 80 kV, 10 mA

Figure 4.3: Exemplary CT images of the human torso phantom taken with different tube voltages and tube currents.

were modified to alter the quality of the phantom CT images. Different voltage-current configurations were used, e. g. 100 kV/50 mA, 120 kV/80 mA and 80 kV/10 mA. Slice thicknesses of 1 mm and 5 mm were used whereas the size of the in-plane pixels was 0.65 mm x 0.65 mm. Figure 4.3 shows two exemplary CT images of the phantom acquired with different tube configurations and a slice thickness of 5 mm. Due to their homogeneous background, the phantom images were used for evaluating the NPS algorithm.

4.4 Modulation transfer function

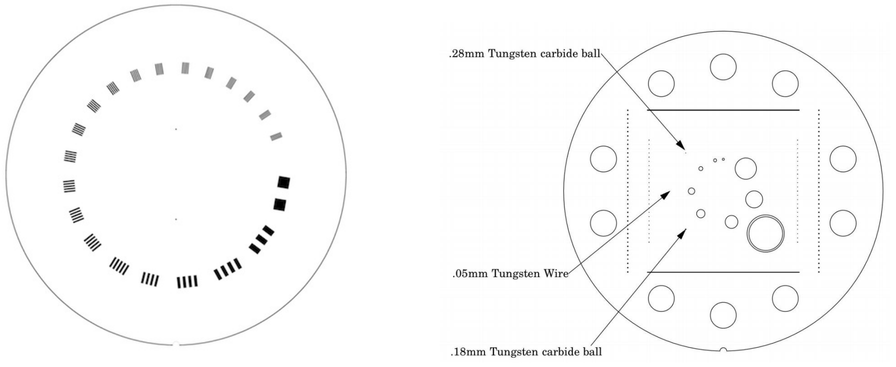
4.4.1 Theory

The MTF is an important metric for image quality assessment. It has various applications in different imaging modalities, ranging from medical imaging to consumer cameras. Every imaging system acts as a low pass filter affecting the sharpness of an object. This effect can be analyzed or described by the MTF which is used to measure the system's response to various input frequencies. It is defined by the ratio of the output to the input modulation. The MTF of an ideal imaging system equals one for all spatial frequencies, i. e. the system's response does not depend on the input frequency. Real imaging systems, however, degrade this ideal input response in a certain manner. The MTF of most imaging systems rapidly decreases at higher frequencies.

Different methods exist for deriving the MTF [Hsieh 09, Bushberg 11]. The computation of a standardized and comparable MTF requires a clearly defined test pattern called phantom. For CT applications, such test phantoms contain a variety of different structures which can be used to obtain a system's MTF. Figure 4.4 shows a typical construction of a CT phantom. The shown line pair structures can be used to compute the contrast along the line pattern for each spatial frequency (see Fig. 4.4a). The MTF for a given spatial frequency is then defined as the contrast at that frequency normalized by the contrast of the lowest frequency component. The phantom slice shown in Fig. 4.4b contains a point bead source which can be used to determine the system's point spread function. The MTF of a system can be computed by the Fourier transform of the acquired point spread function. However, since those methods require a dedicated phantom, they are not applicable in patient CT images.

To overcome the problem of using a dedicated imaging phantom, another approach is required to compute the MTF based on clinical patient CT images. This can be achieved by scanning the edge properties of an object within the image, e. g. the edge between a bone and the surrounding background tissue. However, the determined MTF will then not only characterize the system but also the patient condition.

Figure 4.5 illustrates the basic principle or idea of this edge-based approach. As a first step for computing the MTF, the ESF needs to be obtained from the object to be analyzed. This is achieved by scanning along a line perpendicular to the object's edge along a given number of pixels x resulting in $ESF(x)$ (see Fig. 4.5(b)). While the ideal (or real) edge



(a) Line pair structures **(b)** Point bead structures (and others)

Figure 4.4: Inner construction of widely used phantoms for assessing MTF and other image quality parameters of a CT system. (Image sources: [Catphan 21, Catphan 22])

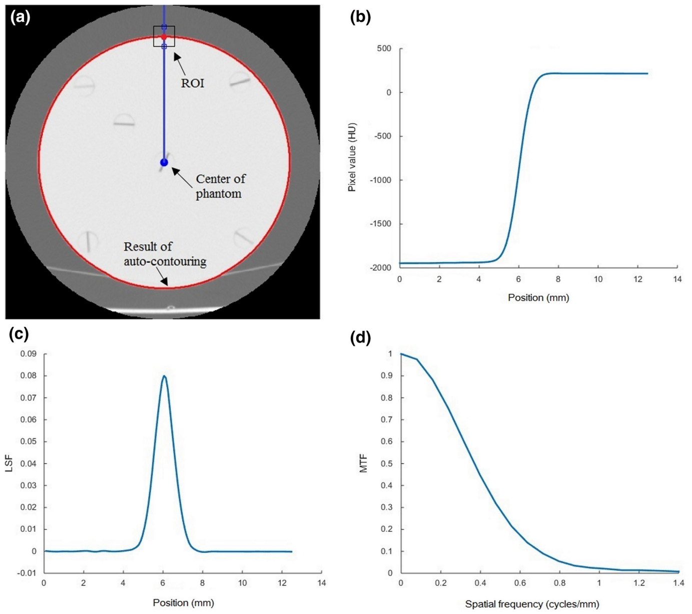


Figure 4.5: Principle of MTF computation based on the analysis of an edge inside the image. An ROI containing the edge needs to be defined (a). Within this ROI, the line profile is estimated along a path perpendicular to the edge, which results in the ESF (b). The first derivative of the ESF leads to the LSF (c) and a subseqeuent Fourier transform of the LSF provides the MTF (d). (Image source: [Anam 18])

of an object can be considered to be a step function, the imaging system introduces a filtering of this edge resulting in a smoothed version of the ideal edge function (compare Fig. 4.5(b)). The transition of the ESF contains information about the filtering introduced by the imaging system, i. e. how it treats different spatial frequency components. The LSF was computed by differentiating ESF(x) with regard to the pixels x:

$$\mathrm{LSF}(x) = \frac{\partial}{\partial x}\mathrm{ESF}(x) \tag{4.1}$$

The computed LSF (compare Fig. 4.5c) provides information about how the imaging process or system spreads out an originally sharp edge. The LSF can be considered as the probability density function of the edge location. In an (theoretically) ideal imaging system, the LSF would be a Dirac delta function. The MTF is defined as the Fourier transform of the LSF:

$$\mathrm{MTF}(k) = \sum_{n=0}^{N-1} \mathrm{LSF}(n) \exp\left(\frac{-i2\pi}{N} kn\right) \tag{4.2}$$

Figure 4.5d shows an exemplary MTF. The MTF is a representation of the LSF in the frequency domain, providing information on the frequency attenuation characteristics of the imaging system.

The determination of the MTF requires the alignment of an edge perpendicular to the scan line. Since the MTF computation within this thesis did not rely on standardized objects or phantoms with known geometries, a different approach was required which is described in Section 4.4.2.

4.4.2 Implementation

The method used for estimating the MTF in clinical CT images was based on the ESF or LSF. The main challenge with this approach was to realize a reliable detection of object's edge (e. g. of the major fissure) inside the CT image. Once the edges of an object were detected, line profiles could be derived perpendicular to the edge. Those lines were used to compute the ESF, LSF and MTF according to the descriptions in Section 4.4.1. The developed algorithm was organized using the following basic steps:

1. Selection of an ROI containing an object (e. g. the major fissure) and background tissue
2. Detection of the object's edges in the ROI
3. Computation of the gradient direction for each pixel of the edge
4. Estimation of a line profile (ESF) along the direction of the edge's gradients
5. Derivation of the LSF from the ESF
6. Computation of the MTF using the LSF
7. Averaging of several MTF functions
8. Extraction of different parameters from the MTF

Figure 4.6 summarizes the different steps to obtain an MTF from a patient CT image. Figure 4.7 gives an exemplary overview of the processing steps required to obtain the LSFs across the major fissure of the left lung. An ROI which contained one of the structures defined in Section 2.4 was chosen manually with the support of a radiologist. Figure 4.7a shows an exemplary, manually selected ROI containing a part of the major fissure. The size of the ROI had to be small (e. g. 18 x 18 pixels) in order to mainly contain the object of interest and, if possible, no other surrounding structures. Prior to the edge detection, a preprocessing was applied to the region of interest to improve the subsequent edge detection algorithm. The following preprocessing was applied:

1. Linear interpolation by a factor of 2
2. Filtering of the ROI using a Gaussian kernel with standard deviation $\sigma = 2$
3. Sharpening of the image using an unsharp masking technique [Polesel 00]

Figures 4.7b and 4.7c show the interpolated and the preprocessed image. It shall be highlighted at this point that the preprocessing steps 2 and 3 were only used to enhance the edge detection process, i. e. the unprocessed image was used to estimate the ESF, LSF and MTF.

Following the ROI selection and the preprocessing, an active contours algorithm was employed for separating the object from the surrounding tissue [Chan 01]. For optimal results, the placement of a seeding area or initialization mask was required to distinguish between the foreground and background areas. This mask was placed in the center of the ROI since the ROI was chosen such that it passes the center of the ROI. The maximum number of iterations for the active contours algorithm was set to 1000. The active contours algorithm resulted in a binary image separating the ROI into two regions (object and background). An edge detection using the Sobel operator was applied to compute the edge pixels required for the following steps. Figures 4.7d and 4.7e show the detected edges and the same edges overlayed on the CT image.

The edge pixels obtained in the previous steps served as a basis for estimating multiple line profiles perpendicular to the object's edge. Therefore, the gradient directions were computed for each edge pixel. Based on the gradient direction of an edge pixel, the line profile was estimated. The total length of the line profile was $5\,\mathrm{mm}$, whereas the center of the line profile was placed at the position of the edge pixel. The length of the line profile

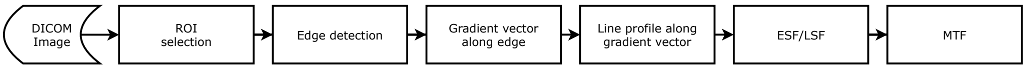
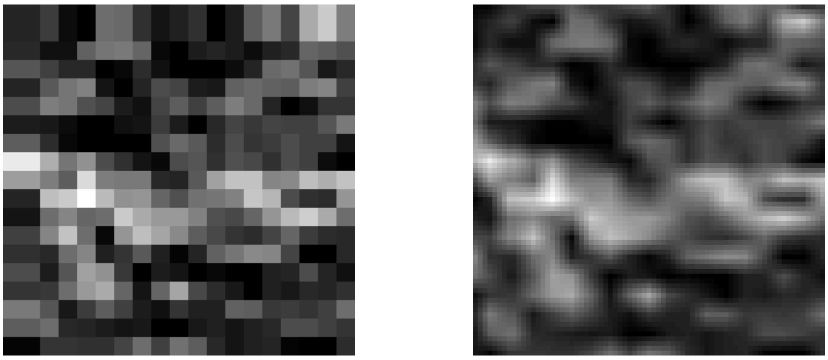
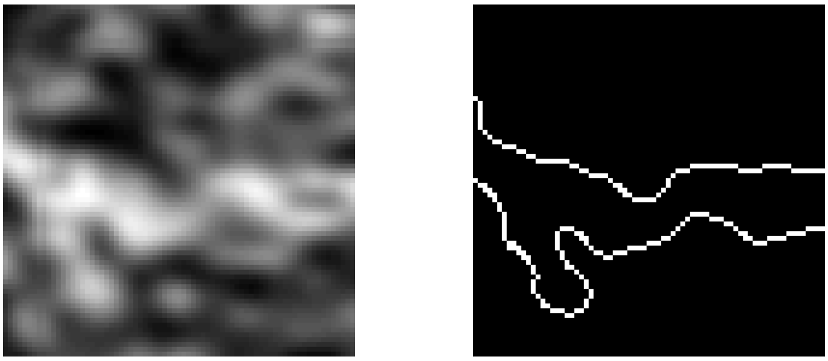


Figure 4.6: Overview of the process which was applied to compute the MTF from a clinical CT DICOM image.



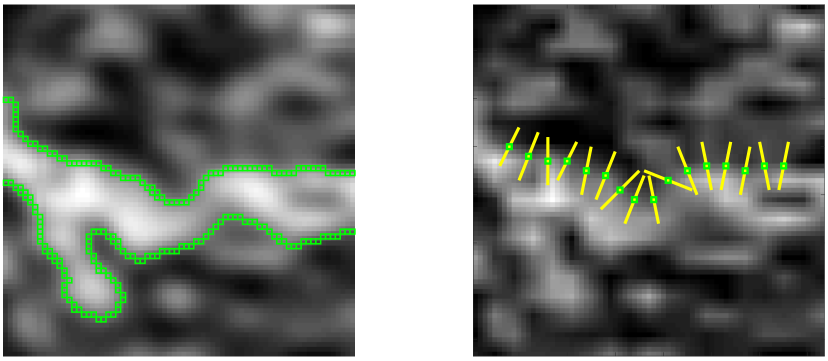
(a) Original ROI

(b) ROI interpolated by a factor of two



(c) Interpolated and processed ROI

(d) Detected edges



(e) ROI (c) with edge overlay

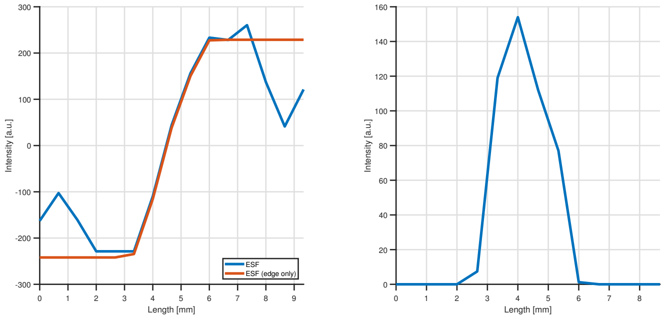
(f) ROI (b) with overlayed paths used for line profile estimation

Figure 4.7: Steps to obtain a line profile from the major fissure in a CT image required for the computation of the MTF. The manually selected ROI (a) was interpolated (b) and preprocessed (c). Edges were detected using an active contours algorithm (d,e). For each edge pixel, a line path was estimated perpendicular to the edge (f), which was used to derive the ESF.

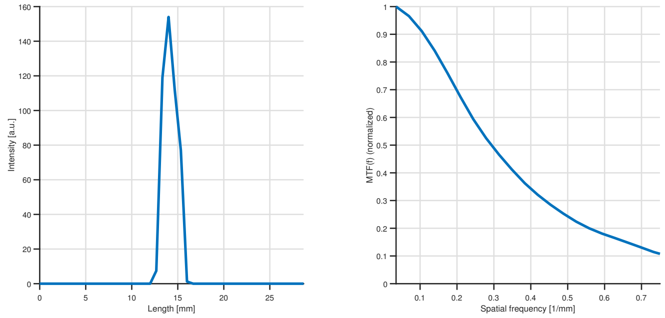
was adjusted experimentally such that it reliably covers parts of the object, its edge and the surrounding (background) tissue. An example is shown in Fig. 4.7f.

Figure 4.8a shows an exemplary line profile or ESF extracted from one of the paths. It can be seen that apart from the edge, various fluctuations exist within the background and the object region. However, the computation of the MTF should solely rely on the properties of the object's edge and not on other signal components outside the edge. Hence, those fluctuations were automatically removed as described in the following. Due to the low number of sampling points, the ESF was interpolated by a factor of 4. The center part of the edge was identified based on the maximum of the ESF's first derivative (i. e. the LSF). Starting from this point, the first local maximum and minimum on each side of the edge was searched within a predefined search window of 9 pixels. Once the maximum and minimum around the edge were located, the remaining parts of the ESF were truncated. The truncated values were replaced with the detected maximum and minimum values.

After those preprocessing steps, the LSF was derived from the ESF by means of the first derivative. Figures 4.8b and 4.8c shows an exemplary LSF and a zero-padded version



(a) ESF **(b)** LSF



(c) Zero-padded LSF **(d)** MTF (average)

Figure 4.8: Processing steps from ESF to MTF: The ESF (a) was processed; signal fluctuation outside the edge were suppressed. The LSF (b) derived from the ESF was zero-padded (c) before computing the MTF (d).

of the LSF used for the computation of the MTF. Zero-padding on the left and right side of the LSF curve was used due to the low number of samples in the LSF. While zero-padding does not add new information to the signal, the increased number of samples increases the resolution in the frequency domain which allows a more detailed analysis of the MTF. The MTF was computed separately for each LSF by the Fourier transform of the LSF (see Section 4.4.1). An average MTF was then estimated based on the computed MTFs. Figure 4.8d shows the average MTF. For a quantitative evaluation of the MTF for an automated image quality analysis, the area under the MTF curve was computed.

4.5 Noise power spectrum

4.5.1 Theory

In general, image noise can be statically defined using the standard deviation of the pixel intensity values or CT numbers in a physical homogeneous region. An increasing variance or standard deviation of the pixel values corresponds to an increasing noise level. The noise in a CT image also depends on the number of photons which are received by the detector elements. This type of noise is referred to as quantum noise. With an increasing number of photons received by the detectors, the noise decreases [Goldman 07b]. The noise occurring in an image has a strong effect on the perceived image quality. Different methods exist to estimate the noise in an image or a certain ROI. In most cases, the variance σ^2 of the CT numbers or pixel values is used to quantify the noise since the noise is proportional to the variation of the CT values within a homogeneous ROI.

In a typical situation where one wants to assess the quality or noise of a CT system given a specific configuration, the noise can be estimated using phantoms with homogeneous regions. Figure 4.9 shows the alignment of different ROIs in a phantom image. The noise estimated within each ROI is averaged to obtain a global noise estimate of the CT image.

Using the variance for estimating the noise, only the noise magnitude is taken into account but not the texture of the noise, i. e. its spatial correlation [Bushberg 11]. Figure 4.10 highlights this aspect. Both images have the same variance, but there is a clearly visible difference in the texture or appearance of the noise distribution. Hence, a more comprehensive approach is required to characterize the noise and the the texture of the noise in a CT image. This can be achieved by computing the NPS of a CT image [Boone 12].

The NPS provides a more extensive noise description than the variance based noise esti-

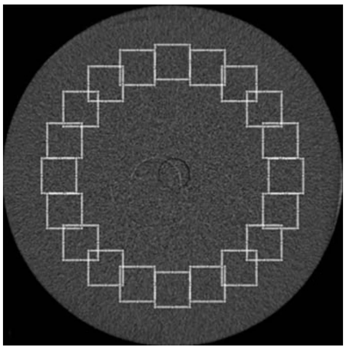


Figure 4.9: Typical positioning of ROIs for estimating the NPS using a phantom. The ROIs are placed in a radial symmetric way. After computing the NPS in each ROI, an averaging is performed. (Image source: [Brombal 19])

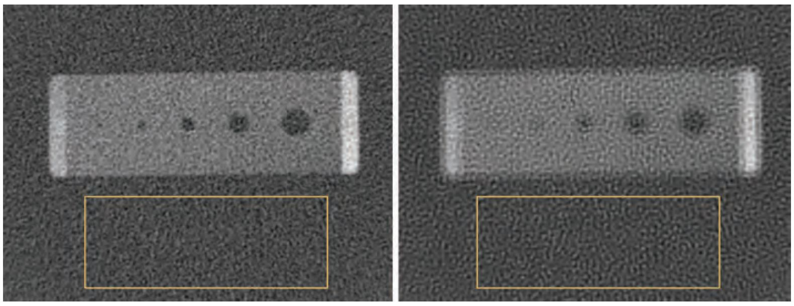


Figure 4.10: The standard deviation in the highlighted ROI of the two CT images is the same, although the texture of the noise is different. The NPS can be used to characterize the frequency distribution of the noise texture. (Image source: [Bushberg 11])

mation. It allows an analysis of the noise variance with respect to the spatial frequency. When the NPS is analyzed together with dosimetric quantities, it enables the comparison between the applied dose and the noise occurring in the image [Solomon 12]. The primary noise correlation source in a CT image is the 3D point-spread function which causes noise correlation in all spatial dimensions [Boone 12].

Due to the fact that CT images are computed by certain reconstruction algorithms, e. g. by iterative reconstruction or filtered back projection, the point-spread function depends – besides acquisition parameters such as dimensions of the detector or the size of the focal spot – on the applied reconstruction technique, i. e. the reconstruction kernel or the slice thickness.

The noise correlation in the spatial domain can be described by the autocorrelation function. The Fourier transform of the autocorrelation function is the NPS, or the Wiener spectrum [Dainty 76]. For a two-dimensional (2D) image or ROI, the 2D-NPS is defined as:

$$\mathrm{NPS}(f_x, f_y) = \frac{1}{N} \sum_{i=1}^{N} \left| \mathrm{FFT}_{\mathrm{2D}} \left[\mathrm{I}_i(x, y) - \bar{\mathrm{I}}_i \right] \right|^2 \frac{\Delta_x \Delta_y}{N_x N_y} \tag{4.3}$$

where $\mathrm{I}_i(x, y)$ is the i-th image or ROI, $\bar{\mathrm{I}_i}$ is the mean of this ROI, N is the number of slices or ROIs used for averaging the NPS values, Δ_x and Δ_y correspond to the pixel spacing, N_x and N_y are the number of voxels in the x- and y-dimension, f_x and f_y are the spatial frequencies and the $\mathrm{FFT}_{\mathrm{2D}}$ is the (discrete) 2D fast Fourier transform. Considering a 2D region of interest, $\mathrm{I}_i(x, y)$ is the signal in the i-th ROI and $\bar{\mathrm{I}_i}$ is the mean of the ROI $\mathrm{I}_i(x, y)$.

Before computing the NPS of an ROI or volume of interest, a detrending of the background can be performed to suppress low frequency fluctuations of the background. This is necessary since real world image data not only contains stochastic noise but also other signal components. These additional (and unwanted) artifactual components lead to an increase of the low-frequency spectral components [Zhou 11]. One way to achieve a detrending of the background is by subtracting a 2D second order polynomial which is fitted to the input image.

4.5.2 Implementation

Overview

As described in Section 4.5.1, the computation of the NPS relies on a homogeneous background without any (anatomical) structures. This requirement makes the computation of the NPS a challenging task in real clinical CT images. The goal was to automatically identify sub-regions of the selected ROI which were dominated by homogeneously distributed tissue.

The general procedure for estimating the NPS in a selected ROI was defined as follows:

1. Selection of a homogeneous ROI close to the target structure (e. g. the major fissure)
2. Partitioning of the ROI in several smaller patches
3. Computation of various features within each patch
4. Clustering of feature values to separate areas of different homogeneity levels (2 clusters)
5. Identification of the homogeneous cluster

6. Detrending of the identified sub-region
7. Computation of the NPS in the homogenous, detrended sub-region

Figure 4.11 summarizes the different processing steps. The first step required the manual selection of an ROI within the desired part of the image. The size of this region was 16 x 16 pixels. It required a careful selection of an area which can be characterized as homogeneous. This area was further analyzed to separate homogeneous and inhomogeneous parts as described in the following. Once a homogeneous area within the selected ROI was identified, a detrending was applied to suppress low frequency variations of the background signals. The 2D-NPS was computed as described in Section 4.5.1. Using radial averaging of the 2D-NPS, a one-dimensional (1D) NPS was derived which was used for the further analysis.

Identification of a homogeneous region using k-means clustering

The reliable computation and analysis of the NPS in the CT image required a homoegenous object or image without any structures, i. e. a region only affected by noise. While one can find such homogeneous areas in CT phantom images (see Section 4.3), clinical patient CT images do not have those areas due to the presence of the inhomogeneous tissue and the anatomical structures. Hence, in order to optimize the ROI in terms of homogeneity, an automated procedure was developed to cope with this problem.

For this purpose, the ROI was subdivided into squared, overlapping patches (shift by 1 pixel). Each patch had a size of 3 x 3 pixels. For each patch, various parameters or features were computed:

- Standard deviation
- Standard deviation of the gradient magnitude and its direction
- Entropy
- Skewness
- Kurtosis

A k-means clustering algorithm was used to separate the patches into two different classes: a homogeneous and an inhomogeneous one. The k-means clustering technique belongs to the family of unsupervised learning algorithms. In general, cluster analysis is a procedure

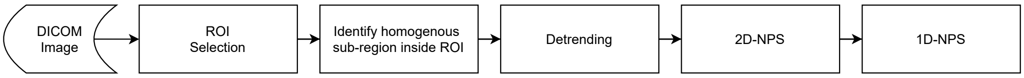


Figure 4.11: Overview of the process which was applied to compute the 1D-NPS from a clinical CT DICOM image.

with which a certain number of objects can be divided into (homogeneous) groups. The aim is that the objects that were allocated to a certain group are as similar as possible. The object properties are divided into dimensions whereas the groups are called clusters.

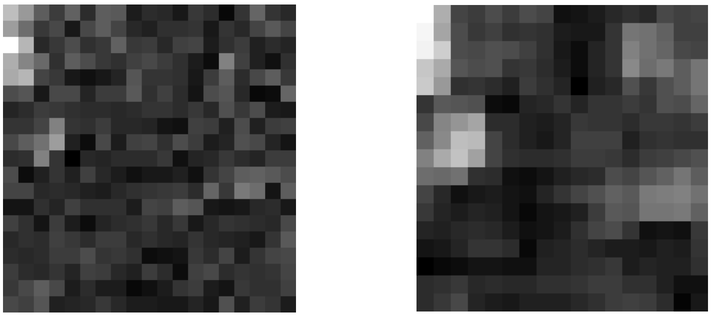
The aim of the k-means algorithm is to divide the observations n into k clusters. The k-means algorithm assigns each new observation to the cluster with the nearest mean. The k-means algorithm can be applied to multidimensional objects and approaches the respective cluster centers by constantly repeating recalculations until no significant change occurs.

The algorithm is composed of three steps. In its first step, initial centroids are defined for each cluster. The simplest approach for defining the initial centroids is to use samples from the dataset. Following this initialization, the k-means algorithm starts to loop between the following two steps. In the first step, each sample is assigned to its closest centroid. In the next step, new centroids are created based on the mean value of all samples which were assigned to the centroid in the previous step. The optimization criterion is to minimize the difference between the previous and the current centroid, i. e. the algorithm loops through the last two steps until a predefined threshold is reached. Whereas the k-means algorithm will always converge, it can result in a local minimum. The found minimum highly depends on how the centroids were initialized. Hence, it might be necessary to repeat the computation using different centroid initializations. The *k-means++* is an algorithm to optimize the selection of the initial values. By initializing the centroids distant from each other, better results can be achieved than using a random initialization.

Figures 4.12 and 4.13 show some of the features used for the k-means clustering as well as the result of the algorithm. Each pixel of the selected ROI was assigned to one of the two clusters. Following an experimental investigation, the cluster in which the ROI pixel values had a lower standard deviation was expected to be the homogeneous cluster. In the shown examples, this was cluster 2 (see Fig. 4.13c). Since the algorithm divided the image into several small regions belonging to the same cluster type, an additional algorithm was implemented to find the largest square of connected pixels belonging to the identified homogeneous region. This sub-region of the ROI was then used for the computation of the NPS. The identified ROI was detrended and the 2D- and 1D-NPS were computed.

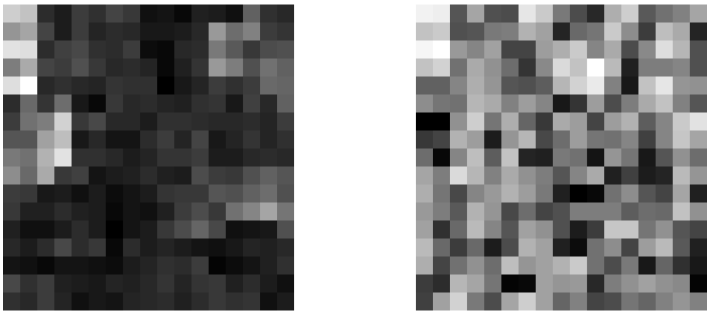
The area under the 1D-NPS curve was defined as:

$$\mathrm{NPS_A} = \int \mathrm{NPS}(f)\ \mathrm{d}f \tag{4.4}$$



(a) ROI

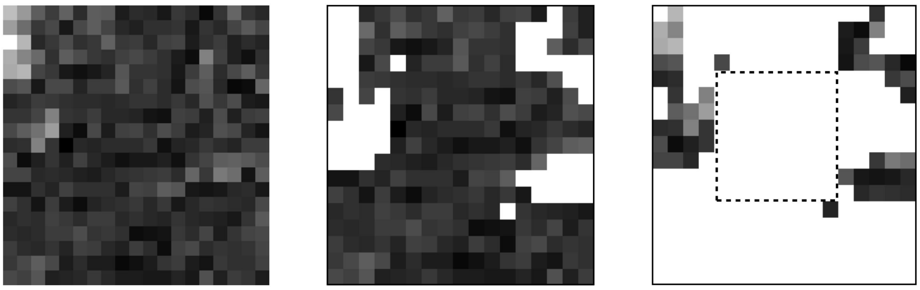
(b) Patch standard deviation



(c) Patch gradient magnitude standard deviation

(d) Patch gradient direction standard deviation

Figure 4.12: Feature computation for the identification of a homogeneous region. Within the chosen ROI (a), smaller patches were defined around each pixel of the ROI. Within these patches, different features were computed. The examples show the standard deviation (b), the gradient magnitude standard deviation (c), and the gradient direction standard deviation (d). The brighter values in (b) and (c) highlight the regions which contain structures while the darker regions are more homogeneous.



(a) ROI

(b) Class 1

(c) Class 2

Figure 4.13: Clustering of the image into two regions: homogeneous region and structure region. Based on the different patch features computed for the selected ROI (a), one cluster marks the structural region (b, in white) and one cluster marks the homogeneous regions (c).

It was also investigated if the average gray value $\bar{\mathrm{I}}_i$ within an ROI had an effect on the NPS. The following parameters were computed within a given ROI $(\mathrm{I}_i(x, y))$:

$$\mathrm{NPS_A}, \qquad \frac{\mathrm{NPS_A}}{\bar{\mathrm{I}}_i}, \qquad \frac{\mathrm{NPS_A}}{\bar{\mathrm{I}}_i^2}\,. \tag{4.5}$$

4.6 Relation of power spectrum and modulation transfer function

Analyzing the PS of an ROI containing "signal" components such as anatomical structures can provide information about the frequency characteristics of an image or a structure. The PS was computed in the ROI which was used for MTF estimation, i. e. the ROI containing the major fissure. The only difference between the NPS and the PS was that for the computation of the PS, there was no further processing of the ROI except the detrending, i. e. the PS was computed using Eq. 4.3.

Figure 4.14 shows the relation between different patterns in an image and their 2D or 1D frequency spectra. It shows how the global distribution of patterns in an image affects the corresponding spectra. In contrast, the MTF (compare Section 4.4.1) was used to characterize the frequency components at a specific object or structure inside the image. Hence, the MTF can be considered as a local frequency analysis whereas the PS of the whole image or ROI can be considered as a global frequency analysis not only containing noise, but noise and structure, i. e.:

- $\mathrm{MTF}(f)$: Local frequency spectrum (at a specific edge)
- $\mathrm{PS}(f)$: Global frequency spectrum containing noise and structure

To obtain a more qualitative ratio between the PS and the MTF, both curves were normalized such that the value at the lowest frequency was equal to one. Following this normalization, the following ratios were computed:

$$\mathrm{Ratio}_1 = \frac{\mathrm{PS}(f)}{\mathrm{MTF}(f)} \tag{4.6}$$

and

$$\mathrm{Ratio}_2 = \frac{\sqrt{\mathrm{PS}(f)}}{\mathrm{MTF}(f)} \tag{4.7}$$

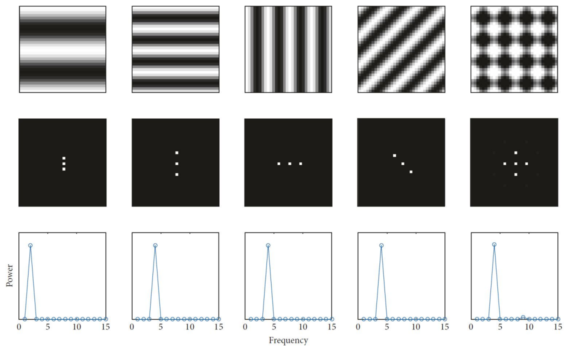


Figure 4.14: Exemplary power spectra for different image patterns. The top row shows the input image, the middle row shows the 2D power spectra and the bottom row shows the radially averaged power spectra. One can observe clearly distinguishable peaks in the 2D spectra depending on the spectral frequencies in the image and their orientation. Radial averaging is used to obtain the 1D spectra which provide a more global estimate of the image's frequency characteristics. (Image source: [Gircys 18])

The area under the 1D-PS curve was defined as:

$$\mathrm{PS_A} = \int \mathrm{PS}(f)\; df \tag{4.8}$$

It was also investigated if the average gray value $\bar{\mathrm{I}}_i$ within an ROI had an effect on the PS. The following parameters were computed within a given ROI $(\mathrm{I}_i(x,y))$:

$$\mathrm{PS_A}, \qquad \frac{\mathrm{PS_A}}{\bar{I_i}}, \qquad \frac{\mathrm{PS_A}}{\bar{I_i}^2}\,. \tag{4.9}$$

4.7 Feature summary and statistical analysis

The different features defined in the previous sections are summarized in Table 4.2. It was tested if there were significant differences for each feature between the two different image quality classes. This was the case for p-values < 0.05. In a first step the distribution of each feature in each class was tested. It was differentiated between a normal and a non-normal ("other") distribution. The following normality tests were used: Kolmogorov-Smirnov test [Massey Jr 51], Anderson-Darling test [Anderson 52], Cramer-Von-Mises test [Anderson 62], Shapiro-Wilk test [Shapiro 65], Shapiro-Francia test [Shapiro 72], D'Agostino-Pearson test [D'Agostino 73] and Jarque-Bera test [Jarque 80]. Depending on the type of the distribution (*normal* or *other*), either an unpaired t-test or a Mann-Whitney U test was applied to test if there was a significant difference for a given feature between the two image quality classes [Fay 10]. For the visual inspection of the distribu-

Table 4.2: List of features including abbreviations and a short description. The features were used to train the SVM.

#	Feature	Explanation
1	MTF_A	Integral under the MTF graph
2	NPS_A	Integral under the NPS graph
3	PS_A	Integral under the PS graph
4	$\text{PS/MTF}_\text{LF/HF ratio}$	Ratio between PS and MTF frequency spectra
5	$\text{PS/MTF}_\text{Slope}$	Average slope of the PS/MTF curve
6	$\text{PS/MTF}_\text{Max loc}$	Frequency with maximum amplitude of the PS/MTF curve
7	σ^2_Fissure	Variance of the ROI containing the major fissure

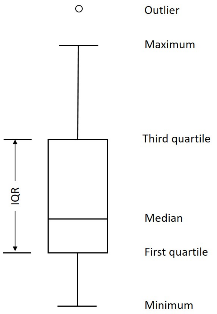


Figure 4.15: Box plot and its parameters (IQR: interquartile range).

tion of the different features between the two classes, box plots were generated for each feature. Figure 4.15 shows the description of the different parameters contained in a box plot.

4.8 Image quality assessment by means of classification

4.8.1 Support vector machines

Support vector machines (SVM) are one method to solve the problem of classification. The SVM is used to define a hyperplane which separates the input data (d-dimensional) into two classes (in the case of this thesis an “acceptable” and “not acceptable” class). In the most simple case, this data is linearly separable. However, in most real world examples, the two classes are not linearly separable. To solve this problem, a “kernel induced feature space” is introduced. Using these kernels, the input data is casted into a higher dimensional space which enables the separation of the two data classes. In general, SVMs are a theoretically well-founded and intuitive method and are practically used for various applications.

Mathematical background of support vector machines

The following paragraphs briefly summarize the mathematical background of SVMs. For further detailed information, the reader is referred to [Boswell 02, Wang 05]. For training the SVM, l training examples $\{\mathbf{x}_i,\ y_i\}, i = 1, \ldots, l$ are given. Each training example has d inputs ($\mathbf{x}_i \in \mathbb{R}^d$), and a binary class label ($y_i \in \{-1, 1\}$). The hyperplanes in $\mathbb{R}^d$ used to separate the two classes are defined by the vector ($\mathbf{w}$) and a constant (b) by the with the following expression:

$$\mathbf{w} \cdot \mathbf{x} + b = 0 \tag{4.10}$$

(The vector w is orthogonal to the hyperplane.) Given the hyperplane $(\mathbf{w}, b)$ to separate the two data classes gives the function:

$$f(\mathbf{x}) = \mathrm{sign}(\mathbf{w} \cdot \mathbf{x} + b). \tag{4.11}$$

This function correctly classifies the training data. Ideally, it also correctly classifies any other unseen data, the "testing data". A given hyperplane which is described by $(\mathbf{w}, b)$ is equally expressed by all pairs $\{\lambda\mathbf{w}, \lambda b\}$ for $\lambda \in \mathbb{R}^+$. The canonical hyperplane, which separates the data from the hyperplane by a "distance" of at least 1, can then be defined [Boswell 02]. For this, the following conditions need to be satisfied:

$$\begin{aligned} \mathbf{x}_i \cdot \mathbf{w} + b &\geqslant +1 \ \text{ when } \ y_i = +1 \\ \mathbf{x}_i \cdot \mathbf{w} + b &\leqslant -1 \ \text{ when } \ y_i = -1 \end{aligned} \tag{4.12}$$

All hyperplanes fulfilling this definition have a "functional distance" $\geqslant 1$, i. e. the the function's value is $\geqslant 1$. This functional distance is different from the "geometric" or "Euclidean distance", which is also known as margin. Given a hyperplane $(\mathbf{w}, b)$, all pairs $\{\lambda\mathbf{w}, \lambda b\}$ define the same hyperplane, but the functional distant between each hyperplane and a given data point is different. To compute the geometric distance between a data point and the hyperplane, the functional distance has to be normalized by the magnitude of $\mathbf{w}$ using the following expression:

$$d\left((\mathbf{w}, b), \mathbf{x}_i\right) = \frac{y_i(\mathbf{x}_i \cdot \mathbf{w} + b)}{||\mathbf{w}||} \geqslant \frac{1}{||\mathbf{w}||} \tag{4.13}$$

The intuitive approach is to maximize the (geometric) distance between a computed hyperplane and the closest data points (see Fig. 4.16). Considering Eq. 4.13, this can be achieved by minimizing $||w||$ (subject to the distance constraints). Lagrange multipliers can be used to solve this problem [Cortes 95, Burges 98]. This minimization problem is referred to as quadratic programming problem [Boswell 02]. Based on the minimization derivation, the optimal hyperplane can be expressed as:

$$\mathbf{w} = \sum_i \alpha_i y_i \mathbf{x}_i. \tag{4.14}$$

α is a vector containing i non-negative Lagrange multipliers which need to be determined. The vector **w** in Eq. 4.14 is a linear combination of the training data. In addition, it can be shown that:

$$\alpha_i(y_i(\mathbf{w} \cdot \mathbf{x}_i + b) - 1) = 0 \quad (\forall i) \tag{4.15}$$

This means that $\alpha_i = 0$ when the functional distance of a sample value is > 1 when $y_i(\mathbf{w} \cdot$

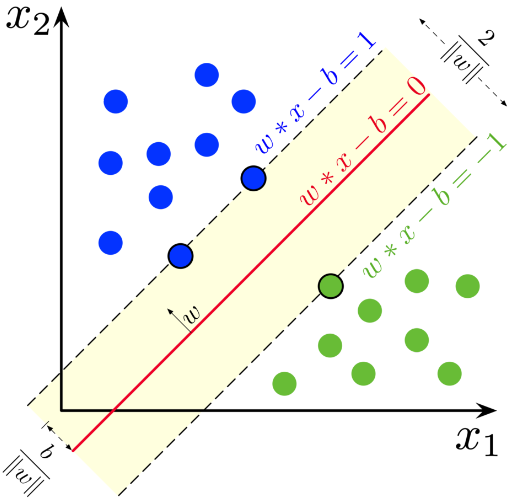


Figure 4.16: SVM trained with samples from two classes with the hyperplane with maximum margin (red) and margins (dashed lines). The samples laying on the dashed margin lines are the support vectors. (Image source: [Larhmam 18])

$\mathbf{x}_i + b) > 1$. Hence, only the nearest data points contribute to **w**. The training data values for which $\alpha_i > 0$ are defined as the support vectors. For defining the optimal hyperplane, only these support vectors are required. The support vectors can be considered as the “edge cases” in the classification or decision function to be learned. α_i can also be seen as a “difficulty rating” for the data sample $\mathbf{x}_i$, i. e. it describes how relevant a specific sample was for the determination of the hyperplane. [Boswell 02]

Once the optimal α (required to construct **w**) was estimated, b must be determined to fully describe the hyperplane. This was done by taking the “positive” and “negative” support vector $\mathbf{x}^+$ and $\mathbf{x}^-$ for which the following applies:

$$\begin{aligned}(\mathbf{w} \cdot \mathbf{x}^+ + b) &= +1 \\ (\mathbf{w} \cdot \mathbf{x}^- + b) &= -1\end{aligned} \tag{4.16}$$

Solving these equations gives:

$$b = -\frac{1}{2}\left(\mathbf{w} \cdot \mathbf{x}^+ + \mathbf{w} \cdot \mathbf{x}^-\right) \tag{4.17}$$

More detailed descriptions can be found in [Cortes 95, Burges 98, Boswell 02]

Kernel usage to map the input samples to other dimensions

When the two classes of a data set are not linearly separable, another approach is required to separate the data. In some cases, circle or polynomial curves might be used for better class separation. However, there are two disadvantages with such an approach. Firstly, it can be difficult to find the optimal curve. Secondly, one could not make use of the method described above to find the optimal hyperplane. The solution to this problem is to preprocess the data such that the data samples are transformed into another space which enables the definition of a hyperplane the same way as for the linear case. Therefore, a mapping function is defined as $\mathbf{z} = \phi(\mathbf{x})$ which transforms the d-dimensional input data vector **x** into a d'-dimensional vector **z**. The kernel is defined based on the mapping function ϕ. Many useful kernels were already discovered. The kernel should be chosen according to the spatial distribution of the data. For further explanation see [Boswell 02]. For the classification problem in this thesis, a cubic kernel was used.

4.8.2 Training, validation and deployment

All features were used to train the SVM classifier with a cubic kernel function. The following parameters were used:

- Acceptable cases (ACC): 16
- Not acceptable cases (NAC): 15
- Total cases: 31
- Cross validation: 4-fold cross validation

The classification results were analyzed using a confusion matrix and by the accuracy which was defined as:

$$\text{Accuracy} = \frac{\text{Correctly classified cases}}{\text{Total cases}} \cdot 100\% \tag{4.18}$$

Due to the low number of available labeled images, a holdout validation was not performed as this is only recommended for larger data sets. Training and evaluation of the SVM was performed using Matlab 2021b (MathWorks, USA).

The final step was to automate the whole IQA process. After an initial manual selection of the ROI, all features were automatically computed and a new image could be classified as either acceptable or not acceptable. Figure 4.17 summarizes the whole process.

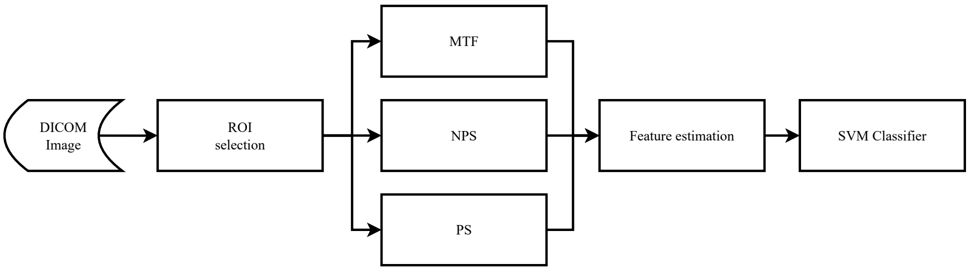
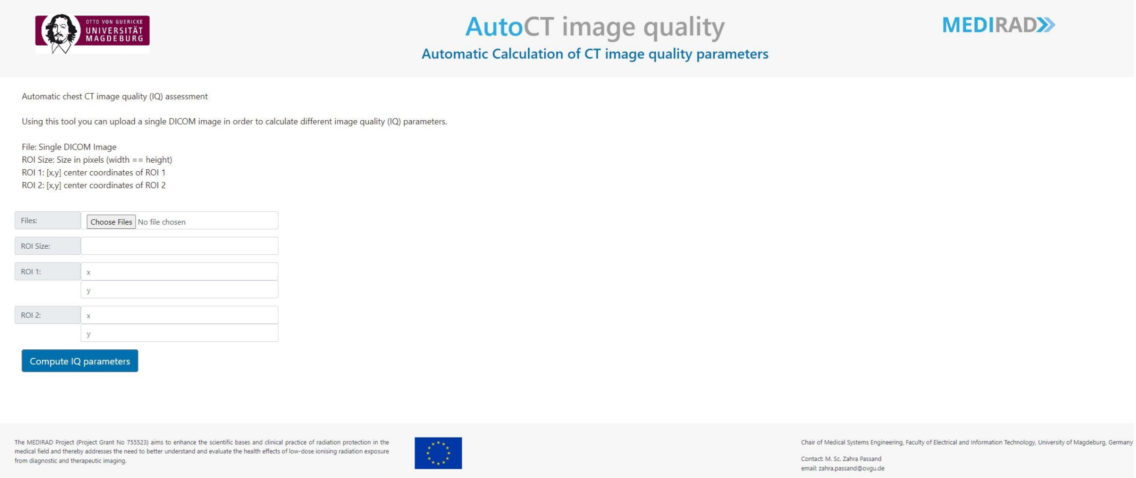


Figure 4.17: Summary of the IQA process. Once the ROIs were selected from the DICOM image, the various features were automatically computed and the image was classified using the SVM.

4.9 Web-based application for image quality assessment

To make the image quality assessment methods available to other users, a web-based application was developed. This application is currently running on a dedicated server. In the back-end of the server, an Octave instance is running which is performing the image quality assessment for an uploaded DICOM image. Once the user has specified the center coordinates of the ROIs to be analyzed, the Octave environment is called by the web server's back-end to compute the image quality parameters. The Python-based back-end of the server generates the web page (front-end), processes the user inputs, triggers Octave to run the scripts for image quality assessment and updates the web page with the results including different numeric values as well as graphs. Figure 4.18 shows a screenshot from the web interface.



AutoCT image quality

Automatic Calculation of CT image quality parameters

Automatic chest CT image quality (IQ) assessment

Using this tool you can upload a single DICOM image in order to calculate different image quality (IQ) parameters.

File: Single DICOM Image
ROI Size: Size in pixels (width == height)
ROI 1: [x,y] center coordinates of ROI 1
ROI 2: [x,y] center coordinates of ROI 2

The MEDIRAD Project (Project Grant No 755523) aims to enhance the scientific bases and clinical practice of radiation protection in the medical field and thereby addresses the need to better understand and evaluate the health effects of low-dose ionising radiation exposure from diagnostic and therapeutic imaging.

Chair of Medical Systems Engineering, Faculty of Electrical and Information Technology, University of Magdeburg, Germany

Contact: M. Sc. Zahra Passand
email: zahra.passand@ovgu.de

Figure 4.18: Screenshot of the web interface used to estimate the different image quality parameters. The center coordinates of the two ROIs needed to be entered for the uploaded DICOM image. ROI1 was used for the MTF and PS estimation and ROI2 was used for the NPS estimation. The back-end computed the different image quality features.

5

Results

This chapter summarizes the thesis' main results. Section 5.1 gives an overview of the image quality parameters obtained from the phantom study described in Section 4.3. In depth results for clinical CT images and the various image quality parameters are described in Sections 5.2 to 5.4. These sections cover the image analysis using MTF, NPS and the combined analysis of PS and MTF. The results of the CT image classification using these and the additionally derived features are described in Section 5.6.

5.1 Phantom image study

Phantom images were used in an initial step for investigating the effect of different CT scanner parameters on the calculated NPS of the image (see Section 4.3). Figure 5.1 shows the NPS derived from different phantom CT images. The effect of slice thickness, tube voltage and tube current on the NPS is visible in this graph. One can observe an increase of the low-frequency noise in the images acquired with lower tube currents ($10\,\mathrm{mA}$ or $50\,\mathrm{mA}$). When keeping the tube current and voltage constant, e. g. at $80\,\mathrm{mA}$ and $100\,\mathrm{kV}$, one can observe the effect of different slice thicknesses of $1\,\mathrm{mm}$ and $5\,\mathrm{mm}$ (green and orange plots), i. e. the noise decreases with increasing slice thickness. The same applied for a fixed tube voltage and slice thickness, e. g. $100\,\mathrm{kV}$ and $5\,\mathrm{mm}$, and tube currents

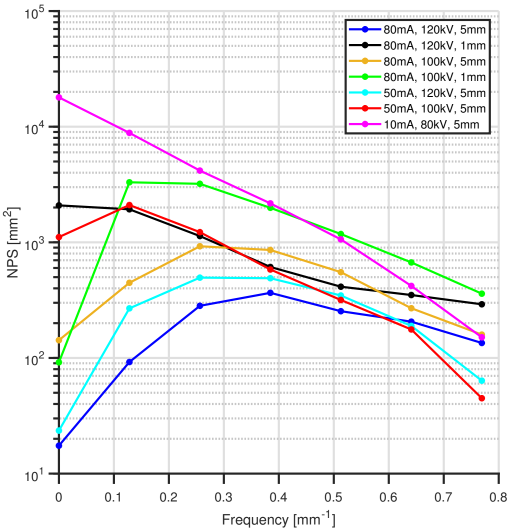
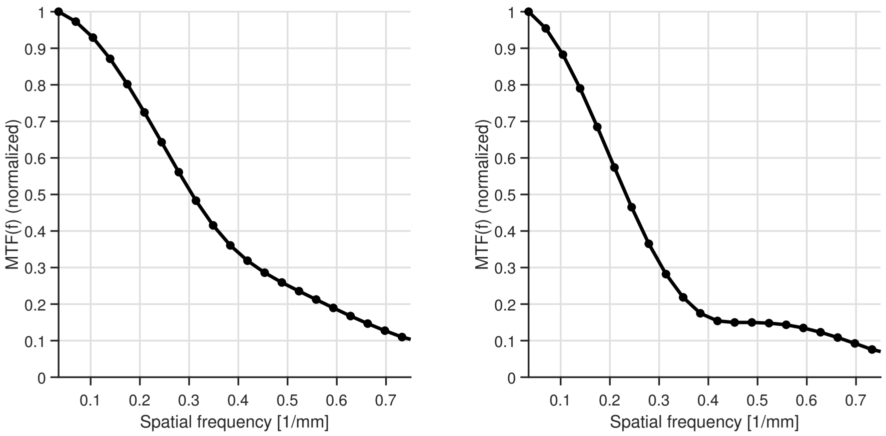


Figure 5.1: NPS spectra of the CIRS 701 Adult male phantom from images acquired with different scanning parameters. Tube current, tube voltage and slice thickness were altered. The lowest NPS values (or area under the NPS curve) were obtained from the measurement with the highest current, voltage and slice thickness.

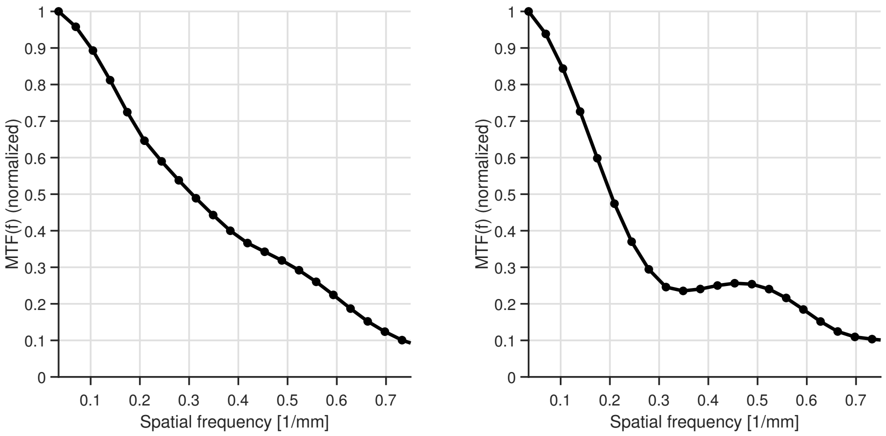
of $50\,\mathrm{mA}$ and $80\,\mathrm{mA}$ (red and orange plots). The effects were more pronounced in the lower frequency range $<0.4\,{}^{1}\!/\mathrm{mm}$ whereas for the higher frequency the expected behavior, e. g. lower NPS at higher tube currents, was partially inverted. The NPS of the image acquired with the lowest tube current ($10\,\mathrm{mA}$) (pink plot) had its maximum in the low frequency range, but it had lower NPS values in the high frequency range than other curves (black, green) which had higher tube voltages and tube currents but a lower slice thickness ($1\,\mathrm{mm}$).

5.2 Modulation transfer function

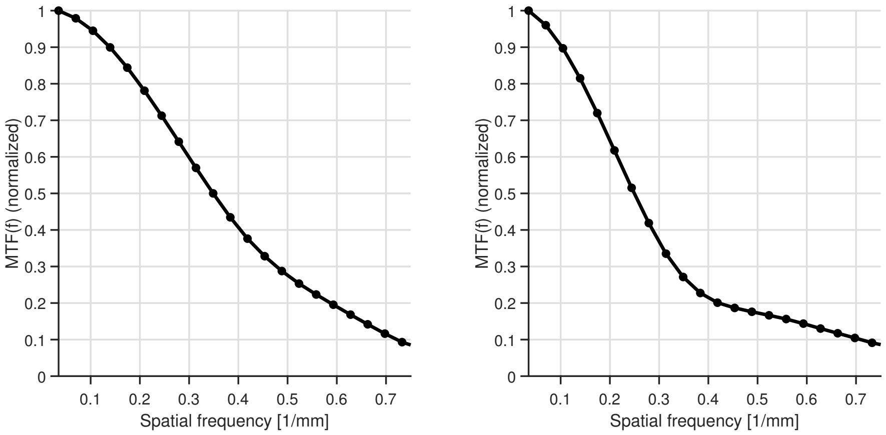
The MTFs of the different clinical CT images were compared by computing the integral under the MTF curves. The area under the MTF curve was computed for all CT images. Figure 5.2 shows six exemplary MTFs, three from images labeled as acceptable and three from images labeled as not acceptable. The MTFs of the acceptable CT images show a slower decay resulting in a larger area under the curve (0.31, 0.32, 0.34) than the not acceptable images (0.24, 0.25, 0.26). It was observed that there is a substantial difference in the MTF area when comparing the two image classes. For the acceptable images, the median of the MTF area was $m = 0.325$ with the first and third quartile being $q1 = 0.318$ and $q3 = 0.339$. For the not acceptable images, values were $m = 0.258$, $q1 = 0.251$ and $q3 = 0.285$. The corresponding box plots are shown in Section 5.5, Fig. 5.9a.



(a) Patient 21 (ACC) **(b)** Patient 18 (NAC)



(c) Patient 64 (ACC) **(d)** Patient 53 (NAC)

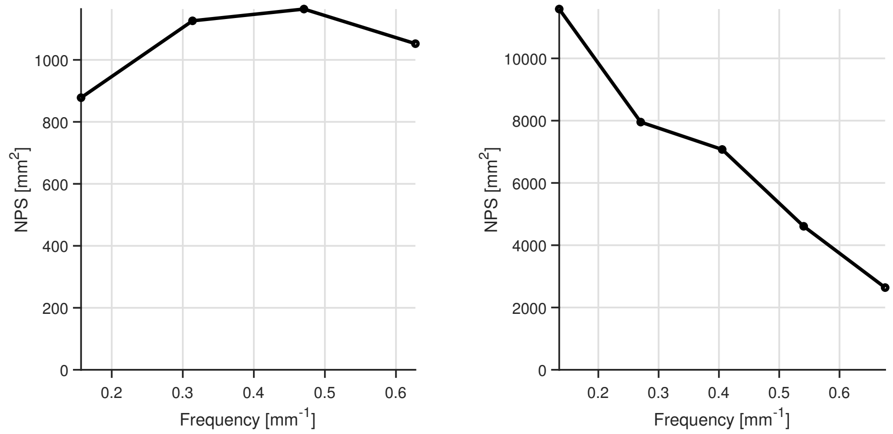


(e) Patient 67 (ACC) **(f)** Patient 56 (NAC)

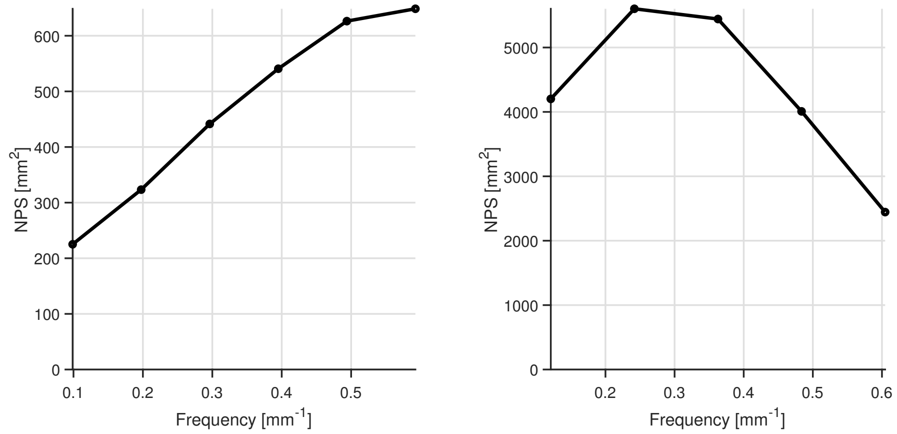
Figure 5.2: MTF for the images which were labeled as acceptable (a,c,e) and not acceptable (b,d,f).

5.3 Noise power spectrum

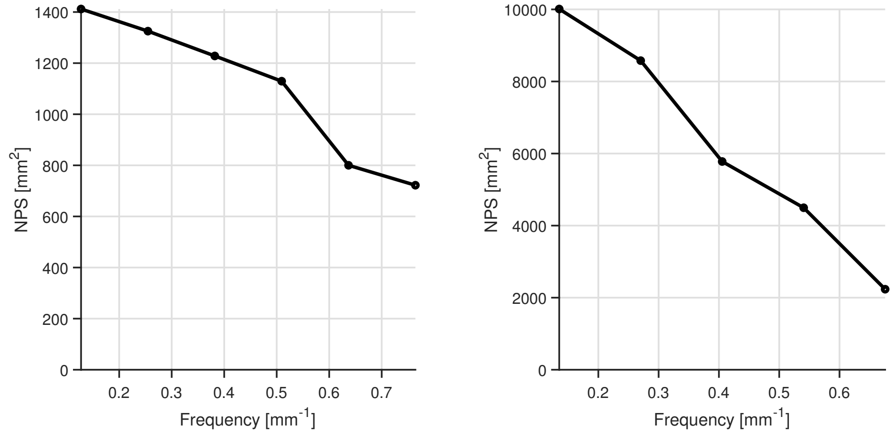
The NPS was computed for each image in a manually selected ROI which was dominated by a homogeneous structure. Before computing the NPS of the selected ROI, remaining structures in the area were separated from the homogeneous regions according to the descriptions in Section 4.5.2. Figure 5.3 shows three NPS graphs for each image class. Comparing these exemplary graphs, one can observe that the area under the curve is lower for the images labeled as acceptable. The NPS graphs shown here have a relatively low frequency resolution. The reason for this low resolution was the limited number of connected pixels (see Section 4.5.2) for a homogeneous region which was used for the computation of the NPS. Considering the NPS of all images analyzed in this work, the area under the NPS curves of the acceptable images had a median of $m = 2288$ with the first and third quartile being $q1 = 1742$ and $q3 = 3448$. For the not acceptable images, the corresponding median was $m = 5004$ and the first and third quartile were $q1 = 2718$ and $q3 = 8506$. The box plots of this analysis are shown in Section 5.5, Fig. 5.9c. This comparison showed that for the majority of the images analyzed in this work, the NPS was higher in the images labeled as not acceptable. Consequently, the NPS was used as another feature for image classification. Figure 5.4 shows the mean gray values in different ROIs and Fig. 5.5 shows the relation between the NPS and gray levels. It was observed that for the acceptable cases, there was a positive correlation between the mean gray values and the NPS within the ROI. For the not acceptable cases, the opposite was the case. Figure 5.11 in Section 5.5 shows the $\mathrm{NPS_A}$ normalized by the mean gray values within the ROI. Figure 5.13 shows the variance distributions in the different image classes.



(a) Patient 21 (ACC) **(b)** Patient 18 (NAC)

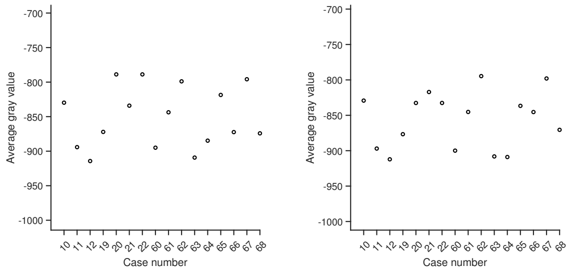


(c) Patient 64 (ACC) **(d)** Patient 53 (NAC)



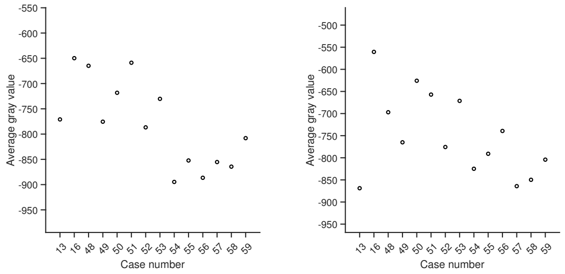
(e) Patient 67 (ACC) **(f)** Patient 56 (NAC)

Figure 5.3: NPS for the images which were labeled as acceptable (a,c,e) and not acceptable (b,d,f).



(a) Acceptable, homogeneous

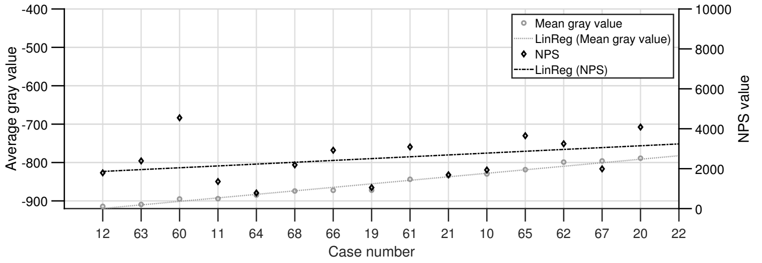
(b) Acceptable, structure



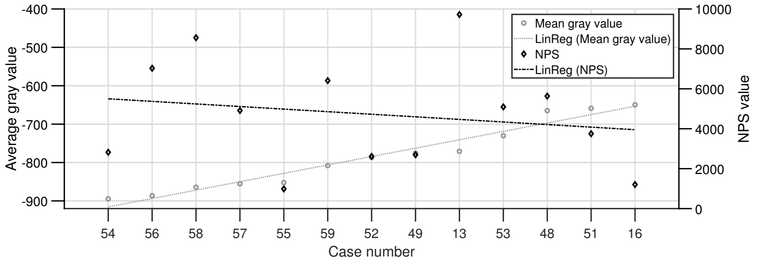
(c) Not acceptable, homogeneous

(d) Not acceptable, structure

Figure 5.4: Mean gray values in the different ROIs (homogeneous ROI and ROI containing major fissure) for the acceptable (a,b) and not acceptable (c,d) cases. The gray values for the acceptable and not acceptable cases are within the same range in the different ROIs. Case 18 (acceptable) is not shown in this plot since it was considered as an outlier.



(a) Acceptable cases

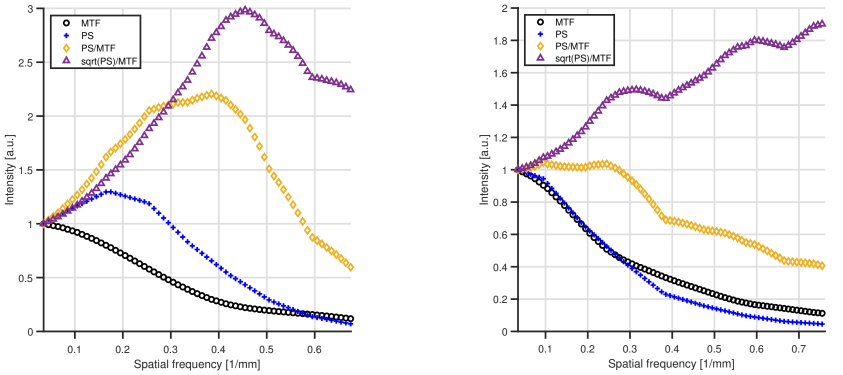


(b) Not acceptable cases

Figure 5.5: NPS and average gray levels computed in the selected ROIs of the acceptable (a) and not acceptable (b) images. The cases were ordered by increasing average gray values. A linear regression was plotted for each parameter to observe relations among them. Cases 18 (acceptable) and 50 (not acceptable) are not shown in this plot because they were considered as outliers.

5.4 Relation of power spectrum and modulation transfer function

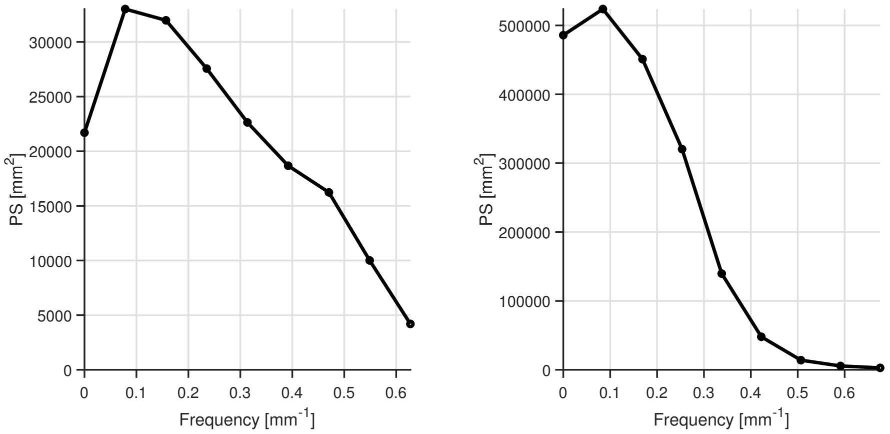
Figure 5.6 shows exemplary MTF, PS and PS/MTF curves for the two different image classes. The plots allow the analysis and comparison of the spatial resolution of a specific object boundary inside the ROI represented by the MTF and the overall spectral characteristics of the ROI represented by the PS. For the shown acceptable case, the MTF has slower decay with increasing frequency, i. e. the spatial resolution is higher than in the not acceptable case. The PS of the acceptable case shows an initial increase at the lower frequency range and decreases at higher frequencies. For the not acceptable, the PS decreases with increasing frequency. Figure 5.7 shows the PS of six different images. The peak location, the average slope and the energy distribution with respect to the frequency are different between the two image classes. Figure 5.8 shows the PS/MTF ratios obtained from images labeled as acceptable and not acceptable. In each of the two classes, the graphs show certain class-specific characteristics. Although the following findings do not hold for all images analyzed, the following main characteristics were observed for the images from the class labeled as acceptable: *1)* The images have more energy in the high frequency range than in the lower frequency range (LF/HF area); *2)* The average slope of the graph is lower; *3)* The location of the maximum is located in the higher frequency range. The corresponding box plots are shown in Section 5.5, Fig. 5.10. Figure 5.12 shows the area under the PS curve (PS_A) normalized by the mean gray values within the ROI.



(a) Parameters for an CT image rated as acceptable

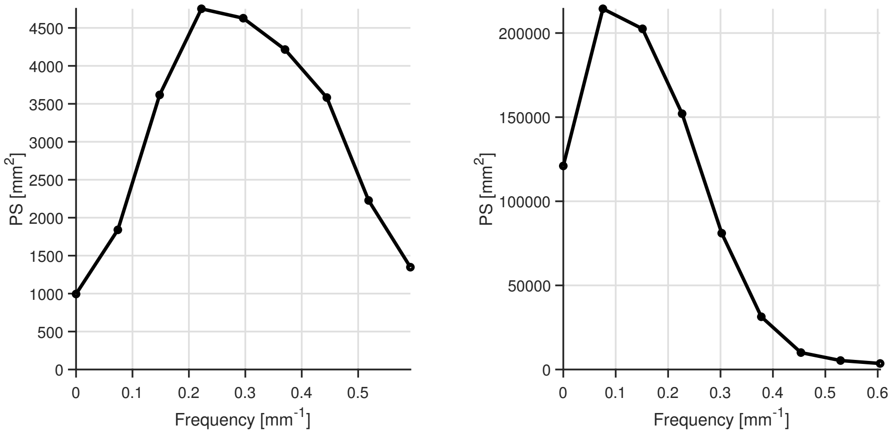
(b) Parameters for an CT image rated as not acceptable

Figure 5.6: MTF, PS and their ratios computed from two different patient CT images in the area of the major fissure. While the MTF of both image classes shows a similar behavior, the PS has different characteristics. The PS of the image rated as acceptable shows higher energies in the central frequency range whereas the PS energies of the not acceptable image decrease with increasing frequency. (All graphs were normalized to one for the lowest frequency value for better comparison within a single plot.)



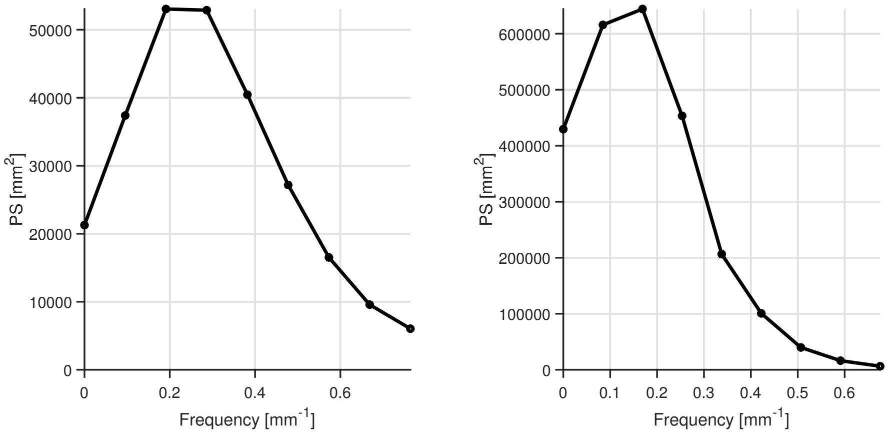
(a) Patient 21 (ACC)

(b) Patient 18 (NAC)



(c) Patient 64 (ACC)

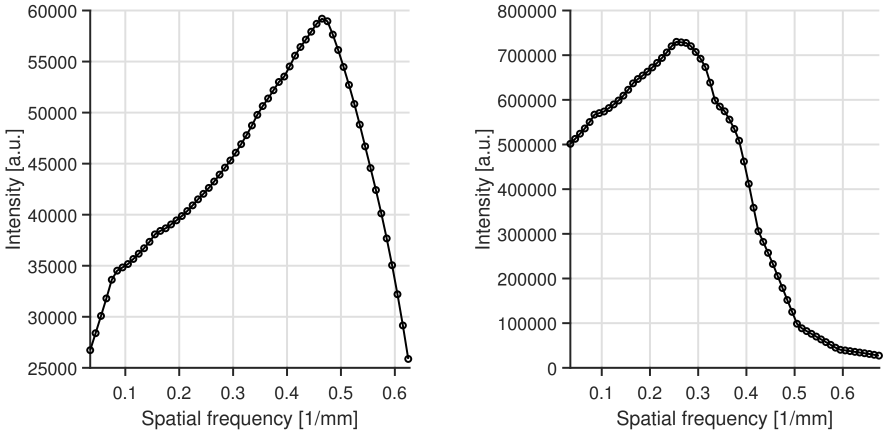
(d) Patient 53 (NAC)



(e) Patient 67 (ACC)

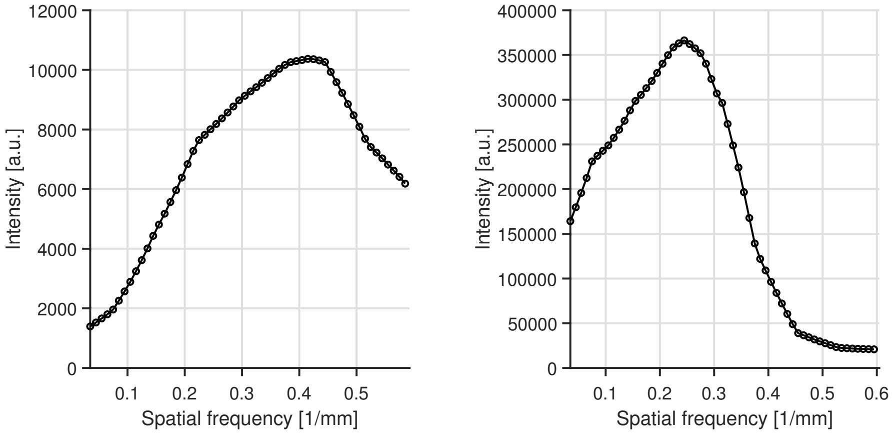
(f) Patient 56 (NAC)

Figure 5.7: PS of the images which were labeled as acceptable (a,c,e) and not acceptable (b,d,f).



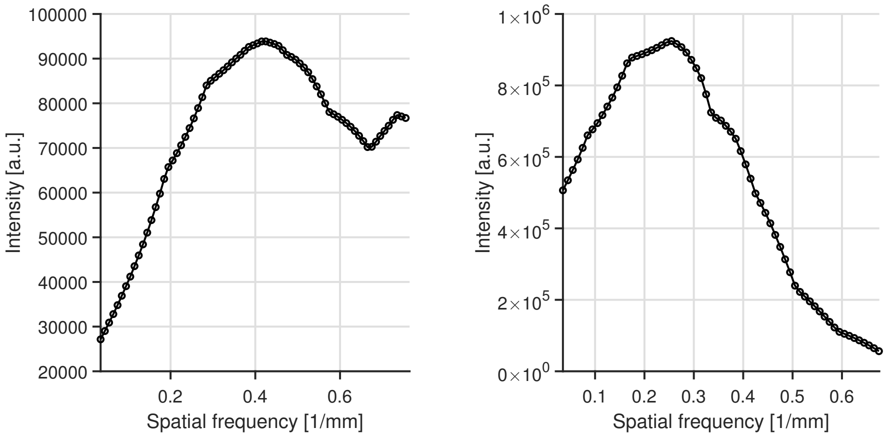
(a) Patient 21 (ACC)

(b) Patient 18 (NAC)



(c) Patient 64 (ACC)

(d) Patient 53 (NAC)



(e) Patient 67 (ACC)

(f) Patient 56 (NAC)

Figure 5.8: PS/MTF ratios of the images which were labeled as acceptable (a,c,e) and not acceptable (b,d,f).

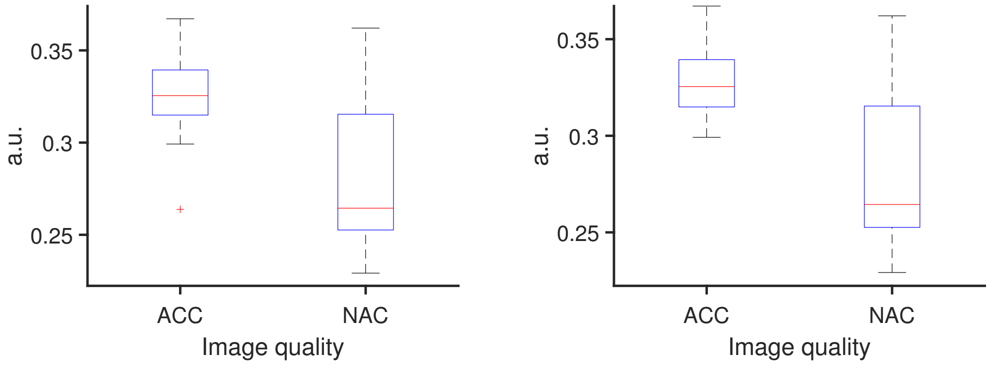
5.5 Statistical analysis

The statistical analysis of the different features shown in Table 4.2 revealed that the features either had a normal distribution or another distribution which is not specified (the distribution of each feature was analyzed separately for each image quality class). The tests revealed that for each feature, there was a significant difference between the two image quality classes. The results of the statistical tests are summarized in Table 5.1.

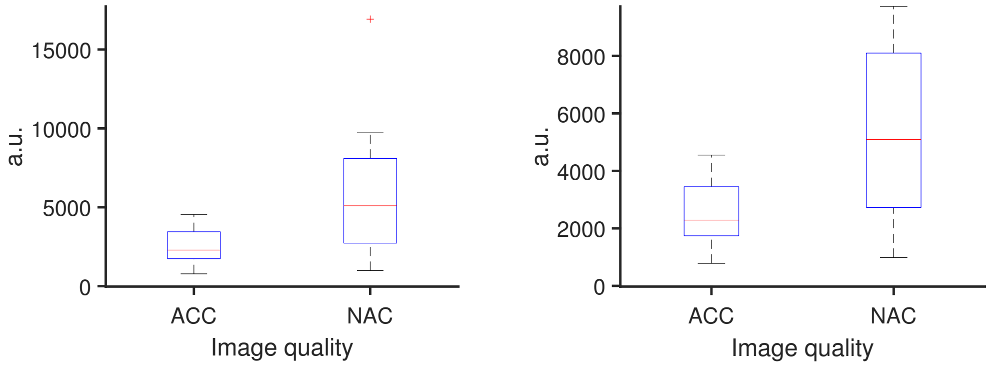
Figures 5.9 to 5.13 show the box plots of the different features for each image quality class. For a better understanding of the plots, the plots are shown with and without the outliers in the feature data. Based on the visual analysis of the plots, e. g. by comparing the median value of a feature between the two classes, it can be observed that the features could be used to separate the two classes, although a single feature was not sufficient for a reliable classification.

Table 5.1: Statistical test results. The distribution of each feature (compare Table 4.2) was tested first to check which method to use for testing the statistical significance. Based on the type of the distribution (normal or other), either an unpaired t-test (for the normal distribution) or a Mann-Whitney U test (for another distribution) was applied to test if a given feature was significantly different between the two image quality classes. All features showed a significant difference for the two classes.

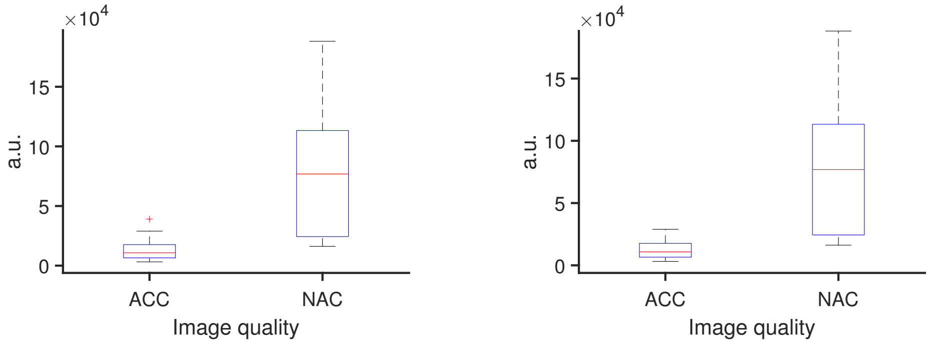
Feature	Name	Distribution	Applied statistical test	p-values	Significant difference
1	$\mathrm{MTF_A}$	Other	Mann-Whitney U test	0.0032	Yes
2	$\mathrm{NPS_A}$	Normal	Unpaired t-test	0.0063	Yes
3	$\mathrm{PS_A}$	Normal	Unpaired t-test	0.00001	Yes
4	$\mathrm{PS/MTF_{LF/HF\ ratio}}$	Normal	Unpaired t-test	0.0001	Yes
5	$\mathrm{PS/MTF_{Slope}}$	Other	Mann-Whitney U test	0.0004	Yes
6	$\mathrm{PS/MTF_{Max\ loc}}$	Other	Mann-Whitney U test	0.0066	Yes
7	$\sigma^2_{\mathrm{Fissure}}$	Normal	Unpaired t-test	0.000008	Yes



(a) MTF$_A$ **(b)** MTF$_A$ without outliers

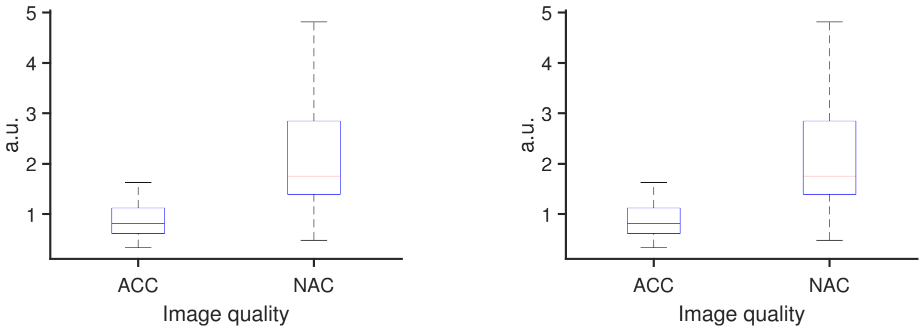


(c) NPS$_A$ **(d)** NPS$_A$ without outliers

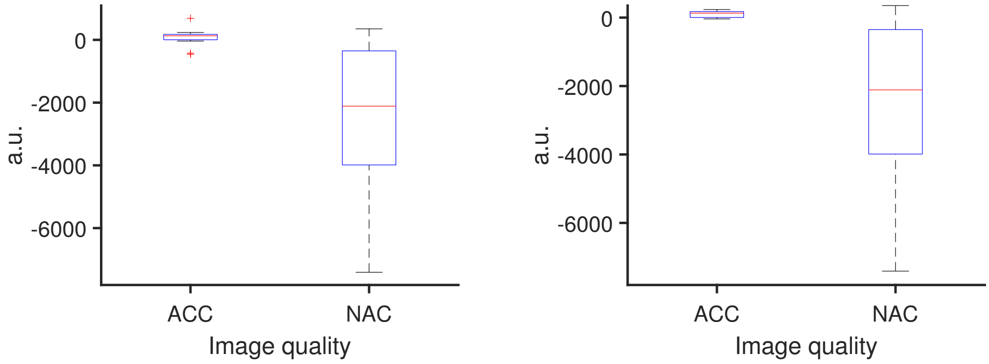


(e) PS$_A$ **(f)** PS$_A$ without outliers

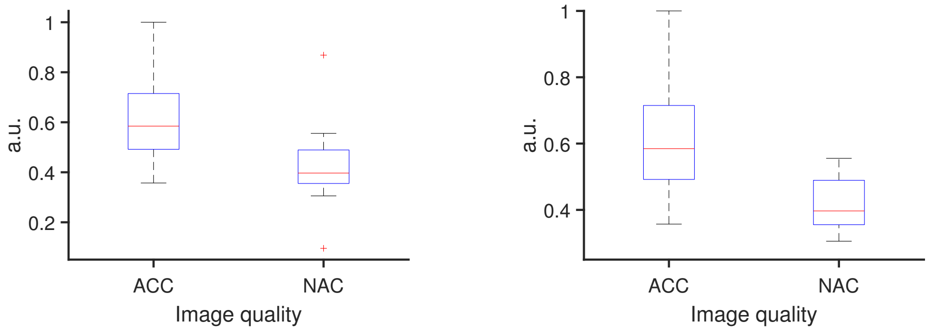
Figure 5.9: Box plots comparing the MTF$_A$, NPS$_A$ and PS$_A$ in acceptable and not acceptable images. (a, c, e) including outliers and (b, d, f) without showing the outliers.



(a) PS/MTF$_{\text{LF/HF ratio}}$ **(b)** PS/MTF$_{\text{LF/HF ratio}}$ without outliers

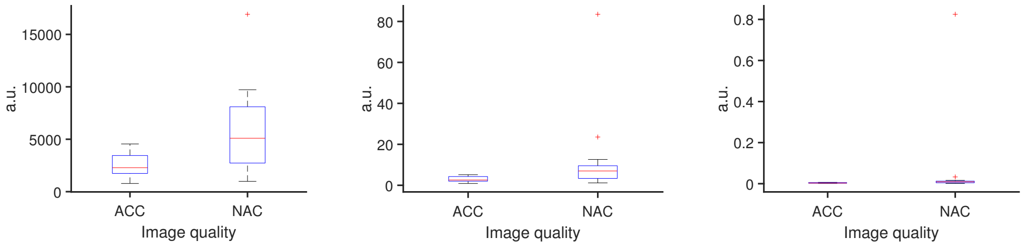


(c) PS/MTF$_{\text{Slope}}$ **(d)** PS/MTF$_{\text{Slope}}$ without outliers



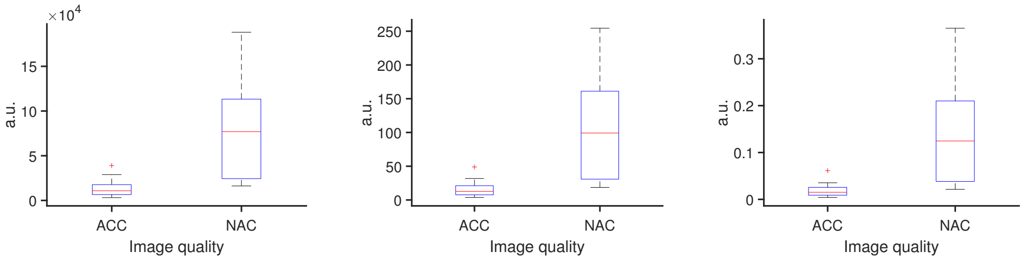
(e) PS/MTF$_{\text{Max loc}}$ **(f)** PS/MTF$_{\text{Max loc}}$ without outliers

Figure 5.10: Box plots comparing PS/MTF$_{\text{LF/HF ratio}}$, PS/MTF$_{\text{Slope}}$ and PS/MTF$_{\text{Max loc}}$ in acceptable and not acceptable images. (a, c, e) including outliers and (b, d, f) without showing the outliers.



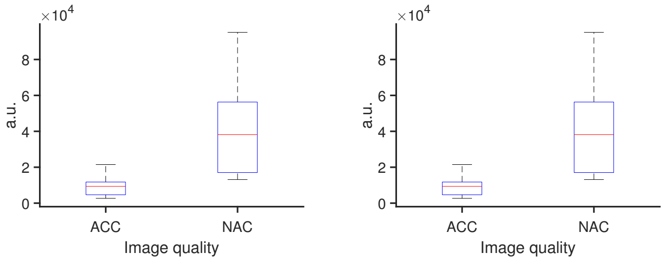
(a) NPS$_\mathrm{A}$ **(b)** $\frac{\mathrm{NPS_A}}{\bar{I}_i}$ **(c)** $\frac{\mathrm{NPS_A}}{\bar{I}_i^{\,2}}$

Figure 5.11: Box plots comparison (a) the area under the 1D-NPS curve NPS$_\mathrm{A}$, (b) the NPS$_\mathrm{A}$ normalized by the gray mean value $\bar{I}_i$ and (c) the NPS$_\mathrm{A}$ normalized by the gray mean value squared $\bar{I}_i^{\,2}$ of the ROI used to compute the NPS. Explanation of the box plot: minimum (lower whisker), first quartile (lower blue line), median (red line in the center of the box), third quartile (upper blue line) and maximum (upper whisker). Outliers are marked by "+"



(a) PS$_\mathrm{A}$ **(b)** $\frac{\mathrm{PS_A}}{\bar{I}_i}$ **(c)** $\frac{\mathrm{PS_A}}{\bar{I}_i^{\,2}}$

Figure 5.12: Box plots comparing (a) the area under the 1D-PS curve PS$_\mathrm{A}$, (b) the PS$_\mathrm{A}$ normalized by the gray mean value $\bar{\mathrm{I}}_i$ and (c) the PS$_\mathrm{A}$ normalized by the gray mean value squared $\bar{I}_i^{\,2}$ of the ROI used to compute the PS.

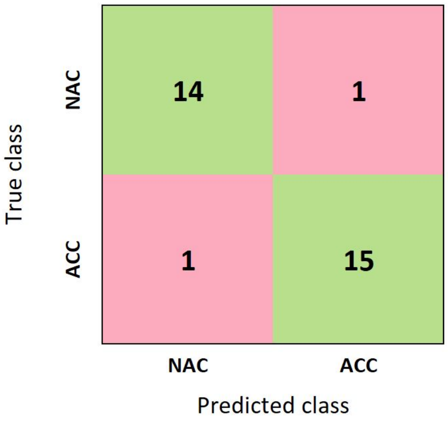


(a) $\sigma^2_\mathrm{Fissure}$ **(b)** $\sigma^2_\mathrm{Fissure}$ without outlier

Figure 5.13: Box plots of the $\sigma^2_\mathrm{Fissure}$ feature in acceptable and not acceptable images (a) with and (b) without showing the outliers.

5.6 Image classification

Based on the different feature results described in Sections 5.2 to 5.5, all features were used to train the classifier described in Section 4.8. An overall accuracy of $93.5\,\%$ was obtained using the 4-fold cross validation. Figure 5.14 shows the confusion matrix using this classifier.



True class \ Predicted class	NAC	ACC
NAC	14	1
ACC	1	15

Figure 5.14: Confusion matrix for the SVM classifier using 4-fold cross validation. The “True Class” corresponds to the annotations made by the radiologists and the “Predicted Class” is the output of the classifier. This means that from the 16 cases labeled as acceptable (True class), 15 were correctly classified and 1 case was wrongly classified. Also, from 15 cases labeled as not acceptable, 14 were correctly and 1 case was wrongly classified.

6

Discussion

This chapter discusses the results and findings made in this thesis. In Section 6.1, the phantom measurements are discussed. The different features used to assess the image quality as well as the classification approach are discussed in Section 6.2. Section 6.3 discusses the manual selection of the ROI which is still required to perform the image quality assessment at a specific region or anatomical structure.

6.1 Quality analysis – phantom CT images

The analysis of the phantom measurement showed the expected behavior. The huge advantages of using a phantom for assessing the quality of an imaging system are the clearly defined structures and the material composition. These characteristics enable the analysis of various imaging parameters used for the image acquisition. Phantoms typically contain homogeneous regions which can be used for assessing the noise in an image. A trend was observable in the analyzed image data that the noise decreased with increasing tube currents. Due to the lower tube current, less photons are present per irradiated volume, leading to an increased quantum noise. With increasing slice thickness, the detected X-rays are averaged over a larger volume while reconstructing a certain voxel of the image, which results in a reduction of the quantum noise, especially at the higher frequency components.

6.2 Quality analysis – patient CT images

While the phantom based image quality analysis is well established for various imaging techniques, the usage of real patient images involves several challenges and problems. First of all, the anatomy of each patient is different and so are the structures visible in the image, i. e. there are no clearly defined structures at certain locations as they exist in phantom images. Due to the well defined construction of a phantom and the known technical parameters, changes in image quality are only related to the imaging system and its configured parameters. Due to the large inter-patient variabilities, the perceived image quality in these real patient images is affected by the imaging system, its parameters, the patient and of course the observer, i. e. the radiologist.

Several patient-related parameters affecting the image quality were described in Section 2.2.2. This can be tissue composition, body mass index, implants, cardiac and respiratory motion or pathological tissue variations. These alterations affect the diagnostic value of the CT image. In such a case, the parameters of the CT scanner can be optimized whereas the pathological tissue alterations such as a fibrosis of the lung can still limit the diagnostic value of the CT image. This needs to be considered when collecting and labeling clinical images, i. e. it needs to be distinguished if the image quality was affected by parameters which are controllable (e. g. beam quality or patient motion) or not controllable (e. g. anatomical or pathological variations).

When compared to phantom images, there is no quantitative image quality metric for clinical patient CT images. The image quality of patient images is mainly defined by the perception of the radiologist rather than quantitative measures. The goal of this thesis was to establish a technique which uses similar methods and metrics as those applied to phantom images to generate a quantitative image quality metric. The analysis was solely performed on the images without considering any metadata such as reconstruction technique, tube parameters or patient properties such as weight or body mass index. The goal was to relate image quality to the subjective perception of the human observer, i. e. the radiologist, rather than analyzing the influence of technical parameters on the resulting image. Furthermore, the amount of labeled image data was not sufficient for separating the data into different sub-sets according to their technical parameters. In the following, these methods and their application to clinical CT images is discussed.

Modulation transfer function

The MTF of an imaging system, e. g. of a CT scanner, is usually estimated by analyzing edges or other structures in dedicated phantoms (compare Section 4.4). However, real clinical patient images neither exhibit such well defined structures nor is their location inside the image known. One clearly defined boundary in a patient CT image is the air-skin boundary. This boundary was used in the work of [Sanders 16]. While this approach was robust with respect to the detection of the boundary and showed reliable results in terms of the estimated MTF, it was limited by the fact that the boundary was far from the anatomical structure which was relevant for the diagnostic purpose. The major fissure which is closer to the center of the image can be affected by various other parameters or artifacts such as noise, scatter or motion. Hence, it was important to compute the MTF as well as the other parameters directly at the anatomical structure.

In order to implement the MTF analysis, the main challenge was to identify the edges of the major fissure. Since the MTF estimation was based on an edge analysis of the major fissure, an initial edge detection was required. The region growing algorithm used to find the boundaries of the object, i. e. the outer edges of the major fissure, produced reliable results. This is remarkable since the major fissure is in general difficult to distinguish due to its low contrast and narrow shape. Once the boundaries of the major fissure were identified, line profiles were automatically estimated (see Section 4.4.2). The semi-automatic approach enabled a convenient way of estimating the MTF of an object with a complex, irregularly shaped structure. This method for detecting object boundaries could be also applicable to other anatomical structures such as bones, vessels or boundaries of adjacent organs.

Comparing the MTFs obtained from the two different image quality classes, differences in the MTF spectra could be observed, i. e. the MTF of the images labeled as acceptable had larger amplitudes in the higher frequency range of the MTF spectrum. As a result of this observation, the area under the MTF curves was significantly higher for the acceptable images when compared to the not acceptable ones.

Although it was not possible to differentiate between the various factors affecting the patient CT image and consequently the MTF, relevant differences were observed in the MTF curves and the derived parameter (area under the curve) in the two different image classes. In general, the results of the MTF analysis were in good agreement with the subjective image annotations performed by radiologists, i. e. there were significant differences between the MTF values of the acceptable and not acceptable with median values of the acceptable cases being higher than the values of the not acceptable cases. This means that the MTF can be used as an indicator for image quality assessment in clinical CT images. However, the area under the MTF curve alone was not sufficient for a classification since there were overlaps of the MTF area values between the two classes.

Noise power spectrum

Computing the NPS from a real patient image containing various types of structures and being mostly inhomogeneous is a challenging task. For the estimation of the NPS, an additional ROI was selected which was close to the analyzed anatomical structure, i. e. the major fissure. Care had to be taken during this manual selection since it was necessary to select a region which contained as few visible structures as possible.

The main difficulty of the approach to compute an NPS from a patient CT image was the fact that there is no completely homogeneous region inside the lung (or other tissue) whose pixel variations are only affected by noise. One common method to obtain a homogeneous region is to subtract two adjacent image slices. In this way, the structures which are very similar in the adjacent slices are removed while the noise remains (with doubled amplitude). However, the subtraction image still contains edges or structures since the anatomy in the two adjacent slices is not exactly the same. This requires further processing steps to detect remaining edge structures in the subtracted images in order to avoid a wrong estimation of the NPS. Because of these reasons, the subtraction technique was not considered and the identification of homogeneous regions was directly performed on the original CT image slice.

The important initial step was to identify a connected region within the manually selected ROI which has homogeneous properties. This was achieved by a k-means clustering approach using various features containing information about the spatial variance inside the ROI and its sub-regions. The largest square shaped region found by the clustering approach was used for computing the NPS. Due to local invariances in the majority of the ROIs of all images, the identified homogeneous regions were relatively small, e. g. 6 x 6 or 7 x 7 pixels. Due to their limited size, the resolutions of the 1D-NPS obtained from the 2D-image region was very low (compare Fig. 5.3).

Even though it was shown that a homogeneous area could be identified using the clustering method, this approach did not guarantee that the identified regions did not contain anatomical structures. This could be one explanation why the NPS in some of the images labeled as acceptable had a high value.

Although the number of pixels and the resulting resolution of the 1D-NPS were very limited, significant differences were observed when the NPS values of the different image classes were compared. This in turn means, that for the images analyzed in this work, the NPS was a helpful and expedient metric for the assessment of CT image quality. Other existent approaches for identifying homogeneous regions in CT images rely on other custom parameters, e. g. on a dedicated structure coherence feature [Chun 15].

The average gray level were compared to the NPS within the ROIs. Different trends or correlations were observed for the two image quality classes, i. e. there was a positive correlation between the gray level and the NPS in acceptable images whereas the opposite was the case for the not acceptable cases. It was not possible to identify a clear reason for this observation but it is possible that due to the low amount of available images, other observation would be made with a larger number of images.

Ratio of power spectrum and modulation transfer function

By computing the PS of the same ROI which was used for the MTF, the frequency characteristics of an area containing anatomical structures as well as noise could be analyzed. Clearly visible differences in the frequency spectra between the acceptable and not acceptable images were observed, with different locations of the maximum amplitude, different energy distributions or slopes of the PS. The PS of the images with an acceptable quality had more energy in the higher frequency range. This could be an indication that the high contrast structures or sharp object boundaries, which are represented by the higher frequency components, were better preserved in the images which were labeled as acceptable. The MTF of the not acceptable images showed a faster decay of the amplitude for higher frequencies. By relating the PS of an ROI to the MTF of the same ROI, the overall spatial frequencies were related to the spatial resolution of the ROI. The parameters derived from the PS to MTF ratio were used as new metrics for image quality assessment.

Image classification

By analyzing the box plots of the different features from the two different quality classes, differences were clearly visible. Using different test approaches it was shown that there were significant differences for each feature between the different classes. Various features derived from the MTF, the PS and the NPS were used to train a classifier. Based on the available and labeled CT images, promising classification results with an accuracy of $93.5\,\%$ were achieved using the 4-fold cross validation. Due to the limited amount of labeled CT images, there was no additional test set available which could have been used to test the classifier on unseen image data. This is a limitation of the current results which should be solved in the future by adding more labeled CT images to the data set. More data would also enable the training and usage of other methods for image quality classification, e. g. deep neural networks.

Another limitation of the image data used for this thesis was the limited number of radiologists annotating the image data. Hence, it has to be assumed that the annotations are biased by the subjective perception or experience of the radiologist. This should be overcome in the future by providing the same images to more radiological experts for annotation.

6.3 Manual selection of a region of interest

A manual ROI selection was still necessary since the image quality analysis was performed in a very specific region of the CT image. Such a region specific, local analysis had the advantage that distortions in other parts of the image, which are not relevant for the diagnosis, did not affect the estimated quality level. Separating the major fissure from the other surrounding structures inside the lung is a challenging task. In recent work, automated methods were proposed for the automated segmentation of the pulmonary lobes or the direct segmentation of the major fissure [Imran 20,Peng 22]. Those or similar approaches might be used in future works for enabling a fully automated image quality assessment based on the analysis of the major fissure. Except of the currently required, manual ROI selection, all other processing steps for analyzing the selected ROI were fully automated and did not require further user interaction.

7

Summary and outlook

7.1 Summary and conclusions

Monitoring the quality of CT images during clinical routine is beneficial for various reasons. Besides maintaining a high level of diagnostic quality, a long-term quality assessment could be used to track the quality of clinical CT images and to identify potential, systematic problems in the image acquisition process. Additional information about the dose and its relation to the estimated image quality could also reveal potential possibilities to reduce the radiation dose used in CT scanning.

For establishing such a system in clinical routine, several challenges need to be overcome. One of the main challenges is to develop fully automatic image quality assessment methods, i. e. methods which do not require any interaction with the user. Another challenge is to obtain and label CT images for the training of such algorithms. To tackle these challenges, this thesis has endeavored to develop methods and approaches for analyzing the quality of CT images in specific regions. The methods were implemented and evaluated on clinical patient CT images annotated by radiologists.

To measure the quality of a CT image, various parameters exist. Commonly used parameters are the MTF, the NPS, contrast measurements, and others. To compute these parameters, certain requirements need to be fulfilled. For example, the MTF computation requires the edges of certain structures or objects within the image. For computing the NPS, homogeneous regions are required. Automatically identifying these regions in clinical CT images for fulfilling the mentioned requirements is a challenging task due to the inhomogenity caused by various anatomical structures, special diseases and the interpatient variability of the size, material composition or location of an organ. In addition to computing the parameters, there is no clear consensus about how to actually rate the quality of an image using these parameters. This is due to the fact that there are no clearly defined thresholds which determine the quality of a CT image.

Within this thesis, different conventional image quality metrics as well as additional derived parameters were computed using clinical CT images. Radiologists annotated the quality of these images in an application specific region (major fissure of the left lung), using different image quality criteria, e. g the noise or clear visibility of the structure. The image quality parameters were computed for the region covering the major fissure. The region containing the major fissure was selected manually. This was the only manual step required to estimate the image quality parameters. The following processing steps were fully automated, e. g. for finding the edges or boundaries of the major fissure or identifying a homogeneous region for the estimation of the NPS. The different parameters are described in Chapter 4.

Using the annotations made by the radiologists and the estimated image quality parameters, a classifier was trained for image quality assessment. The good results on the validation set showed the potential of this approach to combine various image quality parameters to finally assess the quality of clinical CT images.

Within the thesis, the following approaches were developed and results were achieved:

- a dedicated study was initiated in which the subjective image quality of various thorax CT images was evaluated by radiologists for different clinical indications,
- a semi-automated approach for assessing the quality of clinical CT images using different image quality metrics such as NPS, MTF and PS was developed,
- the developed algorithms were evaluated using the annotated, clinical thorax CT images,
- a good correlation between the subjective image quality evaluation of the radiologists and the algorithm based approach was shown,

- the trained classifier showed that IQA was possible using the computed features,
- the developed methods enabled a semi-automatic evaluation of the physics based image quality parameters using patient CT images.

7.2 Future work

Future works should include additional annotated CT images with additional information including the diseases and the ROIs considered for the annotation. Such information can be used to improve the performance of the trained classifiers enabling a more precise characterization of CT image quality. Several expert radiologists should annotate the same images to reduce the subjective bias of individual annotations (inter- and intra- annotator variability). With an increasing number of labeled images, other methods such as deep convolutional neural networks could be trained and deployed for IQA.

One of the remaining challenges for the employment of the developed method is to automatically identify a specific ROI, e. g. a region containing the major fissure. Several research works aim to segment specific anatomical regions or to identify certain landmarks [Zhou 06, Özsavaş 14, Wang 22]. If those methods produce reliable results, a fully automated, region-specific image quality assessment would be possible. Instead of only considering one ROI, multiple ROIs for a certain diagnostic purpose could be considered at the same time to gain additional information about the image quality. In general, the approach of performing a local image quality assessment is helpful if there are distortions in parts of the image which are not relevant for the diagnostic application. This can be the case for cardiac motion artifacts causing image blurring in the region of the heart. Performing a global image quality analysis in such a case could lead to the assumption that the image is of low quality due to the degradation caused by the blurring in a region which is not relevant for this current analysis or diagnosis.

Another challenging task was to identify a region close to the ROI which was suitable for an NPS analysis. The main challenge was to identify homogeneous regions, i. e. regions which did not contain anatomical structures. Due to the high density of anatomical structures it was difficult to identify connected regions containing a sufficient number of pixels to perform meaningful NPS analysis. This approach can make the NPS analysis in such images or regions prone to errors. Hence, future work should be investigate if other, more homogeneous regions within the same image (e. g. larger regions composed of fat tissue) could be used instead.

To gain additional information and performance over time, such a system should be implemented in clinical routine. It should include continuous input from expert radiologists annotating images or correcting the output of such a system to improve the quality of the trained classifier on a long term perspective. The results of such a system could be used for the quality management within the radiology department in order to identify potential ways to reduce dose or to find problems with image acquisition systems or parameter settings.

Bibliography

[Alshipli 17] Marwan Alshipli & Norlaili A Kabir. *Effect of slice thickness on image noise and diagnostic content of single-source-dual energy computed tomography*. In J Phys Conf Ser, volume 851. IOP Publishing, 2017. (Cited on page 10)

[Anam 18] Choirul Anam, Toshioh Fujibuchi, Wahyu Setia Budi, Freddy Haryanto & Geoff Dougherty. *An algorithm for automated modulation transfer function measurement using an edge of a PMMA phantom: Impact of field of view on spatial resolution of CT images*. J Appl Clin Med Phys, vol. 19, no. 6, pages 244–252, 2018. (Cited on pages 35 and 44)

[Anderson 52] Theodore W Anderson & Donald A Darling. *Asymptotic theory of certain goodness of fit criteria based on stochastic processes*. Ann Math Stat, pages 193–212, 1952. (Cited on page 56)

[Anderson 62] Theodore W Anderson. *On the distribution of the two-sample Cramer-von Mises criterion*. Ann Math Stat, pages 1148–1159, 1962. (Cited on page 56)

[Azzam 01] Edouard I Azzam, Sonia M De Toledo & John B Little. *Direct evidence for the participation of gap junction-mediated intercellular communication in the transmission of damage signals from α-particle irradiated to nonirradiated cells*. Proc Natl Acad Sci USA, vol. 98, no. 2, pages 473–478, 2001. (Cited on page 23)

[Bampis 17] Christos G Bampis, Praful Gupta, Rajiv Soundararajan & Alan C Bovik. *SpEED-QA: Spatial efficient entropic differencing for image and video quality*. IEEE Signal Process Lett, vol. 24, no. 9, pages

1333–1337, 2017. (Cited on page 30)

[Barrett 04] Julia F Barrett & Nicholas Keat. *Artifacts in CT: recognition and avoidance*. Radiographics, vol. 24, no. 6, pages 1679–1691, 2004. (Cited on pages 15, 16, 17, 18, 19, 20 and 21)

[Båth 07] M Båth & LG Mansson. *Visual grading characteristics (VGC) analysis: a non-parametric rank-invariant statistical method for image quality evaluation*. Br J Radiol, vol. 80, no. 951, pages 169–176, 2007. (Cited on page 30)

[Båth 10] Magnus Båth. *Evaluating imaging systems: practical applications*. Radiat Prot Dosim, vol. 139, no. 1-3, pages 26–36, 2010. (Cited on page 29)

[Beister 12] Marcel Beister, Daniel Kolditz & Willi A Kalender. *Iterative reconstruction methods in X-ray CT*. Physica Med, vol. 28, no. 2, pages 94–108, 2012. (Cited on page 13)

[Bianco 18] Simone Bianco, Luigi Celona, Paolo Napoletano & Raimondo Schettini. *On the use of deep learning for blind image quality assessment*. Signal Image Video P, vol. 12, no. 2, pages 355–362, 2018. (Cited on page 35)

[Boedeker 07] Kirsten L Boedeker, Virgil N Cooper & Michael F McNitt-Gray. *Application of the noise power spectrum in modern diagnostic MDCT: part I. Measurement of noise power spectra and noise equivalent quanta*. Phys Med Biol, vol. 52, no. 14, page 4027, 2007. (Cited on page 32)

[Boone 12] John M Boone, JA Brink, S Edyvean, W Huda, W Leitz, CH McCollough, MF McNitt-Gray, P Dawson, PLM Deluca, SM Seltzer*et al.* *Radiation dose and image-quality assessment in computed tomography*. J ICRU, vol. 12, no. 1, pages 9–149, 2012. (Cited on pages 49 and 50)

[Bosse 16] Sebastian Bosse, Dominique Maniry, Thomas Wiegand & Wojciech Samek. *A deep neural network for image quality assessment*. In 2016 IEEE International Conference on Image Processing (ICIP), pages 3773–3777. IEEE, 2016. (Cited on page 35)

[Boswell 02] Dustin Boswell. *Introduction to Support Vector Machines*. Departement of Computer Science and Engineering University of California San Diego, 2002. (Cited on pages 58, 59 and 60)

[Brombal 19] Luca Brombal, Fulvia Arfelli, Pasquale Delogu, Sandro Donato, Giovanni Mettivier, Koen Michielsen, Piernicola Oliva, Angelo Taibi, Ioannis Sechopoulos, Renata Longo*et al.* *Image quality comparison between a phase-contrast synchrotron radiation breast CT and a clinical breast CT: a phantom based study*. Sci Rep, vol. 9, no. 1, pages 1–12, 2019. (Cited on page 50)

[Burges 98] Christopher JC Burges. *A tutorial on support vector machines for pattern recognition*. Data mining and knowledge discovery, vol. 2, no. 2, pages 121–167, 1998. (Cited on pages 59 and 60)

[Bushberg 11] Jerrold T Bushberg & John M Boone. The essential physics of medical imaging. Lippincott Williams & Wilkins, 2011. (Cited on pages 43, 49 and 50)

[Carton 05] Ann-Katherine Carton, Dirk Vandenbroucke, Luc Struye, Andrew DA Maidment, Yen-Hong Kao, Michael Albert, Hilde Bosmans & Guy Marchal. *Validation of MTF measurement for digital mammography quality control*. Med Phys, vol. 32, no. 6Part1, pages 1684–1695, 2005. (Cited on page 35)

[Catphan 21] Catphan. *Catphan 605 Product Guide*. The Phantom Laboratory Incorporated, 2021. (Cited on page 44)

[Catphan 22] Catphan. *Catphan 500 and 600 Product Guide*. The Phantom Laboratory Incorporated, 2022. (Cited on page 44)

[Chan 01] Tony F Chan & Luminita A Vese. *Active contours without edges*. IEEE Trans Image Process, vol. 10, no. 2, pages 266–277, 2001. (Cited on page 46)

[Chaudhary 16] SR Chaudhary, LC Smith, V Chidambaram*et al.* *A resident's one-poster guide to common CT artifacts*. In European Congress of Radiology-ECR 2016, Vienna, Austria, 2016. (Cited on page 19)

[Cheng 19] Yuan Cheng, Ehsan Abadi, Taylor Brunton Smith, Francesco Ria, Mathias Meyer, Daniele Marin & Ehsan Samei. *Validation of algorithmic CT image quality metrics with preferences of radiologists*. Med Phys, vol. 46, no. 11, pages 4837–4846, 2019. (Cited on pages 1 and 34)

[Chow 17] Li Sze Chow & Heshalini Rajagopal. *Modified-BRISQUE as no reference image quality assessment for structural MR images*. Magn Reson Imaging, vol. 43, pages 74–87, 2017. (Cited on page 31)

[Christe 10] Andreas Christe, Patricia Flach, Steffen Ross, Danny Spendlove, Stephan Bolliger, Peter Vock & Michael J Thali. *Clinical radiology and postmortem imaging (Virtopsy) are not the same: specific and unspecific postmortem signs*. Legal medicine, vol. 12, no. 5, pages 215–222, 2010. (Cited on page 20)

[Christianson 12] Olav Christianson, Xiang Li, Donald Frush & Ehsan Samei. *Automated size-specific CT dose monitoring program: assessing variability in CT dose*. Med Phys, vol. 39, no. 11, pages 7131–7139, 2012. (Cited on page 28)

[Christianson 15] Olav Christianson, James Winslow, Donald P Frush & Ehsan Samei. *Automated technique to measure noise in clinical CT examinations*. AJR, vol. 205, no. 1, pages W93–W99, 2015. (Cited on page 33)

[Chun 15] Minsoo Chun, Young Hun Choi & Jong Hyo Kim. *Automated measurement of CT noise in patient images with a novel structure coherence feature*. Phys Med Biol, vol. 60, no. 23, page 9107, 2015. (Cited on pages 23, 31 and 81)

[Commission 14] Joint Commission. *New/revised diagnostic imaging standards from Joint Commission*. J Nucl Med, vol. 55, no. 13, 2014. (Cited on page 28)

[Cortes 95] Corinna Cortes & Vladimir Vapnik. *Support-vector networks*. Machine learning, vol. 20, no. 3, pages 273–297, 1995. (Cited on pages 59 and 60)

[Council 06] National Research Council. Health risks from exposure to low levels of ionizing radiation: BEIR VII phase 2. National Academies Press, 2006. (Cited on page 23)

[D'Agostino 73] Ralph D'Agostino & Egon S Pearson. *Tests for departure from normality. Empirical results for the distributions of b_2 and $\sqrt{b}$*. Biometrika, vol. 60, no. 3, pages 613–622, 1973. (Cited on page 56)

[Dainty 76] John Christopher Dainty, Rodney Shaw & LJ Cutrona. *Image science: principles, analysis and evaluation of photographic-type imaging processes*. Phys Today, vol. 29, no. 1, page 71, 1976. (Cited on page 50)

[Doel 15] Tom Doel, David J Gavaghan & Vicente Grau. *Review of automatic pulmonary lobe segmentation methods from CT*. Comput Med Imaging Graphics, vol. 40, pages 13–29, 2015. (Cited on pages 24 and 25)

[Dolly 16] Steven Dolly, Hsin-Chen Chen, Mark Anastasio, Sasa Mutic & Hua Li. *Practical considerations for noise power spectra estimation for clinical CT scanners*. J Appl Clin Med Phys, vol. 17, no. 3, pages 392–407, 2016. (Cited on page 32)

[Dong 12] Frank Dong, William Davros, Jessica Pozzuto & Janet Reid. *Optimization of kilovoltage and tube current–exposure time product based on abdominal circumference: an oval phantom study for pediatric abdominal CT*. AJR, vol. 199, no. 3, pages 670–676, 2012. (Cited on page 9)

[Dost 22] Shahi Dost, Faryal Saud, Maham Shabbir, Muhammad Gufran Khan, Muhammad Shahid & Benny Lovstrom. *Reduced reference image and video quality assessments: review of methods*. EURASIP J Image Video Process, vol. 2022, no. 1, pages 1–31, 2022. (Cited on pages 30 and 31)

[Elgazzar 06] Abdelhamid H Elgazzar. The pathophysiologic basis of nuclear medicine. Springer Science & Business Media, 2006. (Cited on pages 22 and 23)

[EU 13] EU. *Richtlinie 2013/59/Euratom des Rates vom 5. Dezember 2013 zur Festlegung grundlegender Sicherheitsnormen für den Schutz vor den Gefahren einer Exposition gegenüber ionisierender Strahlung und zur Aufhebung der Richtlinien 89/618/Euratom, 90/641/Euratom, 96/29/Euratom, 97/43/Euratom und 2003/122/Euratom*. Guideline NLE 2011/0254, Rat der Europäischen Union, 2013. (Cited on page 2)

[Fay 10] Michael P Fay & Michael A Proschan. *Wilcoxon-Mann-Whitney or t-test? On assumptions for hypothesis tests and multiple interpretations of decision rules*. Stat Surv, vol. 4, page 1, 2010. (Cited on page 56)

[Feuchtner 10] Gudrun M Feuchtner, Daniel Jodocy, Andrea Klauser, Bernhard Haberfellner, Iman Aglan, Alexander Spoeck, Stefan Hiehs, Peter Soegner & Werner Jaschke. *Radiation dose reduction by using 100-kV tube voltage in cardiac 64-slice computed tomography: a comparative study*. Eur J Radiol, vol. 75, no. 1, pages e51–e56, 2010. (Cited on page 9)

[Geyer 15] Lucas L Geyer, U Joseph Schoepf, Felix G Meinel, John W Nance Jr, Gorka Bastarrika, Jonathon A Leipsic, Narinder S Paul, Marco Rengo, Andrea Laghi & Carlo N De Cecco. *State of the art: iterative CT reconstruction techniques*. Radiology, vol. 276, no. 2, pages 339–357, 2015. (Cited on pages 11 and 13)

[Gircys 18] Michael Gircys & Brian J Ross. *Image evolution using 2d power spectra*. Complexity, vol. 2019, 2018. (Cited on page 56)

[Gnannt 12] Ralph Gnannt, Anna Winklehner, Daniel Eberli, Alexander Knuth, Thomas Frauenfelder & Hatem Alkadhi. *Automated tube potential selection for standard chest and abdominal CT in follow-up patients with testicular cancer: comparison with fixed tube potential*. Eur Radiol, vol. 22, no. 9, pages 1937–1945, 2012. (Cited on page 9)

[Goldman 07a] Lee W Goldman. *Principles of CT: radiation dose and image quality*. J Nucl Med Technol, vol. 35, no. 4, pages 213–225, 2007. (Cited on page 28)

[Goldman 07b] Lee W Goldman. *Principles of CT: radiation dose and image quality*. J Nucl Med Technol, vol. 35, no. 4, pages 213–225, 2007. (Cited on page 49)

[Goldman 08] Lee W Goldman. *Principles of CT: multislice CT*. J Nucl Med Technol, vol. 36, no. 2, pages 57–68, 2008. (Cited on page 9)

[Goodall 14] Todd Goodall & Alan C Bovik. *No-reference task performance prediction on distorted LWIR images*. In 2014 Southwest Symposium on Image Analysis and Interpretation, pages 89–92. IEEE, 2014. (Cited on page 31)

[Greffier 19] J Greffier, A Larbi, J Frandon, G Moliner, JP Beregi & F Pereira. *Comparison of noise-magnitude and noise-texture across two generations of iterative reconstruction algorithms from three manufacturers*. Diagn Interventional Imaging, vol. 100, no. 7-8, pages 401–410, 2019. (Cited on page 32)

[Hsieh 09] J. Hsieh. Computed Tomography: Principles, Design, Artifacts, and Recent Advances. SPIE Press, 2nd edition, 2009. (Cited on pages 5, 6, 8, 16, 22, 24 and 43)

[ICRU 12] ICRU. *ICRU Report No. 87: Radiation dose and image-quality assessment in computed tomography*. J ICRU, vol. 12, pages 1–149, 2012. (Cited on page 36)

[Imran 20] Abdullah-Al-Zubaer Imran, Ali Hatamizadeh, Shilpa P Ananth, Xiaowei Ding, Nima Tajbakhsh & Demetri Terzopoulos. *Fast and automatic segmentation of pulmonary lobes from chest CT using a progressive dense V-network*. Comput Methods Biomech Biomed Eng: Imaging Vis, vol. 8, no. 5, pages 509–518, 2020. (Cited on page 82)

[Iyer 00] Rashi Iyer & Bruce E Lehnert. *Factors underlying the cell growth-related bystander responses to α particles*. Cancer Res., vol. 60, no. 5, pages 1290–1298, 2000. (Cited on page 23)

[Jardin 86] M Jardin & J Remy. *Segmental bronchovascular anatomy of the lower lobes: CT analysis*. American Journal of Roentgenology, vol. 147, no. 3, pages 457–468, 1986. (Cited on page 24)

[Jarque 80] Carlos M Jarque & Anil K Bera. *Efficient tests for normality, homoscedasticity and serial independence of regression residuals*. Econ Lett, vol. 6, no. 3, pages 255–259, 1980. (Cited on page 56)

[Kalender 00] W.A. Kalender. Computertomographie. Grundlagen, Gerätetechnologie, Bildqualität, Anwendungen. Publicis MCD 2000, 2000. (Cited on pages 6 and 7)

[Katkar 16] Rujuta Katkar, Douglas D Steffy, Marcel Noujeim, S Thomas Deahl II & Hassem Geha. *The effect of milliamperage, number of basis images,* *and export slice thickness on contrast-to-noise ratio and detection of mandibular canal on cone beam computed tomography scans: an in vitro study*. Oral Surg Oral Med Oral Pathol Oral Radiol, vol. 122, no. 5, pages 646–653, 2016. (Cited on page 10)

[Larhmam 18] Larhmam. *SVM margin.png*. Online, Licensed under CC BY-SA 4.0, `https://en.wikipedia.org/wiki/File:SVM_margin.png`, October 2018. (Cited on page 59)

[Li 10] Chaofeng Li & Alan C Bovik. *Content-partitioned structural similarity index for image quality assessment*. Signal Process Image Commun, vol. 25, no. 7, pages 517–526, 2010. (Cited on page 30)

[Malkus 17] Annelise Malkus & Timothy P Szczykutowicz. *A method to extract image noise level from patient images in CT*. Med Phys, vol. 44, no. 6, pages 2173–2184, 2017. (Cited on page 32)

[Massey Jr 51] Frank J Massey Jr. *The Kolmogorov-Smirnov test for goodness of fit*. J Am Stat Assoc, vol. 46, no. 253, pages 68–78, 1951. (Cited on page 56)

[McCollough 12] Cynthia H McCollough, Guang Hong Chen, Willi Kalender, Shuai Leng, Ehsan Samei, Katsuyuki Taguchi, Ge Wang, Lifeng Yu & Roderic I Pettigrew. *Achieving routine submillisievert CT scanning: report from the summit on management of radiation dose in CT*. Radiology, vol. 264, no. 2, pages 567–580, 2012. (Cited on page 9)

[Mohammadi 14] Pedram Mohammadi, Abbas Ebrahimi-Moghadam & Shahram Shirani. *Subjective and objective quality assessment of image: A survey*. arXiv preprint arXiv:1406.7799, 2014. (Cited on page 28)

[Morgan 03] William F Morgan. *Non-targeted and delayed effects of exposure to ionizing radiation: II. Radiation-induced genomic instability and bystander effects in vivo, clastogenic factors and transgenerational effects*. Radiat Res, vol. 159, no. 5, pages 581–596, 2003. (Cited on page 23)

[Myers 87] Kyle J Myers & Harrison H Barrett. *Addition of a channel mechanism to the ideal-observer model*. J Opt Soc Am A, vol. 4, no. 12, pages 2447–2457, 1987. (Cited on page 35)

[Nagel 07] Hans Dieter Nagel. *CT parameters that influence the radiation dose*. In Radiation dose from adult and pediatric multidetector computed tomography, pages 51–79. Springer, 2007. (Cited on page 10)

[Özsavaş 14] Emin Emrah Özsavaş, Ziya Telatar, Bahar Dirican, Ömer Sağer & Murat Beyzadeoğlu. *Automatic segmentation of anatomical structures from CT scans of thorax for RTP*. Comput Math Methods Med, vol. 2014, 2014. (Cited on page 85)

[Park 17] Ji Eun Park, Young Hun Choi, Jung-Eun Cheon, Woo Sun Kim, In-One Kim, Hyun Suk Cho, Young Jin Ryu & Yu Jin Kim. *Image quality and radiation dose of brain computed tomography in children: effects of decreasing tube voltage from 120 kVp to 80 kVp*. Pediatr Radiol, vol. 47, no. 6, pages 710–717, 2017. (Cited on page 9)

[Passand 20] Zahra Passand & Christoph Hoeschen. *Image quality assessment of real patient thorax CT images using modulation transfer function and noise power spectrum*. In Medical Imaging 2020: Physics of Medical Imaging, volume 11312, page 113122N. International Society for Optics and Photonics, 2020. (Cited on pages 32 and 33)

[Peng 22] Yuanyuan Peng, Pengpeng Luan, Hongbin Tu, Xiong Li & Ping Zhou. *Pulmonary Fissure Segmentation in CT Images Based on ODoS Filter and Shape Features*. arXiv preprint arXiv:2201.09163, 2022. (Cited on page 82)

[Piccini 20] Davide Piccini, Robin Demesmaeker, John Heerfordt, Jérôme Yerly, Lorenzo Di Sopra, Pier Giorgio Masci, Juerg Schwitter, Dimitri Van De Ville, Jonas Richiardi, Tobias Kober*et al.* *Deep learning to automate reference-free image quality assessment of whole-heart MR images*. Radiol Artif Intell, vol. 2, no. 3, 2020. (Cited on page 35)

[Polesel 00] Andrea Polesel, Giovanni Ramponi & V John Mathews. *Image enhancement via adaptive unsharp masking*. IEEE Trans Image Process, vol. 9, no. 3, pages 505–510, 2000. (Cited on page 46)

[Prasad 95] Kedar N Prasad. Handbook of radiobiology. CRC press, 1995. (Cited on page 23)

[Ptrump16 22] Ptrump16. *CT scan Iterative reconstruction (left) versus fil-*

tered backprojection (right).jpg. Online, Licensed under CC BY-SA 4.0, `https://commons.wikimedia.org/wiki/File:CT_scan_Iterative_reconstruction_(left)_versus_filtered_backprojection_(right).jpg`, February 2022. (Cited on page 15)

[Racine 18] Damien Racine, Nick Ryckx, Alexandre Ba, Fabio Becce, Anais Viry, Francis R Verdun & Sabine Schmidt. *Task-based quantification of image quality using a model observer in abdominal CT: a multicentre study*. Eur Radiol, vol. 28, no. 12, pages 5203–5210, 2018. (Cited on pages 35 and 36)

[Raman 13a] Siva P Raman, Pamela T Johnson, Swati Deshmukh, Mahadevappa Mahesh, Katharine L Grant & Elliot K Fishman. *CT dose reduction applications: available tools on the latest generation of CT scanners*. J Am Coll Radiol, vol. 10, no. 1, pages 37–41, 2013. (Cited on page 9)

[Raman 13b] Siva P Raman, Mahadevappa Mahesh, Robert V Blasko & Elliot K Fishman. *CT scan parameters and radiation dose: practical advice for radiologists*. Journal of the American College of Radiology, vol. 10, no. 11, pages 840–846, 2013. (Cited on pages 9 and 10)

[Ramesh 96] R Ramesh, AJ Marrogi, A Munshi, CN Abboud & SM Freeman. *In vivo analysis of the bystander effect: a cytokine cascade*. Exp Hematol, vol. 24, no. 7, pages 829–838, 1996. (Cited on page 23)

[Revel 15] Marie-Pierre Revel, Isabelle Fitton, Etienne Audureau, Joseph Benzakoun, Mathieu Lederlin, Marie-Laure Chabi & Pascal Rousset. *Breast dose reduction options during thoracic CT: influence of breast thickness*. AJR, vol. 204, no. 4, pages W421–W428, 2015. (Cited on page 27)

[RSB 13] RSB. *CT Scanner – Inner Workings Revealed*. `http://www.robotspacebrain.com/ct-scanner-revealed/`, 2013. (Cited on page 7)

[Saini 04] Sanjay Saini. *Multi-detector row CT: principles and practice for abdominal applications*. Radiology, vol. 233, no. 2, pages 323–327, 2004. (Cited on page 8)

[Sanders 16] Jeremiah Sanders, Lynne Hurwitz & Ehsan Samei. *Patient-specific* *quantification of image quality: an automated method for measuring spatial resolution in clinical CT images*. Med Phys, vol. 43, no. 10, pages 5330–5338, 2016. (Cited on pages 34, 36 and 79)

[Schofield 20] R Schofield, L King, U Tayal, I Castellano, J Stirrup, F Pontana, James Earls & E Nicol. *Image reconstruction: Part 1–understanding filtered back projection, noise and image acquisition*. J Cardiovasc Comput Tomogr, vol. 14, no. 3, pages 219–225, 2020. (Cited on pages 11 and 12)

[Schörner 12] Karsten Schörner. *Development of methods for scatter artifact correction in industrial X-ray cone-beam computed tomography*. PhD thesis, Technische Universität München, 2012. (Cited on page 16)

[Schulze 11] Ralf Schulze, Ulrich Heil, D Groβ, Dan Dominik Bruellmann, Egor Dranischnikow, Ulrich Schwanecke & Elmar Schoemer. *Artefacts in CBCT: a review*. Dentomaxillofacial Radiology, vol. 40, no. 5, pages 265–273, 2011. (Cited on page 16)

[Shapiro 65] Samuel Sanford Shapiro & Martin B Wilk. *An analysis of variance test for normality (complete samples)*. Biometrika, vol. 52, no. 3/4, pages 591–611, 1965. (Cited on page 56)

[Shapiro 72] Samuel S Shapiro & RS Francia. *An approximate analysis of variance test for normality*. J Am Stat Assoc, vol. 67, no. 337, pages 215–216, 1972. (Cited on page 56)

[Skornitzke 18] S Skornitzke. *Iterative algorithms for artifact reduction in computed tomography*. Der Radiologe, vol. 58, no. 3, pages 202–210, 2018. (Cited on page 13)

[Smith-Bindman 12] Rebecca Smith-Bindman, Diana L Miglioretti, Eric Johnson, Choonsik Lee, Heather Spencer Feigelson, Michael Flynn, Robert T Greenlee, Randell L Kruger, Mark C Hornbrook, Douglas Roblin*et al.* *Use of diagnostic imaging studies and associated radiation exposure for patients enrolled in large integrated health care systems, 1996-2010*. Jama, vol. 307, no. 22, pages 2400–2409, 2012. (Cited on page 23)

[Solomon 12] Justin B Solomon, Olav Christianson & Ehsan Samei. *Quantitative comparison of noise texture across CT scanners from different manu-*

facturers. Med Phys, vol. 39, no. 10, pages 6048–6055, 2012. (Cited on page 50)

[Soundararajan 11] Rajiv Soundararajan & Alan C Bovik. *RRED indices: Reduced reference entropic differencing for image quality assessment*. IEEE Trans Image Process, vol. 21, no. 2, pages 517–526, 2011. (Cited on page 30)

[Stiller 18] Wolfram Stiller. *Basics of iterative reconstruction methods in computed tomography: a vendor-independent overview*. Eur J Radiol, vol. 109, pages 147–154, 2018. (Cited on pages 13 and 14)

[StrSchG 17] StrSchG. *Gesetz zum Schutz vor der schädlichen Wirkung ionisierender Strahlung*. Bundesamt für Justiz, 2017. (Cited on page 2)

[Suzuki 03] Keiji Suzuki, Mitsuaki Ojima, Seiji Kodama & Masami Watanabe. *Radiation-induced DNA damage and delayed induced genomic instability*. Oncogene, vol. 22, no. 45, pages 6988–6993, 2003. (Cited on page 23)

[Tian 16] Xiaoyu Tian & Ehsan Samei. *Accurate assessment and prediction of noise in clinical CT images*. Med Phys, vol. 43, no. 1, pages 475–482, 2016. (Cited on pages 33 and 34)

[Vendis 18] Tomas Vendis. *Modern CT scanner located at the Lochotín University Hospital in Pilsen, Czech Republic*. Online, Licensed under CC BY-SA 4.0, `https://en.wikipedia.org/wiki/CT_scan`, October 2018. (Cited on page 7)

[Wang 04] Zhou Wang, Alan C Bovik, Hamid R Sheikh & Eero P Simoncelli. *Image quality assessment: from error visibility to structural similarity*. IEEE Trans Image Process, vol. 13, no. 4, pages 600–612, 2004. (Cited on pages 30 and 35)

[Wang 05] Lipo Wang. Support vector machines: theory and applications, volume 177. Springer Science & Business Media, 2005. (Cited on page 58)

[Wang 11] Zhou Wang. *Applications of objective image quality assessment methods*. IEEE Signal Process Mag, vol. 28, no. 6, pages 137–142, 2011.

(Cited on page 28)

[Wang 21] Lanjiang Wang. *A survey on IQA*. arXiv preprint arXiv:2109.00347, 2021. (Cited on page 28)

[Wang 22] Tong Wang, Haiqun Xing, Yige Li, Sicong Wang, Ling Liu, Fang Li & Hongli Jing. *Deep learning-based automated segmentation of eight brain anatomical regions using head CT images in PET/CT*. BMC Med Imaging, vol. 22, no. 1, pages 1–10, 2022. (Cited on page 85)

[Yao 92] Jie Yao & Harrison H Barrett. *Predicting human performance by a channelized Hotelling observer model*. In Mathematical Methods in Medical Imaging, volume 1768, pages 161–168. SPIE, 1992. (Cited on page 35)

[Yao 16] Yuan Yao, Joshua M Ng, Alec J Megibow & Norbert J Pelc. *Image quality comparison between single energy and dual energy CT protocols for hepatic imaging*. Med Phys, vol. 43, no. 8Part1, pages 4877–4890, 2016. (Cited on page 10)

[Zhang 11] Lin Zhang, Lei Zhang, Xuanqin Mou & David Zhang. *FSIM: A feature similarity index for image quality assessment*. IEEE Trans Image Process, vol. 20, no. 8, pages 2378–2386, 2011. (Cited on page 30)

[Zhang 14] Yi Zhang, Anush K Moorthy, Damon M Chandler & Alan C Bovik. *C-DIIVINE: No-reference image quality assessment based on local magnitude and phase statistics of natural scenes*. Signal Process Image Commun, vol. 29, no. 7, pages 725–747, 2014. (Cited on page 31)

[Zhou 06] Xiangrong Zhou, Tatsuro Hayashi, Takeshi Hara, Hiroshi Fujita, Ryujiro Yokoyama, Takuji Kiryu & Hiroaki Hoshi. *Automatic segmentation and recognition of anatomical lung structures from high-resolution chest CT images*. Comput Med Imaging Graph, vol. 30, no. 5, pages 299–313, 2006. (Cited on page 85)

[Zhou 11] Zhongxing Zhou, Feng Gao, Huijuan Zhao & Lixin Zhang. *Techniques to improve the accuracy of noise power spectrum measurements in digital x-ray imaging based on background trends removal*. Med Phys, vol. 38, no. 3, pages 1600–1610, 2011. (Cited on page 51)

Nomenclature

Acronyms

1D	One-dimensional
2D	Two-dimensional
3D	Three-dimensional
ACC	Acceptable
ALARA	As low as reasonably achievable
CNR	Contrast-to-noise ratio
CT	Computed tomography
DNA	Deoxyribonucleic acid
ESF	Edge spread function
FBP	Filtered back projection
FFT	Fast Fourier transform
IR	Iterative reconstruction
IQA	Image quality assessment
LSF	Line spread function
MTF	Modulation transfer function
NAC	Not acceptable
NPS	Noise power spectrum
PS	Power spectrum
RNA	Ribonucleic acid
ROI	Region of interest
SNR	Signal-to-noise ratio
SVM	Support vector machine

Note: For the figures which had been taken from the literature under copyright condition, permission for reprinting had been requested and no objections had been raised.