Digitale Wasserzeichen für MPEG-Videos zur Authentifizierung des Urhebers und Videos

Dissertation

zur Erlangung des akademischen Grades

Doktoringenieur (Dr.-Ing.)

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

von: Dipl-Inf. (FH) Enrico Hauer

geb. am 09.11.1976 in Wittenberg
Darmstadt, den 08.07.2009

Gutachterinnen/Gutachter:
Prof. Dr. Jana Dittmann, Otto-von-Guericke-Universität Magdeburg
Prof. Dr. Stefan Katzenbeisser, Technische Universität Darmstadt
Prof. Dr. Erika Müller, Universität Rostock
Dr. Martin Steinebach, Fraunhofer Institut Darmstadt

Ort und Datum des Promotionskolloquiums:
Magdeburg, den 17.06.2009

Eidesstattliche Erklärung

Ich erkläre hiermit, dass ich die vorliegende Arbeit ohne unzulässige Hilfe Dritter und ohne Benutzung anderer als der angegebenen Hilfsmittel angefertigt habe; die aus fremden Quellen direkt oder indirekt übernommenen Gedanken sind als solche kenntlich gemacht.

Insbesondere habe ich nicht die Hilfe eines kommerziellen Promotionsberaters in Anspruch genommen. Dritte haben von mir weder unmittelbar noch mittelbar geldwerte Leistungen für Arbeiten erhalten, die im Zusammenhang mit dem Inhalt der vorgelegten Dissertation stehen.

Die Arbeit wurde bisher weder im Inland noch im Ausland in gleicher oder ähnlicher Form als Dissertation eingereicht und ist als Ganzes auch noch nicht veröffentlicht.

Darmstadt, 08.07.2009

Enrico Hauer

Danksagung

Für die Aneignung des Wissens, das für die Gestaltung dieser Arbeit notwendig war, möchte ich sämtlichen Personen danken, die mich während der Jahre begleitet und geholfen haben. Ohne ihre Mithilfe wäre es für mich schwierig geworden, diesen Status zu erreichen.

Besonderen Dank gelten dabei besonders Prof. Dr. Dittmann, Prof. Tönnies und Dr. Steinebach. Mit ihren grundlegenden Ideen konnte ich wichtige Bestandteile der Arbeit erfassen, eine Struktur aufbauen und basierend darauf die Verfahren entwickeln und untersuchen.

Einen ebenso wichtigen Dank möchte ich meinen ehemaligen Arbeitskollegen der Fraunhofer Institute IPSI und SIT aussprechen. Dazu gehören Sascha Szmudzinski, Patrick Wolf, Huajian Liu, Stefan Thiemert und Lucilla Croce-Ferri. Durch die Austausch von Konzepten und die Bewertung meiner Ergebnisse hatte ich die Möglichkeit, eine Bewertung der Konzepte durchzuführen. Hervorzuheben ist die immer fröhliche Haltung von Huajian Liu, wodurch auch in schwierigen Zeiten ein Weiterkommen möglich war.

Ebenso möchte ich sämtlichen Studenten für ihre Unterstützung der Untersuchung der Verfahren danken. Die von ihnen durchgeführten Programmierungen bzw. Testuntersuchungen führen zu den Ergebnissen und Bewertungen der Ansätze der Arbeit.

Nicht zu vergessen ist die sprachliche Unterstützung von Martina Zwilling. Ich kann mich kaum glücklicher schätzen, dass sie sich auch in eigenen schwierigen Zeiten bereiterklärt hat, mir ihre Hilfe zur Verfügung zu stellen.

Ebenso danke ich meinen Freunden Michelle und Gideon Rath, Rolf Schäfer, Jörg Müller und Sofiya Raleva. Ihr ständiges Aufmuntern ermöglichte mir, diese Arbeit abzuschließen und nicht vorher abzubrechen.

Inhaltsverzeichnis

1 Einleitung und Motivation 1
1.1 Problemfelder 2
1.2 Zusammenfassung der wissenschaftlichen Erkenntnisse 3
1.3 Aufbau der Arbeit 5
2 Grundlagen 7
2.1 Digitale Videostandards 7
2.2 Sicherheitsmechanismen für Multimediadaten 11
2.3 Anwendungen digitaler Videodaten 14
2.4 Digitale Wasserzeichen 15
3 Anwendungsszenarien für Videowasserzeichen 23
3.1 Nachweis des Urhebers bzw. Empfängers 23
3.2 Überprüfung der Authentizität von Videos 28
3.3 Gegenüberstellung von robusten und reversiblen Wasserzeichen 31
4 Stand der Technik 37
4.1 Einzelbildbasierte Wasserzeichen 37
4.2 Szenenbasierte Wasserzeichen 41
4.3 Wasserzeichen für neue Videostandards 41
5 Analyse existierender Verfahren und Ansätze der Arbeit 43
5.1 Evaluation existierender Wasserzeichen für MPEG Daten 43
5.2 Schwachstellen und Lösungsansätze 50
6 Synchronisation von MPEG-Videowasserzeichen 61
6.1 Konzept der zeitlichen Synchronisation 61
6.2 Evaluation existierender Ansätze 62
6.3 Modularer robuster Videohash zur Frameidentifikation 70
6.4 Evaluation der Ansätze 75
6.5 Zusammenfassung 83
7 Robustes MPEG Wasserzeichen 85
7.1 Konzept des robusten Wasserzeichens 85
7.2 Anforderungen an robuste MPEG-Wasserzeichen 88
7.3 Existierende Ansätze 90
7.4 MPEG Wasserzeichen für Luminanzdaten 91
7.5 Evaluation der Testergebnisse 99
7.6 Zusammenfassung 104
8 Reversibles Wasserzeichen für Videodaten 105
8.1 Konzept für reversible Videowasserzeichen 106
8.2 Anforderungen an reversible Videowasserzeichen 108
8.3 Zusammensetzung der Wasserzeicheninformation 110
8.4 Reversibles MPEG-Videowasserzeichen 112
8.5 Evaluation der Testergebnisse 119
8.6 Zusammenfassung 124
9 Effizientes Watermarking für Transaktionswasserzeichen 125
9.1 Konzept eines Transaktionswasserzeichens 125
9.2 Bewertung von Alternativen 130
9.3 Lösungungsvorschläge für die Anwendungsszenarien 130
9.4 Zusammenfassung 141
10 Bewertung der präsentierten Ansätze 143
11 Zukünftige Arbeit 149
11.1 Markierung des gesamten MPEG-Videos 149

11.2 Anpassung von Wasserzeichen an neue Videostandards 149
11.3 Suche von markierten Videos in Peer-to-Peer Netzwerken 150
11.4 Unterstützung durch Videowasserzeichen im Fernsehumfeld 151
12 Zusammenfassung 155

Abbildungsverzeichnis

Abbildung 1: MPEG-1 Struktur Videostrom [Ti2001] 10
Abbildung 2: MPEG-1 Struktur Makroblockaufbau [Ti2001] 11
Abbildung 3: Darstellung des Video-Online-Shop 24
Abbildung 4: Verbreitungsmethode Try and Buy 26
Abbildung 5: Integration von Mediensicherheit in DILIGENT 28
Abbildung 6: Beispiel einer Modifikation [KuJo2002] 29
Abbildung 7: Auslesen der Wasserzeicheninformation nach MPEG-Reenkodierung [Th2002] 47
Abbildung 8: Veränderung des Frametypes durch Re-Enkodierung 51
Abbildung 9: Konzept des robusten Videohashs 51
Abbildung 10: Beispiel Standard Referenzstruktur einer Group-of-Picture (GoP) 52
Abbildung 11: Erhöhung der Kapazität durch erweiterte Markierung 53
Abbildung 12: Verbesserung der Transparenz 55
Abbildung 13: Effizientes Einbetten durch Aufteilung des Einbettungsmodus 57
Abbildung 14: Verteilung der Markierung auf mehrere Ressourcen 58
Abbildung 15: Zeitliche Synchronisation markierter I Frames 61
Abbildung 16: Exemplarische Demonstration der Sequenzordnung vor und nach dem Angriff 65
Abbildung 17: Beispiel des Synchronisationsschemas 65
Abbildung 18: Pseudozufällige Muster und Glättung nach [FrGo2000] 68
Abbildung 19: Beispiel der Ähnlichkeit benachbarter Bilder 69
Abbildung 20: Konzept des robusten Videohashs 70
Abbildung 21: Videohash zur Synchronisation von Videowasserzeichen 71
Abbildung 22: Merkmalsextraktion durch ein Bildhashverfahren (BHV) 72
Abbildung 23: Hashwertgenerierung durch ein Videohashverfahren 73
Abbildung 24: Identifizierung markierter Frames 74
Abbildung 25: Hash Detektionsrate bei verschiedenem Datenmaterial 76
Abbildung 26: Detektionsrate bei verschiedenen Blockgröße 77
Abbildung 27: Auswirkung unterschiedlicher Sequenzlänge 77
Abbildung 28: Schwellwert zur Frameidentifikation 78
Abbildung 29: Identifikation der Videoframes 80
Abbildung 30: Robustheitsanalyse von Videoset 1 81
Abbildung 32: Grundkonzept für robuste Wasserzeichen 85
Abbildung 33: Übertragung von Referenzwasserzeichen zu einem interkodiertem Frame 87
Abbildung 34: Beispielbewegung zwischen zwei Nachbarframes in einem MPEG-Video 89
Abbildung 35: Nachbearbeitung des DCT-Blocks 94
Abbildung 36: FAR-FFR Gegenüberstellung von [DiSt1998] 96
Abbildung 37: FAR-FFR Gegenüberstellung Median pro Bit 96
Abbildung 38: Detektionsraten robustes Wasserzeichen für Videoset 1 100
Abbildung 39: Detektionsraten robustes Wasserzeichen für Videoset 2 100
Abbildung 40: Detektion des Wasserzeichens in P-Frames von Videoset 1 102
Abbildung 41: Detektion des Wasserzeichens in P-Frames von Videoset 2 102
Abbildung 42: Konzept eines reversibles Wasserzeichen auf Framebasis 106
Abbildung 43: Anpassung von reversiblem Wasserzeichen für MPEG-Videos 115
Abbildung 44: Erweiterter Einbettungskanal 116
Abbildung 45: Einbetten eines reversiblen MPEG-Wasserzeichen 118
Abbildung 46: Optimierung des Einbettungskanals 123
Abbildung 47: Erstellung des Kontainers und Extraktion eines Beispielvideos 132

Abbildung 48: Evaluierung der Echtzeitfähigkeit des Wasserzeichenkontainers 135
Abbildung 49: Integration der fehlenden Videodaten mit Wasserzeichen in das Video 136
Abbildung 50: gLite Netzwerkelemente für Distributed Watermarking [Gl2007] 137
Abbildung 51: Strategie verteiltes Markieren 138
Abbildung 52: Zeitverteilung der Teilprozesse 139
Abbildung 53: Zeitlicher Aufwand der Teilprozesse 139
Abbildung 54: Einbettung von Fernsehwasserzeichen 152
Abbildung 55: Detektion von Fernsehwasserzeichen 152

Tabellenverzeichnis

Tabelle 1: Anwendungsgebiete robuster und reversibler Wasserzeichen 32
Tabelle 2: Verfahrensparameter robuster und reversibler Wasserzeichen 33
Tabelle 3: Verfahrensparameter robuster und reversibler Wasserzeichen 34
Tabelle 4: Resistenz von robusten und reversiblen Wasserzeichen gegenüber Angriffen 35
Tabelle 5: Testparameter [Th2002] 44
Tabelle 6: Evaluationsergebnisse basierend auf dem periodischen Konzept 67
Tabelle 7: Robustheit DivX-Konvertierung 101
Tabelle 8: Robustheit Skalierung 101
Tabelle 9: Durchschnittliche Kompressionsrate pro GoP 116
Tabelle 10: Untersuchte Modifikationen 121
Tabelle 11: Zero-Assumption-Angriff 122
Tabelle 12: Vergleich lokales und distributed Watermarking 140
Tabelle 13: Auslesen der Wasserzeicheninformation 159
Tabelle 14: Auslesen der Wasserzeicheninformation nach dem Auftragen eines zusätzlichen Wertes [Th2002] 160
Tabelle 15: Auslesen der Wasserzeicheninformation nach Videoskalierung [Th2002] 160
Tabelle 16: Auslesen der Kundeninformationen [Th2002] 160
Tabelle 17: Robuster Videohash Detektionsrate Videoset 1 162
Tabelle 18: Robuster Videohash Detektionsrate Videoset 2 163

1 Einleitung und Motivation

Nach dem augenblicklichen Stand der Technik können digitale Multimediadaten mit hoher Qualität vervielfältigt und verbreitet werden, wie mit Hilfe von Online Shops [HaTh2003]. Ein wichtiger Gesichtspunkt ist die Möglichkeit der verlustfreien Vervielfältigung der Medien, d.h. es können ohne Qualitätsverluste Kopien erstellt werden.

Durch den Einsatz verschiedener technischer Geräte, wie DVD-Rekorder bzw. -player werden immer mehr Videos digitalisiert. Was auf der einen Seite ein Vorteil der Verbreitung ist, kann aber auf der anderen Seite zu eingeschränkten Kontrollmöglichkeiten führen. Auch die ständige Verbesserung der Bandbreite der Netzwerke führt einerseits zu einer immer schnelleren Verbreitung des Videos und andererseits können auch immer qualitativ bessere Videos verteilt werden. Dadurch verbessert sich auch die Benutzerfreundlichkeit bei der Verbreitung von Videomaterial über das Internet.

Die vorliegende Arbeit beschäftigt sich mit Sicherheitsmaßnahmen für Videodaten. Bei zunehmender Verbreitung von Videos müssen aber gleichzeitig Mechanismen eingesetzt werden, die die Wahrung sicherheitsrelevanter Aspekte wie das Urheberrecht gewährleisten. Medienproduzenten haben erkannt, dass die digitale Produktion im Vergleich zur analogen zwar qualitative Vorteile hat, sehen sich jedoch gezwungen, die Verbreitungswege entsprechend deutlicher zu kontrollieren.

Durch entsprechende rechtliche Verordnungen wurde eine Basis zur Wahrung des Urheberrechts geschaffen. Eine entscheidende Grundaussage bildet dabei die Erzeugung technischer Sicherheitslösungen, die nicht umgangen werden dürfen. Diese Regel gilt auch für Multimediadaten [Kr2008]. Ein Beispiel für eine Sicherheitslösung ist ein Kopierschutz.

Bei den existierenden Sicherheitslösungen wird zwischen aktiven und passiven Konzepten unterschieden. Aktive Konzepte haben einen direkten Einfluss auf erlaubte und unerlaubte Informationswege, passive Lösungen dienen als Nachweis zur Wahrung entsprechender Rechte. Aktive Sicherheitslösungen, wie z. B. Digital-Rights-Management-Systeme (DRMS) regeln Abspiel- und Kopiervorgänge. Bei passiven Sicherheitslösungen werden zum Beispiel digitale Signaturen für Echtheits- oder Authentizitätsnachweise oder digitale Wasserzeichen für den Urhebernachweis genutzt.

Digitale Wasserzeichen stellen inzwischen eine anerkannte Sicherheitsmethode dar. Sie werden in verschiedenen Anwendungsszenarien, wie Urheberschutz, Authentifizierung, Zugriffskontrolle, Integritätsschutz oder Annotation genutzt. Neben dem alleinigen Einsatz von digitalen Wasserzeichen können sie auch in Kombination mit anderen Sicherheitssystemen, wie z. B. DRM-Systemen, angewandt werden, um eine verbesserte Funktionalität zu ermöglichen [Bo2005].

Wasserzeichen werden in diversen Publikationen in verschiedene Eigenschaften klassifiziert [Di2000], die basierend auf dem jeweiligen Anwendungsszenario unterschiedlich ausgeprägt sind. Dazu gehören z. B. Robustheit gegenüber Modifikationen an Mediendaten, visuelle Transparenz, sowie Sicherheit gegenüber gezielten Angriffen auf das Wasserzeichen. Oft stehen diese Eigenschaften jedoch in Konflikt zueinander [CoMi2002].

In dieser Arbeit werden Bewegtbilder bzw. Videos betrachtet. Mit der Entwicklung neuer Videostandards und Trägermedien ist die Verbreitung deutlich angestiegen. Eine Komprimierung erweist sich als notwendig, um Videodaten effizient verbreiten zu können. Dadurch ist es möglich geworden, Videos mit einem akzeptablen Aufwand und hinreichender Qualität zu vervielfältigen [ISO1995], [ISO2001] und [ISO2003b].

Durch die Entwicklung der neuen Standards für Bewegtbilder haben sich auch neue Herausforderungen für die Entwicklung von digitalen Wasserzeichen ergeben. So hat sich gezeigt, dass die Eigenschaften von Wasserzeichen, wie die Transparenz und Robustheit, neu betrachtet werden müssen. Diese Arbeit beschäftigt sich mit diesen Aspekten und präsentiert entsprechende Lösungen.

1.1 Problemfelder

Die Entwicklung eines digitalen Wasserzeichens ist abhängig vom Anwendungsszenario. So wurden in [Di2000] verschiedene allgemeine Anwendungsgebiete beschrieben, für die die Entwicklung von Wasserzeichen relevant sind. Diese Arbeit konzentriert sich auf Szenarien, in denen die Wahrung der Authentizität von Eigentümern und Kunden als auch von Videodaten wichtig sind.

Dabei haben sich die folgenden Problemfelder beim Einsatz von Videowasserzeichen herausgestellt, wobei diese Arbeit sich auf die folgenden Problemfelder diskutiert:

- Mangelnde Robustheit bei üblichen Modifikationen, wie Formatkonvertierung, Skalierung oder Kompression
- Ineffiziente Nutzung der zur Verfügung stehenden Kapazität
- zeitliche Verzögerung während des Einbettungsvorganges
- Hervorrufen von sichtbaren Artefakten
- Mangelnde Erkennung von Veränderungen am Video

Formatkonvertierungen können genau dann durchgeführt werden, wenn das Problem eines zu gering zur Verfügung stehenden Speicherplatz auftritt. Der weniger notwendige Speicherplatz ist aber auf einen höheren Kompressionsgrad zurückzuführen, was sich wiederrum negativ auf die Auslesequote von eingebetteten Wasserzeichen auswirken kann. Modifikationen wie Skalierung und Kompression sind häufig Bestandteil einer Re-Enkodierung von hoch- in niederfrequentes Videomaterial. Auch bei dieser Modifikation wird einerseits das Video stärker komprimiert, andererseits kann durch die Skalierung ein Teil des Wasserzeichens vollständig verloren gehen. Ein Teil dieser Arbeit entwickelt einen Wasserzeichenansatz, der eine hohe Robustheit gegenüber Modifikation dieser Art erreicht. Dabei wird versucht, einen Kompromiss in der vorgestellten Lösung zu erreichen, die die drei aufgeführten Modifikationen bestmöglichst abdeckt.

Ein weitere Herausforderung ist die häufig zu geringe Kapazität, die vorhandene Wasserzeichenlösungen anbieten. Existierende Wasserzeichenansätze werden im späteren Kapitel vier genauer erläutert. Es kann zwar behauptet werden, dass eine Fokussierung auf

einzelne Frames im Video die beste Robustheit des Wasserzeichens ermöglicht, sie aber gleichzeitig einen Teil der maximal möglichen Kapazität nicht nutzt, da nicht jedes Einzelbild markiert wird. Diese Arbeit betrachtet ebenfalls diesen Aspekt und versucht unter Mitbetrachtung weiterer Eigenschaften des Wasserzeichens die Kapazität so weit wie möglich zu erhöhen.

Weitere Herausforderungen sind die Bewegung innerhalb des Videos und die zeitlich abhängigen Kodierung des Videos eingegangen. Wenn jedes Einzelbild im Video eigenständig mit einem Wasserzeichen markiert werden soll, muss mit visuellen Artefakten gerechnet werden. Das erste Problem liegt in der minimalen Veränderung benachbarter Videoframes. Da, abgesehen von einem Schnitt, der visuelle Unterschied zwischen zwei benachbarten Frames eher gering ist, sollten sich die Wasserzeichen auch nicht deutlich unterscheiden. Werden hingegen zwei unterschiedliche Markierungen in die Nachbarframes eingebettet, kann der Unterschied durchaus sichtbar werden. Das zweite Problem ist die zeitlich basierte Kodierung der Videos. Wie am Beispiel von MPEG-Videos wird es während der Kodierung vermieden, visuell redundanten Inhalt erneut abzuspeichern. Die dadurch erreichte zeitliche Kompression erweist sich bei einer geringen Bewegung als sehr effizient in der Komprimierung des Videos. Das eingesetzte Wasserzeichen muss aber die Kodierung mitbeachten, um keine Mängel in der Robustheit und Transparenz hervorzurufen. Der in dieser Arbeit verfolgte Ansatz versucht innerhalb einer festgelegten Szene im Video eine identische Wasserzeichenmarkierung einzubetten. Dadurch sollten die Markierungen nicht sichtbar werden.

Die Kodierung von Videodaten ist sehr zeitaufwendig. Durch die Einbettung des Wasserzeichens kann es möglicherweise zu einer nicht unbedeutenden Verzögerung der Weiterleitung von Videos kommen. Die durch eine Kodierung möglicherweise auftretende Verzögerung kann bei verschiedenen Online-Shop Einsatzszenarien aber nicht akzeptabel sein. Deshalb entwickelt die Arbeit einen Ansatz, der die zeitlich aufwandstärksten Prozesse der De- und Enkodierung so gering wie möglich hält.

Das zweite Anwendungsfeld dieser Arbeit ist der Nachweis der Authentizität von Videos. Potenzielle Anwendungsfelder sind zum Beispiel Hochsicherheits- oder medizinische Bereiche. Bei diesen Anwendungen haben Veränderungen an Videos bedeutende Auswirkungen auf die Aussage des Inhaltes. Selbst minimale Veränderungen können durchaus die Aussage des Video entscheidend verändern. Da digitale Wasserzeichen mit dem Medium fest „verwebt“ werden, müssen die durch das Wasserzeichen durchgeführten Veränderungen wieder rückgängig gemacht werden. Deshalb muss neben der Authentizitätsüberprüfung auch die Möglichkeit der Entfernung des Wasserzeichens ermöglicht werden. Dies erfordert aber gleichzeitig die Bereitstellung effizienter Sicherheitsmaßnahmen, die in dieser Arbeit diskutiert werden.

1.2 Zusammenfassung der wissenschaftlichen Erkenntnisse

Videos haben einen entscheidenden Unterschied zu anderen Multimediadaten. Die Datenmenge ist um ein Vielfaches höher als bei Audio oder Bild. Durch die Entwicklung entsprechender Videostandards kann die Datenmenge effizient verrringert werden, was ein Übertragung der Videos ohne eine signifikante Zeitverzögerung ermöglicht.

Die Entwicklung digitaler Wasserzeichen für Videos ist zunehmend abhängig vom verwendeten Videostandard. Eine Vielzahl der Wasserzeichen sind für MPEG-Videos entwickelt [HaDi2005], [BiDa2005], [WaPe2006]. Auch in dieser Arbeit liegt der Fokus auf MPEG-Videos. Durch die zunehmende Verbreitung von MPEG-basierenden Anwendungen wie DVD (Digital-Versatile-Disc) und DVB (Digital-Video-Broadcast) kann durchaus davon ausgegangen werden, das für MPEG-Videos entwickelte Wasserzeichen stärkere Anwendung finden werden.

Da MPEG-Videos in einer Vielzahl von Anwendungen genutzt werden, konzentriert sich diese Arbeit auf die beiden Anwendungsbereiche Urheber- und Videoauthentizität. Die Vielzahl der möglichen Anwendungsfälle ermöglicht es uns nicht, das gesamte Spektrum von Wasserzeichenarten zu untersuchen. Die Untersuchung der in dieser Arbeit vorgestellten Anwendungsfälle zeigt aber deutlich, dass ein Optimierungsbedarf der Wasserzeichen-eigenschaften Robustheit, Transparenz, Kapazität und Echtzeitfähigkeit notwendig ist. Ein auf [Di2000] basierender Vergleich verdeutlicht zudem, wie die Eigenschaften bei den in dieser Arbeit betrachteten Wasserzeichen zur Urheber- und Videoauthentizität verschieden ausgeprägt sind.

Um Lücken bei existierenden Wasserzeichen aufdecken zu können, wird in dieser Arbeit eine Evalierung der Wasserzeichen durchgeführt. Sie bildet die Grundlage für die Auflistung der Schwachstellen in den Eigenschaften Robustheit, Transparenz, Kapazität und Echtzeitfähigkeit [CoMi2002]. Dabei sind uns die folgenden Schwachstellen aufgefallen:

- Veränderung der Frametypen bei Markierung von MPEG I-Frames,
- Einbettung verschiedener Wasserzeicheninformationen in benachbarte Frames,
- Geringe Kapazität, da nur einzelne Frames markiert werden und
- Echtzeitfähigkeit ist schwer umsetzbar, da Videos komplett de- und enkodiert werden.

Da viele MPEG-Wasserzeichen nur einzelne Frames markieren, beschäftigt sich diese Arbeit mit der zeitlichen Synchronisation. Ziel ist es, die markierten Frames zu detektieren. Die Synchronisation der markierten Frames gewinnt immer mehr an Bedeutung, da es nicht immer möglich ist, das gesamte Video zu markieren [HaKu2005]. Als Technologie wird ein robuster Hash genutzt. Der Hashwert ist modular aufgebaut, um einen späteren Austausch der Module durch verbesserte Algorithmen zu ermöglichen. Das Verfahren erkennt mit einer hohen Detektionsquote die markierten Frames und schließt nicht markierte Frames aus. Dadurch wird eine Verfälschung des Ausleseergebnisses verhindert.

Gleichzeitig wird eine Lösung präsentiert, die ein Wasserzeichen in das gesamte Video einbettet. Basierend auf den Erkenntnissen diverser Veröffentlichungen [SuKu2002] wird innerhalb der Framesequenz das gleiche Wasserzeichen eingebettet, wodurch dessen Transparenz gewahrt wird. Die Technologie der Bewegungskompensation verhindert eine Mehrfachmarkierung innerhalb der Framesequenz [Ha2000]. Die Ergebnisse zeigen, dass es möglich ist, das Wasserzeichen aus MPEG P-Frames auszulesen, aber die Detektion des Wasserzeichens aus den I-Frames weiterhin der wichtigste Bestandteil des Ausleseprozesses ist.

Für den Anwendungsfall der Videoauthentizität wird ein reversibles Wasserzeichen vorgestellt. Um es auch für hochsensible Videos anwenden zu können, ist das Wasserzeichen so aufgebaut, dass es wieder aus dem Video entfernt werden kann. Die Authentizität wird mittels einer digitalen Signatur nachgewiesen. Als Grundlage dienen Verfahren, die die Integrität des Videos mit einem Hash nachweisen und ihn als Wasserzeicheninformation einbetten [FrGo2001], [FrGo2000]. Die Anwendung einer asymetrischen Signatur ermöglicht zwar eine öffentliche Verifikation der Authentizität, erfordert aber gleichzeitig einen sicheren Aufbau der Wasserzeicheninformation [KaDi2004]. Das in dieser Arbeit entwickelte Wasserzeichen bietet eine ausreichend hohe Kapazität, um augenblicklich als sicher eingestufte Signaturen in das Video einzubetten [Bu2007]. Der Test des Verfahrens zeigt, dass das Video nach jeder der durchgeführten Modifikationen als nicht authentisch eingestuft wird.

Wie oben bereits beschrieben, besteht aufgrund der zunehmenden Verbreitung von Videos über das Internet der Bedarf, Wasserzeichen möglichst schnell in ein Video einbetten zu können. Daher beschäftigt sich diese Arbeit auch mit der effizienten Einbettung von Wasserzeichen, wobei für die folgenden beiden Anwendungsfälle jeweils eine Lösung präsentiert wird:

- Vertrieb über Online Shops und
- Integration von Videosammlungen in digitale Bibliotheken.

Für den Online-Shop-Vertrieb wird ein Wasserzeichenkontainer entwickelt [HaSt2008]. Der Kontainer beinhaltet sämtliche Informationen, die notwendig sind, um eine Kopie des Videos mit einem kundenbasierten Wasserzeichen zu erstellen. Während der Erstellung der Videokopie mit dem kundenspezifischen Wasserzeichen müssen nur noch die notwendigen Videoteile aus dem Kontainer kopiert und in der Kopie abgespeichert werden. Durch den Einsatz des Wasserzeichenkontainers kann während der Einbettung des Wasserzeichens eine 100-fache Echtzeitfähigkeit erreicht werden.

Um während der Integration von Videosammlungen in digitale Bibliotheken digitale Wasserzeichen möglichst effizient einbetten zu können, wird auf verteilte Netzwerke zurückgegriffen [StHa2007]. Die im Hintergrund befindlichen Netzwerke bieten ausreichend Computerkapazität, um eine Vielzahl von Videos ohne Zeitverzögerung zu markieren. Durch eine Aufteilung der Videos in Segmente kann zudem eine Beschleunigung des Einbettungsvorganges erreicht werden, da die Segmente verteilt auf mehrere Computer gleichzeitig markiert werden. Die Evaluation der benötigen Zeit und eine Analyse der Teilvorgänge während der Einbettung verdeutlicht jedoch die Notwendigkeit der Optimierung des Datenflusses innerhalb des Netzwerkes. Der Vorteil der Verteilung der Markierungsvorgänge wird dann besonders deutlich, wenn eine Videosammlung, die aus einer Vielzahl von Videos besteht, markiert werden soll.

1.3 Aufbau der Arbeit

In Kapitel 2 werden die Grundlagen der Arbeit diskutiert. Dies umfasst die Betrachtung von Videostandards (2.1), wobei in unkomprimierte und komprimierte Standards unterschieden

wird. Anschließend folgt eine umfangreichere Beschreibung des MPEG-Standards. Er bildet den Ausgangspunkt für die Entwicklungen der Wasserzeichen in dieser Arbeit (2.1.3). In Kapitel 2.2 werden die Sicherheitsmechanismen für Multimediadaten beschrieben, wobei der Fokus auf Videodaten liegt. Abschließend folgt eine Beschreibung von digitalen Wasserzeichen mit einer Spezialisierung auf Videodaten.

Das dritte Kapitel stellt die Anwendungsszenarien vor, auf welchen die Entwicklungen der Wasserzeichenverfahren dieser Arbeit basieren. Es werden Anwendungen für robuste und reversible Wasserzeichen vorgestellt. Abschließend werden beide Wasserzeichenarten in Form eines Klassifizierungsschema gegenübergestellt.

Das vierte Kapitel diskutiert zu Beginn den Ausgangspunkt bezüglich der existierenden Ansätze bei Videowasserzeichen. Dieser stellt die Basis für die Evaluierung ausgewählter MPEG-Videowasserzeichen im darauffolgenden Unterkapitel. Die während der Evaluierung erkannten Schwachpunkte werden in Kapitel fünf aufgelistet und genauer erläutert. Für jeden der aufgeführten Schwachpunkte werden Lösungsvorschläge beschrieben, die in den Kapiteln sechs bis neun mit entwickelten Ansätzen gelöst werden.

Das Kapitel sechs beschäftigt sich mit dem Problem der Synchronisierung von Videowasserzeichen. Eine Synchronisierung ist erforderlich, wenn das Wasserzeichen nur einzelne Bereiche des Videos markiert. Wir konzentrieren uns dabei auf die Detektion von I-Frames, die sehr häufig als Einbettungsziel für Videowasserzeichen genutzt werden. Der Ansatz der Synchronisierung wird in dieser Arbeit getrennt von den anderen Ansätzen betrachtet, da er universell für die beiden in Kapitel sechs und sieben präsentierten Wasserzeichen eingesetzt werden kann.

In Kapitel sieben werden robuste MPEG-Videowasserzeichen diskutiert. Zuerst wird ein Konzept vorgestellt, das die notwendige unterschiedliche Behandlung intrakodierter Frames (I-Frames) und interkodierter Frames (P-/B-Frames) beschreibt. Das Kapitel beschreibt die Lösung für beide Frametypen.

In Kapitel acht wird ein reversibles Wasserzeichen vorgestellt. Zusätzlich zur Authentizitätsüberprüfung soll durch den Einsatz des Wasserzeichens das Originalvideo wiederherstellbar sein. Das Konzept für reversible Wasserzeichen wird erörtert, eine Lösung für die notwendigen Sicherheitsanforderungen präsentiert und die Anpassungen an MPEG-Videos vorgeführt.

In Kapitel neun wird der Fokus auf die effektive Einbettung des Wasserzeichens gelegt. Hierbei wird das Modell eines Transaktionswasserzeichens definiert und verschiedene technische Lösungen vorgestellt, die unter verschiedenen Gesichtspunkten das Kriterium effizienter Einbettung erfüllen.

Abschließend folgt in Kapitel zehn eine Bewertung der in Kapitel sechs bis neun präsentierten Ansätze. Ziel der Bewertung ist es, zu untersuchen, in wieweit die in Kapitel fünf aufgeführten Schwachstellen gelöst werden konnten. Das 11. Kapitel widmet sich einer perspektivischen Vorschau, indem es noch nicht gelöste Forschungsgebiete und zukünftige Arbeitsbereiche für Videowasserzeichen beschreibt. Abschließend folgt in Kapitel 12 eine Zusammenfassung.

2 Grundlagen

Dieses Kapitel bildet die Grundlage für die gesamte Arbeit. Es behandelt sowohl Sicherheitsaspekte als auch digitale Videostandards.

Bei den Sicherheitsaspekten liegt der Schwerpunkt bei digitalen Wasserzeichen. Wasserzeichen bieten eine Möglichkeit zur Wahrung von Urheberrechten als auch zum Integritäts- und Authentizitätsnachweis bei Multimediadaten, in dieser Arbeit von Videodaten. Hierbei liegt der Fokus wiederrum bei MPEG-Videos, genauer gesagt bei MPEG-1 und 2 Videos, da sie augenblicklich einen hohen Verbreitungsgrad haben.

Das Kapitel untergliedert sich in drei Teile. Der erste Teil behandelt Videostandards, und zwar zunächst nicht komprimierende Standards, um die Notwendigkeit der Weiterentwicklung zu komprimierenden Standards darstellen zu können. Der Schwerpunkt liegt aber weiterhin bei MPEG [ISO1995] (Motion Picture Experts Group). Um die Entwicklungsschritte im Verlauf dieser Arbeit besser verstehen und nachvollziehen zu können, wird der MPEG-Standard an dieser Stelle ausführlicher beschrieben.

Sämtliche Sicherheitstechnologien basieren auf verschiedenen Schutzzielen. Dabei hängt es von der jeweiligen Anwendung ab, wie die Schutzziele im Einzelnen realisiert werden sollen. Der zweite Teil dieses Kapitels beschreibt diese Schutzziele, bevor dann im dritten Teil die digitalen Wasserzeichen als ausgewählte Sicherheitstechnologie behandelt werden. Dieser beinhaltet die Erläuterung von Eigenschaften und Anwendungen von Wasserzeichen, wobei die Eigenschaften auf MPEG-Videodaten abgebildet werden.

Da es inzwischen eine Vielzahl von Publikationen über digitale Wasserzeichen gibt, erweist es zunehmend schwieriger, eine eindeutige Klassifizierung der Technologie zu finden. Zur Vereinfachung und besseren Verständlichkeit stützen sich die Ausführungen dieser Arbeit auf zwei Veröffentlichungen ([CoMi2002] und [Di2000]), die nach Ansicht des Autors eine gute Ausgangsbasis für die Klassifizierung von digitalen Wasserzeichen bilden. In weiteren Publikationen werden noch vereinzelt Eigenschaften, wie Sicherheit und Robustheit erwähnt, sie entsprechen jedoch nicht dem Status einer allgemeinen Klassifizierung [Ec2005].

2.1 Digitale Videostandards

Der Fortschritt in der Entwicklung von digitalen Videostandards wurde in den letzten zehn Jahren stetig verbessert. Videodaten setzen sich aus zwei Teilen zusammen: Audiodaten und eine Bildsequenz. Zusätzliche Systemdaten regeln den zeitlich synchronen Ablauf beider Datenströme. In diesem Kapitel werden nur die Videodaten bzw. Bildsequenzen erklärt. Dabei unterscheiden wir unkomprimierte und komprimierte Datenströme.

2.1.1 Unkomprimierte Videos

Die ersten Anwendungen benutzten unkomprimierte Videodaten. Eins der ersten Videoformate basierte auf unkomprimierte AVI-Videos. In einem AVI-Video der ersten Version werden sämtliche Bilder unkomprimiert abgespeichert. AVI (Audio-Video-Interleave) selbst ist kein Videoformat, sondern ein Kontainerformat. Dadurch wird ermöglicht, verschiedene kodierte Videoströme auf Windows basierten Systemen abzuspielen. Mit Hilfe einer Identifizierungsinformation wird erkannt, wie der Videostrom kodiert und im Kontainer abgespeichert wurde. Dadurch kann der Videoplayer das Video wieder in die Einzelbilder dekodieren [Av2007].

Unkomprimierte Videodaten sind speicherintensiv und benötigen dadurch eine große Bandbreite zur Übertragung der Daten. Somit ist es fast ausgeschlossen, unkomprimierte AVI-Videos über verschiedene Netzwerke, wie z. B. das Internet, zu versenden. Dennoch bieten sie auch einen Vorteil: Da AVI-Videos unkomprimiert abgespeichert werden, gehen keine Bild- und Toninformationen verloren. Folglich entstehen keine Qualitätsverluste.

2.1.2 Komprimierte Videos

Da unkomprimierte Videos eine hohe Menge an Speicherplatz benötigen, wurden Algorithmen entwickelt, die Videodaten effizient komprimieren. Ziel war es dabei, die Datenmenge unter geringen bzw. annehmbaren Qualitätsverlusten zu verringern. Da bei Videodaten im Gegensatz zu Bilddaten auch der zeitliche Verlauf wichtig ist, werden folglich zwei verschiedene Kompressionstechniken angewandt:

(1) Räumliche Kompression
Bei der räumlichen Kompression können Bildinformationen entfernt oder der Datenstrom direkt komprimiert werden. Bei verlustfreier Kompression bleiben die Originaldaten erhalten. Vor und nach der Kompression liegt ein identisches Bild vor. Die am häufigsten verwendete Methode ist die Lauflängenkodierung, womit Bereiche gleicher Farbe effizient komprimiert werden können. Diese Technik ist daher besonders für Bildern mit homogenen gleichfarbigen Flächen geeignet. Da aber viele Bilder auch eine Vielzahl von Farben beinhalten, ist diese Kompressionstechnik nicht immer effizient.

Bei einer verlustbehafteten Kompression wird versucht, Bildinformationen aus dem Datenmaterial zu entfernen, die in der Regel vom Betrachter nicht wahrgenommen werden. Aufgrund dieser Tatsache muss mit dem Verlust von Daten gerechnet werden, die nicht wieder herstellbar sind. Der Grad der Kompression bestimmt dabei die Menge der verloren gegangenen Daten und damit auch die Qualität des komprimierten Bildes. Es wird zwischen mehreren Qualitätsstufen unterschieden, die vom Grad der Kompression abhängen.

Ein effizienter Videostandard versucht beide Kompressionstechniken zu kombinieren. Da durch die verlustbehaftete Kompression Informationen verloren gehen, können durch eine nachfolgende Lauflängenkodierung die „gelöschten“ Bildbereiche effizient zusammengefasst werden [ISO1995].

(2) Zeitliche Kompression
Bei Videomaterial kann auch entlang der zeitlichen Achse komprimiert werden. Je nach Art des Videos kann es sein, dass nur wenige Veränderungen in entsprechenden Szenen auftreten. Damit die Informationen nicht mehrfach abgespeichert werden müssen, werden häufig nur Differenzen zwischen benachbarten Bildern abgespeichert. Framedifferenzierung ist eine effiziente Methode der zeitlichen Kompression. Zeitliche Komprimierung sichert vorher die vollständigen Daten bestimmter Key-Frames und die dazwischen liegenden Änderungen. Je nach Art der angewandten Methode können weitere Mechanismen, wie z. B. die Bewegungskompensation, die zeitliche Kompression optimieren.

2.1.3 Der MPEG-Videostandard

MPEG steht für "Motion Picture Experts Group" - einem internationalen Gremium, das Standards für die Kodierung von bewegten Bildern entwickelt. Um eine hohe Vielzahl von Anwendungsbereichen zu gewährleisten, spezifiziert der MPEG-Standard nur ein Datenmodell zur Kompression von bewegten Bildern und Tonsignalen. Auf diese Weise bleibt MPEG von verschiedensten Computer-Plattformen unabhängig. Prinzipiell kann derzeit zwischen fünf Standards unterschieden werden: MPEG-1 [ISO1993], MPEG-2 [ISO1995], MPEG-4 [ISO2001], MPEG-7 [ISO2003a] und MPEG-21 [ISO2003b].

MPEG-1 [ISO1993] wurde 1993 mit dem Ziel verabschiedet, für Medien mit geringer Datenrate (1 MBit/s bis 1,5 MBit/s) die Übertragung von bewegten Bildern mit zugehörigem Audiosignal bei akzeptabler Bildwiederholfrequenz und möglichst guter Bildqualität zu erreichen. Für die meisten Heimanwendungen (Digitalisierung von Urlaubsvideos) sowie für den Businessbereich (Image-Videos, Dokumentationen) ist MPEG-1 ausreichend.

MPEG-2 [ISO1995] exisitiert seit 1995 und ist in seiner Grundstruktur identisch mit dem MPEG-1-Format. Er ermöglicht Datenraten bis zu 100 MBit/s und kommt bei digitalem Fernsehen (DF1), Videofilmen auf DVD-ROM und in professionellen Videostudios zum Einsatz. MPEG-2 ist in Auflösung und Datenrate über einen weiten Bereich skalierbar. Neben der klassischen Darstellung von MPEG-2 Videos gewinnt es auch immer mehr als Speicherform für aufgenommene Videos im privaten Bereich an Bedeutung.

MPEG-4 [ISO2001] wurde 2001 veröffentlicht. Er verfolgt das Ziel, eine möglichst hohe Videoqualität bei extrem niedrigen Datenraten im Bereich zwischen 10 KBit/s und 1 MBit/s zu gewährleisten. Eine Weiterentwicklung des MPEG-4 besteht in der Gliederung des Bildinhalts in eigenständige Objekte, um sie gezielt anzusprechen oder weiter zu verarbeiten. Damit wurden Forderungen nach Robustheit und fehlerfreier Datenübertragung erfüllt, was sich besonders im Mobilfunk als wichtig erweist. MPEG-4 kommt beispielsweise bei der Videoübertragung über das Internet zum Einsatz und dient als Basis für den Datentransfer bei der Übertragung von Bewegtbildern per Handy.

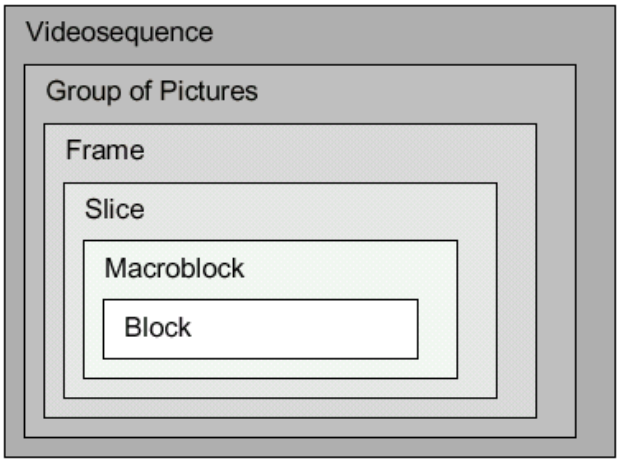

Abbildung 1: MPEG-1 Struktur Videostrom [Ti2001]

Abbildung 1 stellt den hierarchischen Aufbau eines MPEG-1 Videostrom dar. In der obersten Ebene der Videostruktur, der Videosequenz-Ebene, sind allgemeine Informationen des vollständigen Films gespeichert: Bildgröße, Bildseitenverhältnis und Angaben über den zur Verfügung stehenden Speicherplatz während der Dekodierung des Videos.

In einer GoP (Group of Pictures) sind mehrere Frames zusammengefasst. Die Anzahl der Frames pro GoP kann variabel gehalten werden. Für eine Vielzahl von Anwendungen ist die Standardgröße zwischen elf und 13 Frames festgelegt. Die Länge einer GoP muss nicht das gesamte Video über konstant sein. Da es den Kodierungsprozess aber wesentlich vereinfacht, wird sie aber bevorzugt konstant gehalten.

Im MPEG-Format sind vier verschiedene Frametypen definiert. Es wird zwischen Intra Frames (I-Frames), Predicted Frames (P-Frames), Bidirectional Predicted Frames (B-Frames) und Direct Coded Frames (D-Frames) unterschieden.

I-Frames sind die Kernframes im Datenstrom. Sie werden direkt enkodiert und basieren dabei auf JPEG-Bilder. Wichtig ist, dass die Dekompression des Bildes in Echtzeit möglich ist, damit ein flüssiges Abspielen des Videos gewährleistet werden kann. I-Frames dienen als Referenzbilder für P- und B-Frames, d. h. mit ihnen können P- und B-Frames vollständig berechnet werden.

In P-Frames befinden sich nur die Änderungen zu einem zurückliegenden I- oder P-Frame, d.h. um ein P-Frame zu dekodieren, muss bereits das vorherige I- oder P-Referenzframe dekodiert worden sein. Da nur die Differenz zum Vorgänger gespeichert wird, ist die Kompression höher als bei I-Frames. P-Frames können als Referenzbilder für weitere P-Frames und B-Frames dienen.

B-Frames werden aus dem jeweils zurückliegenden I- oder P-Frame und dem zukünftigen I- oder P-Frame interpoliert. Es werden also auch Bildinformationen genutzt, die einige Frames später angezeigt werden. Daher ist die Kompressionsrate bei B-Frames noch effizienter als bei P-Frames. B-Frames können nicht als Referenzbilder für weitere B-Frames dienen, allerdings können zwischen einem I- und einem P-Frame, bzw. zwischen zwei P-Frames beliebig viele B-Frames abgespeichert werden. Bei mehr als drei folgenden B-Frames verschlechtert sich aber die Qualität des Videos erheblich.

Ein Slice ist ein Streifen innerhalb eines Frames. Er beinhaltet mehrere folgende Makroblöcke, die mit den gleichen Parametern quantisiert werden.

Ein Makroblock ist eine Zusammenfassung von sechs DCT-Blöcken. Ein DCT-Block ist ein Bildblock, der aus 8x8 Einzelwerten besteht. Mit Hilfe der „Diskreten Kosinus Transformation“ (DCT) wird der Block in den Frequenzraum transferiert. Es ist zu erwähnen, dass das Ausgangsmaterial im YC_bC_r-Farbformat vorliegt. Das YC_bC_r-Farbformat ist ein Farbformat, das ursprünglich für den Einsatz im Fernsehbereich entwickelt wurde. Es kann mit einer festgeschriebenen Formel aus dem RGB Farbmodell errechnet werden. Bei der am häufigsten verwendeten Kodierung beinhalten vier der sechs Blöcke Helligkeitswerte (Luminanz) und die anderen beiden jeweils die Farbwerte (Chrominanz). Abbildung 2 zeigt den schematischen Aufbau eines Makroblocks bei Verwendung der Abtastrate 4:2:0. Diese Abtastrate kommt im Heimbereich bei der DVD und dem DVB zur Anwendung. Die Luminanzwerte werden im Verhältnis 1:1 und die Chrominanzwerte im Verhältnis 4:1 im Video abgespeichert. Dabei macht man sich entsprechende Schwächen des menschlichen Auges zur Nutze. Ein Makroblock enthält die benötigten Informationen, um einen 16x16 Pixel großen Teil des Bildes darstellen zu können. In der Makroblock-Ebene ist gespeichert, ob der Makroblock Intra oder Non-Intra kodiert ist. Gegebenfalls sind dort auch Bewegungsvektoren gespeichert.

Die DCT-Blöcke bilden die unterste Ebene des MPEG-Datenstroms. Sie enthalten die eigentlichen Bildinformationen, die sich aus Helligkeits- und Farbwerten der Pixel zusammensetzen. Ein Block besteht aus einer 8x8 Bildpixel-Matrix, auf welcher die Kompression und auch die Diskrete Kosinustransformation angewendet werden.

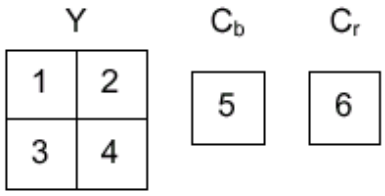

Abbildung 2: MPEG-1 Struktur Makroblockaufbau [Ti2001]

2.2 Sicherheitsmechanismen für Multimediadaten

In der Informationstechnologie können Multimediadaten verschiedenen Risiken ausgesetzt sein. Um einen Überblick über die Gefahren zu erhalten, wurden Sicherheitsziele definiert und ihnen die unterschiedlichen potentiellen Risiken jeweils zugeordnet. Die definierten Sicherheitsziele decken somit jeweils konkrete Sicherheitsrisiken ab. Das Erreichen der Sicherheitsziele erfolgt durch Entwicklung und Einsatz von Sicherheitstechnologien. Für die Entwicklung von digitalen Wasserzeichen sind die Schutzziele Authentizität und Integrität besonders wichtig. Neben diesen beiden Schutzzielen sind noch die Vertraulichkeit, Verfügbarkeit und Verbindlichkeit von Bedeutung. In anderen Sicherheitsanwendungen werden auch Ziele wie Zugriffsschutz und Nachweisbarkeit verfolgt [Ec2001].

Die gewünschten Schutzziele werden in aktive und passive Schutzmechanismen unterschieden [St2003]. Ausschlaggebend ist die Reaktion des Detektors/ Kontrollprogramms, wenn eine Verletzung des Schutzziels nachgewiesen wurde. Dabei lassen sich aktive und passive Schutzziele wie folgt beschreiben:

(1) Aktive Schutzmaßnahmen übermitteln eine erkannte Verletzung des Schutzzieles oder verhindern die Verletzung des Schutzzieles.
(2) Passive Schutzmaßnahmen dienen der Überprüfung möglicher Verletzungen von Schutzzielen.

Aktive Schutzmaßnahmen können bei Multimediadaten Verschlüsselung und Digital-Rights-Management (DRM) sein. Während Verschlüsselung den Zugriff auf die Multimediadaten verhindert, regeln DRM-Systeme die Zugriffsmöglichkeiten.

Digitale Wasserzeichen gehören zur Kategorie der passiven Schutzmaßnahmen. Sie dienen als Nachweismöglichkeit bei einer Verletzung der Schutzziele, wie z. B. Authentizität oder Integrität.

2.2.1 Schutzziele für Videodaten

Die Nutzung und Verbreitung von digitalen Videodaten ist den selben rechtlichen Regeln untergeordnet, wie sie für andere Multimediatypen gelten. Diese Arbeit betrachtet digitale Wasserzeichen, um speziell die Schutzziele Authentizität und Integrität durchzusetzen. Im Folgenden werden aber sämtliche Schutzziele in Relevanz zu Videodaten betrachtet, um einen allgemeineren Überblick zu erhalten [St2003]:

(1) Authentizität
Unter der Authentizität von Videodaten versteht man im Allgemeinen die Echtheit des Videos. Die Authentizität kann dabei Personen oder Informationen, wie Videodaten, zugeordnet werden.
Die Authentizität der Person beinhaltet je nach Anwendungsszenario die Überprüfung des Urhebers oder des Empfängers; mit Authentizität des Videos ist die Überprüfung dessen Originalität gemeint.
In Abhängigkeit zum Anwendungsszenario kann die Authentizität einer Person in die des Urhebers oder des Empfängers unterschieden werden. Die Urheberauthentizität trifft Aussagen über den Eigentümer. Häufig ist das der Hersteller der Videodaten; es besteht aber auch die Möglichkeit, entlang der Produktionskette von Videodaten andere Teilnehmer dem Urheber gleichzusetzen, wie z. B. die Produktionsfirma oder den Vermarkter.
Bei der Empfängerauthentizität wird das Video mit einem Empfänger in Verbindung gebracht. Als Empfänger des Mediums ist dabei der Käufer einer Kopie zu verstehen. Durch seine Individualität kann dadurch von einer personifizierten Kopie gesprochen werden. Sie bietet einerseits den Vorteil einer Bestätigung der Kopie für den Käufer; andererseits ist es aber auch möglich, Missbrauch mit der Kopie klar nachzuweisen.

(2) Integrität

Die Integrität beschreibt die Unversehrtheit von Videodaten. Eine Videodatei gilt als integer oder unversehrt, wenn keine Veränderungen an den Daten vorgenommen wurden [Ec2001].

Durch Verbesserungen der Computerressourcen und Videostandards ist es heute ohne großen Aufwand möglich, Videodaten zu verändern, um ihre ursprüngliche Aussage zu verfälschen. Dabei wird in gezielte und nicht gezielte Veränderungen unterteilt. Gezielte Veränderungen verfälschen bewußt die Aussage des Videos. Unter nicht gezielten Veränderungen sind Modifikationen zu verstehen, die durch übliche Operationen am Video auftreten. So werden z. B. Formatkonvertierungen und Skalierung als nicht gezielte Modifikationen eingestuft.

(3) Vertraulichkeit

Die Vertraulichkeit beschreibt die Geheimhaltung von Informationen gegenüber unbefugten Personen. Bei Videodaten kann es besonders wichtig sein, dass Inhalte aus Videos nicht veröffentlicht werden bzw. kein öffentlicher Zugang zu ihnen möglich ist. Die Vertraulichkeit gewinnt z. B. bei Überwachungs- oder Besprechungsvideos an Relevanz.

Es muss sowohl eine sichere Aufbewahrung der Videodaten als auch ein sicherer Transportkanal bei Übertragungen von Videodaten gewährleistet sein. Bei der Übertragung von Videodaten kann über entsprechende Verschlüsselungsverfahren eine vertrauliche Übertragung gewährleistet werden. Mit einem entsprechend guten Schlüsselmanagement kann der Austausch der Schlüssel für Ver- und Entschlüsselung geregelt werden. Durch die Hinzunahme einer PKI-Struktur wird die Vertraulichkeit zusätzlich erhöht. Eine zusätzliche Signierung der Videodaten vor der Übertragung gibt Sicherheit über die Quelle.

(4) Verbindlichkeit

Die Verbindlichkeit beschreibt die Nichtabstreitbarkeit bzw. Nachweisbarkeit von abgeschlossenen Verträgen. Im Videobereich ist das mit dem Kauf oder der Leihe von Videos über Online Shops zu vergleichen. Die während des Kaufs geschlossenen Vereinbarungen müssen von beiden Seiten eingehalten werden. Dazu gehört die Bereitstellung der Videos vom Verkäufer ebenso wie die Einhaltung von rechtlichen Bedingungen vom Käufer.

(5) Verfügbarkeit

Die Verfügbarkeit beschreibt die Erreichbarkeit von Quellen, von denen Medien erworben werden können. Die Quellen, z. B. Online-Shops, müssen während der Kaufabwicklung und des nachfolgenden Downloads immer erreichbar sein, damit es nicht zu einem Abbruch kommt. Ansonsten können Unregelmäßigkeiten während des Kaufs entstehen oder eine Wiederholung des Downloads notwendig werden. Die Betreiber der Online-Shops müssen garantieren können, dass sie immer erreichbar sind und dass sie im Fall von Übertragungsproblemen mit Hilfe von Aufzeichnungen nachverfolgen können, wer während des betreffenden Zeitpunkts Zugriff auf den Online-Shop hatte.

Besonders bei Videodaten ist aufgrund der hohen Datenmenge die Verfügbarkeit sehr wichtig. Deshalb ist es ratsam, durch parallele Betreibung mehrerer Server oder einer Backup-Lösung die Wahrscheinlichkeit eines Ausfalls so gering wie möglich zu halten.

2.3 Anwendungen digitaler Videodaten

Es gibt inzwischen eine Vielzahl von Anwendungen, bei denen digitale Videos zum Einsatz kommen. In dieser Arbeit werden sie in zwei Kategorien unterteilt: Transport-Streaming und Program-Streaming. Transport-Streaming beschreibt das Versenden von Videos über verschiedenartige Funknetzwerke. Program-Streaming ermöglicht den Zugriff auf das komplette Video.

In die Rubrik Program-Streaming fallen Anwendungen wie DVD-Video [Wi2007] oder ältere Anwendungen wie VCD (Video-CD) oder SVCD (Super-Video-CD). DVD-Video ist eine technische Spezifikation zur Speicherung von Videos auf einem DVD-Datenträger. Gegenüber seinen Vorgängern VCD und SVCD hat es eine deutlich bessere Qualität. Die Videos werden mit einer Bitrate von vier bis acht MBit/s auf der DVD gespeichert. Sie sind MPEG-2 konform und können somit eine Vielzahl der Erweiterungen von MPEG-2 gegenüber seinem Vorgänger MPEG-1 nutzen. Weitere Vorteile sind:

(1) mehrere Tonspuren, die es erlauben, sich das selbe Video jederzeit in anderen Synchronisationsfassungen anzuhören oder auch zu jeder Videostelle Kommentare von Produktionsbeteiligten zu hören,
(2) zuschaltbare Untertitel in bestimmten, auf der jeweiligen DVD vorhandenen Sprachen und
(3) ein Menü, mit dem sich bequem bestimmte Filmstellen, Tonspuren, Untertitel und auch das Bonusmaterial anwählen lassen, wie etwa:
    (a) eine Produktionsdokumentation des Films,
    (b) Interviews und Kommentare mit Produktionsbeteiligten, wie Regisseuren, Schauspielern und
    (c) weitere Extras, wie Kinotrailer, Musikvideos oder gar Computerspiele.

In den Bereich Transport-Streaming fallen Anwendungen wie das digitale Fernsehen DVB (Digital Video Broadcasting) [Wi2007]. Es gibt eine Reihenfolge von speziellen Anwendungen im DVB:

(1) DVB-S und DVB-S2 für die Übertragung durch direktstrahlende Satelliten,
(2) DVB-C für die Übertragung über Kabelnetze,
(3) DVB-T für die Übertragung durch terrestrische Senderketten im VHF- bzw. UHF-Bereich und
(4) DVB-H für die asynchrone Übertragung auf mobile Endgeräte (Handhelds), ebenfalls terrestrisch.

Wie beim DVD-Video werden auch beim DVB-S Datenraten von zwei bis vier MBit/s pro Videokanal erreicht. Allgemeine Vorteile des digitalen gegenüber dem analogen Fernsehen sind:

(1) Die Anzahl der Fernsehprogramme pro Kanal kann vervielfacht werden.
(2) Verschlüsselungsverfahren für Bezahlfernsehen sind einfacher und sicherer zu implementieren.
(3) Eine zusätzliche Verteilung von Rundfunkprogrammen ist möglich.
(4) Es können auch interaktive Datendienste im Kontext der angebotenen Programme übertragen werden.
(5) Die Bild- und Tonqualität kann gesteigert werden, so dass ein Zuschauer, der über ein hochwertiges Fernsehgerät verfügt, auch Sendungen in hochauflösender Qualität auswählen und empfangen kann (HDTV). Auch auf nicht hochauflösenden Fernsehern kann die Digitaltechnik ein viel rauschärmeres Bild und Raumklang ermöglichen.

Zur Zeit ist in Deutschland geplant, das analoge Fernsehen im Jahr 2010 komplett abzuschalten [Wi2007].

2.4 Digitale Wasserzeichen

Digitale Wasserzeichen sind in digitale Multimediadaten eingebettete Informationen. Das Ziel besteht darin, diese Information nicht wahrnehmbar in den Inhalt einzubetten, so dass nur die Personen, die über deren Existenz Bescheid wissen, sie bei Kenntnis über die Auslesemethode wiedererkennen können.

Die Einbettung von Wasserzeichen ist ähnlich dem des „Information Hiding“ [CoMi2002]. Unter „Information Hiding“ wird das Verstecken von Informationen in eine Trägerquelle verstanden. Bei Verwendung von Multimediadaten als Trägerquelle sind die versteckten Informationen digitale Wasserzeichen. Auch wenn die zur Zeit entwickelten Wasserzeichen immer nicht wahrnehmbar versteckt werden, basiert die Idee aber auf wahrnehmbare Wasserzeichen. Sie kommen z. B. in Geldscheinen oder als Fernsehsenderlogo zum Einsatz. Aufgrund ihrer Sichtbarkeit haben sie aber den eindeutigen Nachteil. Sie können gezielt angegriffen und zerstört werden. In dieser Arbeit werden nur nicht-wahrnehmbare Wasserzeichen betrachtet.

Unter einem nicht wahrnehmbaren, digitalen Wasserzeichen verstehen wir ein transparentes, nicht wahrnehmbares Muster, welches in das Datenmaterial (in dieser Arbeit speziell Videodaten) mit einem Einbettungsalgorithmus eingebracht wird [Di2000].

Ein Wasserzeichenalgorithmus umfasst zwei Prozesse:

(1) Einbettung des Wasserzeichens - E (Watermark Embedding)
(2) Detektion des Wasserzeichens - D (Watermark Retrieval)

Während der Einbettung E wird eine vordefinierte Wasserzeicheninformation W (Watermark Message) in das Trägerobjekt C (Original oder Cover) eingebettet. Die

eingebettete Information besteht aus einer Bitsequenz. Sie wird aus der Eingangsinformation gewonnen, welche eine Textinformation oder ein Bild- bzw. Audiosignal sein kann. Die maximal mögliche Kapazität pro kleinste Einheit des Trägermediums *C* bestimmt dabei die Länge der Information.

Basierend auf der Zusammensetzung der Wasserzeicheninformation *W* enthält der Wasserzeichalgorithmus zwei verschiedene Manipulationen, jeweils eine für den Bitwert 1 und eine andere für den Bitwert 0. Das Wasserzeichensignal selbst ist dabei häufig ein pseudozufällig gebildetes Rauschsignals, das basierend auf dem augenblicklichen Wasserzeichenbit in die jeweilige Position des Trägermediums *C* eingebettet wird. Da ein Wasserzeichen eine versteckte Information ist, muss eine Robustheit gegenüber gezielten Angriffen auf das Wasserzeichen garantiert werden. Den Aspekt der Sicherheit deckt dabei ein geheimer Schlüssel *K* ab. Mit dem geheimen Schlüssel *K* wird die Wasserzeichen-information sicher in das Trägerobjekt *C* eingebettet. Die meisten Wasserzeichenverfahren sind symmetrisch. Sowie während der Einbettung als auch des Auslesens der Wasserzeicheninformation wird der gleiche Schlüssel verwendet. Asymmetrische Verfahren sind weniger verbreitet. Eine Herausforderung liegt in der Anwendung des asymmetrischen Schlüsselpaares. Der Wasserzeichenalgorithmus muss beide Schlüssel akzeptieren, einen für die Einbettung und den dazugehörigen anderen Schlüssel für die Detektion des Wasserzeichens.

Viele Verfahren benötigen weitere Parameter zur Einbettung des Wasserzeichens, wie die Wasserzeichenstärke. Generell lässt sich ein Video V als Cover C wie folgt beschreiben:

$$\mathrm{V} = \{\mathrm{F}_1, ..., \mathrm{F}_n\}, n \in N, wobei\ F_i \in Frame \qquad (1)$$

Daraus ist zu erkennen, das ein Video aus einer geordneten Sammlung von Frames bzw. Einzelbildern besteht.

Der Einbettungsprozess kann wie folgt beschrieben werden:

$$\mathrm{V_w} = \mathrm{E}(\mathrm{V}, \mathrm{W}, \mathrm{K}) \qquad (2)$$

Bei symmetrischen Wasserzeichen wird während der Einbettung *E* und Detektion *D* der gleiche Schlüssel *K* benötigt. Aus dem markierten Trägerobjekt C_W werden die Wasserzeicheninformationen ausgelesen. Die Detektion D lässt sich bei Videodaten V_W wie folgt beschreiben:

$$\mathrm{W} = \mathrm{D}(\mathrm{V_w}, \mathrm{K}) \qquad (3)$$

Bei den in Formel (3) beschriebenen Wasserzeichenverfahren wird bei der Detektion nicht das Original *C* benötigt. Man spricht in diesem Zusammenhang von blinden Verfahren. Diese Verfahren lesen die Informationen mit Hilfe von statistischen Verfahren, wie z. B. Korrelationsanalysen aus. Andere Verfahren wiederum benötigen für die Durchführung einer Detektion das Original Trägerobjekt *C* (4). Diese Verfahren werden nicht-blinde Verfahren genannt. Dabei wird häufig das Original vom markierten Dokument abgezogen und das daraus entstandene Vergleichsbild enthält die Wasserzeicheninformationen. Diese

Möglichkeit zieht aber Performanz- und Sicherheitsnachteile mit sich, da zusätzlich zum markierten Objekt auch das Original übertragen und analysiert werden muss. Eine Lösung stellt die Übertragung von Metadaten zur Detektion des Wasserzeichens dar [HaKu2005], [HaBo2007]. Die Metadaten sind zusätzliche Informationen, die die Detektion des Wasserzeichens vereinfachen. Für Videodaten kann die nicht-blinde Detektion *D* wie folgt beschrieben werden:

$$W = D(V, V_w, K) \qquad (4)$$

Auch bei Kenntnis der entsprechenden Schlüssel ist es während der Detektion häufig nicht möglich, das Wasserzeichen komplett zu entfernen, da es sehr fest mit dem Trägermedium *C* verwebt ist. Nur unter deutlichen Qualitätseinbußen kann es entfernt werden. Die Anwendung eines reversiblen Wasserzeichens ermöglicht die Wiederherstellung des Originals.

Wie bereits erwähnt, entspricht die Wasserzeicheninformation einer Binärsequenz. Dadurch werden die durchzuführenden Veränderungen in zwei Mechanismen für Bit 0 und Bit 1 unterteilt. Die Information können vielfältig aufgebaut sein, lassen sich zur Vereinfachung aber in zwei allgemeine Typen unterteilen:

(1) Präsenzwasserzeichen
Ein Präsenzwasserzeichen trifft lediglich eine Aussage über die Existenz eines Wasserzeichens. Es wird auch als 1-Bit-Wasserzeichen bezeichnet.
(2) Informationswasserzeichen
Die Wasserzeicheninformation ist zu einer Binärsequenz kodiert worden. Die Ausgangsdaten können dabei den folgenden Daten entsprechen:
   (a) Urheberdaten zur Kennzeichnung der Urheberrechte,
   (b) kundenspezifische Daten zur Kennzeichnung von legalen und zur Verfolgung von illegalen Kopien, oder
   (c) beschreibende Daten (Metadaten).

2.4.1 Eigenschaften und Anwendungsgebiete digitaler Wasserzeichen

Die Qualität eines digitalen Wasserzeichens kann mit einer Folge mehrerer Eigenschaften beschrieben werden. Da teilweise die Bezeichnungen der Eigenschaften zwischen verschiedenen Veröffentlichungen variieren, konzentriert sich diese Arbeit auf die Angaben in [CoMi2002]. In anderen Veröffentlichungen werden die gleichen Eigenschaften aufgelistet, nur die Bezeichnungen können sich von der hier benutzten Quelle unterscheiden. [CoMi2002] klassifiziert digitalen Wasserzeichen in die folgenden Eigenschaften:

(1) Anwendungsorientierte Transparenz
Die Transparenz eines Wasserzeichens gibt Aussage über die Ähnlichkeit zwischen einem nicht-markierten und einem markierten Medium. Verschiedene Anwendungen erfordern eine unterschiedliche Transparenz. Übertragungen über Satellit können die Qualität von markierten Medien deutlich reduzieren. Das hätte zur Folge, dass aufgrund der Qualitätsminderung ein eingebettetes Wasserzeichen verloren gehen kann. In diesem Fall ist eine teilweise Wahrnehmbarkeit des Wasserzeichens auch akzeptabel, um Robustheitanforderungen ebenfalls gerecht zu

werden. Andere Anwendungen, wie HDTV und DVD, verlangen dagegen eine sehr gute Transparenz des Wasserzeichens, da das Videomaterial in einer sehr hohen visuellen Qualität vorliegt und Änderungen schnell wahrgenommen werden können.

(2) Kapazität des Wasserzeichens
Die Anzahl der Bits, die mittels eines Wasserzeichens in eine festgelegte zeitliche Einheit des Trägermediums eingebettet werden, wird als Datenrate bezeichnet. Im Bildbereich entspricht dies der Anzahl an Bits pro Bild, im Audiobereich dient eine Sekunde als Referenz und im Videobereich kann eventuell die Anzahl an Bits pro Frame oder auch pro Sekunde genannt werden [DiSt2004].
Verschiedene Anwendungsbereiche benötigen eine unterschiedliche Menge an Wasserzeichenbits. So wird z. B. im Bereich Kopierkontrolle eine geringe Menge von vier bis acht Bits benötigt. Anwendungen wie Broadcast Monitoring verlangen aber eine deutlich höhere Datenrate. So können z. B. 24 Bits zur Identifikation sämtlicher Produzenten, die Material für die Zusammenstellung von Nachrichtensendungen liefern, benötigt werden.

(3) Verwendung des Originals bei der Detektion des Wasserzeichens
In einigen Anwendungen liegt während der Detektion des Wasserzeichens das Original des Trägermediums vor. Bei Anwendungen wie dem Nachweis der Urheberschaft liest der Besitzer des Originals das Wasserzeichen aus. Während des Auslesens kann das Original als Unterstützung für den Ausleseprozess mit herangezogen werden. Dies verbessert das Ausleseergebnis, da durch Verwendung des Originals mittels Subtraktion von beiden Medien eine Differenz erstellt wird, welche die Wasserzeicheninformation beinhaltet. Das Original kann auch zur Rückführung von zeitlichen oder geometrischen Veränderungen genutzt werden.

Detektoren, die einen Zugriff auf das Original benötigen, werden als nicht-blinde Detektoren bezeichnet [CoMi2002]. Ein nicht-blinder Detektor benötigt zusätzlichen Informationen während der Detektion des Wasserzeichens. Detektoren, die hingegen ohne das Original das Wasserzeichen auslesen, werden als blinde Detektoren bezeichnet.

(4) Robustheit
Robustheit beschreibt die Fähigkeit eines Wasserzeichen, auch nach üblichen signalverarbeitenden Operationen detektierbar zu sein. Dazu gehören im Videobereich bildveränderende Operationen wie Skalierung, Rotation oder Ausschnitt. Ebenfalls gehören Operationen wie Videoaufzeichnung auf analogen Videokassetten und Veränderungen der Framerate dazu. Die jeweilige Anwendung bestimmt die Anforderungen an die Robustheit. So muss ein Wasserzeichen für den Rundfunk- und Fernsehbereich auch nach Übertragungen detektierbar sein.

Neben robusten Wasserzeichen finden auch fragile Wasserzeichen immer mehr Verwendung. Die Eigenschaft von fragilen Wasserzeichen ist der Nachweis von gezielten Angriffen auf das Trägermaterial. Gezielte Angriffe verändern bewusst den Inhalt des Videos, in dem z. B. einzelne Szenen herausgeschnitten oder verkürzt werden.

(5) Sicherheit

Die Sicherheit beschreibt die Resistenz des Wasserzeichens gegenüber gezielten Angriffen auf selbiges. Dabei kann in unauthorisiertes Detektieren, Entfernen oder Einbetten eines weiteren Wasserzeichens unterschieden werden. Unautorisiertes Entfernen und Einbetten verändern gezielt das Trägermaterial, unautorisiertes Detektieren versucht nur das Wasserzeichen auszulesen.
Nicht jeder Anwendungsbereich benötigt ein sicheres Wasserzeichen. So ist bei Wasserzeichen, die Mehrwertdienste unterstützen, die Sicherheit kein wichtiger Aspekt.

Der Grad der Sicherheit eines Wasserzeichens misst sich an der Verwendung des geheimen Schlüssels, der zum Ein- und Auslesen der eingebetteten Informationen benötigt wird. Ein Wasserzeichenalgorithmus sollte in Bezug auf die Verwendung des Schlüssels folgende Standards bieten:

(a) Die Bekanntheit des Ein- und Ausleseprozesses des Wasserzeichens sollte nicht die Sicherheit des Wasserzeichens negativ beeinflussen.
(b) Die Sicherheit sollte auf der Verwendung des Schlüssels basieren.
(c) Der Schlüssel sollte aus einem großen Schlüsselraum ausgewählt werden.
(d) Ein Großteil der entwickelten Wasserzeichenverfahren benutzt den selben Schlüssel zum Einbetten und Auslesen des Wasserzeichens. Mit der Verwendung eines zweiten Schlüssels zum Verschlüsseln der Nachricht kann die Sicherheit verbessert werden, da nach einer Kompromitierung des ersten Schlüssels zwar ein Wasserzeichen ausgelesen werden kann aber die eingebettete Nachricht nicht ermittelt werden kann.

Mit Wasserzeichen können Aspekte wie z. B. die Individualisierung einer Kopie umgesetzt werden. Dadurch kann der Ursprung einer möglichen unerlaubten Vervielfältigung nachgewiesen werden. Wenn ein Angreifer aber nicht als Ursprung der unerlaubten Verbreitung entdeckt werden will, muss er das eingebettete Wasserzeichen verfälschen. Deshalb muss der Einbettungsprozess eine Sicherheit gegenüber Manipulationen oder Löschen des Wasserzeichens bieten.
Hierbei wird versucht vorzutäuschen, das zweite Wasserzeichen sei das erste, ursprünglich eingebettete Wasserzeichen.

(6) Komplexität des Wasserzeichenalgorithmus’

Die Einbettung des Wasserzeichens in das Video erfordert einen zeitlichen Aufwand. Die Komplexität des Algorithmus ist eine Messgröße, wieviel Zeit benötigt wird, um ein Wasserzeichen in ein Video einzubetten. Bei einer Vielzahl von Anwendungen wird Echtzeitfähigkeit vorausgesetzt. Wenn beim Einbetten und Auslesen eine unakzeptable zeitliche Verzögerung auftritt, führt dies zu wirtschaftlichen Nachteilen. So darf im Broadcast-Monitoring das Auslesen des Wasserzeichens keine Verzögerung im Verarbeitungsprozess hervorrufen.

Ein weitererer wichtiger Bestandteil der Komplexität ist auch die Effektivität des Einbettungsprozesses. Die Effektivität des Einbettungsprozesses beschreibt die Wahrscheinlichkeit, mit der das Trägermedium erfolgreich markiert wurde. Darunter versteht man die Detektionsrate des Wasserzeichens unmittelbar nach dem Einbetten. Eine 100%ige-Detektionswahrscheinlichkeit ist nur schwer erreichbar, da

einzelne Einbettungsversuche nicht erfolgreich sind. So kann ein auf Kontraststrukturen basierendes Wasserzeichen in kontrastarmen Bereichen nicht erfolgreich eingebettet werden.

(7) False-Positive-Rate
Die False-Positive-Rate gibt die Wahrscheinlichkeit an, mit der zufällig erkannte Wasserzeichen das Ausleseergebnis verfälschen. Die Art der Anwendungen und der Medientyp haben direkten Einfluss auf die False-Positive-Rate. So muss bei Anwendungen, wie dem Nachweis der Urheberschaft, die False-Positive Rate, sehr gering sein. Falsche Ausleseergebnisse können zu möglichen falschen Beschuldigungen führen.

Neben den Eigenschaften haben sich auch wichtige Anwendungsgebiete für digitale Wasserzeichen herauskristallisiert. In [CoMi2002] werden verschiedene Anwendungsgebiete beschrieben, wobei sich diese Arbeit auf den Urheberschutz, die Kundenidentifizierung und den Nachweis der Authentizität des Videos konzentriert. Weitere Anwendungsbereiche sind die Annotation des Videos, der Kopierschutz, sowie die Übertragungskontrolle [Di2000].

(1) Identifikation des Eigentümers
Bei digitalen Medien ist es nach der Erstellung der ersten Kopie sehr schwer, den tatsächlichen Eigentümers des „Originals“ zu ermitteln. Deshalb bettet man Informationen des Eigentümers als Wasserzeichen in das Video ein. Da es als sichtbares Wasserzeichen sehr leicht entfernbar wäre, wählt man heute nur noch die nicht-sichtbare Variante. In Videobearbeitungsprogrammen kann der Wasserzeichendetektor integriert werden und beim Öffnen des Videos der Eigentümer automatisch ermittelt werden. Der Vorteil eines solchen Systems wäre die Informierung des augenblicklichen Benutzers, dass das verwendete Material urheberrechtsmäßig geschützt ist, ohne dabei Einschränkungen der visuellen Qualität in Kauf nehmen zu müssen. Diese würde nur dann auftreten, wenn das Wasserzeichen sichtbar in das Medium integriert worden wäre.

Ist der Detektor zudem noch öffentlich zugänglich, muss garantiert werden können, dass er nicht zum Missbrauch genutzt werden kann. Dies würde genau dann geschehen, wenn das ursprünglich eingebettete Wasserzeichen detektiert und gelöscht würde, um eine andere Nachricht einzubetten. Bei Ermittlung des Eigentümers würde der Detektor folglich ein falsches Ergebnis liefern.

(2) Verfolgung und Nachweis von unerlaubten Kopiervorgängen
Durch den kommerziellen Vertrieb digitaler Medien werden eine Vielzahl von Kopien erstellt und verkauft, die aber alle vollkommen identisch sind. Dadurch ist es unmöglich, eine Zuordnung zwischen Kopie und Kunden zu erreichen. Bei unerlaubten Verbreitungen sind Möglichkeiten, die Verbreitungswege erkennen und zukünftig kontrollieren zu können, von großer Bedeutung. Beispiele dafür sind Peer-to-Peer (P2P) Tauschbörsen und frei verfügbares, aber urheberrechtsmäßig geschütztes Videomaterial auf Internetseiten. Um nachzuweisen, welcher Käufer Ursprung der Verbreitung ist, wird ein Wasserzeichen mit kundenspezifischen Daten eingebettet. Wasserzeichen dieser Art, auch individuelle Wasserzeichen genannt, basieren auf Wasserzeichen zur Identifikation des Eigentümers, mit dem

Unterschied, dass hier individuelle Informationen des jeweiligen Käufers eingebettet werden.

Die Tatsache, dass sich diese Käufer-Informationen voneinander unterscheiden, machen sich Käufer zum Teil auch zu nutzen, in dem sie die Wasserzeicheninformationen so verfälschen, dass bei illegaler Verbreitung der ursprünglich legitimierte Käufer und damit die Quelle der unerlaubten Vervielfältigung nicht mehr nachweisbar ist. Dabei schließen sich mehrere Kunden zusammen, identifizieren die Unterschiede zwischen ihren markierten Videokopien und verfälschen diese. Möglichkeiten, wie diesem speziellen Angriff begegnet werden kann, werden in Kapitel 9.1.3 diskutiert.

(3) Überprüfung der Integrität und Authentizität des Videos
Der Inhalt von digitalen Medien kann vielfältig verändert werden. Da aber nicht jede Veränderung wahrnehmbar ist, kann es sich als schwierig erweisen, einen Unterschied zwischen Original und angegriffenem Video zu erkennen.

Zur Überprüfung der Integrität werden Videodaten in Hashwerte abgebildet. Hashwerte sind Einwegfunktionen, mit denen nachgewiwsen wird, ob das Video integer ist. Da sie aber als zusätzliche Information parallel zum Video abgespeichert werden, können sie während des Transports des Videos möglicherweise verloren gehen. Eine alternative Möglichkeit ist das Einbetten der Hashwerte als Wasserzeicheninformation [FrGo2001]. Während der Überprüfung der Integrität wird der eingebettete Hashwert wieder aus dem Video ausgelesen. Ein Vergleich des ausgelesenen mit dem erneut aus dem Video gebildeteten Hashwert ermöglicht die Identifikation von Veränderungen im Video.

Mit Hilfe von digitalen Signaturen kann die Authentizität des Videos nachgewiesen werden. Genau wie bei Hashwerten ist die getrennte Aufbewahrung von Signatur und Video von Nachteil. Eine Lösung stellt auch hier die Einbettung der Signatur als Wasserzeicheninformation in das zu untersuchende Video. Ein Auslesen und die anschließende Überprüfung der Signatur gibt Aussage über die Authentizität des Videos.

3 Anwendungsszenarien für Videowasserzeichen

In diesem Kapitel werden die Anwendungsfälle dieser Arbeit betrachtet. Sie dienen als Ausgangslage zur Identifikation der Schwachstellen existierender Wasserzeichen und zur Präsentation der entsprechenden Lösungen.

3.1 Nachweis des Urhebers bzw. Empfängers

Diese Anwendungsfälle stellen die klassischen Fälle für den Einsatz von robusten Wasserzeichen dar [Di2000]. Ziel ist es, mit Hilfe des Wasserzeichens nachzuweisen, wer der Eigentümer des Originals oder der Erstempfänger der Kopie ist. Aufgrund der stetigen Verbreitung digitaler Videos müssen aber noch weitere Anwendungsszenarien betrachtet werden. Die Verbreitung von Videomaterial geschieht nicht mehr nur über den Verkauf von Trägermedien, sondern immer mehr auch über verschiedene Netzwerke, was den Vorteil bietet, immer auf das Video über das jeweilige Netzwerk zugreifen zu können.

3.1.1 Online Shop

Mit Hilfe eines Online-Shops werden Güter jeglicher Art über das Internet verkauft. Dazu zählen auch Multimediadaten. Online-Shops sind eine zusätzliche Einnahmequelle zu den üblichen Vertriebswegen digitaler Medien bzw. bieten die Möglichkeit, sich verringernde Umsätze bei den konservativen Vertriebswegen wieder auszugleichen. Zudem fallen hier geringere Kosten an. Aufgrund dieser Vorteile etablieren sich Online-Shops immer mehr.

Durch den Anstieg im Verkauf digitaler Medien über Onlineshops sind aber auch parallele Verbreitungswege entstanden, die illegal sind. So werden Online-Tauschbörsen zum unerlaubten Austausch urheberrechtsgeschützter Mediendaten missbraucht. Deshalb bedarf es der Integration von Sicherheitsmaßnahmen wie Wasserzeichen oder DRM, die geeignet sind, um den Missbrauch zu verhindern oder zumindest einzudämmen [StHa2003].

Eine besondere Form von Online-Shops sind sogenannte Data-on-Demand-Systeme. Data-on-Demand-Systeme stellen dem Benutzer gezielt Daten zur Verfügung, die er zuvor angefordert hat (Video-on-Demand bzw. Audio-on-demand Services). Durch diese Dienste werden die angeforderten Daten meist in komprimierter Form zur Verfügung gestellt. Die Komprimierung der Daten bewirkt eine schnelle Übertragung der Daten zum Käufer. Insbesondere bei Videodaten ist das ein wichtiges Kriterium für den Einsatz von Data-on-Demand-Systemen.

Video-on-demand (VoD) ist ein aufkommender Vertriebsweg in der Videobranche. Der Verleih von Videos über das Internet stellt eine realistische Alternative zur konservativen Videothek dar. Zusätzlich kann der Verleih an die Möglichkeit des Kaufes gekoppelt werden. Parallel dazu werden auch die technischen Voraussetzungen zur Versendung und Darstellung des Videos kontinuierlich verbessert.

In dieser Arbeit liegt die Konzentration auf dem klassischen Video-Online-Shop, wo digitale Videos gekauft und heruntergeladen werden. In einem um Wasserzeichen erweiterten VoD Online-Shop können drei Parteien identifiziert werden:

(1) Verkäufer
Der Verkäufer benutzt das Video-Online-Shop Portal, um seine digitalen Mediendaten zu vertreiben.

(2) Kunde
Der Kunde erwirbt aus dem Online-Shop digitale Mediendaten. Während des Vorganges identifiziert er sich mit seiner Kunden-Identifikationsnummer. Diese Kunden-ID wird nach dem Bezahlen während des Downloadvorganges als digitales Wasserzeichen in das Medium eingebettet. Dadurch erfolgt eine Personalisierung des Mediums.

(3) Video-Online-Shop
Neben den Standard-Funktionalitäten eines Online-Shops fungiert dieser Online-Shop auch als Wasserzeichen-Server. Für den Verkäufer dient er als Plattform, um die digitalen Mediendaten zu präsentieren, anzubieten und durch Personalisierung zu schützen. Für den Käufer dient er als Plattform, um Mediendaten zu kaufen.

Abbildung 3 stellt einen auf Wasserzeichen basierenden Video-Online-Shop schematisch dar:

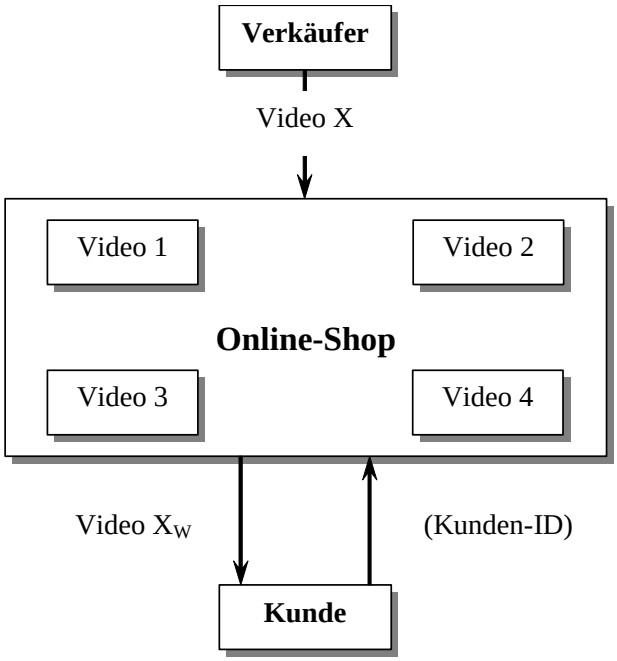

Abbildung 3: Darstellung des Video-Online-Shop

Wie bereits oben erwähnt, wird die Kunden-ID als digitales Wasserzeichen in das Video eingebettet, das dem Kunden zum Download zur Verfügung gestellt wird. Dabei muss das Wasserzeichen während der gesamten Laufzeit des Videos permanent in den Datenstrom eingebettet werden, um eine hohe Sicherheit zu gewährleisten. Da es passieren kann, dass zur selben Zeit mehrere Kunden das gleichen Videos downloaden möchten, müssen parallel mehrere Kopien erstellt werden, die jeweils ein personalisiertes Wasserzeichen enthalten.

Digitale Wasserzeichen, die für dieses Szenario eingesetzt werden sollen, müssen eine hohe Sicherheit aufweisen. Es muss nach Verkauf der Videos immer damit gerechnet werden, dass jegliche Kontrolle über eine Weitergabe verloren geht. Das ermöglicht es Angreifern, mit verschiedenen Angriffen, das eingebettete Wasserzeichen so zu verfälschen, daß es nicht mehr ausgelesen werden kann.

Zusätzlich ist die Transparenz ein wichtiger Aspekt. Wenn über den Online-Shop hochqualitative Video gekauft werden, darf die Transparenz keine sichtbare Verminderung der Qualität des Videos herbeiführen. Eine entsprechende Meßgröße wie der PSNR Wert kann Aufschluß über die visuelle Qualität des Videos geben und ein Anhaltspunkt für die Transparenz des Wasserzeichens sein.

Die Datenrate des Wasserzeichens richtet sich nach der potentiellen Menge an Käufern. Wird mit vielen potentielle Käufer gerechnet, muss die Datenrate entsprechend hoch sein. Zudem sollte mit in Betracht gezogen werden, dass durch ein Vergleich mehrerer Kopien eines Videos die Möglichkeit der Zerstörung der Wasserzeichen besteht [BoSh1995] [WuTr2004].

Der wichtigste Aspekt ist jedoch die Echtzeitfähigkeit des Wasserzeichens. Die Verarbeitung von Videos erfordert ein entsprechend hohes Maß an zur Verfügung stehenden Ressourcen. Zudem muss mit mehreren zeitgleichen Anfragen von verschiedenen Kunden gerechnet werden. Deshalb sollte die Komplexität des Wasserzeichenalgorithmus' eher gering sein. Während der Einbettung ist eine Möglichkeit die Konzentration auf jene Verarbeitungsschritte, die sich zwischen den Kunden unterscheiden [HaSt2008].

3.1.2 Online Shop Szenario mit Vorverteilung von Videos

Um eine gute Qualität des Videos zu gewährleisten, erfordert die Übertragung via Internet ein hohes Maß an Datenübertragung. Es gibt jedoch auch Anwendungsszenarien, bei denen die Menge der übertragenden Daten deutlich geringer ausfällt, ohne dass es zu Qualitätseinbußen kommt [HaSt2005a].

Dabei werden Filme in DivX- oder DVD-Qualität zum Beispiel auf einer DVD als Beigabe zu einer Fernsehzeitschrift ausgegeben. Aus der auf der DVD befindlichen Probeversion sind einzelne Komponenten des Films herausgeschnitten worden. Der Kunde hat die Möglichkeit, sich die Probeversion anzuschauen und kann dann entscheiden, ob er den Film kaufen möchte. Bei einem Kauf werden die fehlenden Daten über das Internet heruntergeladen und in das vorhandene Video integriert. Der Datenstrom enthält eine geringe Anzahl von Daten, da gegenüber dem klassischen On-Demand Szenario nicht das komplette Video vom Shop heruntergeladen werden muss. Dadurch ist der gesamte Downloadvorgang schnell abgeschlossen. In die zu übertragenden Daten wird ein zusätzliches digitales Wasserzeichen integriert, um die Möglichkeit des Nachweises der Urheberrechte zu ermöglichen.

Als Erweiterung dieses Szenarios könnten auch zusätzliche Videos in reduzierter DVD-Qualität verteilt werden. Diese Möglichkeit kann dem Kunden angeboten werden, wenn durch vorherige Käufe ein Trend in bezug zur Filmart, Regisseur oder Schauspieler erkennbar ist. Bei Interesse an einer qualitativ hochwertigen Ausfertigung können dann nach dem gleichen Prinzip wie oben beschrieben die noch fehlenden, mit einem Wasserzeichen

markierten, Videodaten über das Internet erworben werden. Somit bietet sich dem Videoanbieter die Möglichkeit, seinen gesamten Filmbestand besser zu vermarkten.

Das Szenario kann mit dem allgemeinen Konzept „Try and Buy“ verglichen werden [HaSt2005a], nur dass ein zusätzliches personifiziertes Wasserzeichen in die Videodaten integriert wird.

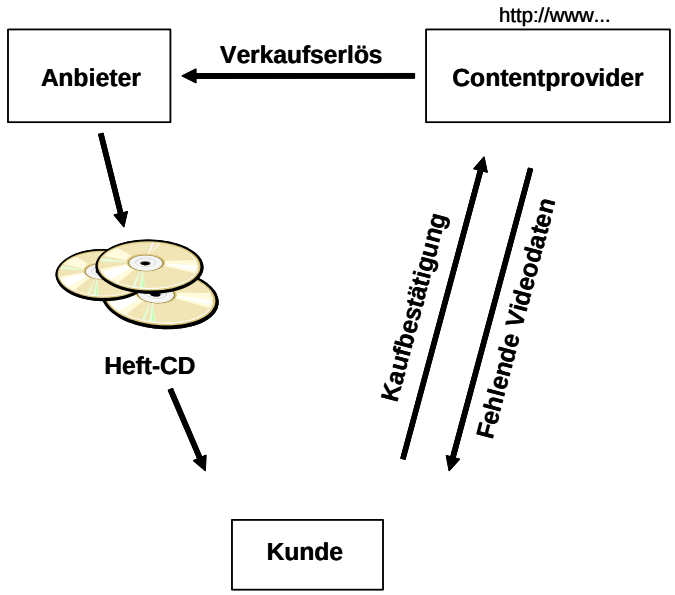

Abbildung 4: Verbreitungsmethode Try and Buy

Abbildung 4 stellt das Szeanrio noch einmal dar. Ein Anbieter vergibt Probevideos an interessierte Kunden, die aber eine niedrige Qualität aufweisen, aber immer noch den Inhalt erkennen lassen. Das Szenario kann sowohl für den Massenmarkt (Fernsehzeitschriften) als auch für spezielleres Videomaterial (Dokumentationen) angewandt werden.

Der Kunde hat mit einem Vorführvideo die Möglichkeit, erst nach einer Einschätzung das Video wirklich zu erwerben. Mit der Kaufbestätigung werden die mit einem Wasserzeichen versehenen Daten von Contentprovider zum Kunden übertragen. Auf der Seite des Kunden werden die übertragenen Videodaten in das Vorführvideo integriert. Das Ergebnis ist dann ein Video mit hoher Qualität, welches zusätzlich ein Wasserzeichen als Sicherheitsmerkmal besitzt.

Sie bietet für beide Seiten des Verkaufsprozesses Vorteile. Der Verkäufer bzw. Anbieter der Videos kann zum einen die Bekanntheit seines angebotenen Filmbestands durch die Vergabe von Vorführvideo steigern und, wie oben beschrieben, diesen gezielt vermarkten, zum anderen hat er auch die Möglichkeit, Sicherheitsmerkmale in die verkauften Videos einzubringen, ohne das gesamte Medium neu verschicken oder zum Kunden übertragen zu müssen. Die Erstellung der markierten Videodaten wird auf einem sicheren Server durchgeführt. Der Kunde hingegen hat die Möglichkeit einer ausgiebigen Bewertung der Videos und benötigt bei einer Kaufentscheidung nur noch die fehlenden Videodaten, um das Video zu komplettieren. Dies kann entscheidende Kostenvorteile beim Übertragen der Videodaten mit sich bringen.

Gegenüber dem vorherigen Szenario ist die effiziente Übertragung der markierten Mediendaten der wichtigste Aspekt. Die Eigenschaften des Wasserzeichen sind identisch mit denen des Online-Shop-Szenarios.

Die übertragenen markierten Videodaten müssen auf der Seite des Kunden in das Vorführvideo integriert werden, ohne dass eine erneute Kodierung des Videos notwendig ist. Sie sollten schon in enkodierter Form beim Kunden ankommen. Dadurch können sie schnell und effizient in das Video integriert werden.

Die Übertragungswege der markierten Videodaten vom Contentprovider zum Kunden müssen durch entsprechende sichere Übertragungskanäle geschützt werden.

3.1.3 Anwendungen in digitalen Bibliotheken

In digitalen Bibliotheken liegen die zu bearbeitenden Medien auf verschiedenen Speicherquellen verteilt vor. Sie bieten gegenüber traditionellen Bibliotheken den Vorteil, dass verschiedene Speicherquellen zusammengeschaltet werden und keine Konzentration auf zentrale Speicherquellen vorliegt. Dabei werden sie einer Vielzahl von Benutzern zur Verfügung gestellt [WiMc1999]. Nach dem Herunterladen können die Medien in verschiedenen Bearbeitungsschritten durch die Benutzer verändert und anschließend wieder in die Bibliothek hochgeladen werden.

Die Basisplattformen, von denen aus auf die digitalen Bibliotheken zugegriffen wird, regeln dabei die Sicherheitsaspekte des Medienzugriffs, d.h. durch die Anwendung von Signaturen wird die Authentifizierung von Benutzer und Computer überprüft. Was aber nicht oder nur teilweise realisiert werden kann, ist die Sicherheit der Medien selbst. Dabei kann es jedoch durchaus vom Eigentümer erwünscht sein, während des Herunterladens oder Einspielens einen Urhebernachweis in die Medien zu integrieren.

Da die Anzahl von Zugriffen auf Medien nicht fest geregelt oder begrenzt ist, wäre es ratsam, über lokale oder verteilte Netzwerke, wie Grid-Netzwerke, genug Ressourcen zur Integration von Mediensicherheit im Hintergrund zur Verfügung zu stellen. Grid-Netzwerke sind aus verschiedenen geographisch verteilten Ressourcen zusammengeschaltete Netzwerke, die ausreichend Computerstärke zur Integration von Wasserzeichen in Medien gewährleisten können. Aufgrund ihrer hohen Kapazität sind sie auch in der Lage, ausreichend Speicherquellen zur Verfügung zu stellen, um als Basisplattform für digitale Bibliotheken genutzt zu werden. Ein Betriebssystem regelt sämtliche Operationen im Grid-Netzwerk.

Im Projekt DILIGENT [Di2007] wurde eine Testplattform für digitale Bibliotheken basierend auf Grid Netzwerken entwickelt [HaSt2007]. Verschiedene Portale ermöglichen den Zugriff auf die Bibliothek und integrieren Anwendungen, wie Feature Extraktion und die Einbettung digitaler Wasserzeichen. Jeder Einbettungsaufruf wird auf die Resourcen des Grid Netzwerkes verteilt, wodurch eine Vielzahl von Einbettungsvorgängen in Echtzeit durchgeführt werden.

Abbildung 5 stellt die technische Integration von Mediensicherheit (ContentSecurity) in diesem Projekt dar. Es ist ein webbasierter Service, der so sehr einfach von anderen Komponenten im Projekt aufgerufen werden kann. Dieser Service kann immer dann

eingesetzt werden, wenn von oder zu verschiedenen Speicherquellen Medien transferiert werden. Dabei ist der Service nicht zwingend, sondern kann vom Benutzer optional angewählt werden.

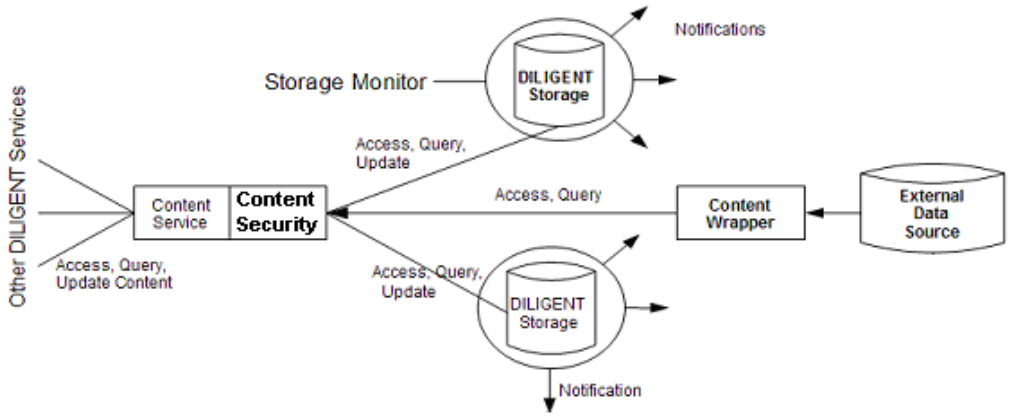

Abbildung 5: Integration von Mediensicherheit in DILIGENT

In diesem Szenario wird auf vorhandene Wasserzeichenverfahren zurückgegriffen [HaSt2005a]. Als Wasserzeicheninformation werden Copyrightinformationen eingebettet.

Da die Videos über verschiedene Quellen verteilt sind, macht es keinen Sinn, sie auf eine zentrale Quelle hochzuladen und dann das Wasserzeichen auf dieser Quelle in die Videos zu integrieren. Vielmehr sollte das zur Verfügung stehende Netzwerk mit seinen Resourcen genutzt werden, um die Einbettungsvorgänge zu verteilen. Wenn mehr Resourcen als Quellen zur Verfügung gestellt werden, besteht zudem noch die Möglichkeit, das Video in unabhängige Abschnitte aufzuteilen und jeden Abschnitt separat innerhalb des Grid-Netzwerkes zu markieren.

Unter Beachtung des Echtzeitanspruchs sollte doch darauf geachtet werden, dass durch die Aufteilung und das spätere Wiederzusammenfügen der zeitliche Aufwand nicht wesentlich höher ausfällt, als dies bei der Markierung des gesamten Videos in einer Einheit der Fall wäre.

3.2 Überprüfung der Authentizität von Videos

Die Authentizität des Videos gibt Aussagen darüber, ob das augenblicklich zu untersuchende Video verändert wurde. Eine Überprüfung der Authentizität verifiziert die Integrität des Videos [Wi2008].

Die Aussage des Videos kann durch einfache und nicht nachzuweisende Manipulationen verändert werden. Besonders wichtig ist die Authentizität bei Anwendungen, wie Videoüberwachung, Medizin oder Juristik. Ein oft vorhandener sichtbarer Zeitstempel im Bild kann nur als unzureichendes Nachweismerkmal hinzugezogen werden.

Abbildung 6 zeigt ein Beispiel einer relevanten Modifikation. Ziel der Veränderung ist eine veränderte Nummer auf dem Nummernschild, welche in Bild c erkannt und visuell dargestellt wird.

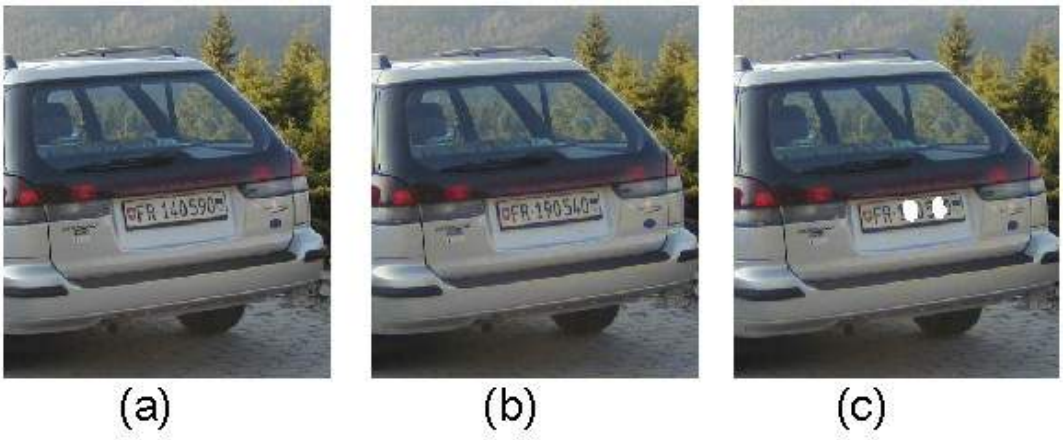

Abbildung 6: Beispiel einer Modifikation [KuJo2002]

Das für die Anwendung entwickelte Wasserzeichen besteht somit aus Daten zur Verifizierung der Authentizität. Dafür bietet sich der Einsatz von fragilen Wasserzeichen in leicht veränderter Struktur an. Ursprünglicherweise wird ein fragiles Wasserzeichen nur für den Nachweis der Integrität genutzt. Um aber den Nachweis der Authentizität zu ermöglichen, müssen Authentizitätsinformationen als Wasserzeichen in das Video eingebettet werden. Authentizitätsinformationen erfordern einen hohen Grad an Sicherheit, da das Wasserzeichen nach Einbettung in das Video gezielter Angriffspunkt sein wird. Nach Durchführung der Modifikation wird versucht, das Wasserzeichen so zu verändern, dass es die Veränderung nicht nachweisen kann und das Video weiter als authentisch einstuft.

Da viele schon bereits entwickelte digitale Wasserzeichen symmetrische Verfahren sind, muss ein sicherer Transport des geheimen Schlüssels gewährleistet werden. Sowohl beim Einbetten als auch beim Auslesen wird der gleiche Schlüssel verwendet. Wenn aber die Einbettung und Verifikation an verschiedenen Orten durchgeführt wird, ist es erforderlich, einen sicheren Transportkanals bereitzustellen. Durch den Einsatz eines asymmetrischen Schlüsselpaares kann eine sichere Übertragung gewährleistet werden. Werden asymmetrische kryptographische Technologien als Bestandteil des Wasserzeichenverfahren eingesetzt, kann die Authentizität direkt nach Erhalt des Videos überprüft werden und es muss keine entsprechende Schlüsselverwaltung durchgeführt werden. So wird das Wasserzeichen mit dem geheimen Schlüssel eingebettet und kann mittels des dazu gehörigen öffentlichen Schlüssels verifiziert werden. Das bietet den Vorteil, dass an jedem Ort das Wasserzeichen verifiziert werden kann. Durch die öffentlich durchführbare Verifikation muss aber die Sicherheit der eingebetteten Wasserzeicheninformation gewährleistet werden, da jede Person Zugriff auf die Daten hat. Durch den Einsatz entsprechender kryptographischen Methoden ist eine Gewährleistung möglich.

Ein Wasserzeichen führt aber immer zu Veränderungen im Trägermaterial. Verschiedene Wasserzeichenparameter können zu einer Verschlechterung des Trägermaterials führen. Wenn die Authentizität durch den Einsatz kryptographischer Methoden abgesichert wird, vergrößert sich auch die notwendige Kapazität des Wasserzeichens. Eine Steigerung der

Kapazität kann zu einer weiteren Verschlechterung der visuellen Qualität des markierten Videos führen. Deshalb muss es bei besonders sensiblen Anwendungsgebieten, wie in den oben aufgeführten Bereichen Juristik, Medizin oder Militär auch die Möglichkeit geben, das ursprüngliche Video wiederherzustellen. Dabei ist aber von der Einschränkung auszugehen, dass nur ausgewählte Personen das Recht der Wiederherstellung haben dürfen. Um eine Wiederherstellung des Originals zu ermöglichen, müssen die Informationen, die durch das Wasserzeichen überschrieben werden, vorher gesichert werden. Sie werden komprimiert und als Bestandteil der Wasserzeicheninformation in das Video eingebettet. Mit Hilfe von symmetrischer kryptographischer Verschlüsselung ist eine Sicherung der eingebetteten Originaldaten garantiert [HaDi2005].

Für das Anwendungsszenario ergeben sich drei Prozesse, die Bestandteil des Wasserzeichenverfahrens sind:

(1) Einbettung
Während der Einbettung wird die gesamte Wasserzeicheninformation in das Video integriert. Der Prozess besteht aus der Extraktion und Kompression der zu schützenden Videodaten des Einbettungskanals und der Zusammenstellung der Authentizitätsinformationen. Mit Hilfe einer symmetrischen Verschlüsselung werden die Videodaten vor unerlaubten Zugriffen geschützt. Die Authentizitätsinformationen werden mit Hilfe einer PKI-Struktur in das Video eingebettet.

(2) Verifizierung
Bei der Verifizierung wird die Authentizität des Videos überprüft. Da während der Einbettung die Sicherheitsinformationen mit dem geheimen Schlüssels in das Video integriert wurden, können sie mit dem öffentlichen Schlüssel verifiziert werden. Dadurch ist es jeder Person gestattet, die Sicherheit zu überprüfen.

(3) Wiederherstellung des Originals
Die Wiederherstellung des Originals kann nur die Person durchführen, die Zugriff auf den geheimen Schlüssel der symmetrischen Verschlüsselung hat. Mit der Entschlüsselung der Originaldaten kann das vor dem Einbetten des Wasserzeichens existierende Video wiederhergestellt werden. Um sicher gehen zu können, dass die symmetrisch verschlüsselten Originaldaten nicht verändert wurden, kann mit Hilfe von kryptographischen Hashfunktionen ihre Integrität separat geprüft werden.

Wasserzeichenverfahren, die für das vorgestellte Anwendungsverfahren entwickelt werden, legen besonderen Fokus auf die Sicherheit und Kapazität.

Die Sicherheit ist deshalb so wichtig, weil einerseits die Authentizität des Videos öffentlich verifiziert wird und andererseits durch die Wiederherstellung des Originals das eingebettete Wasserzeichen entfernt werden kann. Deshalb muss das Wasserzeichen jeden möglichen Versuch der Veränderung nachweisen können.

Die Kapazität ist aufgrund der verwendeten kryptographischen Methoden sehr wichtig. Die ständig fortschreitenden Angriffsmöglichkeiten, um eine kryptographische Methode zu brechen, erfordern deren stetige Verbesserung. Die dabei am häufigsten verwendeten Schritte

sind eine Verlängerung der verwendeten Schlüssel oder die Umstellung auf andere kryptograhische Methoden [Bu2007].

Robustheitsaspekte sind bei dieser Anwendung nicht wichtig, da keine Veränderung erlaubt ist. Deshalb ist der Einsatz von fragilen Wasserzeichen für diese Anwendung sinnvoll. Da das Originalvideo wiederherstellbar sein soll, kommen reversible Wasserzeichen zum Einsatz.

3.3 Gegenüberstellung von robusten und reversiblen Wasserzeichen

Robuste und reversible Wasserzeichen sind zwei unterschiedliche Wasserzeichenarten. Während robuste Wasserzeichen eher zur Authentizität des Eigentümers herangezogen werden, dienen reversible Wasserzeichen der Authentizität des Videos. Bei denen in Kapitel drei vorgestellten Anwendungsfällen werden robuste als auch reversible Wasserzeichen eingesetzt.

In diesem Kapitel werden basierend auf dem Klassifizierungsschema von [Di2000] beide Wasserzeichenarten verglichen. Dabei werden die Klassifikationsmerkmale in Anwendungsgebiet und Verfahrensparameter untergliedert.

3.3.1 Definition

Bevor die Gegenüberstellung beider Wasserzeichenarten durchgeführt wird, ist es sinnvoll, je eine Definition zu verfassen. Sie bilden die Grundlage für das Verständnis und die darauffolgende Klassifizierung.

3.3.1.1 Robustes Wasserzeichen

Ein robustes Wasserzeichen ist ein nicht-wahrnehmbares Muster, das in Mediendaten eingebettet wird und dabei üblichen Modifikationen am Datenmaterial widersteht. Übliche Modifikationen wie Skalierung oder Formatwandlung verändern das Datenmaterial ohne dabei das Ziel zu haben, das eingebettete Wasserzeichen zu zerstören.

3.3.1.2 Reversibles Wasserzeichen

Ein reversibles Wasserzeichen ist ein nicht-wahrnehmbares fragiles Muster, das wieder aus dem Datenmaterial entfernt werden kann. Der Vorgang der Entfernung aus dem Datenmaterial ist eine verlustfreie Rekonstruktion der Originaldaten unter dem Aspekt, dass keine Veränderungen an den markierten Daten vorgenommen wurden. Es bedarf einer hohen Sicherheit, damit kein Angreifer die Möglichkeit hat, dass Original selbst herzustellen.

3.3.2 Anwendungsgebiet

Robuste Wasserzeichen	Reversible Wasserzeichen
- Verfahren zum Urheberschutz Urheber, Produzenten oder Autoren fügen eine eindeutige Kennung in das Videomaterial ein. Die Kennung beschreibt sie als Eigentümer des Videos. Dabei besteht die Kennung eventuell aus einer allgemeinen Information (z. B. Name) oder hat eine Verbindung zu einem Datensatz einer Datenbank.	- Verfahren zur Medienauthentifizierung In das Video werden Informationen eingebettet, die es authentifizieren und zwar in Form einer Signatur. Die Verifikation der Signatur erfolgt während des Auslesevorganges. Wird die Signatur als authentisch eingestuft, ist auch das Video authentisch.
- Verfahren zur Kundenidentifizierung Wenn das Video über verschiedene Quellen vertrieben werden soll und eine Identifizierung des Empfängers notwendig ist, wird ein individuelles Wasserzeichen in das Video eingebettet. Die Identifizierung kann aus einer fortlaufenden Nummerierung bestehen oder auf entsprechende Kundennummern basieren.	- Verfahren zur Wiederherstellung Bei einzelnen Anwendungen kann es notwendig sein, dass das Wasserzeichen wieder entfernbar sein muss. Speziell dafür muss ein Teil der Wasserzeichen-information die Originaldaten beinhalten, die durch das Wasserzeichen verändert werden. Mit der Extraktion und des Zurückschreibens der Originaldaten kann das Original wiederhergestellt werden.

Tabelle 1: Anwendungsgebiete robuster und reversibler Wasserzeichen

3.3.3 Verfahrensparameter

Basierend auf der Anwendung setzen sich die Verfahrensparameter des Wasserzeichens unterschiedlich zusammen.

Parameter	Robuste Wasserzeichen	Reversible Wasserzeichen
Wahrnehmbarkeit	Robuste Wasserzeichen gehören zur Kategorie der nicht-wahrnehmbaren Wasserzeichen. Die durch das Wasserzeichen durchgeführten Veränderungen sind im Video nicht erkennbar. Bei Videodaten ist zudem noch zu beachten, dass die eingebetteten Wasserzeicheninformationen sich nicht nur der räumlichen Verteilung, sondern auch dem zeitlichen Verlauf anpassen müssen.	Reversible Wasserzeichen sind nicht-wahrnehmbare Wasserzeichen. Da das Original wiederherstellbar ist, können unter gewissen Umständen kleine wahrnehmbare Störungen zugelassen werden, die durch die Entfernung des Wasserzeichens auch gelöscht werden.
Robustheit	Bei robusten Wasserzeichen ist die Robustheit die wichtigste Eigenschaft. Das Wasserzeichen muss verschiedenen Modifikationen am Videomaterial widerstehen. Dazu gehören z. B. Formatkonvertierung, Kompression, Skalierung, Frameratenänderung und Szenenschnitte.	Bei reversiblen Wasserzeichen ist die Robustheit nicht wichtig. Sie werden der Kategorie der fragilen Wasserzeichen zugeordnet, die gezielt Angriffe nachweisen sollen.
Verifikation	Robuste Wasserzeichen werden geheim verifiziert, d. h. der Markierer oder eine festgelegte Gruppe von Personen hat das Recht der Verifikation.	Reversible Wasserzeichen können unter dem Aspekt der Medienauthentifizierung auch öffentlich verifiziert werden.

Tabelle 2: Verfahrensparameter robuster und reversibler Wasserzeichen

Komplexität	Robuste Wasserzeichen für Videodaten sind blinde Verfahren, d. h. das Originalvideo wird während der Verifikation nicht benötigt. Wenn hohe Anforderungen an die Robustheit gestellt werden, können auch semi-blinde Verfahren genutzt werden. Semi-blinde Verfahren nutzen Elemente des Originals zur Verbesserung der Detektion.	Reversible Wasserzeichen sind blinde Verfahren. Einzig allein mit Hilfe der detektierten Wasserzeicheninformation aus dem markierten Video wird die Authentizität überprüft.
Kapazität	Robuste Wasserzeichen sollten ausreichend Kapazität besitzen, um Informationen über den Urheber oder Empfänger sicher einbetten zu können. Für eine verbesserte Detektion werden häufig Fehlerkorrekturkodes und Synchronizationsinformationen mit in das Video eingebettet.	Reversible Wasserzeichen benötigen eine sehr hohe Kapazität. Der erforderliche Sicherheitsstandard für Signaturen oder Hashwerte wird kontinuierlich angepasst, was zu steigenden Anforderungen bezüglich der Kapazität führt.
Sicherheit	Die Sicherheit von robusten Wasserzeichen hängt vom geheimen Schlüssel ab. Eine mögliche Weitergabe des Schlüssels an andere Parteien erfordert eine hohe Sicherheit während des Transports.	Die Sicherheit von reversiblen Wasserzeichen lässt sich anhand des Zugriffsschutzes auf die Wasserzeicheninformation bewerten. Bei einer öffentlichen Verifikation ist eine Überprüfung der Integrität der Wasserzeicheninformation sehr wichtig. Die Wiederherstellung erfordert eine sichere Verschlüsselung der Originaldaten. Die Wiederherstellung des Originals ist nur einer vorbestimmten Gruppe von Personen möglich.

Tabelle 3: Verfahrensparameter robuster und reversibler Wasserzeichen

3.3.4 Resistenz gegenüber Angriffen

Eine Vielzahl möglicher Angriffe verfolgt das Ziel, das Wasserzeichen zu verfälschen oder zu zerstören. Dabei können spezifische Angriffe den jeweiligen Wasserzeichen zugeordnet werden. In [Di2000] werden bereits einige Angriffe erwähnt. Die folgenden aufgelisteten Angriffe sind besonders bei Video interessant.

Robuste Wasserzeichen	**Reversible Wasserzeichen**
- Stirmark-Angriff Es ist ein bekanntes Werkzeug, das eine Vielzahl von nicht-linearen Modifikationen im Frame durchführt [PeSt2001].	- Mehrfachmarkierung Ein Angreifer versucht die Authentizität des Videos zu verändern, in dem er eine eigene Signatur in das Video einbettet und so die vorher eingebettete Signatur überschreibt.
- Framekollisionen Ein Vergleich von benachbarten Frames kann Unterschiede in den Wasserzeichen-informationen erkennen und so gezielt verfälschen [SuKu2002].	- Veränderung der Semantik im Frame Es werden einzelne Objekte verändert oder entfernt. Dadurch kann die gesamte semantische Aussage des Videos verändert werden.
- Mehrfachmarkierung Ein zweites Wasserzeichen wird in das schon markierte Video eingebettet. Es entsteht das Problem der Identifizierung welches Wasserzeichen zuerst eingebettet wurde.	- Szenenmanipulation Durch Veränderung des Szenenablaufes kann die zeitliche Aussage im Video manipuliert werden. Es ist nicht mehr möglich, eindeutig nachzuweisen, wann welches Ereignis im Video geschehen ist
- Koalitionsangriff Bei individueller Markierung einer Kopie des Videos ist nur die Wasserzeichen-information unterschiedlich. Durch einen Vergleich der markierten Kopien können die Unterschiede sichtbar gemacht und gefälscht werden. Die Angreifer haben eine Vielzahl von Möglichkeiten, die Wasserzeichen zu verfälschen, was bis zu einer Falschidentifizierung führen kann [TrWu2003], [WuTr2004], [Di2000], [Hau2000].	

Tabelle 4: Resistenz von robusten und reversiblen Wasserzeichen gegenüber Angriffen

4 Stand der Technik

Dieses Kapitel behandelt den Stand der Technik von digitalen Wasserzeichen für Videodaten und erläutert ausgewählte Wasserzeichen. Anhand des Markierungskonzeptes lassen sich Wasserzeichenverfahren für Videodaten gut klassifizieren.

Im Bereich Videowasszeichen zeigen sich folgende Trends [CoMi2002], [SeLa2001]:

(1) Vom Einzelbild zur Bildreihenfolge
(2) Integration der zeitlichen Komponente
(3) Erweiterung auf neue Videostandards
(4) Kombination von Audio und Video

4.1 Einzelbildbasierte Wasserzeichen

Viele Videowasserzeichen greifen auf Konzepte für Bildwasserzeichen zurück. Da MPEG-1 und 2 auf dem JPEG Standard basieren, ist es sinnvoll, auf JPEG Wasserzeichen zurückzugreifen. Dabei können die Wasserzeichenverfahren anhand der angewandten Transformation in DCT (Discrete-Cosinus-Transformation)-, DWT (Discrete-Wavelet-Transformation)- und FFT (Fast-Fourier-Transformation)-Ansätze unterteilt werden. Die DCT-Transformation transformiert einen quadratischen Pixelblock in den Frequenzraum, die FFT-Tranformation transformiert einen vorher festgelegten Bildbereich oder das gesamte Bild in einen Wertebereich von z. B. 1024 Frequenzen und die DWT-Transformation transformiert ein Bild in vier einzelne Subbänder (LL, HL, LH und HH), wobei sie unterschiedlich mit Wasserzeichen markiert werden.

DCT-Wasserzeichenansätze bilden den größten Anteil an der Entwicklung von digitalen Wasserzeichen für Videodaten, da MPEG-Videos auf der DCT-Transformation basieren. Dabei werden verschiedene Methoden zur Markierung der DCT-Koeffizienten angewandt. Häufig werden niederfrequente DCT-Koeffizienten markiert (z. B. [JiYo2000]), einzelne Verfahren verändern den DCT-Block jedoch auch anderweitig. In [SeLa2001] werden zum Beispiel hochfrequente Komponenten weggeschnitten, in dem die Einzelwerte gelöscht wird.

[ZhCh2001] beschreibt ein nicht blindes Wasserzeichenverfahren, das die Wasserzeichen-information in DC und AC-Koeffizienten unterschiedlich einbettet. Der DC-Koeffizient beschreibt den Gleichanteil der Frequenzverteilung im DCT-Block. Die AC-Koeffizienten beschreiben die Differenzen zum DC-Wert. Das Wasserzeichen setzt sich aus zwei unterschiedlichen Informationen zusammen, wobei das erste Wasserzeichen ein Urheberlogo ist und das zweite Wasserzeichen aus Urheberdaten besteht. Für das zweite Wasserzeichen wird zusätzlich eine Textur- und Helligkeitsanalyse durchgeführt, wodurch die lokale Markierungsstärke gemessen wird. In [ChXu2006] wird ein LSB-Wasserzeichen in die DC-Koeffizienten eingebettet. Dabei werden verschiedene Fehlerkorrekturkodes (ECC) untersucht und ein hybrides Schema gebildet, welches neben den Wasserzeicheninformationen zwei ECC Kodes beinhaltet. Durch den Einsatz der ECC Kodes kann die Robustheit des Verfahrens verbessert werden. Die Markierung des DC-

Koeffizienten wird einerseits durch die schwache Quantisierung des DC-Wertes begründet und andererseits sind nach Modifikationen am Video die Veränderungen bei den AC-Koeffizienten signifikanter als bei DC-Koeffizienten. Deshalb wird das Wasserzeichen nur in die DC-Werte eingebettet. [Ha2000] markiert Gruppen von DCT-Koeffizienten von MPEG-2 Sequenzen pro Wasserzeichenbit. Dabei werden die DCT-Koeffizienten der Markierungspunkte in Gruppen aufgeteilt und so manipuliert, dass sie während der Detektion mit Hilfe einer Korrelationsentscheidung das Wasserzeichenbit bestimmen. So wird das Verhältnis der Energie beider Gruppen so verändert, dass in Abhändigkeit vom jeweiligen Wasserzeichenbit eine Manipulation der Energiestärke durchgeführt wird. Basierend auf der Auswahl der DCT-Koeffizienten ist das Wasserzeichen auch in den nachfolgenden Frames auslesbar.

[DiHa2001a] und [BiDa2005] markieren hingegen AC-Koeffizienten. [DiHa2001a] bettet kundenspezifische Informationen in Videosequenzen ein. Kundenspezifische Informationen können fortlaufende Kundenidentifikationsnummern oder digitale Fingerabdrücke sein. In [DiHa2001a] wurden zwei verschiedene Fingerabdrücke getestet. Ein Koalitionsangriff ist eine spezielle Form einer Beurteilung der Sicherheit von digitalen Fingerabdrücken. Mit verschiedenen Angriffn können die eingebrachten Informationen teilweise oder vollständig zerstört werden. In [WuTr2004] werden verschiede Konzepte von Koalitionsangriffen beschrieben. Es wird erläutert, dass nicht nur die Einbettungsmethode des Wasserzeichens wichtig ist, auch der Aufbau der Fingerabdrücke ist mit entscheidend, um resistent gegenüber Koalitionsangriffen zu sein. Ein beispielhafter Angriff verfälscht die erkannten Differenzen in den Videos und wird in [WuTr2004] als „average attack“ bezeichnet. In einer weiteren Methode werden einzelne GoPs zwischen unterschiedlich markierten Videos ausgetauscht. Das Ergebnis ist der Austausch einzelner Informationen zwischen den Videos, was bei kundenspezifischen Wasserzeichen im schlechtesten Fall zu Beschuldigungen falscher Kunden führen kann. Das Wasserzeichen wird auf die mittelfrequenten AC-Koeffizienten aufaddiert. Der mittlere Frequenzbereich ist geeignet für Wasserzeichen, da einerseits die Veränderungen nicht sichtbar sind und andererseits wird der Bereich nicht stark quantisiert. Bei [BiDa2005] wird das Wasserzeichen auch auf den gleichen Frequenzbereich aufgetragen. Als Wasserzeichen wird ein Bild in acht Teilbilder zerlegt, wobei jedes Teilbild eine Bitebene das Wasserzeichenbildes darstellt. Jedes der Teilbilder, auch als Planes bezeichnet, wird in die mittelfrequenten AC-Koeffizienten sämtlicher Frames einer Videoszene eingebettet. Ein visuelles Modell misst den Grad der Bildaktivität anhand des Verlaufes von Ecken und Konturen im Block. Ein hoher Aktivitätsgrad ermöglicht ein stärkeres Einbetten des Wasserzeichens. Während des Detektion kann nach Auslesen der Planes das Wasserzeichenbild wieder rekonstruiert werden.

[WaPe2006] markiert das Video auch im DCT-Frequenzraum, aber nicht im 8x8 Bildpixelblock, sondern nach Durchführung einer DCT-Tranformation des gesamten Bildbereiches. Dabei werden aber nicht die I-Frames direkt transferiert. Es wird eine DCT-Transformation beider Nachbarframes durchgeführt und ein interpoliertes Frame aus beiden transferierten Nachbarframes erstellt. Das führt zu besseren Ergebnissen, da nicht nur der lokale Inhalt eines Frames als Basis zum Einbetten des Wasserzeichens dient. Die AC-Koeffizienten werden basierend auf dem Quantisierungsfaktor des augenblicklichen DCT-Blockes angepasst. Je höher die Frequenz des augenblicklich betrachteten AC-Koeffizient ist, desto stärker wird er markiert. Der PSNR Wert des untersuchten Videomaterials liegt zwischen 43 und 45 dB und das Verfahren ist robust gegenüber einem leichten Rotationsgrad und einer Skalierung von 20%. Der PSNR Wert ist das Verhältnis zwischen dem Bildsignal

und dem durch das Wasserzeichen hinzugefügten Rauschsignal. Der Wert wird üblicherweise in Dezibel angegeben. Liegt ein hoher PSNR Wert vor, dann hat das Wasserzeichen eine hohe Transparenz und ist nicht sichtbar.

[ThVo2004] verfolgt ein ähnliche Strategie wie [WaPe2006]. Das Wasserzeichen ist nicht von einzelnen Blöcken abhängig. Es werden Regionen aus DCT-Blöcken gebildet und anhand der Intensität des Kantenverlaufes sortiert. Man geht davon aus, dass dieses Merkmal auch nach verschiedenen Modifikationen weitgehend erhalten bleibt. Während der Einbettung werden zwei Schemen angewandt. Im ersten Schema wird aus jeder Region jeweils ein Block ausgewählt und einer Blockgruppe zugeordnet. Ziel ist es, je nach Wasserzeichenbit eine Mehrzahl von Blöcken mit einem Blockdurchschnitt zu haben, der über bzw. unter dem Durchschnitt der gesamten Blockgruppe liegt. Der Blockdurchschnitt wird anhand der Luminanzwerte im Block gebildet. Ist dies nicht gegeben, werden die Blöcke verändert, die am nächsten zum Blockgruppendurchschnitt liegen. Im zweiten Schema wird das Verhaltnis zwischen ausgewählten AC-Koeffizienten verändert. Nach der Subtraktion beider Werte ergibt das mod2 Ergebnis das Wasserzeichenbit. Das Verfahren wurde während einer Testreihenfolge als transparent bewertet und ist robust gegenüber Filteroperationen und Re-Enkodierung. Dagegen besitzt das Verfahren ein hohe Fehlerrate nach Durchführung einer Skalierung.

[SeLa2001] dagegen führt keine Veränderung der DCT-Frequenzwerte durch, sondern entfernt Frequenzbereiche in den DCT-Blöcken. Das Verfahren ist besonders für Videos mit niedriger Bitrate spezifiziert. Der DEW (Differential Energy Watermarking) Algorithmus stellt zwischen zwei Bereichen ein Verhältnis in den Frequenzen her. Dabei wird die Summe der DCT-Koeffizienten beider Regionen berechnet und als Verhältnis zueinander gegenübergestellt. Die angewandte Einbettungsmethode ist die Entfernung hochfrequenter Werte in den DCT-Blöcken. Besonders beachtete Aspekte sind die Kapazität, die visuelle Qualität und die Robustheit des Wasserzeichens. Die Anzahl der zu markierenden DCT-Blöcke bestimmt die Kapazität, die minimale Energiedifferenz die Robustheit und der minimale Schnittpunkt die visuelle Qualität des Wasserzeichens. [LiZo2004] erweitert den DEW Algorithmus mit einer Anpassung der Energiedifferenz und des minimalen Schnittpunktes. Die Energiedifferenz wird jetzt ins Verhältnis zur globalen Energie des Frames gesetzt. Dadurch wird die Fehlerrate gesenkt, da das Wasserzeichen jetzt nicht mehr vom Energieverhältnis zweier Blockgruppen abhängig ist. Lokale Eigenschaften bewirken, dass das Wasserzeichen nicht erfolgreich eingebettet werden kann. Im ursprünglichen DEW Algorithmus muss die Energie beider Blockgruppen über der vorher festgelegten Energiedifferenz liegen. In [LiZo2004] wird aber nur die Energie einer Gruppe manipuliert. Dadurch ist die Veränderung innerhalb des Frames geringer.

Entgegen der Anwendung der DCT-Transformation zur Einbettung der Wasserzeichen werden auch andere Frequenztransformationen genutzt.

In [FuSh2007] und [Fu2007] werden objektbasierte Videowasserzeichen vorgestellt. Durch Einbetten des Wasserzeichens in das Videoobjekt ist das Wasserzeichen resistenter gegenüber Rotation und Skalierung. Das Videoobjekt wird normalisiert und mit Hilfe der DWT in den Frequenzraum transferiert. Das Wasserzeichen wird auf die mittelfrequenten DWT-Koeffizienten aufgetragen. Es verändert das Verhältnis zwischen zwei Koeffizienten im LH und HL Subband, die sich an der gleichen Position im Subband befinden. Dabei wird der Durchschnitt beider Koeffizienten auf einen Koeffizienten addiert und den jeweiligen

anderen Koeffizienten subtrahiert. Während der Detektion wird zuerst analysiert, ob das Frame markiert ist und danach das Verhältnis beider Koeffizieten untersucht. Eine niedrig festgelegete FAR (False-Acceptance-Rate) verhindert die Identifikation von unmarkierten Videos als markiert.

[KoYa2007] hingegen bettet das Wasserzeichen zwar in die räumlichen Komponenten ein, nutzt aber die DWT-Transformation auf der Suche nach Blöcken mit hoher Helligkeit, textueller Struktur und hohem Bewegungsgrad. Die in diese Kategorie fallenden Blöcke werden stärker markiert als die anderen Blöcke. Für die Untersuchung nach den Eigenschaften Helligkeit, Textur und Bewegung werden einzelne DWT-Subbänder untersucht. Der JND-Wert (Just-noticable-difference) dient als visuelles Modell des Wasserzeichens. Die Wasserzeicheninformation ist ein Einzelbild, von dem vor der Einbettung in das Videoframe ein Hashbild berechnet wird. Dadurch ist nur bei Vorhandensein des Schlüssels, mit dem während der Einbettung das Hashbild gebildet wurde, eine Überprüfung möglich. Während der Detektion wird durch einen Vergleich des neu gebildeten und ausgelesenen Hashbildes die Wasserzeicheninformation rekonstruiert.

Die DWT weist aber auch einzelne Mängel auf, wie z.B. dass sie keine Bewegung dargestellen kann. Deshalb wird eine erweiterte Variante, die Dual-Tree-Complex-Wavelet-Transformation (DC-CWT), die in [MaEl2006] genutzt wird. Das Wasserzeichen wird durch einen Vergleich der Nachbarpixel in das transferierte Frame eingebettet. Die Pixel des Wasserzeichenbildes werden vor der Einbettung pseudo-zufällig neu angeordnet. Die Einbettungsmethode kann einem Spread-Spektrum-Wasserzeichen gleichgesetzt werden, d.h. die Pixel werden über das gesamte Frame verteilt. Während der Detektion wird durch Betrachtung der Nachbarpixel das Bit wieder ausgelesen und das Wasserzeichenbild wieder in seinen ursprünglichen Zustand zurückgeführt.

In [KaDe1999] wird ein Wasserzeichen präsentiert, das speziell für den Fernsehbereich entwickelt wurde. Es ist robust gegenüber einer Vielzahl von Verarbeitungsschritten, die bei der Übertragung des Videosignals auftreten. Das Wasserzeichen weist eine geringe Komplexität auf und hat eine geringe False-Positive-Rate und False-Negative-Rate. Die geringe Komplexität ist notwendig, da es nicht zu einer Verzögerung während der Bearbeitungsschritte kommen darf. Die geringe False-Positive-Rate verhindert eine Identifikation nicht markierter Video und die geringe False-Negative-Rate das Auslesen falscher Wasserzeicheninformationen.

Einzelne Konzepte betten das Wasserzeichen auch direkt in die Bilddaten ein. Häufig werden hierbei die Luminanzdaten markiert, da Veränderungen nicht bzw. nur schwach erkannt werden können. [DiSt1998] bringt ein niederfrequentes Muster in die Helligkweitswerte ein. Es wird während des Einbettungsprozesses ein additives Rauschmuster auf die Luminanzwerte aufaddiert. Das Wasserzeichenmuster ist ein 8x8 Bildpixel großes Helligkeitsmuster, dass auf die Luminanzwerte des DCT-Blockes addiert wird. Das Verfahren ist besonders robust gegen Tiefpassfilter, Ausschnittbildung und Rauschaddition.

Während der Entwicklung bildbasierter Wasserzeichen haben sich aber die Unterschiede von Video- zu Bilddaten bemerkbar gemacht. In einzelnen Veröffentlichungen wird das Problem der Framekollision vorgestellt [SuKu2002], eine allgemeine Schwäche von Bildwasserzeichen diskutiert und Regeln für Videowasserzeichen aufgestellt. Es kann davon ausgegangen werden, dass unterschiedliche Wasserzeichen in Nachbarbildern sichtbar

werden können bzw. dass es möglich ist, mit Hilfe von Framevergleichen die Unterschiede aufzuspüren. Deshalb wird das Wasserzeichen auf extrahierte Merkmale der Frames aufgetragen und innerhalb eines Bereiches, wie einer Szene, identisch gehalten [HaKu2005].

4.2 Szenenbasierte Wasserzeichen

[DeCs1999] bettet zwei Wasserzeichen in das Video ein. Das erste Wasserzeichen basiert auf der Diskrete-Fourier-Transformation (DFT) und markiert eine Videoszene mit dem Wasserzeichen. Das zweite Wasserzeichen ist ein Template, das zur Erkennung verschiedener Attacken auf das markierte Videomaterial dient. [SwZh1998] nutzt dagegen die Diskrete Wavelet Transformation (DWT) zur Markierung von Videoszenen.

[LiHu2006] bettet zwar das Wasserzeichen in I-Frames ein, zieht aber zur Berechnung des Wasserzeichensignals das zeitlich vorher abgespeicherte Frame hinzu. Ein „Block Matching“ Algorithmus analysiert die Bewegung zwischen den beiden Frames und berechnet die Bewegungsvektoren. Pro Block werden je zwei Wasserzeichenbits eingebettet, in dem die Bewegungs des Blockes so verändert wird, dass sie die Wasserzeichenbits ergibt. So muss bei den Bits „00“ die Bewegung von unten rechts nach oben links erfolgen. Ist das nicht gegeben, dann wird der Blockinhalt so manipulitert, das diese Bitfolge ausgelesen werden wird.

Neben der Entwicklung von szenenbasierten Videowasserzeichen wurden auch Lösungen für die zeitliche Synchronisation entwickelt [LiDe2003], [RoVl2005], [HaKu2005]. Mit Methoden wie Schlüsselmanagement oder robusten Hashwerten als Metainformation werden die Frames oder Szenen im Video wiedergefunden, die mit dem Wasserzeichen markiert wurden. Dies ist besonders hilfsreich, wenn vorausgewählte Szenen mit Bildwasserzeichen markiert werden.

4.3 Wasserzeichen für neue Videostandards

Auch wenn die meisten Verfahren für MPEG-1 und -2 entwickelt wurden, wurden auch für neue und weiterentwickelte Videostandards, wie h.264 Verfahren entwickelt [WiSu2003]. [Ha2000] bettet das Wasserzeichen in MPEG-4 Videos ein, in dem es Kopf und Schulter von Menschen markiert. In [WuWa2005] wird ein Verfahren für h.264 vorgestellt, das auch nach einer Kompressionsrate von 40:1 gegenüber MPEG I-Frames auslesbar ist. Es wird ein 1-Bit-Wasserzeichen in 4x4 DCT-Blöcke eingebettet. In [ZhHo2007] wird ein weiteres h.264 Wasserzeichen vorgestellt. Das eingebettete Wasserzeichenmuster wird zuerst h.264 DCT-Blockdaten angepasst und in die komprimierte Domaine eingebettet. Das Wasserzeichen besitzt hohe Robustheit, gute Transparenz und erhöht dabei nicht die Bitrate des Videos. Verschiedene Test haben bewiesen, dass das Wasserzeichen robust gegen Formatkonvertierung und Signalmodifikationen, wie z. B. Kompression und Konstrastveränderung, ist.

In [NoMe2006] wird ein Wasserzeichen vorgestellt, das nach einer Bewegungsanalyse geeignete DCT-Koeffizienten findet, die zur späteren Markierung des Wasserzeichens

genutzt werden. Dabei wird davon ausgegangen, dass bei einem hohen Bewegungsgrad eingebettete Wasserzeichen weniger wahrgenommen werden, als bei geringer Bewegung. Es wird die Bewegung vom I-Frame zum letzten P-Frame analysiert. In dem Bewegungshistogramm werden die Pixel erkannt, die auf einer hohen Bewegung basieren. Die Wasserzeicheninformation wird über die geeigneten Koeffizienten verteilt.

5 Analyse existierender Verfahren und Ansätze der Arbeit

Videodaten haben in ihrem Aufbau eine große Ähnlichkeit mit Bilddaten. Deshalb ist es durchaus möglich und sinnvoll, Ansätze für Bildwasserzeichen auf Videodaten zu adaptieren. Dennoch sollten Unterschiede, die bei der Kodierung und Darstellung von Video zu Bild vorhanden sind, betrachtet und mit in die Entwicklung von Videowasserzeichen integriert werden.

Die Ergebnisse der Evaluation dienen der Identifikation möglicher existierender Schwachpunkte, die anschließend genauer erläutert werden. Dabei wird für jeden erkannten Schwachpunkt ein Konzept vorgestellt und auf das entsprechende spätere Kapitel in der Arbeit hingewiesen, das das jeweilige Konzept ausführlich behandelt.

5.1 Evaluation existierender Wasserzeichen für MPEG Daten

Während des Projekts H2O4M [DiSt2000] wurden verschiedene Wasserzeichenverfahren untersucht, die im vorherigen Kapitel beschrieben worden sind. Die Untersuchung sollte die Qualität der Verfahren ermitteln und mögliche Schwachstellen bestimmen. Die erkannten Schwachstellen dienen als Vorlage für die weitere Vorgehensweise in dieser Arbeit.

Während der Untersuchung wurden die Wasserzeichen so selektiert, dass sie verschiedene Markierungsstrategien abdecken. Dabei liegt der Fokus auf drei verschiedene Verfahren [DiHa2001a], [DiSt1998] und [SeLa2001]. Während [DiHa2001a] ein Wasserzeichen auf einzelne DCT-Werte aufträgt, baut [SeLa2001] ein Verhältnis zwischen zwei DCT-Bereichen im Frame auf, in dem es hochfrequente DCT-Werte wegschneidet. In [DiSt1998] werden durch das Auftragen eines niederfrequenten Muster die Luminanzen innerhalb eines DCT-Blocks verändert. Eine ausführlichere Beschreibung der Algorithmen kann im Kapitel zum Stand der Technik gefunden werden.

5.1.1 Auswahl des zu untersuchenden Videomaterials

Als Testmaterial dient MPEG-1 Videomaterial vom Deutschen Rundfunkarchiv (DRA) [DRA2006]. Es ist ein Referenzvideosatz von 20 Testvideos, für die zusätzlich auch eine Kopie mit einer festen GoP-Struktur von zwölf Frames pro GoP erstellt wurde.

Das Videomaterial selbst wurde so ausgewählt, dass verschiedene Eigenschaften abgedeckt werden. Das Videomaterial besteht aus einzelnen Standbildvideos bis zu einem schnellen und gleichmäßigen Bewegungsgrad. Zusätzlich wurden auch Videos herangezogen, wo Szenenwechsel auftreten. Bei der Vielfalt der verwendeten Farben wurde darauf geachtet, daß sowohl Schwarz-Weiß-Videos als auch Farbvideos ausgewählt wurden. Die Farbstärke und der Kontrast in den Farbvideos sind auch breit gestreut. Diese Eigenschaften können durchaus wichtige Aspekte zur Bestimmung der Qualität von digitalen Wasserzeichen sein.

Tabelle 5 zeigt eine Liste der Testparameter für die drei evaluierten Verfahren. Sie bilden die Verfahrensparameter, mit denen das Wasserzeichen in die Videos eingebettet wird. Dadurch entsteht ein Vielfalt von Möglichkeiten, die die gesamte Spanne der einzelnen Verfahrensparameter abdeckt.

Verfahren	Parameter				
alle	R1	25%	50%	75%	100%
[DiHa2001]	R2	5	10	15	
	R3	3	5	9	
	Range	5	10		
	WS	1	1,5	2	4
[DiSt1998]	WS	1	1,5	2	
[SeLa2001]	min. Cutoff-Point	3	10	20	
	Differenz	5000	10000	20000	

Tabelle 5: Testparameter [Th2002]

5.1.2 Verfahrensparameter

Der Parameter *R1* drückt die prozentuale Anzahl der zu markierenden DCT-Blöcke im Verhältnis zur gesamten Blockanzahl aus. Es wird in Schritten von je 25% erhöht. Mit Hilfe des Parameters kann untersucht werden, wie sich das Verhältnis zwischen der Anzahl der zu markierenden Blöcke im Frame und die visuelle Qualität auswirkt. Der Parameter *R2* ist nur für [DiHa2001a] wichtig und gibt an, wieviele Frequenzwerte im DCT-Block markiert werden. Der dabei bevorzugte Frequenzbereich ist der mittlere Frequenzbereich. *R2* hat eine maximale Größe von 15 Werten, da längere Wasserzeichensequenzen hochfrequente Werte markieren, die aber aufgrund der stärkeren MPEG-Kompression stärker verfälscht werden. Dadurch kann das Ausleseergebnis verfälscht werden. Der Parameter *R3* bestimmt die Anzahl der GoPs pro Wasserzeichenbit. Aufgrund der Kapazität des Videos wurde eine maximale Größe von neun GoPs pro Wasserzeichenbit festgelegt. Die Range gibt die maximal mögliche Änderung eines DCT-Frequenzwertes an und wird auf einen maximalen Wert von zehn festgelegt, um eine ausreichend gute Transparenz des Wasserzeichens zu gewährleisten. Der Frequenzwert kann maximal um zehn erhöht bzw. verringert werden. Die Wasserzeichenstärke wird maximal auf vier festgelegt, da davon ausgegangen werden kann, dass bei höheren Wasserzeichenstärken sich die Transparenz des Wasserzeichens verschlechtern wird. In der Veröffentlichung wird die obere Begrenzung der Wasserzeichenstärke auf vier festgelegt, da jeder auf die mittleren DCT-Frequenzwerte zu addierende Wert vervierfacht wird und somit die Transparenz sich deutlich verschlechtert.

Im Verfahren [DiSt1998] liegt der Wertebereich der Wasserzeichenstärke aufgrund der Transparenz zwischen eins und vier.

Die Parameter von [SeLa2001] setzen sich aus dem minimalen Cutoff-Punkt und der Differenz zusammen. Laut [SeLa2001] verzeichnet ein minimaler Cutoff-Punkt größer als 20 Einbußen in der Robustheit. Die Differenz gibt das Verhältnis zwischen beiden Gruppen an. Die Differenz zwischen 5000 und 20000 ist der optimale Bereich für diesen Parameter. Bei höheren Werten wird der minimale Cutoff-Punkt häufig unterschritten und der Block kann nicht genutzt werden. Dadurch kann das Wasserzeichen seltener eingebettet werden.

5.1.3 Durchführung der Evaluation

Zuerst werden die markierten Videos erstellt und die Effektivität des Einbettungsalgorithmus' gemessen. Die Quote gibt an, aus wievielen Videos nach Einbetten des Wasserzeichens es auch wieder erfolgreich ausgelesen werden kann.

Die nächste Stufe beinhaltet eine Untersuchung der Eigenschaften Robustheit und Sicherheit. Dabei werden die markierten Videos verwendet, bei denen das Wasserzeichen vorher erfolgreich detektiert wurde. Dies Videos werden verschiedenen Angriffen unterzogen, die einerseits die Robustheit und andererseits die Sicherheit des Wasserzeichenverfahrens betreffen.

Die Robustheitsangriffe umfassen eine MPEG-Re-Enkodierung mit verschiedenen Bitraten. Die Bitraten liegen innerhalb des Bereiches 0,6 bis 1,5 Mbit/s.

Parallel dazu werden die gleichen Testvideos mit einer konstanten GoP-Struktur enkodiert und markiert. Es wird untersucht, ob ein unterschiedlicher Aufbau der GoP-Struktur Auswirkungen auf das Wasserzeichenverfahren hat. Ziel der Re-Enkodierung ist es, die Robustheit der Verfahren gegen die MPEG-Kompression zu untersuchen.

Weitere Robustheitsangriffe sind eine Skalierung des Videos und das Auftragen eines additiven Rauschens. Die Skalierung ist sehr oft Bestandteil von Veränderungen der Framegrößen, das Auftragen eines Rauschmusters hingegen bei in einer Vielzahl von Filteroperationen am Video.

Für eine Beurteilung der Qualität der Sicherheit werden die in [DiHa2001a] vorgestellten Konzepte von digitalen Fingerabdrücken getestet. Die eingebettete Information basiert z. B. mit dem Boneh-Shaw-Verfahren [DiHa2001a] auf folgenden Parametern:

(a) Anzahl der Wiederholungen: 2
(b) Anzahl der Kunden: 3
(c) Kundennummer: 2

Die Wasserzeicheninformation ist der Bitvektor „1001“. Je nach Anzahl vorhandener GoPs wird der Vektor mehrfach in das Video eingebettet. Es wurde eine kurze Wasserzeicheninformation benutzt, um die Information möglichst sehr oft in das Video einbetten zu können. Dadurch kann [DiHa2001a] umfangreich getestet werden, da der Vektor mit bis zu neunfacher Redundanz in ein Video eingebettet werden kann.

Ziel der Testumgebung ist es, zu untersuchen, in wieweit die drei getesteten Verfahren für spätere Forschungen geeignet sind. Mit Hilfe der Evaluation sollen ebenfalls mögliche Schwachstellen der untersuchten Wasserzeichenalgorithmen erkannt werden.

5.1.4 Testresultate und daraus gewonnene Erkenntnisse

Die ersten Testergebnisse befinden sich in Tabelle 13 im Anhang A. Sie beinhalten die False-Bit-Rate (FBR) (5) und die Ausleseraten (A-Rate) (6) des Wasserzeichens der ausgelesenen Fingerabdrücke. Die A-Rate gibt an, wieviel Prozent der eingebetteten Wasserzeicheninformation wieder erfolgreich detektiert werden konnte. Sind die eingebetteten Informationen hingegen Fingerabdrücke, gibt die FBR an, wieviel Prozent der eingebetteten Fingerabdrücke falsch ausgelesen wurden. Dabei wird nicht der gesamte Fingerabdruck betrachtet. Es werden sämtliche Bits der Wasserzeichennachricht betrachtet.

$$FBR = \frac{\sum \overline{W_i}}{\sum W}, \; wenn \quad \overline{W_i} \neq W_i \qquad (5)$$

$$A - Rate = \frac{WZBit_k}{N} \qquad (6)$$

Zuerst werden die Verfahren [DiHa2001a] und [DiSt1998] anhand der Wasserzeichenstärke (WS) verglichen. Für beide Verfahren liegen zwar verschiedene Einbettungskanäle vor, trotzdem können beide Verfahren aufgrund des gleichen Parameters miteinander verglichen werden. [DiHa2001a] zeigt bei einer niedrigeren Wasserzeichenstärke die größten Auslesefehler. Es hat sich gezeigt, dass besonders bei den Markierungspositionen, an denen das Bit 1 als Information eingebettet wird, eine hohe Fehlerquote ausgelesen wird. Durch eine Re-Enkodierung sind die Veränderungen des Videos zu signifikant, sodass die mit [DiHa2001a] eingebetteten Wasserzeicheninformationen nicht mehr korrekt ausgelesen werden können.

Zusätzlich zu der Wasserzeichenstärke werden noch die in [DiHa2001a] aufgeführten verfahrensspezifischen Parameter untersucht. Ziel der Untersuchung ist es, herauszufinden, ob einer der Parameter für die hohen Fehlerraten verantwortlich ist. Wie die Testergebnisse zeigen, hat das Verfahren bei jeder Kombination der Parameter R1, R2, R3 und Range generell eine schlechte Ausleserate. Es kann davon ausgegangen werden, dass keiner der Parameter ausschlaggebend für die schlechten Ergebnisse im Ausleseprozess ist.

Dagegen zeigen die Testergebnisse von Verfahren [DiSt1998], dass das Einbetten eines Musters auf die Luminanzdaten gute Ausleseergebnisse liefert. Das Verfahren [SeLa2001] weist gute Ergebnisse auf und ist durchaus zum Markieren von Videos geeignet. Es zeigt im gesamten ausgewählten Wertebereich der Parameter positive Ergebnisse.

Abbildung 7 zeigt die Ausleseraten nach einer MPEG-Reenkodierung. Damit wird die Robustheit der drei Verfahren gegenüber der MPEG-Kompression untersucht. Das Testmaterial wird mit vier verschiedenen Bitraten (0.6, 0.8, 1.0, 1.5 Mbit/s) re-enkodiert. Dabei werden nur die Videos re-enkodiert, bei denen vorher die Wasserzeicheninformationen erfolgreich ausgelesen werden konnten. Die Testvideos weisen eine variable GoP-Struktur auf. Der Testvorgang wird zusätzlich auch mit einer konstanten GoP-Struktur durchgeführt, d. h. vom gesamten Testmaterial wird eine Kopie erstellt, wobei die GoP-Struktur im gesamten Video konstant gehalten wird. Es ist davon auszugehen, dass bei einer variablen GoP-Struktur die Wasserzeicheninformationen verloren gehen, da nur I-Frames markiert

werden. Durch die Re-Enkodierung kann sich der Typ des Frames ändern und durch die veränderte Kodierung können die Informationen verloren gehen. Um trotzdem die Robustheit der Verfahren zu testen, wurden die Testvideos mit konstanten GoP-Struktur kodiert. Dadurch werden die Frametypen nicht verändert und die eingebetteten Informationen bleiben erhalten.

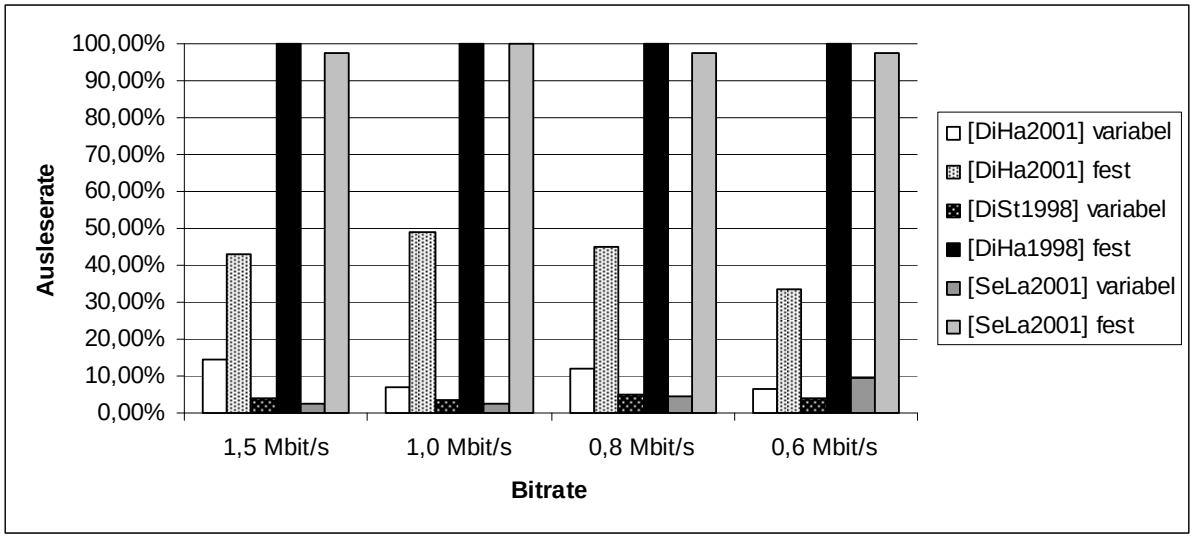

Abbildung 7: Auslesen der Wasserzeicheninformation nach MPEG-Reenkodierung [Th2002]

In Abbildung 7 werden die Ergebnisse der Robustheit der Verfahren gegenüber MPEG-Kompression dargestellt. Dabei wurde jedes Verfahren mit einer variablen und festen GoP-Struktur untersucht. Die Ausleseergebnisse sind bei einer festen GoP-Struktur deutlich besser, als bei einer variablen GoP-Struktur. Ein Balken im Diagramm stellt dabei den Durchschnitt der 20 markierten Testvideos dar. Eine ausführlichere Auflistung der Ausleseergebnisse kann aus Tabelle 13 im Anhang A dieser Arbeit entnommen werden.

Die unterschiedlichen Ausleseergebnisse bei Vergleich der Verfahren, sind auf Verfahrensnachteile zurückzuführen. [DiHa2001a] zeigt nur durchschnittliche Ausleseresultate, die im Durchschnitt bei 50% liegen. Dagegen liegen bei den anderen Verfahren die Ergebnisse mit konstanter GoP-Struktur zwischen 90% und 100%.

Weitere getestete Robustheitsangriffe sind die Skalierung des Videos und das Auftragen eines gleichmäßigen Rauschmusters auf das Video. Tabelle 14 in Anhang A erläutert die Ergebnisse der Robustheit der Verfahren nach dem Auftragen eines additiven Wertes. Es ist zu erkennen, dass im Gegensatz zu anderen Verfahren lediglich [DiHa2001a] nicht robust gegen den „Add-Value“ Angriff ist. Bei dem Skalierungsangriff (Tabelle 15 in Anhang A) wird zuerst das markierte Video auf eine festgelegte Größe skaliert. Im zweiten Fall wird das markierte Video wieder auf Originalgröße re-skaliert. Die Ergebnisse verdeutlichen, dass die Verfahren, außer Verfahren [DiHa2001a], bessere Ergebnisse liefern, wenn die skalierten Videos wieder in die ursprünglichen Framegrößen re-skaliert werden.

Die in [DiHa2001a] vorgestellten Fingerabdrücke wurden hinsichtlich der Resistenz gegen Koalitionsangriffe untersucht. Dabei werden je zwei mit verschiedenen Fingerabdrücken markierte Videos verglichen. Der Vergleich soll die Unterschiede zwischen den

Fingerabdrücken finden und verfälschen. Die Videos wurden basierend auf [DiHa2001a] und [DiSt1998] markiert, um untersuchen zu können, ob allein der Fingerabdruck der entscheidende Aspekt für die Resistenz gegenüber Koalitionsangriffen ist. Es wurden beide Ansätze für digitale Fingerabdrücke aus [DiHa2001a] getestet. Die Testergebnisse befinden sich in Tabelle 16 in Anhang A. Zwei Videos (Werbevideo und Zeichentrickvideo) dienen als Ausgangsbasis. Die unterschiedlichen Eigenschaften der Videos sollen Aufschlüsse darüber geben, ob Unterschiede beim Videomaterial zu unterschiedlichen Ergebnissen bei den Fingerabdrücken führt.

Dass nur [DiBe1999] eine Liste der Angreifer bzw. Piraten ermittelt, liegt im Aufbau des Fingerprintalgorithmus'. [BoSh1995] liefert mit hoher Wahrscheinlichkeit immer einen der Angreifer. Als negativer Aspekt liefert [BoSh1995] aber teilweise einen falschen Kunden. Hier besteht unbedingter Verbesserungsbedarf, da bei diesen Ergebnissen unschuldige Kunden beschuldigt werden könnten.

5.1.5 Vergleich mit anderen Wasserzeichenverfahren

Eine Vielzahl weiterer Veröffentlichungen markiert ebenfalls I-Frames. Aufgrund der hohen Ähnlichkeit von MPEG I-Frames mit JPEG-Bildern kann auf die Entwicklung von Bildwasserzeichen zurückgegriffen werden. Der Einbettungsmechanismus basiert sehr oft auf der „Spread Spektrum" Technologie [Ha2000]. Dabei wird das Wasserzeichen über das gesamte Video verteilt. Durch Hinzunahme des geheimen Schlüssel werden die Positionen pseudo-zufällig ausgewählt.

Zuletzt hat man sich aber von der Strategie des Einbettens der Wasserzeicheninformationen an fest vorausgewählten Positionen entfernt. Besonders bei üblichen Modifikationen am Frame, wie Skalierung und Ausschnitt, besteht die Möglichkeit, dass die Einbettungspositionen nicht mehr korrekt identifiziert werden. Bevor das Wasserzeichen in das Frame eingebettet wird, werden robuste Merkmale des Frames identifiziert [BaCh2001], [ThVo2004]. So wird das Wasserzeichen an Positionen im Frame eingebettet, die nach einer Vorfilterung im gefilterten Frame vorkommen. Beispielsmerkmale sind Ecken oder Kantenverläufe im Frame. Der Vorteil liegt darin, dass sie nach Modifikationen am Video weiter erkennbar sind. Verfahren, die robust gegenüber Rotation, Skalierung und Transformation sind, werden auch als RST Verfahren bezeichnet.

Außerdem werden aufgrund der Anfälligkeit der DCT-Transformation gegenüber Videomodifikationen [LiHu2006] die neuesten Videowasserzeichen nicht mehr in den DCT-Frequenzraum eingebettet. Einzelne Frequenzen sind derartanfällig, dass eine sichere Detektion des Wasserzeichens erschwert ist. Als Alternative werden die Framedaten in den DWT-Frequenzraum transformiert [FuSh2007]. Einzelne spezielle Wavelettransformationen seien robust gegenüber dem „Shiften" des Frames um einzelne Pixel. Das Verschieben des Frameinhaltes um einige Pixel horizontal oder vertikal ist ein sehr effizienter Angriff.

Auch hat sich gezeigt, dass mit Hilfe des Einsatzes von Fehlerkorrekturkodes eine gewisse Anzahl von Auslesefehlern des Wasserzeichens wieder korrigiert werden [ChXu2006]. Im Allgemeinen kann die detektierte Wasserzeichennachricht als richtig oder falsch eingestuft werden. Zur Anwendung kommen dabei Fehlerkorrekturkodes. Durch den Einsatz von

Fehlerkorrekturcodes kann zwar das Ausleseergebnis leicht verbessert, jedoch die Robustheit des Wasserzeichens nicht signifikant verbessert werden.

Ein Aspekt der Ergebnisanalyse zeigt, dass das Einbetten des Wasserzeichens in jedes Frame vom Aufbau des Videoinhaltes abhängt. So werden entweder die gleichen Wasserzeicheninformationen in eine gesamte Szene eingebettet oder einzelne Videoobjekte verändert. Verschiedene Wasserzeicheninformationen in benachbarten Frames rufen einerseits visuelle Artefakte vor, andererseits können sie durch den Vergleich der Frames identifiziert und gelöscht werden [SuKu2002]. Wenn ein Frame markiert wird, werden die nachfolgenden Frames auf mögliche visuelle Artefakte untersucht und kompensiert [Ha2000]. Dadurch steigt die Komplexität der Verfahren signifikant. Wenn einzelne Verfahren als echtzeitfähig eingestuft werden, dann ist es eher eine sehr einfache Markierung. Ziel ist es, das Wasserzeichen in so wenige Frames wie möglich einzubetten. Bei komplexen Verfahren ist die Echtzeit meistens keine Anforderung.

5.1.6 Schlüsse und Herausforderungen

Es wurden die Verfahren [SeLa2001], [DiHa2001a] und [DiSt1998] untersucht. Dabei arbeiten [SeLa2001] und [DiHa2001a] im Frequenzraum und [DiSt1998] im Bildbereich. Sie wurden aufgrund der positiven Testergebnisse in den jeweiligen Publikationen verwendet. Zudem wurden sie auch an JPEG-Bildern erfolgreich getestet, die eine hohe Ähnlichkeit zum MPEG-Standard aufweisen.

Außer [DiHa2001a] liefern die Verfahren durchaus positive Testergebnisse und eignen sich für das Markieren von Videomaterial aus dem Archivbereich. Dabei können mit [SeLa2001] und [DiSt1998] Urhebermerkmale in digitale Videos eingebettet werden. [DiHa2001a] bettet kundenspezifische Merkmale in MPEG-Videos ein.

Die beiden Verfahren [SeLa2001] und [DiSt1998] können zum Nachweis der Urheberschaft genutzt werden. Die Anforderungen wurden mit Robustheitstests untersucht. Die Konzentration lag dabei auf der MPEG-Kompression, dem Auftragen eines zusätzlichen Wertes und der Skalierung des Videos. In Bezug auf MPEG-Kompression bieten die getesteten Wasserzeichen eine gute Robustheit. Auch beim Auftragen von Rauschmustern auf das Video sind die Wasserzeichen resistent. Dadurch besteht eine gute Resistenz gegen Filteroperationen, wie sie z. B. bei der Aufarbeitung von altem Filmmaterial in Videoarchiven angewendet werden. Einzig bei der MPEG-Kompression bei Videos mit einer variablen GoP-Struktur wurden Mängel entdeckt. Bei diesem Robustheitstest können die Wasserzeicheninformationen mit hoher Wahrscheinlichkeit nicht mehr korrekt ausgelesen werden.

Die beiden untersuchten Fingerprintalgorithmen [DiHa2001a] liefern gute Ergebnisse, nur besteht augenblicklich der Nachteil, dass sehr viel Kapazität zum Erstellen der Fingerabdrücke notwendig ist.

5.2 Schwachstellen und Lösungsansätze

Anhand der Testergebnisse konnten Problemfelder identifiziert werden. Im diesem Kapitel werden einige dieser Problemfelder erläutert und die in dieser Arbeit entwickelten Ansätze präsentiert. Die Arbeit deckt nicht alle möglichen Problemfelder ab, als Basis dient einerseits die Evaluation im vorherigen Kapitel und andererseits die Anwendungsszenarien in Kapitel drei. Es ist durchaus möglich, dass Videowasserzeichen weitere Lücken haben, aber um die Arbeit nicht zu sprengen, liegt der Fokus auf den folgenden Lücken:

5.2.1 Veränderung der Frametypen durch MPEG-Re-Enkodierung führt zur mangelnder Robustheit des Wasserzeichen

Dieses Kapitel beschreibt das Problem, dass durch Veränderung der Frametypen die Ausleserate des Wasserzeichens fällt.

5.2.1.1 Problembeschreibung

Ein MPEG-Video wird mittels Zuteilung von Frametypen sowohl räumlich als auch zeitlich komprimiert. Speziell dafür werden Referenz- und Differenzframes erstellt und im Video zeitlich angeordnet. Intrakodierte Frames dienen dabei immer als Referenzframes. Aufgrund ihrer Eigenschaft als Referenzframes werden sie nicht so stark komprimiert wie interkodierte Frames.

Da aber laut Kapitel 4 viele Wasserzeichen nur intrakodierte Frames markieren, besteht der allgemeine Nachteil, dass mit einer Re-Enkodierung des Videos davon ausgegangen werden kann, dass sich die Frametypen im Video ändern. Nach fast jede Bearbeitung eines MPEG-Videos folgt eine abschließende Enkodierung. Wenn bei der Bearbeitung einzelne Frames oder Framesequenzen hinzugefügt oder entfernt werden oder während des abschließenden Enkodiervorganges ein anderes Enkodierschema als vor der Bearbeitung des Videos benutzt wird, kann der Fall auftreten, dass sich der Frametyp entsprechender Frames ändert. So besteht immer die Möglichkeit, dass die Detektion des Wasserzeichens aus nicht markierten Frames erfolgt. Abbildung 8 illustriert dieses Problem. Bei Frame sieben und zehn werden durch eine erneute Enkodierung die Frametypen geändert. Die Detektion eines I-Frame-basierten Wasserzeichens würde aus dem siebten und nicht wie ursprünglich aus dem zehnten Frame das Wasserzeichen auslesen.

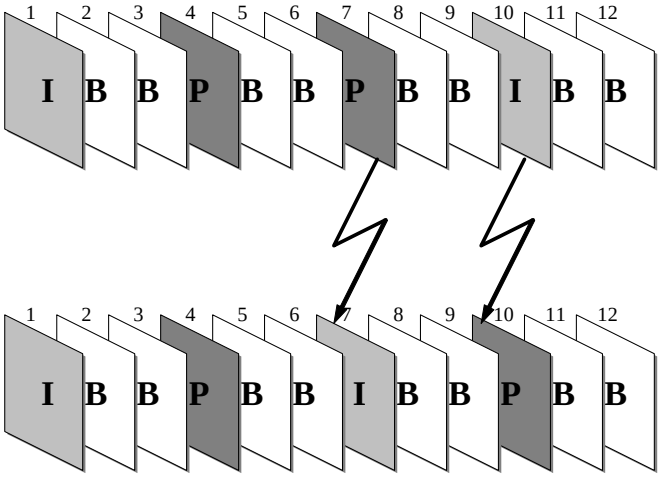

Abbildung 8: Veränderung des Frametypes durch Re-Enkodierung

5.2.1.2 Designansatz

Wenn nur I-Frames markiert werden, benötigt das Wasserzeichen den zusätzlichen Mechanismus der zeitlichen Synchronisierung [HaKu2005]. Ziel dieser Methode ist es, die ursprünglich markierten Frames wieder aufzuspüren und anschließend die eingebetteten Wasserzeicheninformationen wieder auszulesen. Dabei wird das Video in die Einzelframes zerlegt und abschließend resynchronisiert. Nach der Analyse der Synchronisationsinformation werden die ursprünglich markierten Frames erkannt und das Wasserzeichen aus diesen Frames ausgelesen. Da diese Methode aber sehr zeitaufwendig sein kann, muss der angewandte Algorithmus eine geringe Komplexität aufweisen.

Der Ansatz dieser Arbeit beschäftigt sich mit der zeitlichen Synchronisation. In Kapitel sechs wird das Konzept der zeitlichen Synchronisation vorgestellt. Die dafür verwendete Technologie ist der robuste Hash [HaBo2007], der zur Identifikation der Frames benutzt wird, bevor aus ihnen das Wasserzeichen ausgelesen wird.

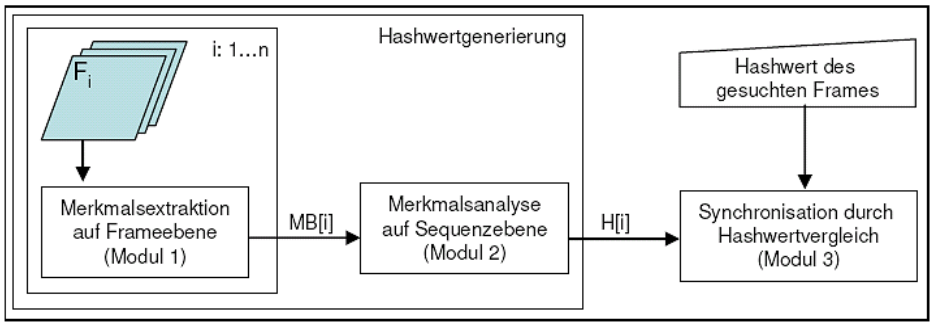

Abbildung 9: Konzept des robusten Videohashs

Abbildung 9 stellt den in dieser Arbeit verfolgten Ansatz dar. Das Hashverfahren teilt sich in drei Module auf. Ziel ist es, vor dem Auslesen der Wasserzeicheninformation die ursprünglich markierten Frames wiederaufzufinden. Da die generierten Hashwerte robust sind, werden die Hashbits basierend auf extrahierten Merkmalen der Frames gebildet, d.h. aus den zu markierenden Frames werden Merkmalsinformationen extrahiert und an das Modul 2 weitergegeben. Im 2. Modul werden die Merkmale analysiert und die Hashwerte generiert.

Besonderheit dieses Verfahrens ist die Generierung der Hashwerte aus Sequenzebene. Das erweist sich als notwendig, da die extrahierten Merkmalsinformationen benachbarter Frames sehr oft identisch sind und so eine framegenaue Identifikation schwer durchführbar ist. Die Unterschiede in den Hashwerten entstehen erst bei Betrachtung einzelner Sequenzen, die das markierte Frame beinhalten. Das 3. Modul wird vor dem Auslesen der Wasserzeicheninformation angewandt. Durch einen Vergleich der generierten Hashwerte aus dem zu untersuchenden Video mit den während der Einbettung des Wasserzeichens gebildeten Hashwerte ist ein Wiederauffinden der ursprünglich markierten Frames möglich. Dieses Verfahren wird im Kapitel 6.3 beschrieben und anschließend evaluiert, um die Anwendbarkeit zu untersuchen und auf mögliche Schwachstellen bzw. offene Punkte zu stoßen.

5.2.2 Kapazitätsmängel

Dieses Kapitel beschreibt das Problem, dass durch die Markierung vorausgewählter einzelner Frames nicht die maximale Kapazität des Videos genutzt wird.

5.2.2.1 Problembeschreibung

Die Enkodierung von MPEG-Videos basiert auf GoPs. Eine GoP ist die Anordnung von I-, P- und B-Frames. Das Video wird in GoPs aufgeteilt und jedes Frame innerhalb der GoP wird anhand eines vorher festgelegten Schemas enkodiert. Trotz einer freien Auswahl besteht eine GoP sehr oft aus zwölf bis 15 Frames. Diese Anordnung bietet eine gute visuelle Qualität. Abbildung 10 zeigt das Beispiel einer Referenzstruktur mit einer Größe von zwölf Frames pro GoP.

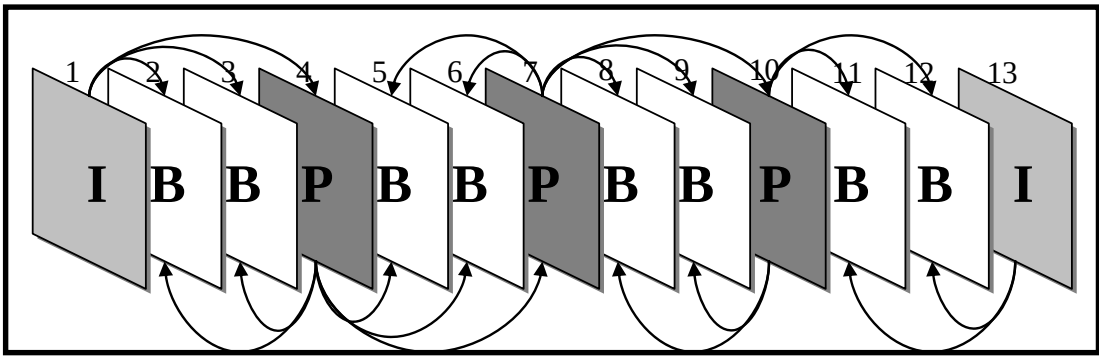

Abbildung 10: Beispiel Standard Referenzstruktur einer Group-of-Picture (GoP)

Da nur das erste Frame ein I-Frame ist, können von einem Video, basierend auf einer GoP-Struktur wie in Abbildung 10, mit einem I-Frame basierten Wasserzeichen nur ungefähr 8% der gesamten Framemenge markiert werden. Wasserzeichen dieser Art können auch als intrakodierte Wasserzeichen bezeichnet werden, da sie ausschließlich intrakodierte DCT-Blöcke markieren. Ein I-Frame besteht nur aus intrakodierten DCT-Blöcken. So befinden sich zum Beispiel bei einer Framerate von 25 Frames pro Sekunde, einer Spieldauer von 60 Sekunden und einer GoP Größe von zwölf Frames insgesamt nur 125 I-Frames im Video. Wenn das Wasserzeichen beispielsweise 13 Bits pro Frame einbettet, können in das gesamte Video nur eine Gesamtmenge von 1625 Bits eingebettet werden.

Da bei digitalen Fingerabdrücken eine hohe Kapazität notwendig ist, sind Kapazitäten dieser Größe gewünscht. Ausschlaggebende Parameter sind dabei die Anzahl der möglichen Kunden, für die eine individuelle Kopie erstellt wird und das Sicherheitslevel, das den Grad der Sicherheit des Fingerabdruckes gegen gezielten Angriffen widerspiegelt. So steigt z. B. beim Schwenk-Überberg-Verfahren [DiBe1999] die Gesamtlänge des Fingerabdruckes exponentiell gegenüber dem Sicherheitslevel an. Auch wenn kein digitaler Fingerabdruck zum Einsatz kommt, sollte versucht werden, die Wasserzeicheninformation redundant einzubetten, d.h. mehrfach verteilt über das gesamte Videomaterial. Auch besteht die Möglichkeit, dass nicht jeder Videoabschnitt aufgrund spezieller Inhaltseigenschaften markiert werden kann, ohne zu deutliche visuelle Einbußen in Kauf zu nehmen. Deshalb sollte versucht werden, jede Möglichkeit zum Einbetten des Wasserzeichens zu nutzen, z. B. auch P- und B-Frames.

5.2.2.2 Designansatz

Abbildung 11 demonstriert den Ansatz dieser Arbeit. Entgegen der in Kapitel 4 vorgestellten Verfahren werden in dem in dieser Arbeit präsentierten Verfahren nicht nur I-Frames markiert, sondern es wird eine Verbesserung der Kapazität dahingegen erreicht, in dem auch P- und B-Frames markiert werden. Dabei müssen aber die Besonderheiten der basierend auf dem MPEG-Standard kodierten Video betrachtet werden. Die in P-und B-Frames zur Verfügung stehende Datenmenge ist häufig um ein Vielfaches geringer als in I-Frames. Dadurch müssen am Wasserzeichen Anpassungen getroffen werden, damit sie robust entegegen der in P- und B-Frames stärkeren MPEG-Kompression sind.

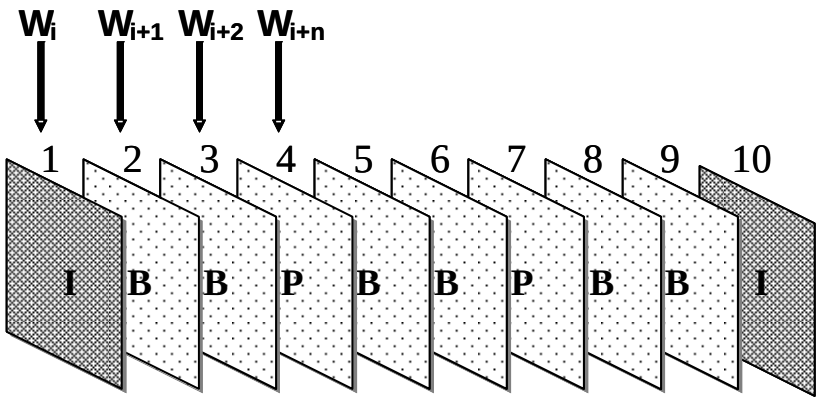

Abbildung 11: Erhöhung der Kapazität durch erweiterte Markierung

In Kapitel sieben wird das Wasserzeichen, das den Aspekt der erweiterten Markierung mitbetrachtet, genauer vorgestellt. Dabei werden auch die Ausleseergebnisse nach Einbettung des Wasserzeichens in P-Frames analysiert. Die Ergebnisse sollen Aufschluß darüber geben, in wieweit das Verfahren auch für B-Frames anwendbar ist.

5.2.3 Referenzen innerhalb des Videos verschlechtert die Transparenz des Wasserzeichens

Diese Kapitel beschreibt den Zusammenhang zwischen den zeitlichen Referenzen im Video und die daraus resultierenden Nachteile für die Transparenz des Wasserzeichens.

5.2.3.1 Problembeschreibung

Videostandards analysieren auch die Bewegung zwischen den Frames. Ziel ist, Informationen, die in Nachbarbildern identisch sind, nicht mehrfach abzuspeichern. MPEG kodiert auch die Bewegung von Frameinhalt zwischen den Frames. Da der zeitliche Abstand zwischen zwei Nachbarframes sehr gering ist, kann auch davon ausgegangen werden, dass der Inhalt auch fast identisch ist.

Diese Eigenschaft hat auch Auswirkungen beim Einbetten von Wasserzeichen. Wenn in ein Frame ein Wasserzeichen eingebettet wird, muss damit gerechnet werden, dass es aufgrund der Referenzen innerhalb der GoP auch in Nachbarbilder übertragen wird. Digitale Wasserzeichen für intrakodierte Frames haben dabei die größte Auswirkung. Da es Bestandteil des Referenzframes ist, wird es in alle Frames der GoP übertragen.

Aufgrund der vorliegenden minimalen Veränderungen zwischen Nachbarbildern muss die Transparenz des Wasserzeichens angepasst werden, damit es nicht zu visuellen Artefakten kommt. Visuelle Artefakte können z. B. auftreten, wenn sich die Wasserzeicheninformation zwischen zwei Nachbarbildern deutlich unterscheidet, sowohl bei den Wasserzeichenbits als auch an den Positionen, wo die Wasserzeichenbits eingebettet werden sollen.

5.2.3.2 Designansatz

Wie in Abbildung 12 dargestellt, wird in dieser Arbeit die Strategie verfolgt, daß innerhalb einer GoP die gleichen Wasserzeichenbits an den gleichen Markierungspositionen eingebettet werden. Dadurch wird das Wasserzeichen nicht pro Frame sondern pro GoP eingebettet. Zusätzlich werden vor dem Einbetten von Wasserzeichen in interkodierte Frames die übertragenen Wasserzeichensignale aus den Referenzframes kompensiert. In [Ha2000] wurde das Prinzip des Drift Compensation Signals entwickelt. Ziel ist es, zuerst die übertragenen Wasserzeichensignale zu entfernen, bevor das für das augenblickliche Frame berechnete Wasserzeichensignal eingebettet wird.

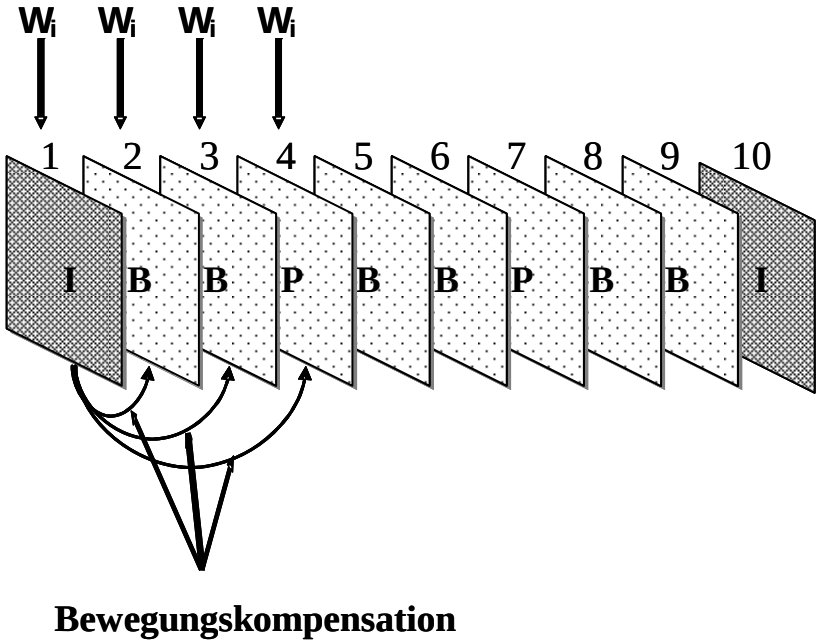

Abbildung 12: Verbesserung der Transparenz

Die aufgeführte Aspekte fließen zusammen mit dem Erkenntnissen aus Kapitel 5.2.2 in die Entwicklung des robusten Wasserzeichens dieser Arbeit mit ein und sind Bestandteil des siebten Kapitels dieser Arbeit.

5.2.4 Mangelhafte Echtzeitfähigkeit durch zu hohe Komplexität des Wasserzeichens

Diese Kapitel beschreibt die zu betrachtende Notwendigkeit der Verbesserung der Komplexität des Wasserzeichens zur Verbesserung der Echtzeitfähigkeit des Algorithmus'.

5.2.4.1 Problembeschreibung

Bei diversen Anwendungsfällen, wie z. B. den Urheberschutz, ist die Komplexität sehr wichtig. Es ist davon auszugehen, dass zur gleichen Zeit mehrere Kopien des Videos mit je einer dem Empfänger spezifizierten Wasserzeichennachricht erstellt werden. Einsatzgebiete sind zum Beispiel Videoportale, die Videos zum Herunterladen anbieten.

Aufgrund der Komplexität des MPEG-Videostandards erfordert die Enkodierung eines MPEG-Videos eine ausreichende Kapazität an Ressourcen. Während die Dekodierung nur aus einer Interpretation des MPEG Bitstroms und der Speicherung der Videobilder besteht, weist der Enkodierungsvorgang auf eine höhere Komplexität hin. Der dabei anspruchvollste Teilprozess ist die Bewegungsanalyse. Während der Bewegungsanalyse wird der optimalste Bewegungsvektor innerhalb eines festgelegten Suchraums bestimmt. Je nach Größe des Suchraumes verzögert die Bewegungsanalyse die vollständige Enkodierung.

Während der Einbettung eines Wasserzeichens muss das Video komplett re-enkodiert und, je nach Art der Einbettungsmethode, bis in seine Einzelframes zerlegt werden.

5.2.4.2 Designansatz

Ziel der Wasserzeichenlösungen muss also sein, den Aufwand für den Kodiervorgang so gering wie möglich zu halten. Die Integration des Wasserzeichens ist nur ein kleiner Anteil der gesamten Videobearbeitung. Deshalb verfolgen die Ansätze dieser Arbeit die Strategie, den Videokodiervorgang zu optimieren [Di2007], [StHa2007], [HaSt2008].

Dabei werden für den Einbettungsmodus die folgenden Ansätze verfolgt:

(a) Vorbereitung der zu markierenden Videodatei
(b) Verteilung des Markierungsaufwandes auf verteilte Ressourcen

Da bei verschiedenen individualisierten Einbettungsvorgängen meistens nur die Wasserzeichennachricht unterschiedlich ist, bestünde ein effizienter Ansatz darin, die zu markierende Videodatei so vorzubereiten, dass der jeweilige De- und Enkodiervorgang während des Einbettens vom Wasserzeichens nicht mehr notwendig ist. Bei identischem Schlüssel ist nur die Wasserzeicheninformation unterschiedlich. So werden an entsprechenden Markierungspositionen je eine Änderung für die Wasserzeichenbits null oder eins durchgeführt. Wenn diese Änderungen zusammen mit den Originaldaten abgespeichert werden könnten, wäre die Erstellung des markierten Videos effizient durchführbar, da nur die notwendigen Daten zusammengeführt werden müssen. Dies kann mit einem Kopieren von ausgewählten Datenbereichen in das markierte Video verglichen werden.

Abbildung 13 demonstriert den Ansatz. Ziel ist es, den Einbettungsvorgang so aufzuteilen, dass die rechenintensiven Schritte einmal durchgeführt werden. Dafür wird im Erstellungsvorgang (Preparation Stage) das allgemein markierte Video erstellt, aus dem im Finalisierungsschritt (Finalization Stage) jedes mögliche markierte Video erstellt werden kann. In der Abbildung wird das allgemein markierte Video Container genannt, da es sämtliche Möglichkeiten von markierten Frames beinhaltet. Während der Erstellung des Containers werden die rechenintensiven Schritte, wie MPEG De- und Enkodierung durchgeführt. Da dieser Schritt nur einmal angewandt wird, ist nur einmal der Nachteil einer mangelhaften Echtzeit gegeben. Der eigentliche Schritt, wo das markierte Video erstellt, ist der Finalisierungsschitt. Während dieses Vorganges müssen nur die markierten Teile aus dem Container extrahiert werden, die aufgrund der vorgegebenen Wasserzeicheninformation notwendig sind. Da im Container die Bestandteile schon komplett MPEG-kodiert vorliegen, müssen sie jetzt nur noch zusammenkopiert werden. Dadurch wird eine sehr gute Echtzeit erreicht.

In Abbildung 13 wird beispielhaft ein 1-Bit-Wasserzeichen pro Frame in das Video eingebettet. Im Kontainer werden dann jeweils ein markiertes Frame mit demWasserzeichenbit null und eins abgespeichert. Das markierte Video ist das ein Zusammentragen der Frames, wobei hier als Beispielinformation „10011“ angegeben ist.

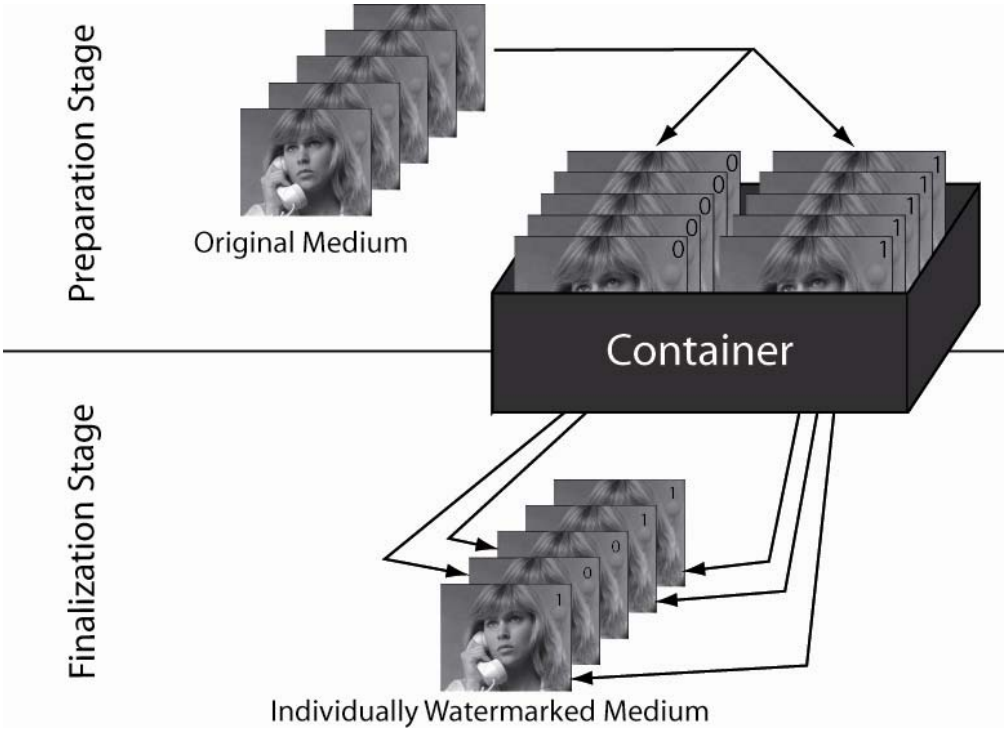

Abbildung 13: Effizientes Einbetten durch Aufteilung des Einbettungsmodus

Eine weitere Möglichkeit ist die Verteilung der Wasserzeichenintegration auf verschiedene Computerresourcen. Dies kann sich insbesondere in solchen Situationen als sinnvoll erweisen, in denen eine Vielzahl von Videos mit unterschiedlichen Wasserzeichen oder eine Vielzahl von Kopien erstellt werden sollen. Eine parallele Einbettung würde auf einer Computerresource einen hohen und ineffizienten Rechenaufwand hervorufen. Die Lösung besteht in einer Verteilung der Einbettungsversuche auf mehrere Computer.

Abbildung 14 demonstriert den zweiten Ansatz. Auf dem Computer A wird in einem ersten Prozess das Video in mehrere Teilsequenzen zerlegt, in der Abbildung in beispielhafte drei Teile. Jedes dieser Teile wird separat an eine Ressource gesendet und dort markiert. Nach dem Einbetten der Wasserzeichen in die Teilsequenzen werden die markierten Sequenzen zu Computer A zurückgesendet wo sie in einem dritten abschließenden Prozess wieder zu einem markierten Video zusammengefügt werden. Durch die parallele Einbettung des Wasserzeichens auf verschiedenen zur Verfügung stehenden Ressourcen sollte es möglich sein, dass Video effizienter zu markieren als nur auf Computer A.

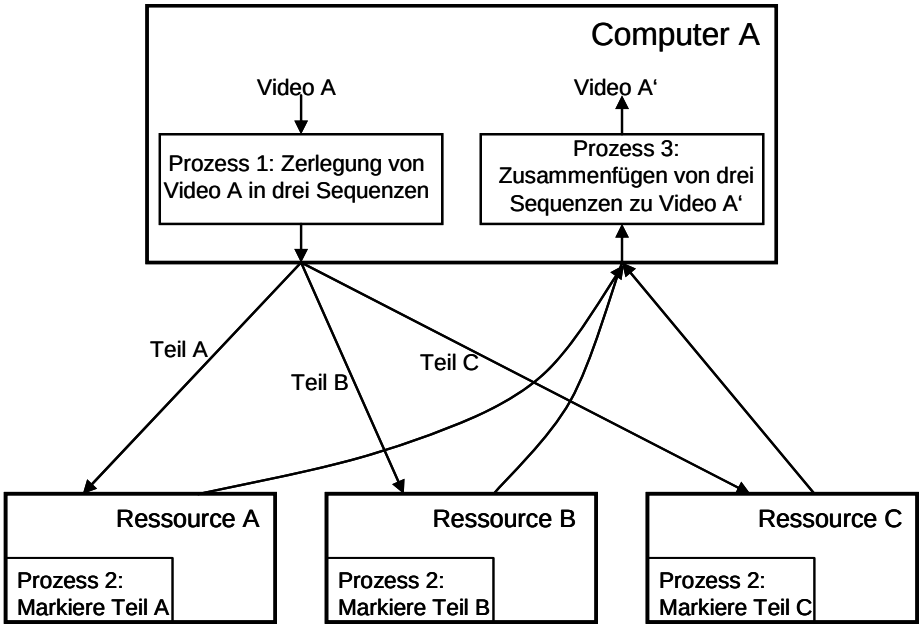

Abbildung 14: Verteilung der Markierung auf mehrere Ressourcen

Die vorgestellten Lösungen werden in Kapitel neun dargestellt. Dafür wird das Konzept eines Transaktionswasserzeichens definiert. Die vorgestellten Lösungen richten sich nach den in Kapitel 3.1 vorgestellten Anwendungsfällen. Jede der entwickelten Lösungen ist für je einen Anwendungsfall anwendbar.

5.2.5 Videoeditierung führt zur Veränderung der Integrität bzw. Echtheit des Videos

Dieses Kapitel beschreibt das Problem, wenn durch standardisierte Bearbeitungsprogramme die Integrität des Video verändert wird. In dem verfolgten Ansatz werden die Aspekte aufgelistet, die zur Generierung der Wasserzeicheninformation notwendig sind und wie das Wasserzeichen in das Video eingebettet wird, um die Integrität und Echtheit des Videos nachzuweisen.

5.2.5.1 Problembeschreibung

Da Veränderungen an digitalem Videomaterial ohne qualitativen Verlust durchgeführt werden können, erweist es sich dementsprechend auch als sehr aufwendig, die Veränderungen nachzuweisen. Die Änderungen haben aber verschiedene Bedeutungen in bezug zur Veränderung der Aussage des Videos. So werden durch Videoeditoren verschiedene Veränderungen, wie Rotation oder Skalierung durchgeführt, ohne dabei aber die Aussage des Videos zu verändern. Andererseits kann aber auch mit Veränderungen, wie Schnitte, die Bedeutung des Videos gezielt verändert werden.

Mit Videoeditoren werden gängige Veränderungen am Video durchgeführt, die Auswirkungen auf das eingebettete Wasserzeichen haben können. Cut-and-splice und cut-insert-splice sind zwei übliche Modifikationen in Videoeditoren. Cut-insert-splice wird z. B. bei Fernsehanstalten für das Einfügen von Werbung in Filmen verwendet. Für den

Wasserzeichendetektor haben Operationen, wie die Entfernung von Videoabschnitten jedoch Auswirkungen auf das Ausleseergebnis.

Veränderungen dieser Art können am Videomaterial nur schwer nachgewiesen werden. Deshalb muss das Wasserzeichen robust gegenüber Angriffe dieser Art sein. Wenn das Wasserzeichen redundant eingebettet wird, können durch Schnitte im Video verlorengegangene Wasserzeicheninformationen von ausgelesenen Wasserzeichen in anderen Abschnitten korrigiert werden. Wenn die gesamte Wasserzeicheninformation über einen Zeitraum im Video verteilt werden muss, können durch das Einbetten einer Startsequenz als Teil der Wasserzeicheninformation veränderte Teile des Videos übersprungen und wieder bei einer unverfälschten Videoabschnitt fortgefahren werden.

Durch Veränderungen des Videoinhaltes kann aber auch die Bedeutung dieses Inhalts verändert werden. Angriffe dieser Art verändern den Inhalt, in dem sie Elemente in den Frames verändern. Besonders in Anwendungsgebieten wie Überwachungsvideos erweist sich der Nachweis der Authentizität des Videos als notwendig.

Um die Integrität bzw. Echtheit von Videos nachzuweisen, werden fragile Wasserzeichen angewendet. Je nach erlaubten Veränderungen kann der Grad der Fragilität des Wasserzeichens bestimmt werden. Sind keine Veränderungen erlaubt, erweist sich ein fragiles Wasserzeichen als beste Lösung für diese Anwendung.

5.2.5.2 Designansatz

Die Sicherung der Authentizitä des Videos erfordert die Berücksichtigung folgender Aspekte:

(a) Sicherung der Integrität des Einzelbildes
(b) Sicherung der Integrität der Sequenzanordnung der Einzelbilder innerhalb des Videos

Der Nachweis von Integrität wird gängigerweise mit Hilfe von Hashverfahren durchgeführt. Der Vorteil eines Hashwertes liegt darin, dass es sich um eine nicht rückführbare Operation handelt. Es ist nicht möglich, aus ihm den abgebildeten Inhalt zu ermitteln. Nur ein Vergleich zwischen der erneuten Generierung des Hashs und dem vorliegenden Hashwert ermöglicht eine Kontrolle der Integrität.

Die Position des Bildes innerhalb des gesamten Videos ist eine Erweiterung zur Integrität des Videos gegenüber dem Einzelbild. Da es ohne großen Aufwand möglich ist, einzelne Bilder im Video auszuschneiden oder ihre Position innerhalb des Gesamtvideos zu verändern, muss eine zusätzliche Framenummer als Integritätsinformationen mit in das Wasserzeichen integriert werden. Damit wird sichergestellt, dass sich die Position des Einzelbildes nicht verändert hat.

Zum Nachweis der Authentizität kann eine digitale Signatur herangezogen werden. Durch Kenntnis des Schlüssels und der Signatur kann das Video durch Überprüfen der Signatur als authentisch eingestuft werden. Mit der Anwendung einer asymmetrischen Signatur besteht zusätzlich noch die Möglichkeit, die Überprüfung der Authentizität durch Einzelpersonen auf Gruppen zu erweitern.

Durch das Einbetten eines Wasserzeichens muss oft mit qualitativen Einbußen gerechnet werden. Bei höchst sensiblen Videodaten können geringe Veränderungen zu einer verfälschten Inhaltsaussage führen. Deshalb muss bei speziellen Anwendungsfällen auch die Möglichkeit gegeben sein, das Wasserzeichen wieder zu entfernen. Mit Hilfe von reversiblen Wasserzeichen können diese Anwendungsfälle realisiert werden. Die Videodaten, die durch das Wasserzeichen überschrieben werden, müssen vorher gesichert und als Bestandteil der Wasserzeicheninformation mit in das Video eingebettet werden. Durch eine Extraktion der Originaldaten aus dem detektierten Wasserzeichen kann das Originalvideo wiederhergestellt werden.

Die Herausforderungen des Wasserzeichens liegen in der Sicherheit und in der Bereitstellung der hohen Kapazität. Die Sicherheit ist ein wichtiger Aspekt, der durch die Anwendung der Kryptographie realisiert werden kann. Damit ist auch die Möglichkeit der Zugriffskontrolle gewährt. Das Wasserzeichen muss auf gezielte Veränderungen an den Videodaten reagieren können. Dies betrifft sowohl Veränderungen im Einzelbild als auch in der Sequenz der Bilder.

Eine entsprechend hohe Kapazität muss ebenfalls garantiert werden, damit die als sicher eingestuften kryptographischen Informationen eingebettet werden können. Bei gezielten Angriffen auf Hash- und Signaturverfahren hat sich gezeigt, dass Hashwerte mindestens 128 Bit und eine Signatur mindestens 1024 Bit Länge aufweisen sollten [Bu2007]. Eine größere Länge erweist sich dabei jedoch als sicherer. Zusätzlich wird bei der Anwendung eines reversiblen Wasserzeichen eine höhere Kapazität benötigt, da Originaldaten des Videokanals, der zum Einbetten der kryptographischen Daten genutzt wird, mit in die Wasserzeicheninformation integriert werden.

In Kapitel acht der Arbeit wird der Aspekt der Überprüfung der Authentizität des Video genauer diskutiert. Es wird das Konzept des reversiblen Wasserzeichens vorgestellt. Weitere diskutierte Bestandteile sind einmal die Erstellung einer kryptographisch sicheren Wasserzeicheninformation als auch notwendige Anpassungen am Wasserzeichenalgorithmus. Die Wasserzeicheninformation muss sicher sein, da sie die Informationen zur Wiederherstellung und die Signatur beinhaltet. Es werden auch Anpassungen dargestellt, da Verfahren wie [FrGo2001] nicht die notwendige Kapazität anbieten können, um eine zum heutigen Stand als sicher eingestufte Signatur einzubetten.

6 Synchronisation von MPEG-Videowasserzeichen

Das Einbetten von digitalen Wasserzeichen in intrakodierte MPEG Frames ist die zurzeit am häufigsten genutzte Methode digitaler MPEG-Videowasserzeichen [HaSt2005b]. Wie schon in Kapitel 4.1 aufgeführt, können zur Markierung von intrakodierten Frames durchaus Einzelbildwasserzeichen genutzt werden. Neben Vorteilen wie der guten Transparenz haben intrakodierte Wasserzeichen auch Mängel. Dieses Kapitel beschäftigt sich mit der Re-Synchronisierung der markierten I-Frames.

Es untergliedert sich in die folgenden Rubriken. Zuerst wird das grobe Konzept der zeitlichen Synchronisation vorgestellt, das das Ziel verfolgt, ursprünglich markierte I-Frames wiederzufinden. Anschließend werden verschiedene Vorgehensweisen diskutiert und evaluiert, wie das markierte Video wieder synchronisiert werden kann. Die Evaluation bildet die Basis für den in diesem Kapitel entwickelten und vorgestellten Ansatz des robusten Hashs. Dabei wird der Fokus auf die Modularität des Ansatzes gelegt, damit später auch Weiterentwicklungen universell eingesetzt werden können. Die abschließende Testevaluierung soll Aufschluss über die Anwendbarkeit des Ansatzes geben.

6.1 Konzept der zeitlichen Synchronisation

Intrakodierte Frames haben einen hohen Ähnlichkeitsgrad zu JPEG-Bildern. Die einzige Erweiterung ist der Skalierungssfaktor, der eine variable Quantisierung ermöglicht. Damit wird ein möglicher Über- oder Unterlauf des Videodekodierbuffers verhindert.

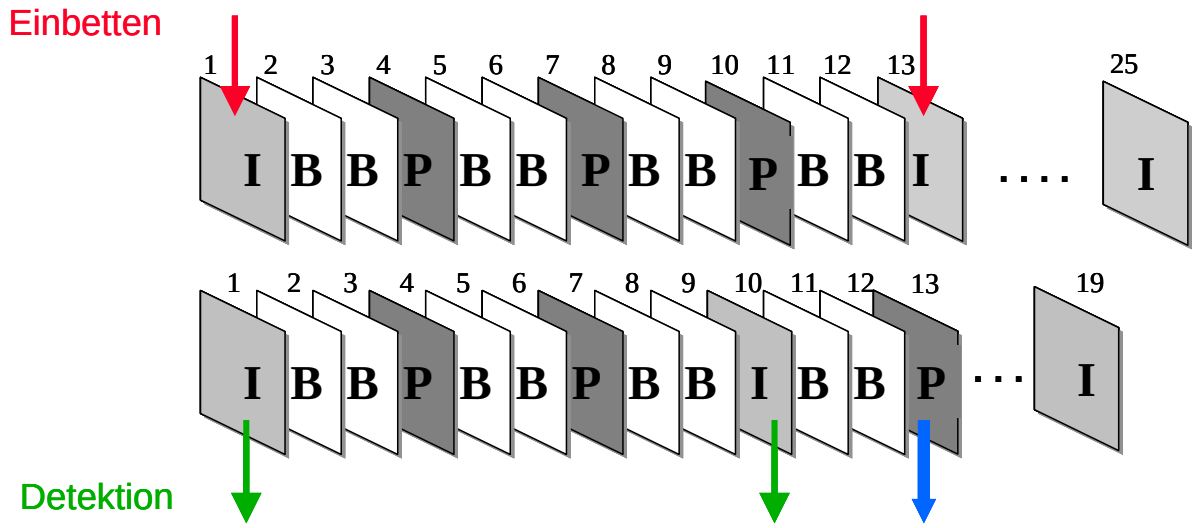

Abbildung 15: Zeitliche Synchronisation markierter I Frames

Die besondere Herausforderung bei intrakodierten Wasserzeichen ist die Synchronisation der markierten intrakodierten Frames. Da das gesamte Video nicht nur aus intrakodierten Frames besteht, müssen die markierten Frames vor dem Auslesen des Wasserzeichens detektiert werden. Nur die Analyse des Frametyps ist eine unzureichende Möglichkeit, um die

markierten Frames zu detektieren. In dieser Arbeit wird speziell der Fall der zeitlichen Synchronisation behandelt. Dabei werden entlang der Zeitachse durch Manipulationen entstandene Veränderungen erkannt und das manipulierte Video wieder in den ursprünglichen Zustand zurückgeführt. Abbildung 15 zeigt ein Beispiel, bei dem die Frametypen zweier Frames verändert wurden (grüne Pfeile). Ziel ist es, die ursprünglich markierten intrakodierten Frames wieder aufzufinden und so eine Detektion des Wasserzeichens aus ursprünglich nicht markierten Frames zu verhindern. In Abbildung 15 besteht die Aufgabe der Synchronisation darin, das Wasserzeichen aus Frame 13 und nicht aus Frame 10 auszulesen (blauer Pfeil). Die Lösungen, die speziell dafür entwickelt werden, müssen ebenfalls robust gegenüber den Veränderungen entlang der Zeitachse sein. Dabei kann die Wasserzeicheninformation zusätzliche Synchronisationsinformationen beinhalten oder es werden durch eine Vorfilterung des Videos die ursprünglich markierten Frames detektiert [HaSt2006].

6.2 Evaluation existierender Ansätze

Das Problem der Re-Synchronisation von Videodaten wurde schon in verschiedenen Publikationen betrachtet, die in den nächsten beiden Unterkapiteln aufgeführt werden. Im Allgemeinen lässt es sich in zwei Rubriken untergliedern: mit oder ohne Analyse des Frameinhaltes. Da aber aus den Publikationen nicht eindeutlich ersichtlich ist, welche Methode die bessere ist, werden zunächst beide Methoden evaluiert. Anschließend folgt aus den Ergebnissen ein Schluss über die entgültige Verwendung zur Synchronisation.

6.2.1 Inhaltsunabhängige Synchronisation

Das Prinzip besteht darin, ohne Analyse des Frameinhalts das Wasserzeichen wieder aufzufinden. Dabei werden zwei Methoden angewandt. Bei der ersten Methode wird zum eigentlichen Informationswasserzeichen ein weiteres Wasserzeichen in das Video eingebracht. Ziel dieser Methode ist es, das zweite Wasserzeichen, auch als Synchronisationsmuster bezeichnet, zu erkennen und mit dessen Hilfe die Veränderungen am Frame zurückzuführen, um danach das eigentliche Informationswasserzeichen wieder auslesen zu können.

Da das Synchronisationsmuster selbst für die Synchronisation zuständig ist, kann es selbst ein bevorzugtes Ziel von Angriffen sein. Durch Fälschung oder Zerstörung des Musters kann auch das Informationswasserzeichen nicht mehr ausgelesen werden kann. Um Schwachstellen dieser Art entgegen zu wirken, wird das Synchronisationsmuster selbst sehr robust eingebracht, was aber auch negative Auswirkungen auf die Transparenz haben kann [LiDe2002].

In [NiSc2002] wird ein temporäres Synchronisationsmuster eingebettet. Dabei wird davon ausgegangen, dass eine geometrische Veränderung in sämtlichen Frames eine zeitlich begrenzte Auswirkung hat. Deshalb wird das Videomaterial periodisch aufgeteilt und zusammen mit der Cronecker-Delta-Funktion werden einzelne Pixel verändert, so dass sie in einem Zeitabschnitt voneinander abhängig werden. Im Ausleseprozess wird nach entsprechender Vorfilterung die Abhängigkeit wiedererkannt.

Entgegen der Trennung von Synchronisation und Wasserzeicheninformation durch zwei verschiedene Wasserzeichen können diese auch miteinander verknüpft werden. Im Informationswasserzeichen befindet sich als Teil der Gesamtheit auch die Information über das Synchronisationsmuster. Der Auslesealgorithmus des Informationswasserzeichens muss sehr effizient sein, da er jetzt auch das Synchronisationsmuster auslesen muss. Nur anhand der ausgelesenen Information kann über deren Richtigkeit entschieden werden, da ein Teil der Information das Synchronisationsmuster beinhalten. [SeKa2002] beschreibt ein Videowasserzeichen, in dem die Wasserzeicheninformationen über ein räumlich aufgebautes Synchronisationsmuster verstreut werden.

6.2.2 Inhaltsabhängige Synchronisation

Entgegen der vorherigen beschriebenen Strategie der inhaltsunabhängigen Synchronisation werden bei der inhaltsabhängigen Synchronisation speziell dafür extrahierte Framemerkmale zur Synchronisation genutzt. Dabei ist es egal, ob die Eigenschaften mit dem Wasserzeichen verknüpft werden oder vor der Detektion des Wasserzeichens ein Vergleich zwischen den Werten der Merkmale während der Einbettung und Detektion des Wasserzeichens durchgeführt wird. Der wichtigste Aspekt ist die notwendige hohe Robustheit der zur Synchronisation extrahierten Merkmale. Man kann davon ausgehen, dass die gleichen Anforderungen an die Robustheit der extrahierten Merkmale gestellt werden wie beim Wasserzeichen, also wie. z. B. die Re-Enkodierung oder Formatkonvertierungen.

In [FrGo2000] und [Fr2000] wird eine robuste Hashlösung für Bilddaten vorgestellt. Es wird zwar von „visual hash“ gesprochen: da der Hash jedoch robust gegenüber geometrischen Transformationen ist, wird er auch als robuster Hash bezeichnet [OoKa2001]. Der generierte Hashwert ist robust gegenüber Rotation und Skalierung. Abschließend wird der Hash mittels eines Wasserzeichens in das Bild eingebettet.
Eine Vielzahl existierender robuster Hashverfahren basiert auf Bilddaten [MiVe2001], [RoVl2005]. In [RoVl2005] werden die Merkmale mit Hilfe der Radeon Transformation [LeCz2003] aus den Bilddaten extrahiert. In [SkUh2003] werden verschiedene Ansätze zur robusten Hashgenerierung speziell gegenüber einer JPEG- und JPEG2000-Komprimierung getestet. Aufgrund der hohen Korrelation der JPEG-Kompression zur MPEG-Kompression ist das Verfahren auch für MPEG geeignet.

Es werden aber auch speziell für Videodaten Hashverfahren entwickelt, da hier zusätzlich die zeitliche Achse bei der Verifizierung der Daten mitbeachtet werden muss [OoKa2001]. In [OoKa2001] werden zeitliche Veränderungen der Bilddaten für die Hashwertgewinnung herangezogen. Dabei werden die Durchschnitte von festgelegten Luminanzblöcken gebildet und ihre Veränderung zum lokalen und temporalen Vorgängerblock berechnet. Das Vorzeichen des Ergebnisses entscheidet über das Hashbit. In [ZhSc2006] wird eine Erweiterung zu [OoKa2001] beschrieben. Es werden Ähnlichkeiten der Luminanzdaten von räumlichen und zeitlichen Bereichen herangezogen und zur Hashwertberechnung benutzt. In [RoVl2005] werden bei der Merkmalsextraktion ähnliche Videosequenzen durch einen separaten Merkmalsvektor dargestellt; dagegen werden in [CoSa2004] die Merkmale basierend auf einer 3D-DCT-Transformation aus den Videodaten extrahiert.

In [OoKa2001] wird ein Videohash zur Identifikation des Mediums vorgestellt. Dabei ist es notwendig, dass die Eigenschaften für die Medienidentifikation ohne großen Aufwand extrahiert werden können. Mit ihnen kann das Medium identifiziert werden. Basierend auf den Helligkeitswerten werden der Durchschnitt und die Varianz in Bezug zu den Nachbarblöcken innerhalb des augenblicklichen Frames und zu den Nachbarframes ermittelt und daraus die Hashbits berechnet. Eine Erhöhung der Robustheit wird mit Hilfe einer Verringerung der Hashlänge erreicht. Ein im Verfahren auftretendes Problem, sind die Ergebnisse einzelner Hashbits. Ein festgelegter Schwellwert bestimmt das Hashbit. Ist das Ergebnis zu nahe am Schwellwert kann sich der Wert des Hashbits schon nach einfachen Operationen ändern. Um diesem Schwachpunkt entgegen zu wirken, werden im Verfahren nur die Hashbits herangezogen, die eine eindeutiges Ergebnis haben und so ein robustes Hashbit liefern. Dadurch kann behauptet werden, dass die Bitwerte sich auch nach einzelnen Operationen am Video, wie einer Re-Enkodierung, nicht ändern. Im Verfahren selbst wird von den glaubwürdigsten Hashbits gesprochen.

6.2.3 Vergleich existierender Ansätzen zur zeitlichen Synchronisation

Um zur späteren Anwendung für die Synchronisation des Wasserzeichens die richtige Methode auszuwählen, müssen die beiden Synchronisationsansätze zunächst evaluiert werden. Der entsprechend bessere Ansatz wird dann weiter verfolgt.

6.2.3.1 Anwendungsfall

Während der Evaluation liegt die Konzentration auf Veränderungen der Anordnung von Sequenzen im Video. Es gibt auch weitere Angriffsmöglichkeiten, die bis zur Manipulation einzelner Frames heruntergebrochen werden können. Die Manipulation von Szenen spiegelt ein breites und allgemeines Angriffsfeld wieder. Dabei liegt die Fokussierung auf die folgenden drei Angriffe:

(a) Sequenzaustausch
(b) Sequenzausschnitt
(c) Sequenzaddition

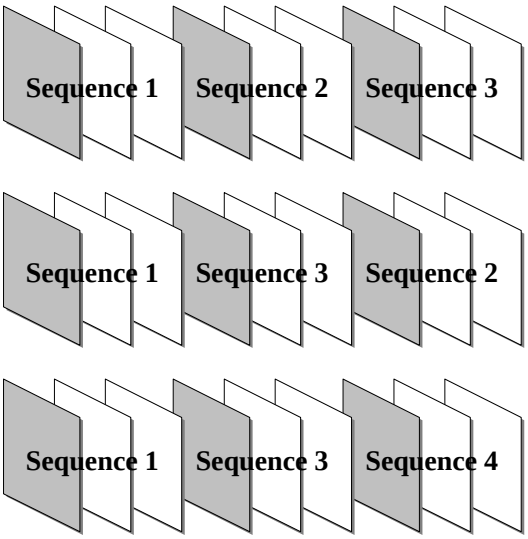

Abbildung 16: Exemplarische Demonstration der Sequenzordnung vor und nach dem Angriff

Abbildung 16 soll die Notwendigkeit der Betrachtung der Angriffe verdeutlichen. Es zeigt die Reihenfolge von jeweils drei Sequenzen, wobei die erste Anordnung die ursprüngliche Anordnung zeigt. Das jeweils grau dargestellte Frame stellt das intrakodierte Frame dar, das mit dem Wasserzeichen versehen ist. In der zweiten Anordnung wurden zwei Sequenzen ausgetauscht. In der dritten Anordnung wurde die zweite Sequenz gelöscht und eine weitere hinzugefügt. Gängige Wasserzeichen würden die Wasserzeicheninformationen an den intrakodierten Frames auslesen. Wenn sie aber über eine Reihe von intrakodierten Frames aufgeteilt wären, könnte das Ergebnis nur verfälscht dargestellt werden.

6.2.3.2 Inhaltsunabhängige Synchronisierung

Ausgehend von den Anforderungen und der Aussage in [LiDe2002], dass ein eigenständiges Synchronisationsmuster ein gezieltes Angriffsobjekt wäre, wird die zweite im vorherigen Kapitel beschriebene Methode bevorzugt, d.h. das Synchronisationsmuster wird mit der Wasserzeicheninformation verknüpft und als gemeinsames Wasserzeichen in das Video eingebettet [HaTh2004]. Das bevorzugte Synchronisationsschema setzt sich unter Berücksichtigung der MPEG-Referenzstruktur und der Aufteilung der Wasserzeicheninformation über mehrere intrakodierte Frames wie folgt zusammen:

Periodic ID	Sequence ID
1111100000	0000000000
1010110101	0000000001
0101010101	0000000010
•••	•••
1111100000	0000001000

Abbildung 17: Beispiel des Synchronisationsschemas

Die Synchronisationsinformation setzt sich aus einer periodischen (*PeriodicID*) und einer sequentiellen (*SequenceID*) Teilinformation zusammen. Die periodische ID wird nach einer festen Anzahl von Frames wiederholt und bildet somit eine festgeschriebene Periode, die immer wieder wiederholt wird. Die Größe der Periode richtet sich nach der Kapazität der Wasserzeicheninformation, d. h. über wie viele intrakodierte Frames das Wasserzeichen verteilt wird. Die *PeriodicID* ist so aufgebaut, dass sie zwischen den benachbarten intrakodierten Frames eine Hamming-Distanz von 50% hat. Dieser deutliche Unterschied verhilft zu einer sicheren Aussage über den Beginn eines weiteren ursprünglichen markierten Frames. Während der Detektion werden sämtliche Frames untersucht; genau der Fall, in dem sich die ausgelesene Wasserzeicheninformation deutlich von der des vorherigen Frames unterscheiden, deutet es auf ein ursprünglich markiertes Frames hin. Nachdem das gesamte Video gefiltert wurde und eine Menge von Frames identifiziert wurden, werden genau die Frames selektiert, deren ausgelesene Synchronisationsinformation der entsprechenden *PeriodicID* am ähnlichsten ist. Die *SequenceID* regelt die Reihenfolge der markierten Frames.

Dieser Informationsaufbau ist deshalb notwendig, weil die Kapazität von intrakodierten Frames bei robusten Wasserzeichen häufig limitiert sein kann. Verschiedene Faktoren beeinflussen die Datenrate des Wasserzeichens. Während die Größe des Frames und die Anzahl der Markierungseinheiten die Kapazität erhöhen, haben die Kompressionsrate im Frame, sowie angewandte Redundanz und Fehlerkorrekturkodes eher negative Auswirkungen auf die Kapazität. Da aber robuste Wasserzeichen diese Technologien verwenden, muss auch mit dem Zustand gerechnet werden, dass die Wasserzeicheninformation häufig über mehrere zu markierende intrakodierten Frame verteilt werden.

Als Wasserzeichen wird ein DCT-Block basiertes Luminanz-Wasserzeichen [DiSt1998] genutzt. Es bettet eine niederfrequentes Helligkeitsmuster in die Luminanzen der DCT-Blöcke des MPEG-Videos ein. Die eingebrachten Informationen verteilen sich über mehrere I-Frames.

Zur Evaluation werden vier Videos mit unterschiedlichem Bewegungsgrad herangezogen. Bei einer Länge von 70 bis 90 Sekunden und einer Bitrate von 1,7 Mbit/s haben sie eine Framerate von 25 Frames pro Sekunde (fps). Die Information zur Synchronisierung besteht aus 20 Bits und jedes Bit wird fünffach redundant eingebettet. Die Robustheit des Synchronisationsschemas wurde mit Hilfe einer Re-Enkodierung untersucht. Während der Evaluation wurde die GoP-Größe von ursprünglich zwölf auf fünf Frames pro GoP verändert.

Die Testergebnisse setzten sich aus der Anzahl der positiv ermittelten Perioden vor und nach der Re-Enkodierung zusammen. Der Slash trennt dabei die Testergebnisse vor und nach dem Angriff.

Video	Positive Perioden
video 1 (WMS 1.0)	62/40
video 1 (WMS 1.5)	62/43
video 2 (WMS 1.0)	59/20
video 2 (WMS 1.5)	59/24
video 3 (WMS 1.0)	27/27
video 3 (WMS 1.5)	27/27
video 4 (WMS 1.0)	13/5
video 4 (WMS 1.5)	19/10

Tabelle 6: Evaluationsergebnisse basierend auf dem periodischen Konzept

Die Testergebnisse aus Tabelle 6 demonstrieren, dass nicht sämtliche eingebetteten Synchronierungsperioden erfolgreich detektiert werden konnten. Auch eine Erhöhung der Wasserzeichenstärke (*WMS*) verbessert das Ergebnis nicht signifikant, was aber auch gleichzeitig zu einer Verschlechterung der visuellen Qualität des Video führt und nicht immer anwendbar ist.

6.2.3.3 Inhaltsabhängige Synchronisation

Bei diesem Teil der Evaluation ist eine spezielle Betrachtung des Aufbaus des Wasserzeicheninformation nicht notwendig. Die markierten Frames werden einzeln identifiziert. Abschließend kann anhand der bekannten ursprünglichen Anordnung der Frames die Reihenfolge automatisch wiederhergestellt werden.

Bei der inhaltsabhängigen Synchronisierung wird eine Identifikation der markierten Frames mit Hilfe von Hashverfahren durchgeführt. Dabei wird zur Evaluation der Anwendbarkeit das Bildhashverfahren von [Fr2000], [FrGo2000] herangezogen. Von jedem intrakodierten Frame wird ein Hashwert gebildet. Während der Entwicklung des Verfahrens ging man in [FrGo2000] davon aus, dass bei niederfrequenten DCT-Koeffizienten keine großen Änderungen der Einzelwerte vorgenommen werden können, ohne dass das Bild signifikant verändert wird.

Im Verfahren werden pseudozufällig n verschiedene Muster gebildet. Aufgrund einer anschließenden Glättung der Muster durch einen Tiefpassfilter wird eine entsprechende visuell akzeptable Qualität der Muster erreicht. Dadurch können z. B. Muster entstehen, wie sie in Abbildung 18 dargestellt werden. Anschließend wird der Durchschnitt des Musters abgezogen und so das Muster gleichanteilsfrei gemacht. Abschließend wird das Muster auf den Bildblock projeziert und der Schwellwert T berechnet, so dass 50% der Hashbits unter dem Schwellwert und 50% über dem Schwellwert liegen.

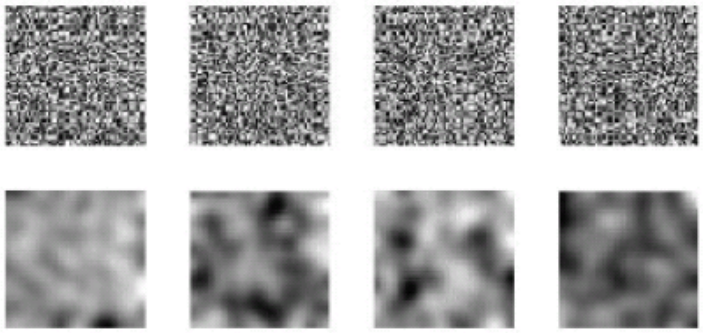

Abbildung 18: Pseudozufällige Muster und Glättung nach [FrGo2000]

Durch den existierenden Durchschnitt von null hängt die Projektion nur von den Variationen im Block ab. Zusätzlich ist die Projektion auch bildabhängig und Veränderungen sollten damit erkannt werden. Der Hashwert erweist sich als robust gegenüber verschiedenen Grauwertoperationen. Da das Verfahren aber sehr anfällig gegenüber geometrischen Operationen wie Rotation und Skalierung ist, wurde auch eine andere Methode vorgeschlagen. Um die Extraktion der Hashbits weniger sensitiv gegenüber Veränderungen am Graustufenbild zu machen, wird ein Kantenbild mit Hilfe des Laplacefilters vom Histogramm der Bilddaten erstellt.

Die Strategie in [SuCh2004] ist gegenüber dem in [Fr2000] vorgestellten Hash etwas verändert. Die zu extrahierten Informationen werden dem Wasserzeichendetektor mit übergeben, damit er vor dem Auslesen des Wasserzeichens die eigentliche Synchronisation durchführen kann. Der allgemeine Vorgang entspricht dann einem Vergleich zwischen den übermittelten und den erneut extrahierten Merkmalen der Videoframes, um auf die Bereiche zu stoßen, in denen das Wasserzeichen eingebettet wurde. Dieser Ansatz macht die Verwendung von robusten Hashwerten zur Synchronisation möglich, da durch einen Vergleich die Identifikation von markierten Frames oder Videosequenzen gelingt.

Das Verfahren wurde an einer Kollektion von MPEG-2-Videos getestet. Die Videos haben dabei eine Bitrate von 4 Mbit/s, einer Framerate von 25 fps und einer Bildgröße von 720 x 576 Bildpixel. Die Maße entsprechen ungefähr einem Video wie es auch auf einer DVD zu finden ist.

Da eine Anforderung an das Hashverfahren die Detektion von markierten intrakodierten Frames ist, hat sich relativ deutlich der Schwachpunkt von Bildhashverfahren in Bezug zur Ähnlichkeit von Nachbarframes herausgestellt [HaSt2006]. Bei einem robusten Hashverfahren wird entgegen der überlicherweise verwendeten fragilen Hashverfahren die Sensibilität der extrahierten Merkmalen heruntergesetzt, da sie auch nach einer Manipulation am Trägermedium noch korrekt ausgelesen werden sollen. Dadurch kann aber das Problem entstehen, dass benachbarte Frames einen identischen Hashwert aufweisen und es deshalb nicht immer möglich ist, das markierte intrakodierte Frame wieder zu bestimmen. Abbildung 19 zeigt benachbarte Bilder aus einem Beispielvideo. In der Abbildung sind jeweils die Frames mit der Nummer eins, zwei, drei und zehn von oben links nach unten rechts aufgeführt. Da das Video mit einer Framerate von 25 fps kodiert wurde, liegt der zeitliche Abstand zwischen den Nachbarbildern bei 0,04 Sekunden.

Abbildung 19: Beispiel der Ähnlichkeit benachbarter Bilder

Um eine mögliche Fehlerrate bei der Identifikation zu verringern, erweist es sich als effizienter, für die Identifikation der Frames keinen Bildhash, sondern einen Videohash zu verwenden. Durch die Betrachtung einer Sequenz von nachfolgenden Einzelframes erweist es sich als effizienter, die Eindeutigkeit der markierten Frames wieder zu erreichen.

Für die Evaluation wurde das Videohashverfahren von [OoKa2001] benutzt. Es berechnet die lokalen Differenzen des Durchschnitts bzw. die Varianz von einem Bildblock zu seinem Nachbarblock als auch die zeitliche Differenz oder Varianz des Bildblockes zu seinem Vorgängerblock im davor liegenden Frame. Auch dafür werden die Luminanzdaten herangezogen. Letztendlich entscheidet das Vorzeichen über den Wert des jeweiligen Hashbits.

Bei der Untersuchung dieses Videohashverfahren wurden die gleichen Videos wie beim Bildhashverfahren benutzt. Die Untersuchungen wurden aber mit einer Bitrate von 18 Mbit/s durchgeführt. Es kann davon ausgegangen werden, dass bei dieser Bitrate kein signifikanter Unterschied zu der angewandten Bitrate während der Tests des Bildhashs besteht. Einzig die Quantisierung ist etwas schwächer.

Der jeweils gebildete Hashwert beinhaltet das zu beachtende intrakodierte Frame jeweils an erster Stelle. Nach einer durchgeführten Re-Enkodierung mit einer Veränderung der GoP-Struktur war es fast immer möglich, das ursprüngliche intrakodierte Frame wieder aufzufinden. In den ursprünglichen Testergebnisse von [OoKa2001] wurde die Hashwertlänge auf 30 Frames erweitert, um so die gewünschte Robustheit zu erreichen.

Erste Untersuchungen haben gezeigt, dass ab einer Framelänge von fünf Frames pro Hashwert eine eindeutige Identifizierung der I-Frames möglich ist. Die identifizierten Framesequenzen enthielten die I-Frames jeweils an der ersten Stelle der Sequenz. Zudem ist

es nicht zu einer Framekollision gekommen, d.h. die ermittelten Hashwerte konnten eindeutig der Framesequenz zugeordnet werden. Die Fehlerquote der extrahierten Hashbits lag zwischen 5 bis 30 % und die Fehlerquote der extrahierten Wasserzeichenbits lag bei maximal 5% der Gesamtmenge der eingebetteten Wasserzeichenbits.

Zusammenfassend kann also gesagt werden, dass die Synchronisierung des Wasserzeichens weitaus erfolgreicher ist, wenn ein robuster Hash genutzt wird, als wenn die Synchronisierung über das Wasserzeichen selbst durchgeführt wird.

6.3 Modularer robuster Videohash zur Frameidentifikation

Die aus der Evaluierung gewonnenen Erkenntnissen verdeutlichen, dass die Verwendung eines robusten Hash für die Synchronisierung am effizientesten ist. Aufgrund der Eigenschaften von Videodaten wird ein videobasierter Hash entwickelt [HaBo2007].

Bei den Videodaten haben sich zwei besondere Gesichtspunkte herausgestellt, die einen entscheidenden Anteil bei der Entwicklung des Hashverfahrens haben:

(1) Die visuellen Unterschiede zwischen benachbarten und auch entfernten Frames können minimal sein.
(2) Die Höhe der visuellen Unterschiede zwischen Frames ist nicht proportional zur zeitlichen Distanz zwischen diesen Frames.

Um eine ausreichende Robustheit für die Einzelbilder zu erlangen, erweist es sich als notwendig, Ideen von Bildhashverfahren zu übernehmen. Speziell für Bilder entwickelte Hashverfahren weisen eine hohe Robustheit auf, z. B. gegenüber geometrischen und nicht-geometrischen Angriffen [FrGo2000], [Fr2000], [RoVl2005]. Deshalb werden die Merkmale auf Frameebene extrahiert. Da aber die zeitliche Ebene nicht vernachlässigt werden darf, müssen die extrahierten Merkmale auf Videoebene analysiert werden. Dadurch soll die Veränderung der framebasierten Merkmale mit betrachtet werden. Ziel ist es, eine Individualisierung von Videoabschnitten zu erreichen, da, wie schon in der Evaluierung erkannt, Frames nicht immer individuell erfolgreich identifiziert werden können. Abschließend müssen zur Synchronisation der Videoframes die Hashwerte verglichen und die markierten Frames identifiziert werden.

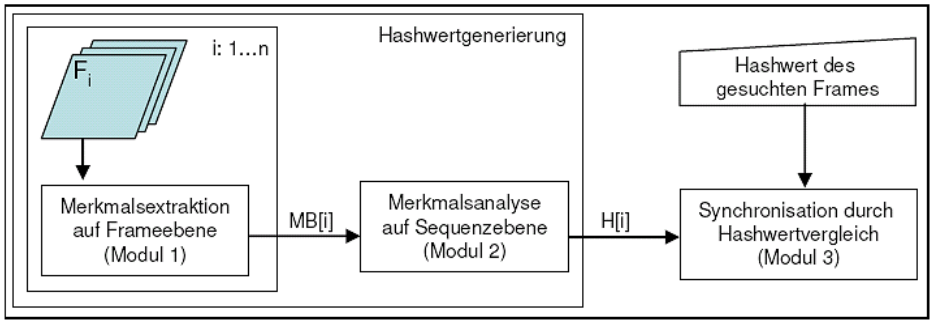

Abbildung 20: Konzept des robusten Videohashs

Abbildung 20 zeigt das Konzept des Videohashs. Dabei werden die vorherigen beschriebenen drei Prozesse in drei Module unterteilt. In Modul 1 werden die Merkmale von den Frames F_i extrahiert und in Merkmalsblöcke MB_i abgespeichert. Im zweiten Modul werden die Merkmalsblöcke MB_i analysiert und daraus der Hashwert H_i gebildet. Im dritten Modul wird der Vergleich zwischen den generierten Hashwerten H_i und der Hashwerte der gesuchten Frames durchgeführt.

Der nächste Schritt ist eine Kombination des in Abbildung 20 vorgestellten Hashverfahrens mit den Synchronisationsschritten des Wasserzeichens. Dabei ergibt sich ein allgemeines Synchronisationskonzept für die Synchronisation von Videoframes, dargestellt in Abbildung 21.

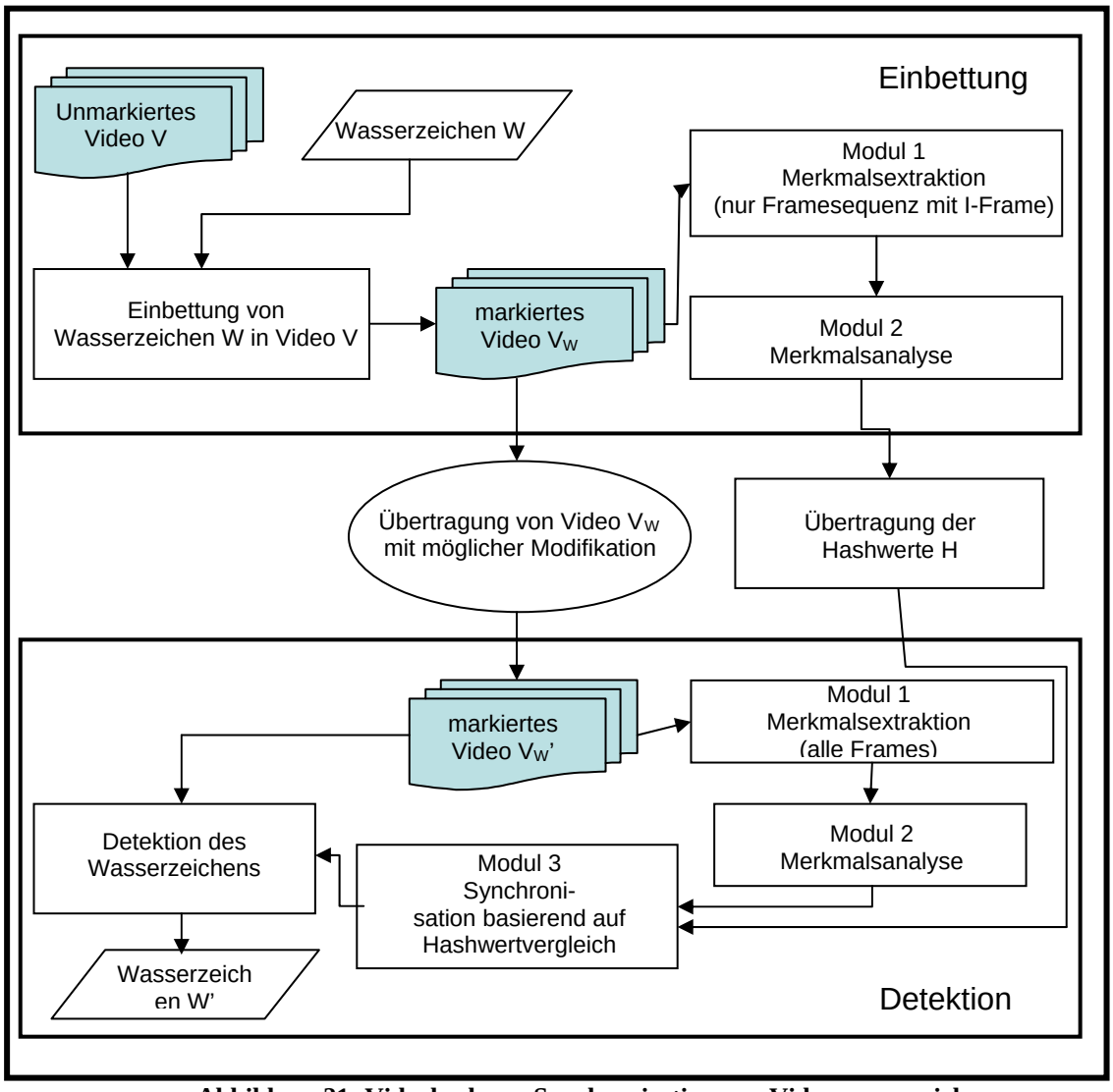

Abbildung 21: Videohash zur Synchronisation von Videowasserzeichen

In den Kapiteln 6.3.1 bis 6.3.3 werden die drei Module genauer erläutert. Sie werden zuerst allgemein erläutert und werden dann mit einem jeweiligen für das Modul geeignete Verfahren spezifiziert. Die Module sind allgemein aufgebaut, damit die in dieser Lösung genutzten Verfahren auch durch andere geeignete Verfahren ersetzt werden können.

6.3.1 Merkmalsextraktion auf Frameebene

Für die Extraktion der Merkmale kann jeder Hashalgorithmus genutzt werden, der es ermöglicht, die Framedaten auf die Merkmalsmatrix *MB* abzubilden. Die Merkmalsmatrix *MB* besteht dabei aus l_{MBX} x l_{MBY} skalaren Werten. Die Größe der Werte von l_{MBX} und l_{MBY} sind dabei frei wählbar. Die Merkmalsmatrix wird von Bildhashverfahren gebildet.

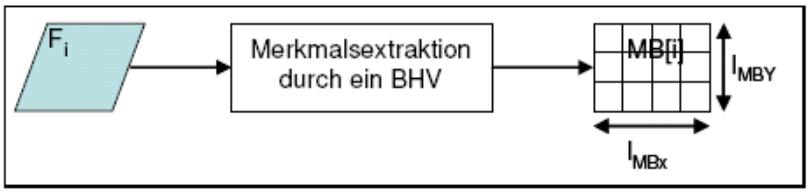

Abbildung 22: Merkmalsextraktion durch ein Bildhashverfahren (BHV)

Jedes Hashverfahren bildet das jeweilige Hashbit mit einer Schwellwertoperation. Da dadurch die Merkmalsinformationen auf jeweilige 1-Bit-lange Codewörter quantisiert werden würden, entfällt diese Operation bei der Merkmalsextraktion. Dadurch wird sichergestellt, dass die extrahierten Merkmale komplett der Merkmalsanalyse übergeben werden. Zudem erweist es sich auch als sinnvoll, die Werte der Merkmalsmatrix anhand ihres Robustheitsgradres zu ordnen. So befindet sich der robusteste Wert an Position [0][0] und der sensitivste Wert an Position[l_{MBX}][l_{MBY}].

Zusammenfassend kann die Funktion auch als eindeutige Funktion mit der folgenden Formel dargestellt werden:

$$f_{BHV} :\mapsto R^{l_{MBX}} \times R^{l_{MBY}} \qquad (7)$$

6.3.2 Merkmalsanalyse auf Sequenzebene

Während dieses Prozesses werden die Unterschiede bzw. Gemeinsamkeiten zwischen den benachbarten Frames analysiert und in jeweilige Hashwerte umgewandelt. Dafür lassen sich Aspekte von Videohashverfahren (VHV) übernehmen, die genau die zeitliche Veränderung mit betrachten.

Dafür werden die Merkmalsmatrizen *MB* der Frames i bis $i + j$ analysiert und in einen Video-Merkmalsmatrix $MV[i]$ umgewandelt (Abbildung 23). Dabei können sich aufgrund der Merkmale verschiedener angewandter Hashverfahren $MB[i]$ und $MV[i]$ voneinander unterscheiden. Die Dimension von $MV[i]$ wird dabei vom verwendeten Videohashverfahren bestimmt. Abschließend wird basierend auf der Schwellwertoperation des Videohashverfahrens der Hashwert $H[i]$ gebildet. Die Länge des Hashwertes $H[i]$ richtet sich dabei nach der Menge der Elemente der Merkmalsmatrix $MV[i]$. Die Anzahl der

verwendeten Frames ist dabei der maßgebliche Aspekt zur Identifikation der markierten Frames des Videos.

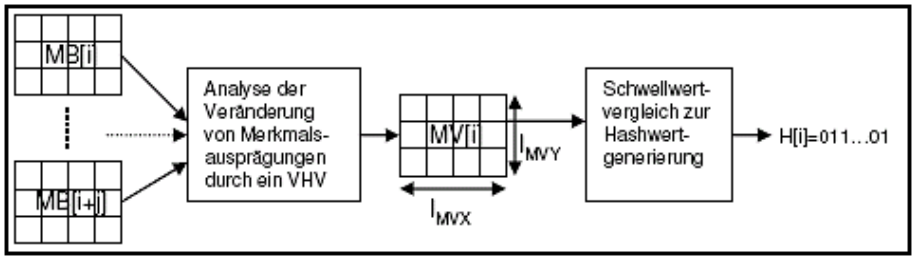

Abbildung 23: Hashwertgenerierung durch ein Videohashverfahren

Obwohl der Hashwert $H[i]$ über eine Framesequenz gebildet wurde, wird er nur zur Bestimmung der markierten Frames F_i genutzt. Jeweils ein Frame der gesamten Sequenz ist das jeweilige markierte Frame.

Der Algorithmus zur Generierung der Videomatrizen wird wie folgt beschrieben:

$$f_{VHV} : \left\{ MB \mid MB \in R^{l_{MBX}} \times R^{l_{MBY}} \right\} \mapsto R^{l_{MVX}} \times R^{l_{MVY}} \qquad (8)$$

Der Vergleich des Schwellwertes wird wie folgt beschrieben:

$$f_{SH} : \left\{ MV \mid MV \in R^{l_{MVX}} \times R^{l_{MVY}} \right\} \mapsto \{0,1\}^{l_H} \qquad (9)$$

6.3.3 Hashwertvergleich

Während des Detektionsprozesses des Wasserzeichens dient ein vorgezogener Hashwertvergleich zur Identifikation der ursprünglich markierten Frames. Dabei wird ein Vergleich der Hashwerte $H[i]$ von F_i mit den Hashwerten $H'[\ ]$ der Frames F' eines gegebenfalls manipulierten Videos durchgeführt. Mit Hilfe eines festgelegten Schwellwertes werden die Elemente der Hashwerte $H'[\ ]$ auf ihre Zugehörigkeit zu $H[i]$ hin überprüft und die markierten Frames identifiziert.

$$f_{\Delta H} : \left\{ (H, H') \mid H \in \{0,1\}^{l_H}, H' \right\} \mapsto \{0,1\}^{l_H} \qquad (10)$$

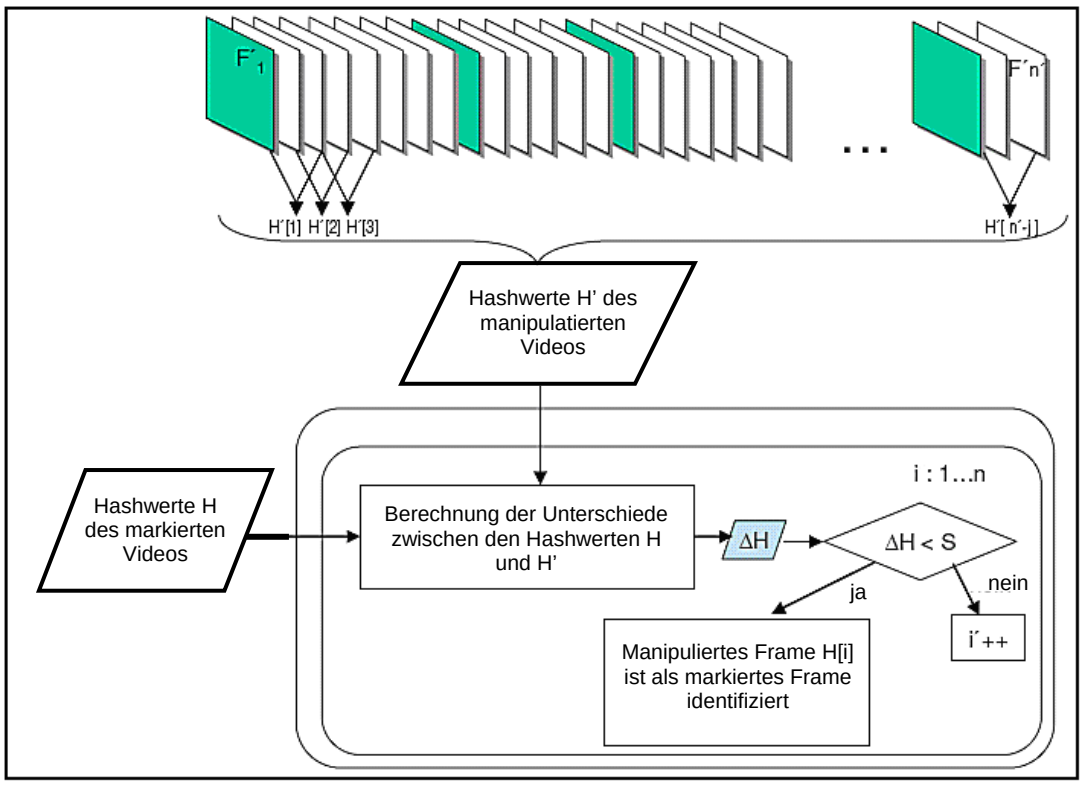

Abbildung 24: Identifizierung markierter Frames

Abbildung 24 zeigt die Identifizierung der markierten Frames aus der Menge von Hashwerten H. Nach der Bewertung der Unterschiede zwischen den Hashwerten wird mit einer anschließenden Schwellwertoperation die Zugehörigkeit der Frames F' zu den ursprünglich markierten Frames F überprüft. Am Ende wird aus den detektierten Frames das Wasserzeichen ausgelesen.

6.3.4 Eingesetzte Verfahren für die Module 1 bis 3

Für die Merkmalsextraktion (Modul 1) dient das Hashverfahren von [OoKa2001], das in Kapitel 6.2.2 vorgestellt wurde. Es werden lokale Differenzen der Luminanzwerte für die Merkmalsextraktion herangezogen. Entgegen dem ursprünglich vorgestellten Algorithmus mit einer festen Blockgröße wird in Relation zur Größe des Frames eine relative Blockgröße angewendet. Dadurch soll eine Robustheit gegenüber einer linearen Skalierung des Frames erreicht werden. Parallel zur Anwendung des Durchschnitts zur Merkmalsextraktion wird n einer weiteren Variante die Varianz angewendet. Die Varianz kann aufgrund ihrer Eigenschaft als Moment zweiter Ordnung Diskontinuitäten im Helligkeitsverlauf robuster als der Durchschnitt abbilden. Diese Erkenntnis wurde auch in [RoVl2005] angewandt, nachdem man auch hier zuerst den Durchschnitt genutzt hatte. Mit der Anwendung einer zusätzlichen Quantisierungsfunktion ist es möglich, die Genauigkeit der extrahierten Merkmale mit zu beeinflussen. Dadurch konnte in [Ne2005] eine Verbesserung der Robustheit erreicht werden.

Für die Merkmalsanalyse (Modul 2) wurden zwei bestehende Ansätze überprüft und dem Videohash angepasst. Auch hier werden wieder die Unterschiede bzw. Ähnlichkeiten der Merkmale herangezogen. Während dieser Bewertung dient aber der zeitliche Verlauf der

Framesequenz als Bewertungskriterium für die spätere Hashwertgenerierung. Die Größe der Gruppe ist dabei stark abhängig vom zeitlichen Bewegungsverlauf im Video.

Der erste Ansatz basiert auch auf [OoKa2001]. Durch einen 2x2-Haar-Filter werden zeitliche und räumliche Differenzen der extrahierten Merkmale berechnet. Der Haar-Filter bietet selbst eine gute Robustheit gegenüber nicht-geometrischen Merkmalen.

Im zweiten Ansatz wird entgegen der Differenz die Korrelation der extrahierten Merkmale berechnet [ZhSc2006]. Speziell dafür wird eine Autokorrelationsfunktion zur Analyse der extrahierten Merkmale angewandt.

Zusätzlich bestünde die Möglichkeit der Anwendung von sogenannten Grenzframes. In [RoVl2005] werden Grenzframes berechnet, um Videos in Sequenzen zu unterteilen und nach der Merkmalsanalyse einen Hashwert zu generieren. Das Verfahren kann aber nur als optionales Verfahren genutzt werden, wenn es gelingt, auch nach der Einteilung des Videos in Sequenzen eindeutig die markierten Frames zu analysieren. Eine Vielzahl von MPEG-Enkodern bietet die Möglichkeit der Schnittanalyse an. Dabei beginnt ein neuer Schnitt mit einem intrakodierten Frame, wodurch mögliche visuelle Artefakte vermieden werden können. Durch diese Möglichkeit würden zwar weniger Frames für die Synchronisation des Wasserzeichens benutzt werden, die generierten Hashwerte können jedoch als robuster eingestuft werden.

Für den Vergleich der Hashwerte und die Identifizierung der Frames (Modul 3) wird der Hamming-Abstand benutzt. Es wird die Anzahl der mit dem gleichen Wert festgelegten Hashwertbits bestimmt. Die abschließende Identifizierung des Frames wird mit einem festgelegten Schwellwert durchgeführt. Liegt das Ergebnis des Vergleichs über dem Schwellwert, gehört das aktuell analysierte Frame $F'[i]$ zu den ursprünglich markierten Frames F'.

6.4 Evaluation der Ansätze

Zuerst werden die Wahlparameter der drei Module dahingegen untersucht, um die optimalen Parameter für die Detektion der intrakodierten Frames zu finden. Basierend auf den Ergebnissen werden abschließend die Detektionsraten an zwei verschiedenem Videosets in SVCD- und DVD-Qualität dargestellt.

Während der Detektion der Frames werden folgende Kategorien unterschieden:

(1) Richtig erkannt: Anzahl der erkannten intrakodierten Frames
(2) Nicht erkannt: Anzahl der nicht erkannten intrakodierten Frames
(3) Falsch erkannt: Anzahl der falsch erkannten Frames

Für die Analyse wird ein Referenzdatensatz von MPEG-2 Videos genutzt. Die Videos haben DVD Qualität in Bildgröße und Bitrate und decken verschiedene Eigenschaften wie Bewegung, Verteilung der Farben und Inhaltsmenge ab. Sie beinhalten 38 intrakodierte Frames bei einer Gesamtmenge von 375 Frames.

Zuerst wird die Auswirkung des Datenmaterials auf die Framedetektion untersucht. Dabei werden die Videos mit einer Datenrate von 12 MBit/s in 8x6 Pixelblöcke unterteilt. Abschließend werden die Videos mit 8 MBit/s re-enkodiert.

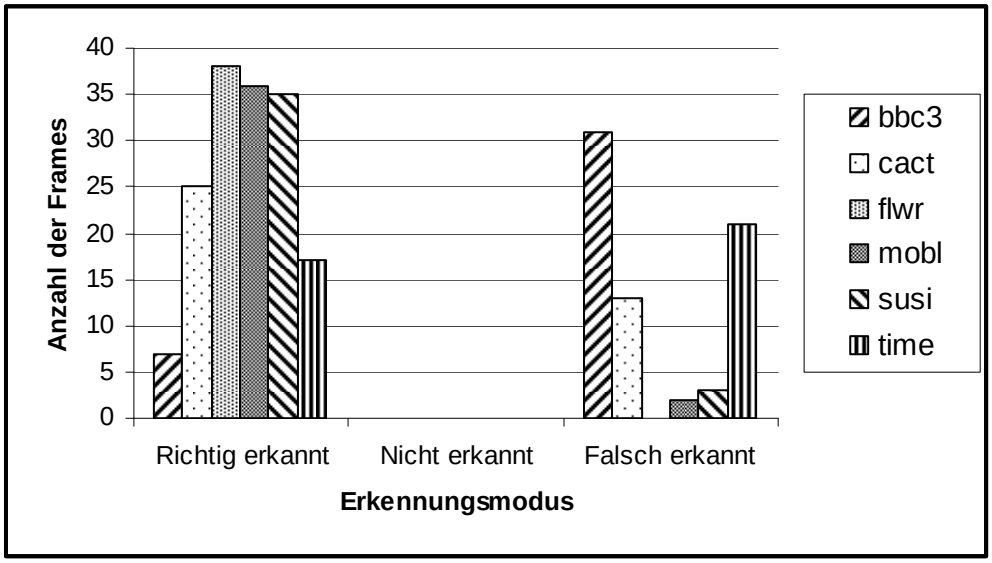

Abbildung 25: Hash Detektionsrate bei verschiedenem Datenmaterial

Die Ergebnisse zeigen, dass Videos mit konstanter Bewegung (flwr, mobl, susi) die besten Ergebnisse zeigen. Bei einem geringen Bewegungsgrad (bbc3 und time) werden hingegen eine Vielzahl von Frames falsch erkannt. Das stellt einen Schwachpunkt des Hashverfahrens dar. Bei geringer bis keiner Bewegung ist das Verfahren nicht anwendbar. Zusätzlich ist aber auch zu erkennen, dass nicht der Fall aufgetreten ist, dass ein ursprünglich markiertes Frame nicht erkannt wurde,. Dies deutet darauf hin, dass die Hashwerte wieder detektiert wurden, nur mit der Besonderheit, dass sie manchmal falsche Frames identifiziert haben.

In einem weiteren Test wurde die optimale Blockgröße evaluiert. Dabei wurde sowohl der Durchschnitt als auch die Varianz als eindeutiges Hashmerkmal für Bilddaten analysiert.

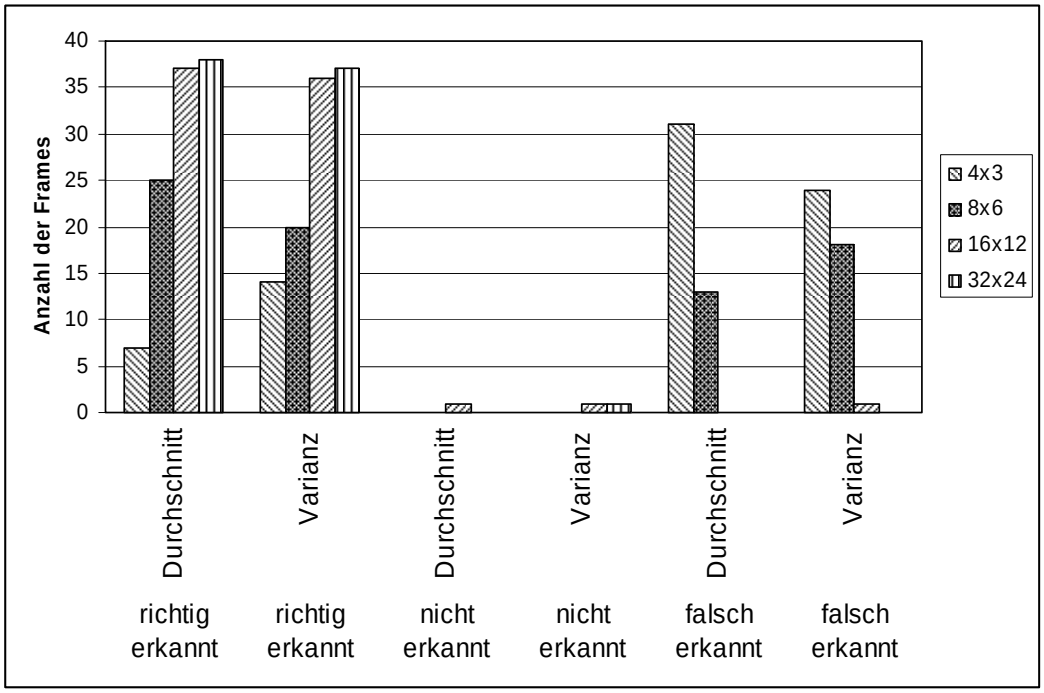

Abbildung 26: Detektionsrate bei verschiedenen Blockgröße

Bei den Blockgrößen 16x12 und 32x24 wurden entsprechend gute Auslesequoten erreicht. Da sie relativ identisch sind, ist eine Blockgröße von 16x12 ausreichend.

Danach wurde die Robustheit des Hashs bei verschiedenen Framesequenzlängen untersucht. Dabei wurden Längen von einer bis zehn nachfolgenden Framesequenzen untersucht. Als Merkmal wurde auch wieder der Blockdurchschnitt und die Varianz untersucht.

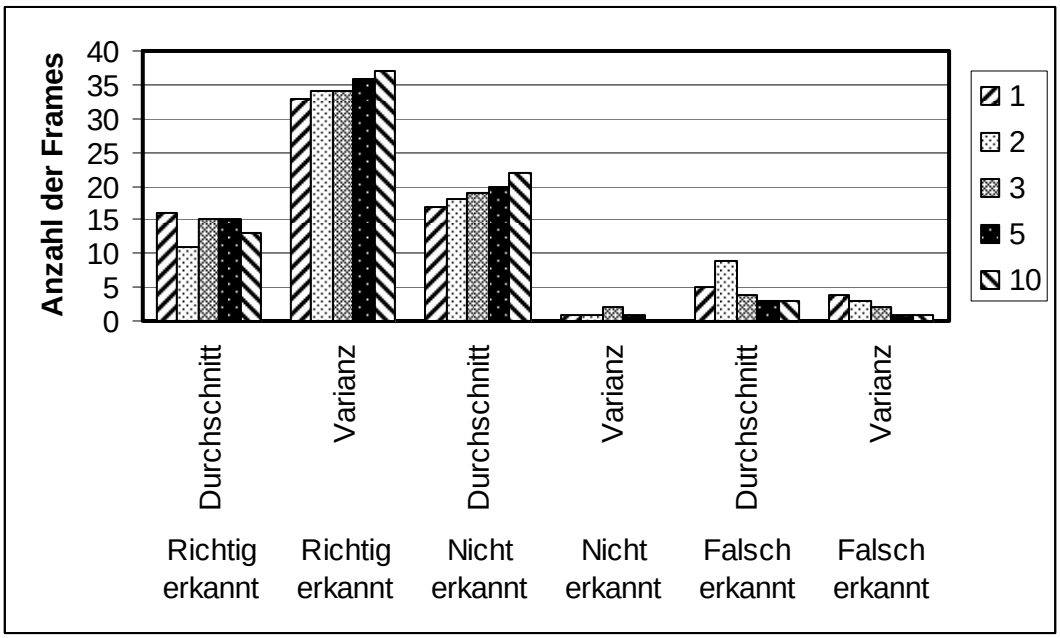

Abbildung 27: Auswirkung unterschiedlicher Sequenzlänge

Es ist besonders zu erkennen, dass die Varianz als Merkmal für die Länge der Videosequenz pro Hashwert deutlich robustere Ergebnisse liefert, als der Durchschnitt.

Zuletzt wurde noch untersucht, wie hoch der Schwellwert T der Hamming-Distanz sein soll, um die markierten Frames zu detektieren.

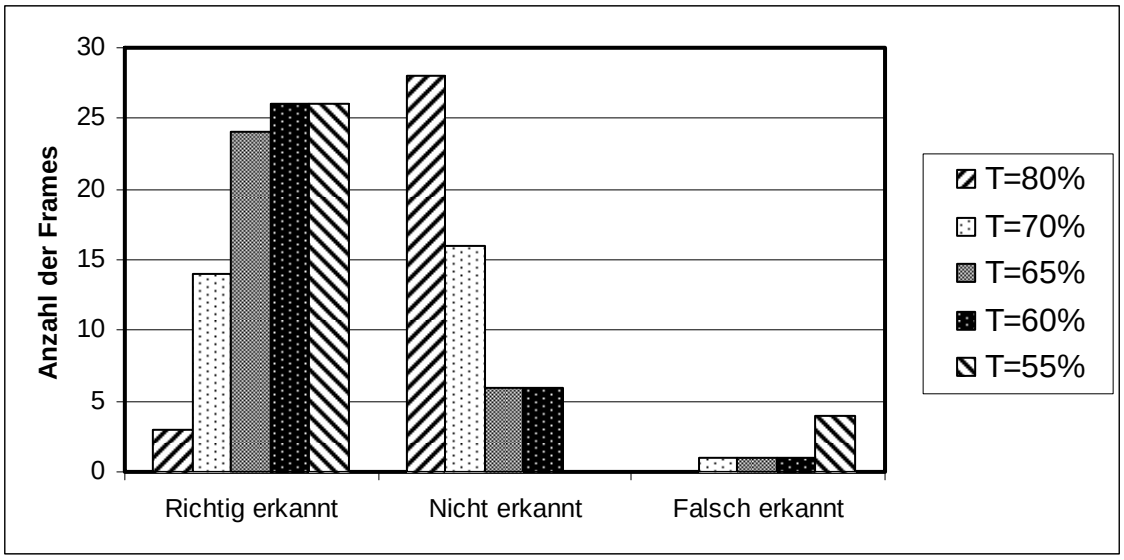

Abbildung 28: Schwellwert zur Frameidentifikation

Anhand der Ergebnisse ist zu erkennen, dass es ausreichend ist, wenn der Schwellwert zwischen 60 bis 70% liegt. Ein hoher Schwellwert führt zu einer höheren Anzahl an nicht identifizierten Frames, ein niedriger zu einer höheren Anzahl von falsch erkannten Frames.

Basierend auf den Testerkenntnissen wurden die Parameter für die Detektion der markierten Frames wie folgt festgelegt:

(1) Modul 1:
   (a) Blockgröße 16 x 12 Pixel
   (b) Bildmerkmalsextraktion basiert auf arithmetischem Durchschnitt
(2) Modul 2:
   (a) Merkmal der Framesequenzextraktion basiert auf Varianz
   (b) Parameter j = 5 (Framesequenz hat die Länge von 5 Frames)
   (c) Selektion von 90% der extrahierten Hashwerte (die 10% der schwächsten Hashwerte werden nicht berücksichtigt)
(3) Modul 3:
   (a) 60% der Bits der Hashwerte zwischen H[i] and H'[i] müssen identisch sein, um Frame F` als ein original markiertes Frame zu identifizieren.

Basierend auf den festgelegten Parametern wurde das Hashverfahren an zwei verschiedenen Videosets angewandt und verschiedenen Angriffen unterzogen. Der Fokus liegt dabei auf Videomaterial, wie es im analogen Fernsehbereich ausgestrahlt wird, und auf DVD-Material. Sie unterlegen verschiedenen Qualitätsstufen und aufgrund der weitverbreiteten Anwendungsbereiche ist die Fokussierung auf diese beiden Anwendungsbereiches durchaus sinnvoll. Videoset 1 besteht aus 40 Dokumentarvideos, aufgenommen aus dem übertragenen

analogen Fernsehprogramm. Videoset 2 besteht aus 20 Videos in DVD-Qualität. Die Videos bestehen aus den folgenden Parametern:

(1) Videoset 1: SVCD-Video
    (a) Bitrate 1500 kBit/s
    (b) Framerate 25 fps
    (c) GoP 12 Frames pro GoP
    (d) Super-Video-CD-Format
(2) Videoset 2: DVD-Video
    (a) Bitrate 4000 kBit/s
    (b) Framerate 25 fps
    (c) GoP 12 Frames pro GoP
    (d) DVD-Format

Die durchgeführten Angriffe sind:

(1) Veränderung der GoP-Größe von 12 auf 9
(2) Veränderung der GoP-Größe von 12 auf 15
(3) Re-Enkodierung mit einer Bitrate von 1500 auf 1000 kBit/s (Videoset 1) bzw. von 4000 auf 2000 kBit/s (Videoset 2)
(4) Re-Enkodierung mit einer Bitrate von 1500 auf 750 kBit/s (Videoset 1) bzw. von 4000 auf 1000 kBit/s (Videoset 2)
(5) Veränderung der Framerate auf 24 fps
(6) Veränderung der Framerate auf 30 fps

Abbildung 29 stellt die Testergebnisse als prozentualen Vergleich zwischen den drei jeweils möglichen Kategorien dar. Dabei werden drei Detektionsraten herangezogen:

(1) Right Dedection Rate (RDR): Sie entspricht dem prozentualen Anteil der detektierten Frames in Relation zur Menge der ursprünglich mit dem Wasserzeichen markierten Frames.
(2) Non Dedection Rate (NDR): Sie entspricht dem prozentualen Anteil der nicht detektierten Frames in Relation zur Menge der ursprünglich mit dem Wasserzeichen markierten Frames.
(3) False Dedection Rate (FDR): Sie entspricht dem prozentualen Anteil der falsch detektierten Frames in Relation zur Menge der ursprünglich mit dem Wasserzeichen markierten Frames.

Da sämtliche drei Detektionsraten 100% ergeben, werden sie in der folgenden Graphik in gestapelter Form gegenübergestellt. Dabei wird ein Durchschnitt der Detektionsraten der sechs durchgeführten Angriffe gebildet. Mit Hilfe des Gesamtdurchschnittes kann ein Überblick über die Qualität der Robustheit des Hashwertes erkannt werden.

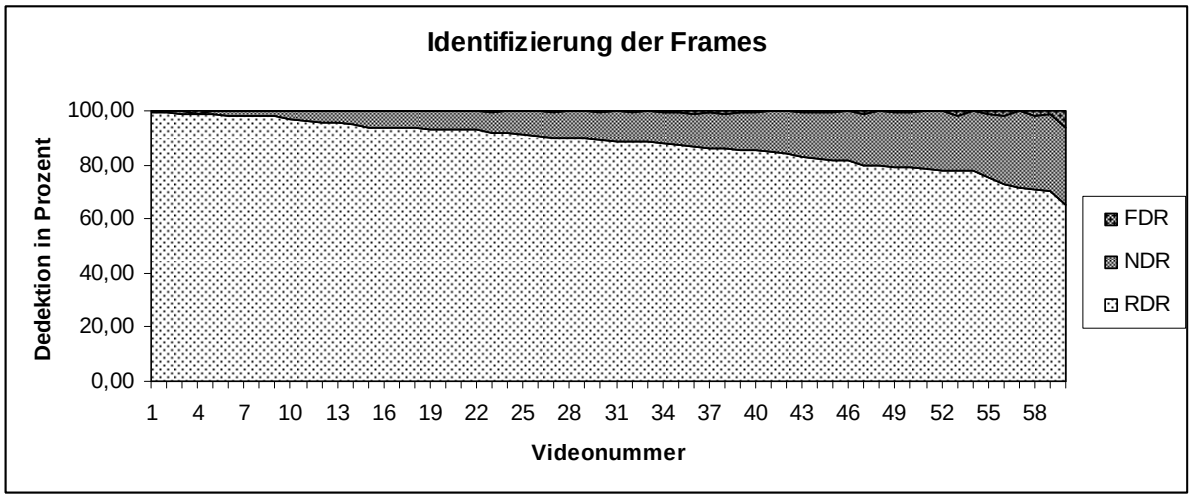

Abbildung 29: Identifikation der Videoframes

Aus den Erkenntnissen von Abbildung 29 kann geschlossen werden, dass eine allgemein hohe Detektionsrate erreicht werden kann. Zudem ist die FDR-Detektionsrate fast nicht existent. Dadurch kann der schlechtere Fall, nämlich dass das Wasserzeichen aus den falschen Frames ausgelesen wird, fast ausgeschlossen werden. Allerding ist auch zu erkennen, dass die RDR-Dedektionsrate bereits deutlich nachlässt. Das liegt am qualitativen Unterschied zwischen beiden Videosets. Videoset 2 scheint eine bessere RDR-Detektionsrate zu haben als Videoset 1. Um das genauer zu untersuchen, werden beide Videosets noch einmal separater untersucht.

Die separaten Testergebnisse der Untersuchung von Videoset 1 werden in Abbildung 30 dargestellt. Zur besseren Übersicht werden nur 20 Videos aus Videoset 1 untersucht. Da die Testergebnisse von Angriff 1 und 2 sowie von Angriff 3 und 4 relativ identische Werte geliefert haben, wird nur jeweils einer der beiden Angriffe im Diagramm dargestellt. Nur ein Vergleich von Angriff 5 und 6 liefert teils deutliche Unterschiede, wobei eine leichte Verlangsamung des Videos von 25 auf 24 fps zu teils gravierenden Einbußen der Framedetektion führt. Die kompletten Ergebnisse können aus Anhang B entnommen werden.

Die deutlich schlechten Ergebnisse von Angriff 5 können darauf zurückgeführt werden, dass im neu erstellten Video benachbarte Frames zu einem gemeinsamen Frame interpoliert werden. Durch die Interpolation besteht die Möglichkeit, dass sich die Informationen im Frame so stark verändern können, dass das Hashverfahren veränderte Featurewerte extrahiert.

Zusätzlich ist bei Vergleich der beiden Diagramme noch zu bemerken, dass das DVD-Datenmaterial generell eine bessere Detektionsrate gegenüber dem SVCD-Material hat.

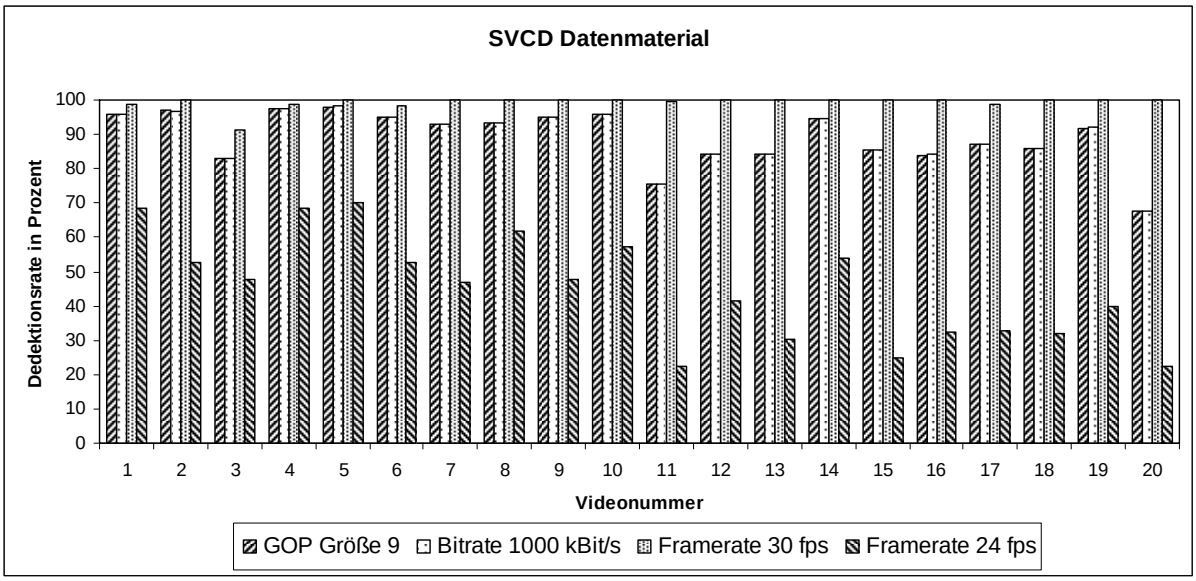

Abbildung 30: Robustheitsanalyse von Videoset 1

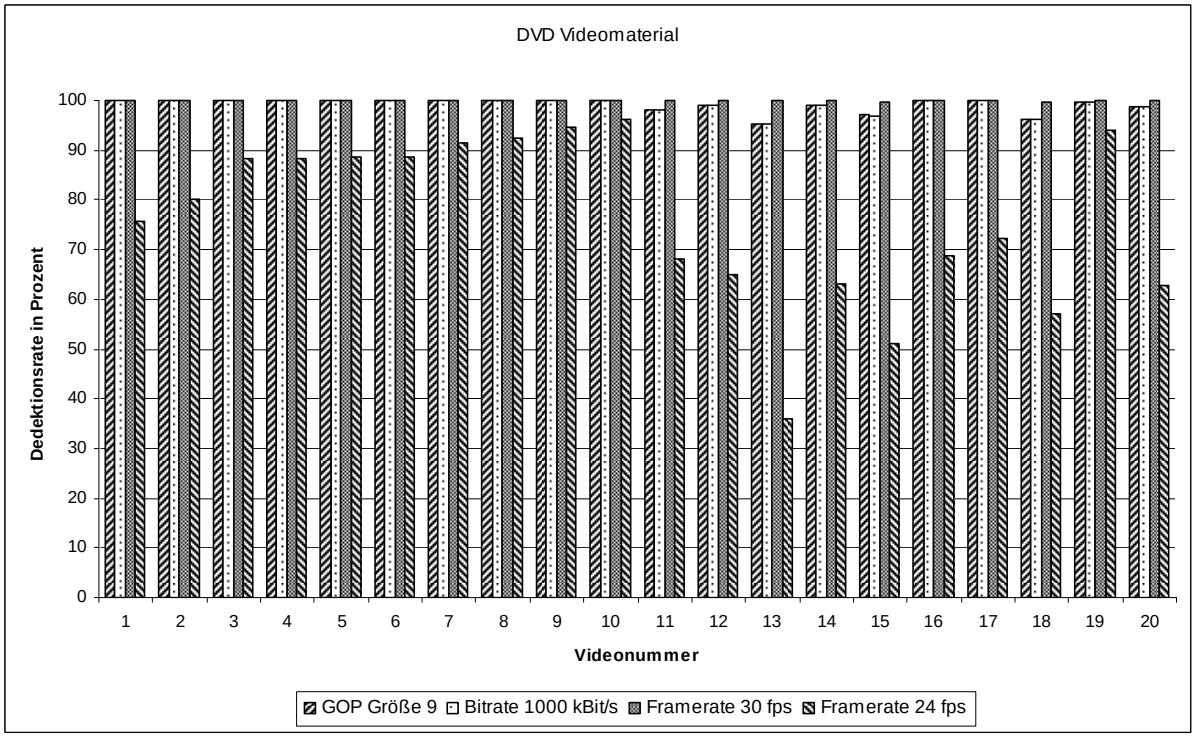

Abbildung 31: Robustheitsanalyse von Videoset 2

6.4.1 Folgerungen und Optimierungsvorschläge

Im Allgemeinen kann man sagen, dass das entwickelte Hashverfahren eine hohe Robustheit bei verschiedenen Modifikationen aufweist. Nur bei Verlangsamung des Videos ergeben sich noch signifikante Schwächen: Durch die Interpolation benachbarter Frames ändert sich der Inhalt im Frame so deutlich, dass die extrahierten Features nicht ausreichend robust dagegen

sind. Trotzdem weist das Verfahren eine ausreichend gute Sensibilität auf. Die Sensibilität verhindert, dass falsche Frames identifiziert werden, was wiederrum zu einer signifikanten Verschlechterung der Detektion des eingebetteten Wasserzeichens führen würde.

Weiterhin wird deutlich, dass das Hashverfahren abhängig von der Videobitrate ist. So ist die RDR beim DVD-Material generell höher als beim SVCD-Material. Die Extraktion eines Merkmal zur Hashbitbestimmung scheint bei einer hohen Bitrate effizienter zu sein. Der Hashwert kann somit genauer und robuster berechnet werden.

Es gibt aber noch einzelne Punkte, an denen Optimierungsbedarf am Verfahren besteht. Dazu zählen die drei folgenden Aspekte:

(a) Rechenaufwand verringern
(b) Übertragung der Hashwerte als Metainformationen
(c) Konzentration auf zentrale Bereiche im Frame

In der augenblicklichen Version ist das Hashverfahren so aufgebaut, dass drei separate Module jeweils separat Eingabedaten benötigen und sie in Ausgangsdaten umwandeln. Da sie unabhängig voneinander aufgebaut sind, werden die temporären Zwischendaten zwischen den Modulen separat abgespeichert. Der durch den Speicheraufwand benötigte Zeitbedarf verzögert die Hashwertberechnung sowie die Identifikation der Frames signifikant. Die Lösung besteht in der Integration der einzelnen Module in ein gesamtes Modul mit der gleichzeitigen Bereitstellung eines Puffers, der die gesamte notwendige Datenmenge der zu betrachtenden Frames beinhaltet. So wären z. B. fünf aufeinanderfolgende Frames als eine Einheit zu betrachten und sie nach dem Dekodieren zuerst nach der Abspielreihenfolge zu ordnen, danach würden die framebasierten Merkmalsmatrizen MB , anschließend die sequenzbasierten Merkmalsmatrizen MV und zum Schluss die Hashwerte berechnet. Während der Hashwertbildung ist dies nur zu Beginn eines weiteren erkannten I-Frames notwendig, während der Frameidentifikation muss der Schritt für sämtliche Frames separat gemacht werden. Dadurch würden keine Zwischendaten gespeichert und die gesamte Berechnung schneller abgeschlossen werden.

Augenblicklich müssen die ermittelten Hashwerte für den Prozess der Frameidentifikation an den Wasserzeichendetektor übermittelt werden. Das ist nur möglich, wenn ein sicherer Übertragungskanal vorhanden ist oder wenn der gesamte Prozess auf einem sicheren Server als Service angeboten wird. Dann muss nur das zu untersuchende Video hochgeladen werden. Eine andere Möglichkeit wäre die Fokussierung auf Hashwerte, die sehr häufig im Videomaterial vorkommen. Wenn es möglich ist, Hashwerte zu erkennen, die oft im Video vorkommen und die bei der Identifikation der markierten I-Frames helfen, dann wäre eine Übertragung der berechneten Hashwerte nicht mehr notwendig. Dabei müssen aber Hashwerte mit einer geringeren Länge als in der augenblicklichen Version festgelegt werden, da die Wahrscheinlichkeit von mehrfach auftretenden Hashwerten mit einer hohen Länge sehr gering ist. Dies kann auf [OoKa2001] zurückgeführt werden, da genau durch die Verlängerung der zu betrachtenden Videosequenz eine Eindeutigkeit erreicht werden soll. Zudem muss garantiert werden, dass die festgelegten Hashwerte jeweils ein I-Frame identifizieren.

Die untersuchte Implementierung unterteilt das jeweilige Frame in gleich große Blöcke vom oberen linken bis zum unteren rechten Teil des Frames auf. Bei der Generierung des

Hashwerts können aber nach verschiedenen Bearbeitungsschritten am Video Probleme auftreten. Besonders relevant wird es bei den verschiedenen Seitenverhältnissen, wie z. B. 16:9, 14:9 oder 4:3. Dabei werden von verschiedenen Bearbeitungsprogrammen die an den Rändern des Frames auftretenden schwarzen Streifen entfernt und das Video neu enkodiert. Eine mögliche robuste Lösung ist die Konzentration auf den zentralen Bereich des Frames. Innerhalb des Framezentrums wird häufig der wichtigste Teil der Aussage des Videos abgebildet. Dabei liegt der Fokus besonders auf dem ersten Modul. Die extrahierten Features bilden die Grundlage für die Generierung der Hashwerte. Sie müssen sich als besonders robust gegenüber den vorher erwähnten Angriffen erweisen. Dabei kann es durchaus sinnvoll sein, für die Extraktion des Hashbits keine gleich großen Flächen festzulegen, sondern sie anhand einer separat durchgeführten Extraktion von sogenannten markanten Punkten zu bilden. Eine bestimmte Menge von markanten Punkte bildet eine Einheit, aus der dann später die Features für die Hashbits generiert werden können. Die markanten Punkte können auch nach Schnitten wiedergefunden werden. Die Extraktion von markanten Punkten bildet auch häufig die Basis von objektbasierten Wasserzeichen. Ein Beispiel unterteilt mit Hilfe der Delaunay-Triangulierung das Bild in Dreiecke und bettet ein Wasserzeichen in die Dreiecke ein [BaCh2001].

6.5 Zusammenfassung

Da die I-Frames die effizienteste Markierungsquelle darstellen, wurde mit Hilfe des robusten Videohashs ein Konzept zur Identifikation der markierten I-Frames entwickelt. Dadurch ist es möglich, auch nach einer Frametypänderung die Identifikation der Frames durchzuführen und das Wasserzeichen erfolgreich auszulesen.

Das Hash wurde modular konzeptioniert, um die einzelnen Module einfach austauschen zu können. Da existierender Algorithmen optimiert oder durch neue Algorithmen ersetzt werden können, ermöglicht das Prinzip des modularen Aufbaus einen einfachen Austausch. Einzige Bedingung ist aber die Einhaltung der Schnittstelle zwischen den Modulen.

Das Konzept wurde an zwei unterschiedlichen Videoklassen getestet: hochqualitative DVD-Videos und niederqualitative SVCD-Videos. Das Hashverfahrens ist bei Änderung der Videobitrate und GoP-Größe sehr robust. Die positiven Ergebnisse verdeutlichen die Anwendbarkeit des Hashs für hoch- und niederqualitative Videos. Nur bei einer Verringerung der Framerate, wie z. B. von NTSC zu PAL, weist der Algorithmus Schwächen auf. Das kann auf Mängel der Framemerkmalsextraktion zurückgeführt werden. Diese Modifikation führt zu einer Interpolation benachbarter Frames und dadurch verändern sich die Framemerkmale deutlich.

7 Robustes MPEG Wasserzeichen

Schwachstellen existierender Wasserzeichen für MPEG-Videos sind zum einen die mangelnde Robustheit gegenüber einer Re-Enkodierung mit einer Veränderung der GoP-Struktur und zum anderen die Tatsache, dass nur einzelne Bilder zum Markieren genutzt werden. Der Grund für diese Mängel liegt in der Markierung intrakodierter Frames und in der fehlenden Markierung interkodierter Frames. Um Schwachstellen dieser Art zu beseitigen, erweist es sich als notwendig, auch interkodierte Frames mit einem Wasserzeichen zu markieren. Da interkodierte Frames je nach Art der MPEG-Kodierung jedoch teilweise gravierende Unterschiede zu intrakodierten Frames aufweisen, müssen digitale Wasserzeichen für intrakodierte Frames an interkodierte Frames entsprechend angepasst werden. Dieses Kapitel befasst sich mit den Anforderungen an Wasserzeichen für das gesamte MPEG-Video, die mit in die Entwicklung des Wasserzeichens einfließen, welches dann im weiteren Verlauf präsentiert wird.

In diesem Kapitel wird zuerst ein grobes Konzept für ein robustes Videowasserzeichen vorgestellt, welches anschließend an MPEG-spezifische Parameter angepasst wird. Der vorgestellte Ansatz bettet das Wasserzeichen sowohl in I- als auch P- und B-Frames ein. Die Evaluation des Ansatz gibt Aufschluss über die Anwendbarkeit des Ansatz für das komplette MPEG-Video.

7.1 Konzept des robusten Wasserzeichens

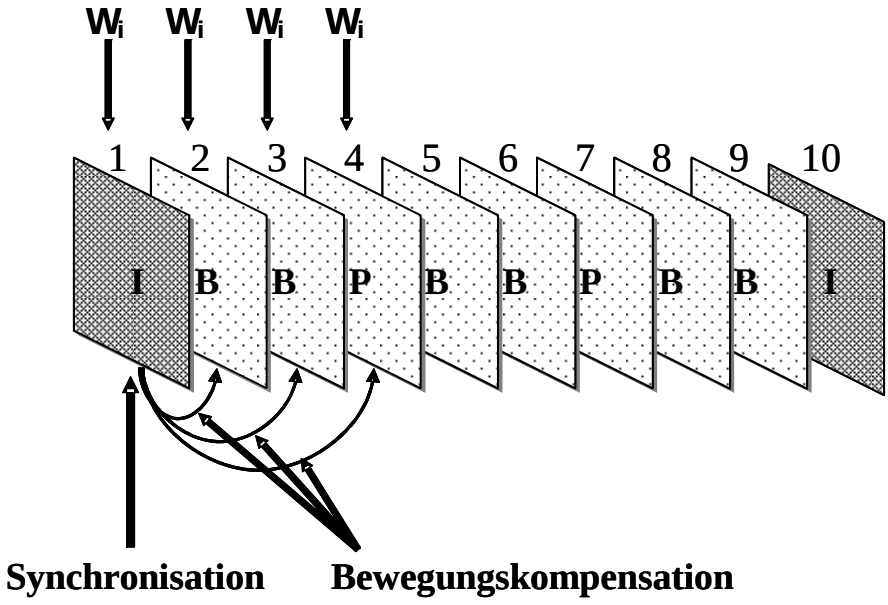

Abbildung 32: Grundkonzept für robuste Wasserzeichen

Im Konzept dieser Arbeit (Abbildung 32) werden intrakodierte Frames weiterhin als die vielversprechendste Möglichkeit zum Einbetten von digitalen Wasserzeichen in MPEG-Videos betrachtet. Aufgrund der Robustheit gegenüber Bildmanipulationen und der visuellen Qualität existierender intrakodierter Wasserzeichen wird hier ebenfalls ein intrakodiertes

Wasserzeichen genutzt, um die Authentizität des Urhebers bzw. Empfängers zu ermöglichen. Die in Abbildung 32 als stark gepunkteten Frames dargestellten intrakodierten Frames dienen also weiterhin als Ausgangsbasis für den gesamten Wasserzeichenalgorithmus. Einzige Aufgabe besteht in der Re-Synchronisation der intrakodierten Frames (Kapitel 6), um die ursprüngliche Framestruktur wiederherstellen zu können. Dadurch wird die Detektion des eingebetteten Wasserzeichen erheblich vereinfacht.

Bei interkodierten Frames existieren bisher noch nicht sehr viele Lösungen [HaSt2005b]. Die Darstellung der interkodierten Frames in Abbildung 32 unterscheidet sich deshalb von den intrakodierten Frames.

Das vorliegende Konzept soll es ermöglichen, neben den vorher markierten intrakodierten Frames auch die übrigen interkodierten Frames mit einem Wasserzeichen zu markieren. Der Wasserzeichenansatz besteht dabei aus drei Hauptschritten:

(1) Einbettung des Wasserzeichens in intrakodierte Frames,
(2) Berechnung und Subtraktion des Bewegungskompensationssignals von den interkodierten Frames und
(3) Einbettung der gleichen Wasserzeicheninformation in die interkodierten Frames.

Schritt 1 ist die Einbettung des Wasserzeichens in das intrakodierte Frame. Das dabei verwendete Wasserzeichen wurde ursprünglich für JPEG-Bilder entwickelt [DiSt1998]. Der JPEG-Standard diente als Basis für die Entwicklung von MPEG-1 und –2. So weisen I-Frames und JPEG-Bilder eine hohe Ähnlichkeit auf. Das Wasserzeichen wird auf die Helligkeitswerte aufgetragen, indem ein pseudo-zufälliges Muster mit der Größe von 8x8 Bildpixel die Helligkeitswerte des I-Frames modifiziert. Wie der Ausgangsalgorithmus das Wasserzeichen einbettet und verifiziert, wird in den Kapiteln 7.4.1.1 und 7.4.1.2 ausführlich beschrieben. Danach folgen noch Optimierungen, um das Ausleseergebnis zu verbessern.

Der nächste Schritt ist die Berechnung des Bewegungskompensationssignals [Ha2000] des zuvor markierten Referenzframes. Das Bewegungskompensationssignal ist das Differenzsignal eines markierten intrakodierten DCT-Blocks vor und nach dem Einbetten des Wasserzeichensignals. Durch das Einbetten des Wasserzeichens verändert sich der Inhalt des DCT-Blockes. Da der Inhalt des Referenzframes in die von ihm abhängigen Bilder übertragen wird, werden die eingebetteten Wasserzeichendaten auch mit übertragen. Das hat zur Folge, dass die übertragenen Wasserzeichendaten auch in den nächsten Videoframes mit abgespielt werden. Abbildung 10 in Kapitel 5.2 zeigt die Abhängigkeiten zwischen den einzelnen Frames in der GoP.

Interkodierte Bilder beinhalten nur die Veränderungen zwischen dem augenblicklichen Bild und seinem Referenzbild. Da die in den Referenzframes eingebetteten Wasserzeichen übertragen werden, sind sie nach einer kompletten Dekodierung der interkodierten Videoframes auch vorhanden. Ohne eine Kompensation des übertragenden Wasserzeichens würde die Einbettung eines eigenständigen Wasserzeichens in die interkodierten Bilder einer Mehrfachmarkierung des Bildes gleichkommen. Da in einer GoP mehrere interkodierte Frames abgespeichert werden, ist dieser Effekt entlang der gesamten GoP zu beobachten. Dadurch ist mit einer deutlichen Verschlechterung der Transparenz des Wasserzeichens zu rechnen. Um diesen Effekt zu kompensieren, muss das übertragene

Bewegungskompensationssignal vom jeweiligen interkodierten Frame subtrahiert werden, bevor das eigentliche Wasserzeichen für das interkodierte Frame eingebettet wird.

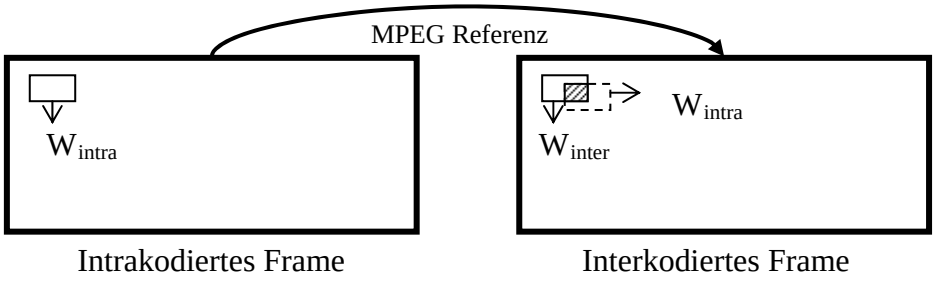

Abbildung 33: Übertragung von Referenzwasserzeichen zu einem interkodiertem Frame

Abbildung 33 demonstriert den Effekt, wenn ein in ein intrakodiertes Frame eingebettetes Wasserzeichen auf ein folgendes interkodiertes Videoframe übertragen wird. Das intrakodierte Frame ist das Referenzframe für die folgenden interkodierten Frames. Es ist mit dem Wasserzeichen W_{Intra} markiert, in Abbildung 33 dargestellt im linken Videoframe. Das in das Referenzframe eingebettete Wasserzeichen W_{Intra} wird in das nachfolgende interkodierte Frame übertragen. Ohne Veränderungen im interkodierten Frame ist das übertragene Wasserzeichen W_{Intra} nicht sichtbar. Wenn aber ein zusätzliches Wasserzeichen W_{Inter} in das interkodierte Wasserzeichen eingebettet wird, muss mit möglichen Überlagerungen gerechnet werden (in der Abbildung geriffelt dargestellt), welche folglich zu visuellen Einbußen führen.

Als Alternative könnten die Wasserzeichen an unterschiedlichen Positionen zwischen den benachbarten Videoframes eingebettet werden. Dadurch würden sich die Wasserzeichenmuster nicht überlagern, jedoch müsste der Wasserzeichendetektor dann auch Kenntnis über die Framereihenfolge haben, ansonsten würde das Wasserzeichen an den falschen Positionen ausgelesen. Deshalb wird es als wichtiger angesehen, die gleichen Markierungspositionen in der gesamten GoP zu benutzen [SuKu2002]. Durch die Subtraktion des Bewegungskompensationssignals wird der Effekt von sich überlagernden Wasserzeichen verhindert.

Der dritte Schritt ist das Einbetten des Wasserzeichens in das interkodierte Frame. Basierend auf den Möglichkeiten der Übertragung der Videodaten zwischen Videoframes kann davon ausgegangen werden, dass die Änderungen des Inhaltes innerhalb einer GoP meistens sehr gering sind. Deshalb erweist es sich als sinnvoll, die gleichen Wasserzeicheninformationen innerhalb der GoP einzubetten. Innerhalb der GoP sind die Veränderungen nicht sehr signifikant, abgesehen vom einen Schnitt im Video. Eine Vielzahl von MPEG Enkodern erkennen aber Schnitte im Video selbstständig und starten mit dem Beginn der neuen Szene die Enkodierung einer neuen GoP. Möglichkeiten der einfachen Schnitterkennung sind z. B. die Berechnungen von Histogrammen, die sich während eines Schnitt zwischen zwei Nachbarbildern signifikant unterscheiden.

7.2 Anforderungen an robuste MPEG-Wasserzeichen

Bevor das digitale Wasserzeichens auch in das komplette MPEG-Videos eingebettet wird, müssen zunächst die relevanten Anforderungen diskutiert werden, die an die Entwicklung eines robusten Wasserzeichens für sämtliche Frames eines MPEG-Videos gestellt werden. Der Fokus liegt dabei auf den Aspekten der MPEG-Quantisierung, der Bewegungskompensation und der geringen Kapazität in interkodierten Frames.

7.2.1 Anforderungen an die Robustheit

Da die Konzentration der Arbeit auf MPEG-Videodaten liegt, ist der Aufbau einer Robustheit des Wasserzeichens gegenüber der MPEG-Datenkompression die primäre Aufgabe zur Bereitstellung eines robusten Wasserzeichens. Dabei muss auf die Unterschiede bei der Kodierung zwischen intra- und interkodierten Frametypen geachtet werden.

Für beide Frametypen werden zwar die gleichen Quantisierungsformeln angewendet, jedoch können die eingesetzten Werte deutlich unterschiedlich sein. Der DC-Wert der 8x8-Koeffizientenmatrix wird bei intrakodierten Blöcken gesondert quantisiert. Da er Aussage über die gleichverteilte Frequenz im Block macht, haben Veränderungen an ihm Auswirkungen auf den gesamten Block. Deshalb wird er auch von einer Vielzahl von Wasserzeichen gesondert betrachtet oder nicht markiert [DiHa2001a]. Bei interkodierten Blöcken wird der DC-Wert hingegen nicht gesondert betrachtet, da kein vollwertiger DCT-Block vorliegt, sondern nur Differenzdaten.

Die Quantisierungsformel ist wie folgt aufgebaut:

$$F[u][v] = \left((2 \times QF[u][v] + k) \times W[w][u][v] \times quantizer\_scale\right) / 32 \qquad (11)$$

$F[u][v]$	DCT-Koeffizient
$QF[u][v]$	quantisierter DCT-Koeffizient
$W[w][u][v]$	Wert der Quantisierungsmatrix
$k \begin{cases} 0 \\ Sign(QF[u][v]) \end{cases}$	intra kodierter DCT Block inter kodierter DCT Block

Die Formel (11) wird bei intrakodierten Blöcken für die AC-Koeffizienten und bei interkodierten Blöcken für den gesamten DCT-Block benutzt. Für intra- und interkodierte DCT-Blöcke werden zwei auf die Quantisierung abgestimmte Quantisierungsmatrizen *QF* verwendet. Es besteht aber die Möglichkeit, eine frei gewählte Quantisierungsmatrix in das Video zu integrieren. Ein Großteil von MPEG-Enkodern verwendet aber festgelegte Standardmatrizen. Der Skalierungsfaktor („quantizer_scale“) regelt den Schutz vor einem Über- oder Unterlauf des Videopuffers, wodurch es zu Störungen der Synchronität zwischen Video und Audio kommen könnte. Bei niedrigen Bitraten werden höhere Werte angewandt, die wiederum zu einer stärkeren Quantisierung führen.

Durch die variable Quantisierung müssen bildbasierte Wasserzeichen an MPEG-Videos angepasst werden. Während ein JPEG-Bild mit gleichmäßiger Stärke quantisiert wird, kann sich die Quantisierungsstärke in einem MPEG-Video ändern. Eine Lösung ist die Koppelung

der Wasserzeichenstärke an den Skalierungsfaktor. So sollte die Wasserzeichenstärke bei einem hohen Skalierungsfaktor zusätzlich erhöht werden.

7.2.2 Anforderungen an die Transparenz

Die Herausforderung in Bezug auf die Transparenz ist der Grad der Bewegung zwischen den Nachbarframes. Der Grad der Bewegung des Aufnahmegerätes kann zwischen Stillstand und schneller Bewegung liegen. Bei Stillstand der Kamera beinhalten entlang des zeitlichen Verlaufs einzelne Bereiche in einem Frame die gleichen Daten wie die Nachbarbilder. Bei einem Schwenk bewegt sich die Kamera gleichmäßig, wodurch zwischen den Nachbarbildern eine konstante und über das gesamte Frame gleichmäßige Bewegung vorliegt. Komprimierte Videostandards analysieren die Bewegung zwischen den Bildern und speichern sie in Bewegungsvektoren, um so eine erneute Speicherung der Daten zu verhindern.

Abbildung 34 zeigt zwei benachbarte Bilder, bezeichnet als Bild 1 und Bild 2 eines Beispielvideos. Das Video ist durch eine gleichmäßige Bewegung gekennzeichnet. Zwischen den beiden Bildern 1 und 2 ist auf den ersten Blick kein Unterschied zu erkennen. Im dritten Bild wird wiederum Bild 2 dargestellt, jedoch mit der Visualisierung der Bewegungsvektoren. Das vierte Bild ist das Differenzbild nach der Kompensation der Bewegung. Hier ist deutlich zu erkennen, dass zwischen den Bildern 1 und 2 einzelne Bereiche deutlichere Unterschiede aufweisen, als andere Bereiche. An diesem Beispiel ist auch zu erkennen, dass in den Bereichen mit stärkeren Bewegungsvektoren auch ein stärker ausgeprägtes Differenzbild vorhanden ist.

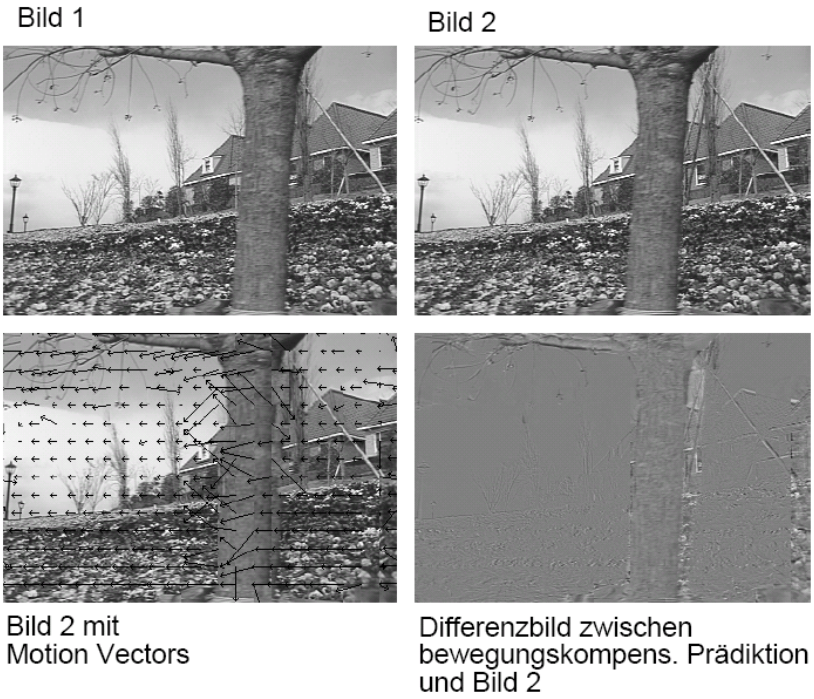

Abbildung 34: Beispielbewegung zwischen zwei Nachbarframes in einem MPEG-Video

Abbildung 34 deutet die Auswirkungen auf einzubettende Wasserzeichen an, wenn der Inhalt in Abhängigkeit zu anderen zeitlich versetzten Inhalten steht. Da der Inhalt einzelner Frames in andere Frames übertragen wird, werden auch Wasserzeichen übertragen. Dieser Aspekt

muss bei der Entwicklung einer Wasserzeichenlösung für alle Frame beachtet werden. Da zwischen den Bildern häufig nur minimale Unterschieden vorliegen, dürfen auch die in die Frames einzubettenden jeweiligen Wasserzeichen keine großen Unterschiede aufweisen. Unterschiedliche Wasserzeichen in benachbarten Bildern können zu visuellen Artefakten führen [SeLa2001].

7.2.3 Anforderungen an die Kapazität

Der MPEG-Videostandard beeinträchtigt auch die Kapazität des Wasserzeichens. Speziell interkodierte Frames werden mit verschiedenen Methoden kodiert, um Platz zu sparen. So wird innerhalb des Frames stärker komprimiert und es wird zwischen mehreren Frames nach identischen Inhalt gesucht und dieser nur in einem Frame abgespeichert, wobei die restlichen Frames dann auf ihn verweisen.

Sind zudem die Daten zwischen einem interkodierten Block und seinem Referenzblock identisch, so wird in dem Block nur der Bewegungsvektor abgespeichert. In diesem Block liegen keine Daten vor, die möglicherweise mit einem Wasserzeichen versehen werden könnten. Je identischer der Inhalt von Nachbarframes ist, desto höher ist die Anzahl von Blöcken, die auf diese Weise aufgebaut sind. Dadurch wird die Kapazität des Wasserzeichens stark eingeschränkt. Wenn die Möglichkeit besteht, muss der Aufbau dieser Blöcke so verändert werden, dass ein Wasserzeichen eingebettet werden kann. Falls es nicht möglich ist, müssen die Blöcke als nicht markiert eingestuft werden.

Es ist deutlich geworden, dass bei der Entwicklung des Wasserzeichens sowohl die Aspekte der Robustheit, der Transparenz als auch der Kapazität berücksichtigt werden müssen. Dabei sind besonders die Aspekte der MPEG-Kompression, des Einflusses zeitlich benachbarter Bildern und die eingeschränkte Kapazität bei interkodierten Frames zu beachten und mit in die Entwickelung des Wasserzeichens einzufließen.

7.3 Existierende Ansätze

Andere Veröffentlichungen präsentieren bereits verschiedene Strategien. So wird in [KaDe1999] das Video komplett in die Bildsequenz zerlegt und ein niederfrequentes Muster in die Bilder eingebettet. Ein zusätzlich lokaler Skalierungsfaktor des verwendeten visuellen Modells regelt die Transparenz des Wasserzeichens. Er misst lokale Aktivitäten im Videoframe und ordnet Regionen mit geringer Aktivität (gleichmäßige Farbverteilung) einer geringen Wasserzeichenstärke und Regionen mit hoher Aktivität (Texturen und Ecken) einer hohen Wasserzeichenstärke zu. Experimente haben gezeigt, dass die Verwendung des Laplace-high-pass-Filter sehr gut für die Bestimmung des Skalierungsfaktor ist. Die verwendete Korrelationsmethode ist die SPOMF Methode (Symmetrical Phase Only Filtering Method) [Br1992]. Die Wasserzeichen konnten erfolgreich aus sämtlichen Videoframes ausgelesen werden.

In [SeLa2001] wird eine erweiterte Methode des DEW (Differential Energy Watermarking) Algorithmus vorgestellt. Der DEW-Algorithmus bettet das Wasserzeichen in das Videoframe ein, indem es eine gezielte Manipulation des Energieverhältnisses von zwei verschiedenen

Regionen durchführt. Ursprünglich wurde die Methode für intrakodierte Videoframes entwickelt, da sie ausreichend Inhalt bzw. Energie beinhalten, um das Verhältnis für das Wasserzeichen herzustellen. In [SeLa2001] wird die Methode auf P-Frames erweitert. Innerhalb des P-Frames gibt die Varianz des Frames an, wie viele Blöcke für ein Wasserzeichenbit benötigt werden. Wenn z. B. die Varianz innerhalb eines intrakodierten Frames 50 mal höher ist, als die im P-Frame, dann werden im P-Frame 50 mal so viele Blöcke zum Einbetten des Wasserzeichens benötigt. Für B-Frames kann die Methode nicht angewandt werden, da B-Frames fast immer zu wenig Energie haben, um ein Verhältnis aufzubauen.

In [SuKu2002] werden vor dem Einbetten per Voranalyse Bereiche selektiert, die sich gut zum Einbetten des Wasserzeichens eignen. Die Voranalyse soll keine effiziente Framekollision ermöglichen. Eine Framekollision tritt immer dann auf, wenn durch das Einbetten von Wasserzeichen in benachbarte Frames die Unterschiede sich soweit verstärken, dass durch einen Vergleich der Frames die Wasserzeichen erkannt werden können. Das Wasserzeichen ist ein DFT Muster, das in die Regionen räumlich verteilt wird.

Eine Bewertung zeigt, dass die drei vorgestellten Methoden nur bedingt nutzbar sind. Die Methode in [KaDe1999] ist nur geeignet, wenn das Wasserzeichen keine Echtzeitansprüche haben muss. Bei verschiedenen Szenarien, wie Online-Shops darf aber während des Downloadvorganges keine deutliche Verzögerung durch das Wasserzeichen stattfinden. [SeLa2001] ist auch nur bedingt anwendbar, da nur I- und P-Frames markiert werden können. Die Überlegungen in [SuKu2002] sollten besonders dann herangezogen werden, wenn durch die Erweiterung der Einbettungsmethode mögliche Framekollisionen verhindert werden sollen.

7.4 MPEG Wasserzeichen für Luminanzdaten

Das Wasserzeichen basiert auf einem Markierungskonzept für JPEG-Bilder [DiSt1998]. Im folgenden Kapitel werden zuerst der Einbettungs- und Detektionsprozess vorgestellt und darauf die Optimierungen für intrakodierte Frames sowie das Einbetten und Detektieren in interkodierten Frames.

7.4.1 Luminanz-Wasserzeichen für intrakodierte Frames

Das Wasserzeichen, das in die intrakodierten Frames eingebettet werden soll, bildet die Basis für die Markierung des gesamten Videos. Die ursprüngliche Idee für das Wasserzeichen stammt von einem Bildwasserzeichen für JPEG-Bilder ab [DiSt1998].

7.4.1.1 Einbettungsprozess

Das zuerst in [DiSt1998] vorgestellte Wasserzeichen ist ein Luminanz-Wasserzeichen, das ein pseudo-zufällig generiertes Muster auf 8x8-Pixelblöcken aufträgt. Dabei ist es auch möglich, die Größe des zu markierenden Blocks im Frame um ein Vielfaches von acht zu

vergrößern. Wichtig ist dabei nur, dass garantiert werden kann, dass das Wasserzeichenmuster normalverteilt ist.

Der erste Schritt ist die Bildung einer Sequenz S von Markierungspositionen im Frame. Die Bildung der Sequenz erfolgt pseudo-zufällig auf Basis des geheimen Schlüssels K. Neben der Sequenz S wird auf Basis des geheimen Schlüssels für jeden zu markierenden Block ein pseudo-zufällig generiertes hochfrequentes Muster M generiert.

Für die Bildung des niederfrequenten Musters werden in fünf Schritten die Einzelwerte des Musters berechnet. Während des ersten Schrittes wird ein pseudo-zufälliges normalverteiltes Rauschmuster erstellt. Die anschließenden beiden Schritte führen eine Glättung des Musters durch. Dabei werden die Werte des Musters in Abhängigkeit zu ihren Nachbarwerten $N_{x,y}$ berechnet. Im darauf folgenden Schritt wird die Farbtiefe des Musters erhöht. Dies ist notwendig, um eine Robustheit gegenüber der MPEG-Quantisierung zu erreichen; nur so kann es auch nach der Enkodierung des Videos weiter ausgelsen werden. Da aber mit einer verstärkten Sichtbarkeit des Musters zu rechnen ist, werden im letzten Schritt die Werte des Musters $\overline{M}$ durch den Durchschnitt der Nachbarwerte $N_{x,y}$ des augenblicklichen Musterwerts erneut geglättet.

Nach der Mustergenerierung wird das Muster in den jeweiligen Luminanzblock B_i aus der Blocksequenz S eingebettet. Basierend auf dem Wert des jeweiligen Informationsbits W_i wird das Wasserzeichenmuster wie folgt eingebettet:

$$\overline{B_i} = \begin{cases} B_i - s \bullet \overline{M_{x,y}}, wenn\ W_i = 0 \\ B_i + s \bullet \overline{M_{x,y}}, wenn\ W_i = 1 \end{cases} \qquad (12)$$

Der Wert s ist die Wasserzeichenstärke s, welche die Stärke des einzubettenden Wasserzeichenmusters regelt. Um eine akzeptable Wasserzeichenstärke zu erreichen, kann die Wasserzeichenstärke auch mit Hilfe eines visuellen Modells berechnet werden.

7.4.1.2 Detektionsprozess

Der Detektionsprozess kann im Allgemeinen als eine Korrelation zwischen dem Originalmuster und dem ausgelesenen markierten Block beschrieben werden.

Der erste Schritt ist auch hier wieder die Bildung der pseudo-zufälligen Sequenz S der Markierungspositionen im Frame auf Basis des geheimen Schlüssels K.

Der nächste Schritt zur Bildung der Korrelation besteht in der Berechnung zweier Mittelwerte. Der erste ist der Mittelwert M_{pos} der Luminanzwerte im Block, die einem positiven Musterwert zugeordnet werden, der zweite der Mittelwert M_{neg} derer Luminanzwerte, die einem negativen Musterwert zugeordnet sind. Die Korrelation ist ein Vergleich der beiden Mittelwerte. Das Verhältnis entscheidet über das Wasserzeichenbit:

$$W_i = \begin{cases} 0, \textit{ wenn } M_{pos} < M_{neg} \\ 1, \textit{ wenn } M_{pos} \geq M_{neg} \end{cases} \qquad (13)$$

7.4.1.3 Optimierungen für Einbettung und Detektion

Die Einbettung und Detektion sind bei [DiSt1998] nur erfolgreich, wenn der Inhalt des zu markierenden Blocks auch normalverteilt ist. Ist er hingegen nicht so aufgebaut, muss eventuell eine Optimierung des Einbettungsprozesses durchgeführt werden oder während der Detektion die markierten Blöcke vorselektiert werden. Die Optimierung des Einbettungsprozesses kann nur dann durchgeführt werden, wenn andere Eigenschaften, wie die Transparenz oder Sicherheit, gewahrt werden.

Während der Einbettung des Wasserzeichens müssen die folgenden zwei Punkte beachtet werden:

(1) Einbettung des Wasserzeichen in größere Markierungsblöcke
(2) Kontrolle und möglicherweise Nachbearbeitung des Markierungsblocks

Ursprünglich wird das Wasserzeichenmuster auf einen 8x8-Bildpixelblock aufgetragen. Um aber das Ausleseergebnis sicherer zu gestalten, wird das Wasserzeichen in einen 16x16-Pixelblock eingebettet. Die in Kapitel 5.1 durchgeführte Evaluierung befasste sich auch mit diesem Verfahren. Es stellte sich heraus, dass öfters das falsche Ergebnis ausgelesen wurde. Einzelne Pixel in dem 8x8-Pixelblock verfälschten das Ausleseergebnis signifikant. Um die Fehlerquote zu senken, wird der Einbettungsbereich vergrößert. Das für das Verfahren verwendete visuelle Modell [Di2000] wird auch angepasst. Es untersucht den Verlauf von Ecken und Kanten und misst die Verteilung von gleichmäßigen Flächen im 8x8 Pixelblock. Da der Markierungsbereich jetzt die vierfache Größe hat, berechnet das visuelle Modell den Durchschnitt der vier einzelnen 8x8-Pixelblöcke.

Zusätzlich wird unmittelbar nach dem Einbetten des Wasserzeichens eine Überprüfung durchgeführt. Hierbei wird kontrolliert, ob das richtige Bit ausgelesen wird. Ziel ist es dabei,

durch die nachfolgende Kontrolle ein mögliches falsches Ergebnis zu korrigieren. Zudem muss mit einer Verfälschung des Ausleseergebnisses gerechnet werden, wenn an dem markierten Video verschiedene Manipulationen durchgeführt werden. Fast jede Manipulation an Videos führt zu einer abschließenden Re-Enkodierung. Während der Untersuchung von 32000 Testblöcken in einem MPEG-Video zeigte sich, dass sich nach einer Re-Enkodierung des Videos mit einer gleichzeitigen Verringerung der Bitrate von 1,2 auf 1 MBit/s beide Mittelwerte angenähert haben. Das kann damit begründet werden, dass durch die erneute Durchführung der verlustbehafteten MPEG-Kompression Blockinhalt verloren geht. Wenn während des Einbettens das Verhältnis zwischen den beiden Mittelwerten nicht überprüft und nachgebessert wird, können Fehler während der Detektion auftreten und sich das Ausleseergebnis verschlechtern. Das Ziel ist der Ausschluss falscher Ausleseergbnisse.

Die Korrelation ist die Differenz von M_{Pos} und M_{Neg}. Sie wird nur nachgebessert, wenn sie unter dem Grenzwertes T liegt. Es wird also genau dann eine Nachbearbeitung des Blocks durchgeführt, wenn Folgendes gilt:

$$\begin{aligned} \left|M_{Pos}\right| - \left|M_{Neg}\right| < \left|T\right| \quad & Nachbearbeitung\ des\ Blockes \\ \left|M_{Pos}\right| - \left|M_{Neg}\right| > \left|T\right| \quad & keine\ Nachbearbeitung\ des\ Blockes \end{aligned} \tag{14}$$

Abbildung 35 zeigt einen Beispielblock; hier wird bei dem Versuch, ein Wasserzeichenbit eins einzubetten, trotzdem das verkehrte Bit ausgelesen. Während der Nachbearbeitung werden die Luminanzwerte mit einem grünen Pfeil erhöht und die mit einem blauen Pfeil verringert.

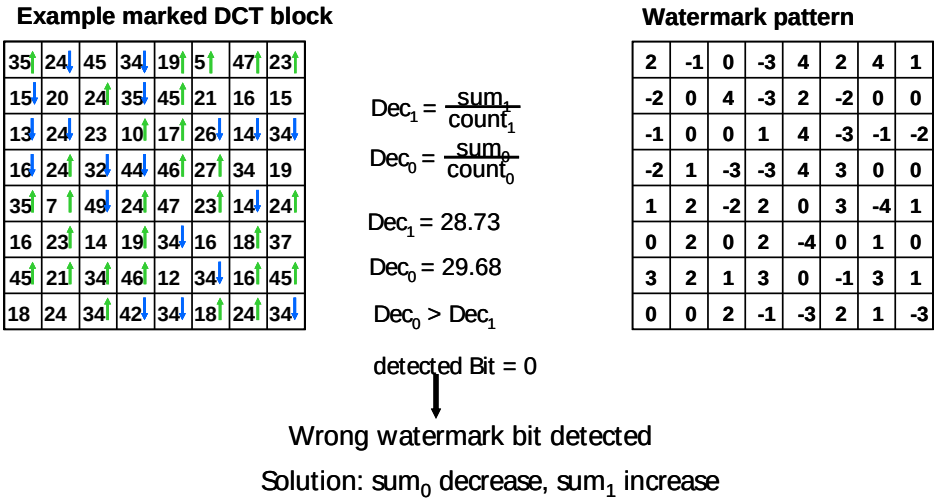

Abbildung 35: Nachbearbeitung des DCT-Blocks

Während der Untersuchungen hat sich erwiesen, dass eine Nachbearbeitung des DCT-Blocks genau dann notwendig ist, wenn die Korrelation nahe bei null liegt. Deshalb wird der Threshold für die Nachbearbeitung auf 1 festgelegt; liegt die Korrelation also zwischen -1 und 1, so wird der Block nachgearbeitet. Dabei wird auf jeden Wert des Markierungsblocks B_i ein zusätzlicher Wert t addiert bzw. von ihm subtrahiert, wenn der dazugehörige Wasserzeichenmusterwert P_i und das Wasserzeichenbit W_i die folgenden Bedingungen erfüllen:

$$\overline{\overline{B_i}} = \begin{cases} \overline{B_i} + t, wenn \;\; P_i < 0 \wedge W_i = 0 \\ \overline{B_i} - t, wenn \;\; P_i > 0 \wedge W_i = 0 \\ \overline{B_i} - t, wenn \;\; P_i < 0 \wedge W_i = 1 \\ \overline{B_i} + t, wenn \;\; P_i > 0 \wedge W_i = 1 \end{cases} \qquad (15)$$

Während der Detektion des Wasserzeichens werden folgende Optimierungsstrategien angewandt:

(1) Varianz des markierten Blocks eingrenzen
(2) Auswahl der Medianwerte aus den Korrelationsergebnissen pro Wasserzeichenbit

Die Eingrenzung der Varianz des Markierungsblocks ist aufgrund des Aufbaus des Wasserzeichenmusters notwendig. Das Wasserzeichenmuster ist ein gleichmäßig verteiltes Muster. Um ein sicheres Korrelationsergebnis zu erreichen, sollte auch der Markierungsblock gleichmäßig verteilt sein. Das Ausleseergebnis kann besonders dann verfälscht werden, wenn einzelne Werte vom gesamten restlichen Blockinhalt abweichen. Deshalb werden bei der Detektion Markierungsblöcke nicht betrachtet, die eine zu hohe Varianz haben.

Die in [DiSt1998] vorgestellte Lösung erkennt nicht, ob ein Wasserzeichen eingebettet wurde oder ob das Video nicht markiert ist. Ein Vergleich der FAR (False-Acceptance-Rate) mit der FRR (False-Rejection-Rate) in Abbildung 36 zeigt, dass selbst beim idealen Schnittpunkt noch immer eine FAR von ungefähr 15% vorliegt. Abbildung 36 zeigt die Verteilung der Korrelation bei DVD-Videos, die mit dem Bit 0 markiert wurden (rote Linie) und die Korrelation bei nicht markierten Videos (grüne Linie). So stufen 15% der ausgelesenen Wasserzeichensignale ein nicht markiertes Videos als markiert ein und wiederum 15% ein markiertes Video als nicht markiert. Dadurch wird deutlich, dass sich beide Signale nicht deutlich genug unterscheiden, um eine sichere Entscheidung zwischen unmarkiert und markiert treffen zu können.

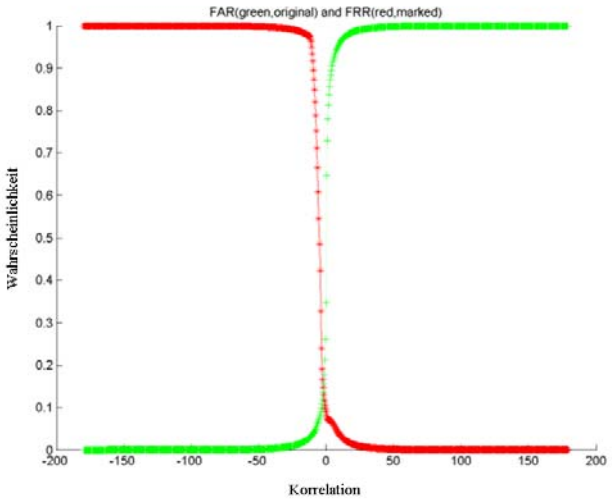

Abbildung 36: FAR-FFR Gegenüberstellung von [DiSt1998]

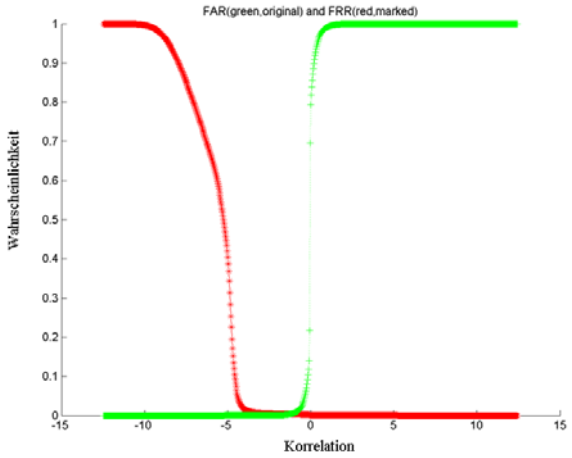

Abbildung 37: FAR-FFR Gegenüberstellung Median pro Bit

Da die Fehlerraten zu hoch sind, wird die Detektionsmethode geändert. Die geänderte Methode sucht den Median sämtlicher Korrelationsergebnisse pro Bit. Dadurch wird die Verteilung der Auslesergebnisse eingegrenzt, da die Streuung der Korrelationsergebnisse nicht mehr relevant ist. Abbildung 37 zeigt bei den gleichen Videos wie in Abbildung 36 die Gegenüberstellung von FAR und FFR, wenn der Medianwert als Entscheidungskriterium für das Wasserzeichenbit genutzt wird. Das Ergebnis führt zu geringeren Fehlerraten am Schnittpunkt beider Graphen. Basierend auf den Ergebnissen wird innerhalb des Wertebereichs von [-1, +1] das Video als nicht markiert eingestuft. Die ausgelesenen Medianwerten aus nicht markierten Videos befinden sich zu 99% innerhalb des Wertebereichs. Deshalb liegt die FAR Fehlerrate bei 1%.

7.4.2 Luminanzwasserzeichen in interkodierte Frames

Da die Änderungen zwischen den Frames innerhalb einer GoP meistens sehr gering sind, erweist als sich als ratsam, das Wasserzeichen, das in das intrakodierte Frame eingebettet

wurde, auch in die nachfolgenden Frames der GoP einzubetten. Durch die Erweiterung auf interkodierte Frames kann das Wasserzeichen aus dem gesamten Video ausgelesen werden. Mit dieser Strategie soll es möglich werden, die mangelhafte Robustheit von I-Frame basierten Wasserzeichenlösungen zu verbessern. Speziell die Robustheit gegenüber der Re-Enkodierung kann verbessert werden, da das Wasserzeichen aus der kompletten GoP ausgelesen werden kann.

Die Einbettung des Wasserzeichens in interkodierte Frames setzt sich dabei aus zwei Schritten zusammen:

(1) Subtraktion des Bewegungskompensationssignals
(2) Einbetten des Wasserzeichens

7.4.2.1 Berechnung des Bewegungskompensationssignals

Interkodierte Frames beinhalten DCT-Blöcke, die aus Bewegungsvektoren und Differenzdaten zu ihren Referenzframes bestehen. Die Menge der Daten pro Block hängt von der Bewegung zwischen dem Frame und seinem Referenzframe ab. Dadurch entstehen verschiedene Effekte während der Detektion des Wasserzeichens:

(1) Keine Bewegung und Differenzdaten zwischen dem interkodierten Frame und seinem Referenzframe
In diesem Fall werden die kompletten DCT-Blockdaten vom Referenzframe zum interkodierten Frame übertragen. Dadurch kann das transferierte Wasserzeichen an der gleichen Position im interkodierten Frame detektiert werden.

(2) Keine Bewegung aber Differenzdaten zwischen dem interkodierten Frame und seinem Referenzframe
Auch in diesem Fall wird der DCT-Blockinhalt des Referenzframes an die gleiche Position im interkodierten Frame transferiert. Die Detektion des transferierten Wasserzeichens kann aber im interkodierten Frame verfälscht werden, da die Differenzdaten mit in die Detektion einfließen. Die zusätzlichen Daten können die Detektion des Wasserzeichens erschweren.

(3) Bewegung zwischen dem interkodierten Frame und seinem Referenzframe
Im interkodierten Frame ist das Wasserzeichen nur detektierbar, wenn die Bewegung berücksichtigt wird. Werden aber wie im Referenzframe die gleichen Markierungspositionen benutzt, wird nur ein Teil des transferierten Wasserzeichensignals ausgelesen.

Aufgrund dieser Möglichkeiten erweist es sich als sinnvoll, das von den Referenzframes transferierte Wasserzeichensignal zuerst zu entfernen, bevor ein Wasserzeichen in das interkodierte Videoframe eingebettet werden kann. Das zu entfernende Wasserzeichensignal ist die Differenz zwischen dem Referenzframe im markierten Zustand $F(i,j)_{\mathrm{Re}f\_mark}$ und dem Originalzustand $F(i,j)_{\mathrm{Re}f\_orig}$.

$$F(i,j)_{,comp} = F(i,j)_{\mathrm{Re}f\_mark} - F(i,j)_{\mathrm{Re}f\_orig} \qquad (16)$$

Während der Durchführung der Bewegungskompensation kann es möglich sein, dass nach der Subtraktion des Referenzwasserzeichensignals eine Nachbearbeitung des markierten Blocks im Frame notwendig ist und Schema und Typ des Macroblockes neu erstellt werden müssen.

7.4.2.2 Einbetten des Wasserzeichens

In die interkodierten Frames wird das gleiche Wasserzeichen wie in den davor liegenden intrakodierten Frames eingebettet. Nur vereinzelte Verfahrensparameter, wie Wasserzeichenstärke, werden angepasst. Formel 17 stellt den Einbettungsprozess dar. Dabei ist B_i der augenblicklich zu markierende Block und M_I das zuvor gespeicherte Wasserzeichenmuster vom intrakodierten Frame I . W_i ist das augenblickliche Wasserzeichenbit.

$$\overline{B_i} = \begin{cases} B_i - s \bullet \overline{M_I}, wenn\ \ W_i = 0 \\ B_i + s \bullet \overline{M_I}, wenn\ \ W_i = 1 \end{cases} \qquad (17)$$

Zur Bestimmung der Wasserzeichenstärke und der Nachuntersuchung des markierten Blocks erweist es sich als notwendig, den DCT-Block im interkodierten Frame komplett zu dekodieren und den Referenzblock aus dem Referenzframe aufzuaddieren. Die Daten des jeweiligen Referenzblocks $B_{\mathrm{Re}\,f}$ werden auf die Blockdaten B_i addiert, bevor die Wasserzeichenstärke berechnet wird. Der zu markierende Block wird ein vollständiger intrakodierter Block und die Einbettungsmethode aus den I-Frames kann vollständig angewandt werden. Dadurch ändert sich der Einbettungsprozess wie folgt:

$$\overline{B_i} = \begin{cases} B_i + B_{\mathrm{Re}\,f} - s \bullet \overline{M_I}, wenn\ \ W_i = 0 \\ B_i + B_{\mathrm{Re}\,f} + s \bullet \overline{M_I}, wenn\ \ W_i = 1 \end{cases} \qquad (18)$$

Nach dem Einbetten des Wasserzeichens wird der Referenzblock wieder abgezogen und eine entsprechende Kontrolle der Blockstruktur durchgeführt. Dabei müssen wie bei der Bewegungskompensation Typ und Schema des Makroblocks überprüft werden, damit auch nach dem Einbetten des Wasserzeichens das Video MPEG-konform bleibt.

Der Detektionsvorgang in P-Frames wird identisch zu einem I-Frame aufgebaut. Dafür wird das Video vollständig zu einem Motion-JPEG-Video dekodiert. Jedes Frame liegt bei diesem Videotyp als I-Frame vor.

7.5 Evaluation der Testergebnisse

Die Untersuchung der Qualität des Wasserzeichens untergliedert sich in zwei Teile. Der erste Teil ist die Analyse der Robustheit des Wasserzeichens für I-Frames. Im zweiten Teil wird die Erweiterung für P-Frames analysiert. Dieser Teil befasst sich mit der Detektionsrate der Wasserzeichen in interkodierten Frames. Die Erfolgsquote der Detektion soll Aufschluss darüber geben, unter welchen Bedingungen das Wasserzeichen anwendbar für die Erweiterung auf B-Frames ist.

7.5.1 Testsuite

Als Testdaten dienen je zwei Videosets als Ausgangsbasis. Dabei werden zwei Qualitätsstufen untersucht, in Super-VCD- und in DVD-Qualität. Dadurch soll es möglich sein, die Robustheit des Verfahrens bei Videos mit hoher und geringer Qualität zu untersuchen. Beide Sets haben jeweilige Parameter:

(1) Videoset 1: 60 SVCD-Videos
   (a) Bitrate 1500 kBit/s
   (b) Framerate 25 fps
   (c) GoP 12 Frames pro GoP
   (d) Super Video CD Format
(2) Videoset 2: 40 DVD-Videos
   (a) Bitrate 4000 kBit/s
   (b) Framerate 25 fps
   (c) GoP 12 Frames pro GoP
   (d) DVD Format

7.5.2 Testergebnisse und Evaluation für I-Frames

Das Wasserzeichen für intrakodierte Frames wird einer Reihenfolge von Videomodifikationen unterzogen, danach wird die Detektionsrate ermittelt. Zuerst werden die Videos mit unterschiedlichen Bitraten erneut enkodiert. Während des Tests wurde die Right-Detection-Rate (RDR) und False-Detection-Rate (FDR) für jede untersuchte Bitrate berechnet. Die Right-Detection-Rate gibt den prozentualen Anteil aller korrekt ermittelten Bits gegenüber der gesamten Menge von eingebetteten Bits an. Die False-Detection-Rate ist das Verhältnis der falsch ausgelesenen Bits zur Gesamtbitmenge.

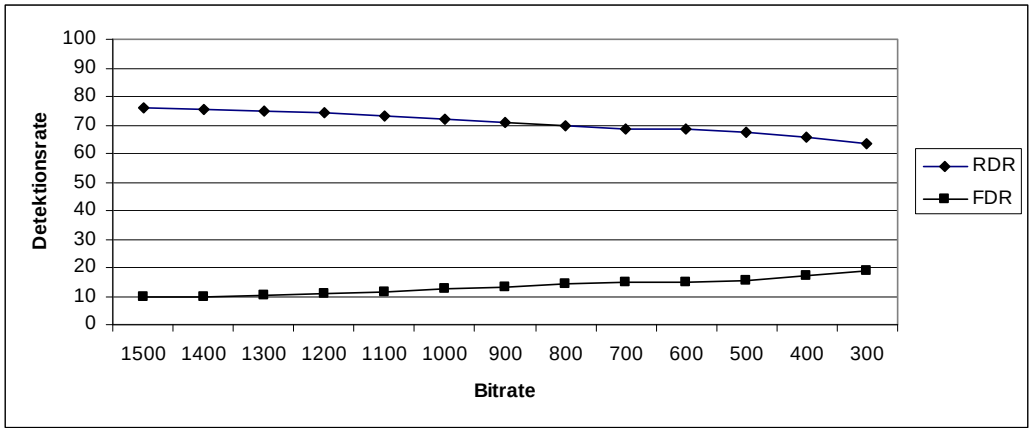

Abbildung 38: Detektionsraten robustes Wasserzeichen für Videoset 1

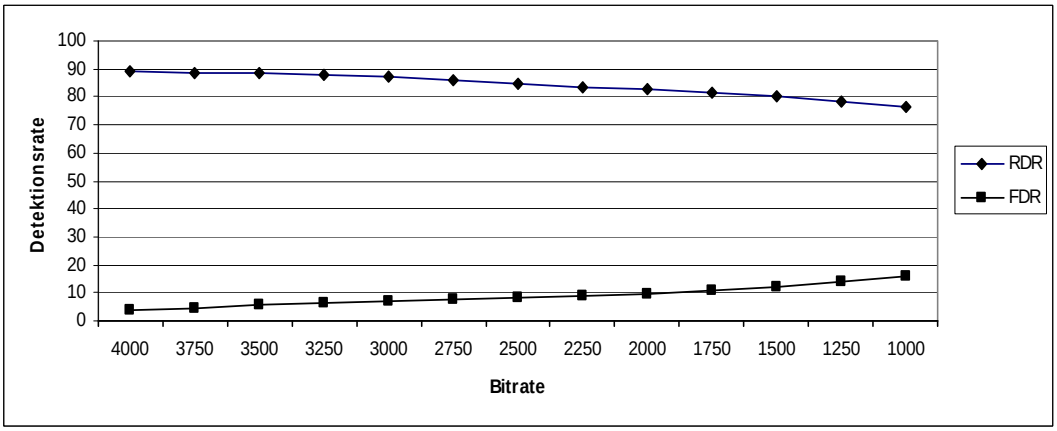

Abbildung 39: Detektionsraten robustes Wasserzeichen für Videoset 2

Abbildung 38 und Abbildung 39 zeigen die Ergebnisse der Robustheit gegenüber der MPEG-Kompression beider Videosets. Videoset 1 wurde von 1500 kBit/s bis auf 300 kBit/s und Videoset 2 von 4000 kBit/s bis auf 1000 kBit/s komprimiert. In beiden Abbildungen ist zu erkennen, dass die RDR auch bei entsprechend niedrigen Bitraten immer noch höher als die FDR war. Das bedeutet, dass auch nach einer MPEG-Kompression mit einem hohen Kompressionsgrad das Wasserzeichen immer noch erfolgreich ausgelesen werden konnte.

In weiteren Tests wurde die Robustheit des Wasserzeichens gegenüber Formatkonvertierung und Skalierung untersucht. Bei der Durchführung dieser Tests wurde zuerst die jeweilige Modifikation durchgeführt und anschließend das Video wieder zurück in das Ausgangsformat konvertiert, um das Wasserzeichen auszulesen. Die beiden Tests wurden mit Videoset 2 durchgeführt, da eher DVD-Videos zu DivX-Videos umgewandelt und dabei skaliert werden, als Super-Video-CD-Videos. Das Format, in das die Videos konvertiert werden, ist DivX 4 und die Bitrate liegt zwischen 600 und 1200 kBit/s in 100 kBit/s Schritten. Die Ergebnisse werden in Tabelle 7 dargestellt.

Bitrate	RDR	FDR
1200	65,4	27,0
1100	63,7	28,7
1000	62,0	30,5
900	60,1	32,3
800	58,1	34,4
700	55,7	36,9
600	52,8	39,8

Tabelle 7: Robustheit DivX-Konvertierung

Wie aus der Tabelle entnommen werden kann, ist es möglich, dass das Wasserzeichen bis zu einer Bitrate von 600 kBit/s erfolgreich ausgelesen werden kann. Die RDR-Rate ist stets höher als die FDR-Rate, was für den Erfolg der Wasserzeichendetektion notwenig ist.

Bei der Skalierung wurden Höhe und Breite des Frames gleichmäßig skaliert. Dabei wurde das mit dem Wasserzeichen markierte Video bis zu 50% der Höhen- und Breitenlänge in jeweils 10%igen Schritten skaliert. Bevor das Wasserzeichen wieder ausgelesen werden kann, wurde das modifizierte Video wieder auf die Originalgröße zurückskaliert. Tabelle 8 stellt die Ergebnisse dar.

Skalierung		
	WS 0.6	WS 0.6
Faktor	RDR	FDR
90	79,5	13,9
80	79,4	14,1
70	75,6	16,7
60	74,0	19,6
50	70,7	23,1

Tabelle 8: Robustheit Skalierung

Wie aus der Tabelle zu entnehmen ist, kann das Wasserzeichen auch bei einer Skalierung auf 25% der Fläche - d. h. die Seitenverhältnisse werden halbiert - erfolgreich ausgelesen werden. Das ist dann besonders wichtig, wenn das Video auf SVCD-Qualität konvertiert wird. Nach der Konvertierung haben die Frames nur noch 25% der Originalgröße.

7.5.3 Testergebnisse und Evaluation für P-Frames

Während dieser Tests wird untersucht, wie hoch die Detektionsraten in den einem I-Frame folgenden P-Frames sind. Wir gehen dabei von der Standardkodierung von 12 Frames pro GoP aus, wobei drei Frames ein P-Frame sind und einen Abstand von je drei Frames zueinander haben.

In Abbildung 40 und 41 werden die Detektionsergebnisse für die P-Frames beider Videosets dargestellt. Sie sind mit der Right-Detection-Rate (RDR) dargestellt. Dabei werden die Ergebnisse denen der I-Frames gegenübergestellt (RDR I-Frames), um einen Vergleich der Detektionsqualität zu ermöglichen. Zuletzt wird noch die durchschnittliche False-Detection-

Rate (FDR) dargestellt. Sie ist der Durchschnitt der vier gemessenen Fehlerraten, beinhaltet somit die Fehlerraten der I-Frames und der drei P-Frames und dient damit der Gegenüberstellung der verschiedenen RDR-Raten. Wenn die FDR niedriger als die RDR-Raten ist, kann das Wasserzeichen erfolgreich ausgelesen werden.

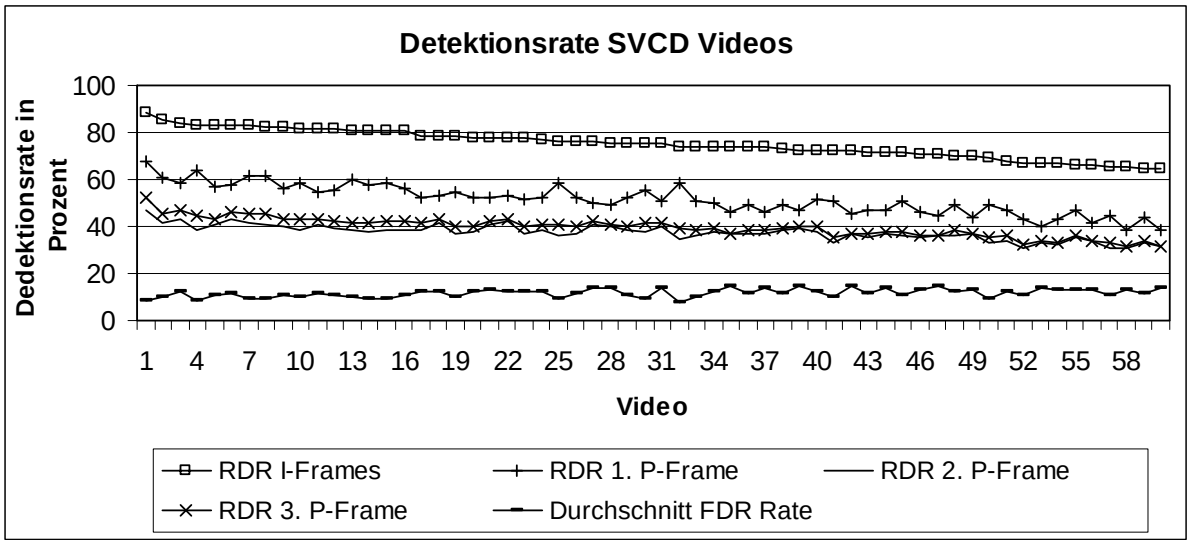

Abbildung 40: Detektion des Wasserzeichens in P-Frames von Videoset 1

Aus Abbildung 40 kann entnommen werden, dass die Detektionsraten für die P-Frames zwar nicht die Qualität der I-Frames aufweisen, sie aber immer noch deutlich über der durchschnittlichen FDR liegen. Ein weiteres Ergebnis ist der Unterschied zwischen der RDR des ersten P-Frames und der beiden nachfolgenden P-Frames. Der Unterschied der Detektionsraten liegt im Durchschnitt bei 15 bis 20 Prozent. Die Detektionsraten des zweiten und dritten P-Frames befinden sich fast immer in einer engen Spanne. Die durchschnittliche FDR-Rate liegt für das gesamte Videoset unter 20 Prozent.

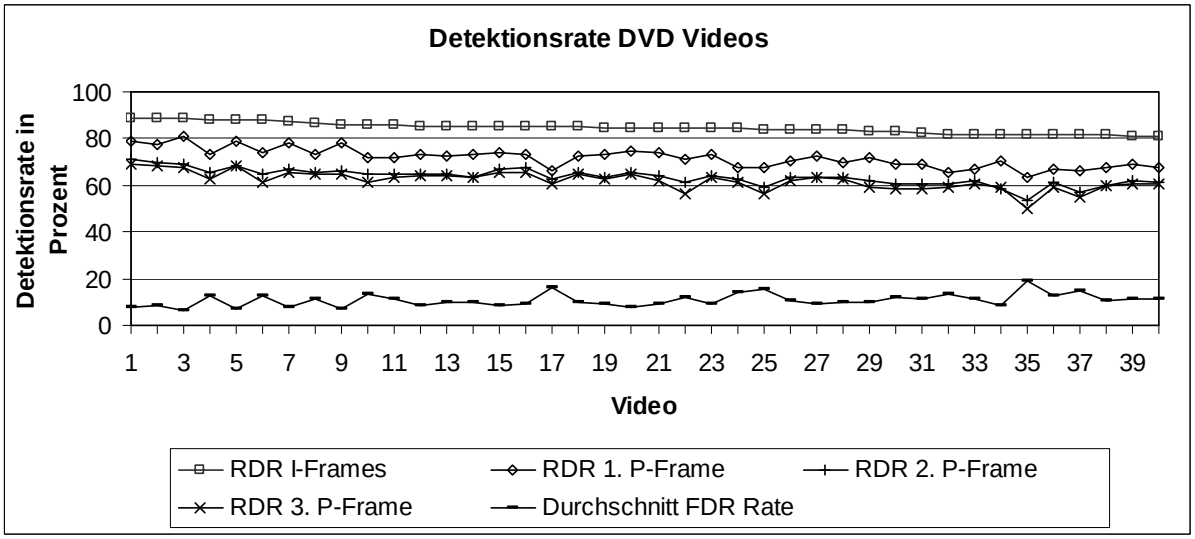

Abbildung 41: Detektion des Wasserzeichens in P-Frames von Videoset 2

In Abbildung 41 sind die Ergebnisse ähnlich zu Abbildung 40. Was aber im Gegensatz zu Abbildung 40 auffällt, ist, dass die RDR Raten besser sind. Die durchschnittliche FDR ist hingegen weiter im Schnitt geringer als 20 Prozent.

Beide Abbildungen verdeutlichen die erfolgreiche Detektion der Wasserzeichen aus den P-Frames. Für die drei P-Frames pro GoP ist die jeweilige RDR Rate konstant höher als die FDR Rate. Die durchschnittliche FDR Rate von unter 20 Prozent lässt darauf schließen, dass die Wasserzeicheninformation einerseits robust gegenüber der MPEG-Kompression in P-Frames ist und andererseits auch die Kontrollmechanismen während der Detektion des Wasserzeichens erfolgreich sind. Es lässt sich dadurch begründen, dass besonders bei Videoset 1 die RDR des jeweiligen P-Frames und die FDR nicht 100 % ergeben. Der Detektionsprozess erkennt eine Vielzahl falsch detektierter Wasserzeichenbits und beachtet sie nicht im Ausleseergebnis.

Die verschiedenen Ergebnisse der Detektionsrate zwischen dem ersten und den beiden nachfolgenden P-Frames kann mit der Erhöhung des Skalierungsfaktors der MPEG Quantisierung ab dem zweiten P-Frame begründet werden. Einzelne Tests haben ergeben, dass sich nach dem ersten P-Frame der Skalierungsfaktor im Durchschnitt um ungefähr 10 Prozent erhöht hat. Zudem werden nach dem ersten P-Frame deutlich mehr Blöcke als invalid eingestuft, d. h. es konnte kein Wasserzeichenbit eingebettet werden. Es kann ebenfalls davon ausgegangen werden, dass die verstärkte Quantisierung zu einer deutlicheren Veränderung des Blockinhaltes führt, wodurch kein Wasserzeichen mehr ausgelesen werden kann.

7.5.4 Schlussfolgerungen und Optimierungsvorschläge

Basierend auf den gewonnenen Erkenntnissen ist das Wasserzeichen sehr gut zum Markieren von I-Frames geeignet. Die Detektionsraten können auch nach üblichen Modifikationen, wie Skalierung und Formatkonvertierung, als vielversprechend bezeichnet werden.

Das Markierungskonzept kann auch erfolgreich auf P-Frames übertragen werden. Zwar sind die Detektionsraten nicht so gut wie in I-Frames, aber trotzdem noch besser als die False-Detection-Rate. Dadurch ist es durchaus auch möglich, die Methode für B-Frames anzuwenden. Wenn Teile der B-Frames nicht markiert werden können, muss der Detektionsprozess des Wasserzeichens sie herausfiltern.

Es werden folgende Optimierungen vorgeschlagen:

(1) Anpassung des verwendeten zwei-dimensionalen visuellen Modells an ein drei-dimensionales Modell

Das eingesetzte visuelle Modell ist ein für Bilddaten entwickeltes Konzept. Es berechnet die Wasserzeichenstärke anhand der Verteilung von Ecken/Kanten und gleichmäßigen Flächen im Frame. Da Videodaten aber zeitlich verteilt sind, ist es ratsam, während der Berechnung der lokalen Wasserzeichenstärke auch die Nachbarframes mit zu beachten. So können Videoabschnitte und Flächen anhand des Bewegungsgrades und der Veränderung entlang der Zeitachse klassifiziert werden. Eine schnelle Bewegung und signifikante Richtungsänderung ermöglicht es, das Wasserzeichen in diesen Bereichen verstärkt einzusetzen. Dagegen ist die Transparenz in ruhigen Bereichen mit wenig Bewegung eher sensibler [OoKa2001].

(2) Wasserzeichenstärke an den MPEG Skalierungsfaktor koppeln
Der MPEG-Skalierungsfaktor ist der maßgebliche Wert für den Videobuffer. Er verhindert einen Über- oder Unterlauf des Buffers, wodurch das Video asynchron abgespielt werden könnte.

Die Wasserzeichenstärke sollte an den Skalierungsfaktor gekoppelt werden. Die Tests haben gezeigt, dass ein stärkerer Skalierungsfaktor zu einer Verschlechterung des Ausleseergebnis führt. Eine Möglichkeit wäre die lineare Erhöhung der Wasserzeichenstärke. Mit einem simulierten Durchlauf des kompletten MPEG-Quantisierungsprozesses kann untersucht werden, welche der Markierungsblöcke eine Verstärkung der Wasserzeichenstärke benötigen. Das erhöht zwar die Komplexität des Wasserzeichenalgorithmus, führt aber gleichzeitig nicht zu einer signifikanten Verschlechterung der Transparenz des Wasserzeichens.

7.6 Zusammenfassung

Das in diesem Kapitel vorgestellte Wasserzeichen ist so aufgebaut, dass es die Wasserzeicheninformationen sowohl in intrakodierte als auch interkodierte Frames einbetten kann. Für intrakodierte Frames wurde das Wasserzeichen den MPEG-Eigenschaften so angepasst, dass es nach typischen Modifikationen, wie Skalierung, Formatkonvertierung oder MPEG-Kompression weiter auslesbar ist.

Zusätzlich wurde neben den I-Frames als Einbettungsquelle die Markierungsmethode auf P-Frames ausgeweitet. Dadurch verbessert sich die Robustheit des Wasserzeichens deutlich und die Abhängigkeit von den I-Frames wird verringert. Da aufgrund des geringen zeitlichen Unterschieds zwischen benachbarten Frames die Möglichkeiten des Einbettens verschiedener Wasserzeichen gering sind, werden sowohl in I- als auch in P-Frames die gleichen Wasserzeicheninformationen eingebettet. In dem präsentierten Ansatz wurde eine Anpassung auf interkodierte Frames vorgestellt, die in weiteren Optimierungen auf das gesamte Video erweitert werden sollte.

Die Testergebnisse verdeutlichen, dass es möglich ist, das Wasserzeichen aus P-Frames erfolgreich auszulesen. Dabei sind die Detektionsraten ausreichend hoch, um eine falsche Detektion zu verhindern. Die Transparenz kann durch eine entsprechende Analyse des Frameinhaltes mittels eines visuellen Modells und der danach festgelegten Wasserzeichenstärke gewahrt werden.

Augenblicklich muss das präsentierte Wasserzeichen noch an den MPEG-Quantisierungsprozess angepasst werden. So kann bei einer entsprechend starken Quantisierung das Wasserzeichen nicht mehr komplett ausgelesen werden. Ein weiteres Problem liegt in der Einbettung des Wasserzeichens an festen Positionen im Frame. Nach eventuellen Verschiebungen der Markierungspositionen ist es für den Detektor relativ schwer, das Wasserzeichen wiederzufinden. Das erfordert die Integration von Synchronisationsmechanismen im Frame.

8

Die Authentizität von Videomaterial ist besonders in jenen Anwendungsfällen wichtig, wenn sie wichtige Aussagen über den Inhalt beinhalten. So erweist es sich bei Überwachungsvideos als wichtig, nachweisen zu können, dass keine Veränderungen am Video durchgeführt wurden. Bei der augenblicklich zur Verfügung stehenden Soft- und Hardware ist es inzwischen ohne großen Aufwand möglich, Veränderungen im Video durchzuführen. Beispiele sind der Austausch von Szenen oder das Löschen einzelner Frames. Die veränderten Videos sind dann nicht mehr authentisch, da die Aussage des Videos durch die Editierung verändert wurde.

Digitale Wasserzeichen sind auch für diese Anwendung eine mögliche Lösung [CoMi2002]. Um die Veränderungen nachzuweisen, können fragile Wasserzeichen genutzt werden [LaTh2003].

Fragile Wasserzeichen erkennen also die Veränderungen am Videomaterial. Dabei müssen die eingebetteten Wasserzeichendaten aber geschützt werden, um einen Missbrauch zu verhindern. Auch die Einschränkung der Authentizitätskontrolle kann Bestandteil der Anwendung sein.

Kryptographische Methoden schränken den Zugriff auf Daten ein. Sie können mit dem Wasserzeichen kombiniert werden, in dem sie den Zugriff auf die Wasserzeicheninformationen einschränken. Zusätzlich zur Einschränkung des Zugriffes können sie auch zur Überprüfung der Authentizität genutzt werden.

Normalerweise sind digitale Wasserzeichen mit dem Trägermaterial fest verwebt. Wenn sie robust eingebettet werden, können sie entweder nicht mehr oder nur mit einem Qualitätsverlust aus dem Video entfernt werden. Da aber durch das Einbetten des Wasserzeichens zu einem minimalen Qualitätsverlust am Trägermedium kommen kann, ist es in vereinzelten Anwendungsfällen notwendig, die ursprüngliche Qualität des Video wiederherzustellen. Für Anwendungen dieser Art werden reversible Wasserzeichen entwickelt, die die Möglichkeit der Entfernung des Wasserzeichen aus dem Video und dadurch die Wiederherstellung des Originalvideos ermöglichen.

In diesem Kapitel wird ein reversibles Wasserzeichen vorgestellt. Es kombiniert die Wiederherstellung des Originalvideos und mit der Überprüfung der Authentizität des Videos. Im ersten Teilkapitel wird das Konzept eines reversiblen Wasserzeichens erläutert. Nach der Diskussion der Anforderungen folgt ein Wasserzeichenalgorithmus für MPEG-Videos [HaDi2005] mit der abschließenden Evalierung der Lösung und der Zusammenfassung.

8.1 Konzept für reversible Videowasserzeichen

Ein reversibles Wasserzeichen wurde zuerst in [FrGo2001] und [StDi2003] vorgestellt. Das Konzept wird dabei in die einzelnen Teilprozesse Einbettung, Verifizierung und Wiederherstellung unterteilt.

8.1.1 Einbettung

Die Einbettung eines reversiblen Wasserzeichens kann allgemein beschrieben werden. Nach der Auswahl des Einbettungskanals werden die Daten des Einbettungskanals komprimiert und als Teilinformation der Wasserzeicheninformation zugeordnet. Durch die Komprimierung der Daten des Einbettungskanals wird ein Platz zur Verfügung gestellt, der von der eigentlichen Wasserzeichennachricht genutzt wird. Basierend auf dem Anwendungsszenario kann die Wasserzeichennachricht unterschiedlich aufgebaut sein. Die einzige zu erfüllende Bedingung ist, dass die Wasserzeichennachricht nicht länger als der durch die Komprimierung geschaffene Platz sein darf. Ist es nicht möglich, muss eventuell eine effzientere Komprimierung angewandt oder der Einbettungskanal vergrößert werden.

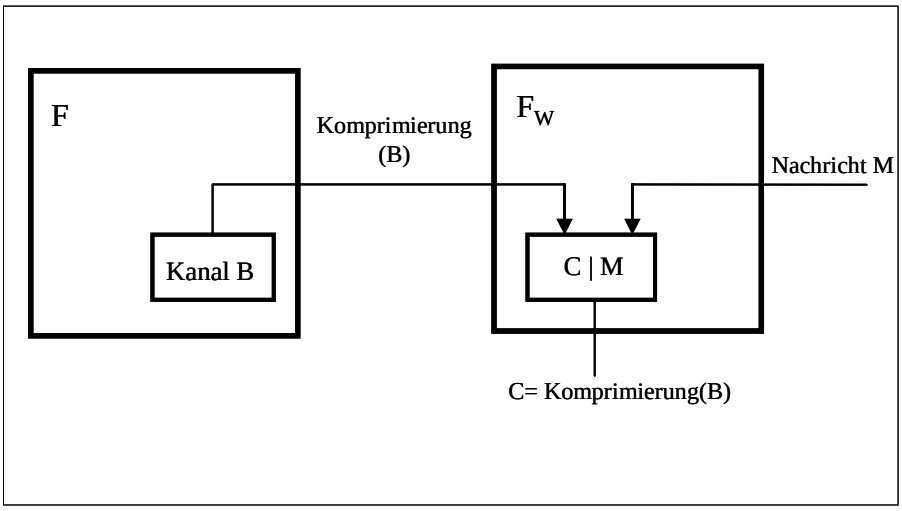

Abbildung 42: Konzept eines reversibles Wasserzeichen auf Framebasis

Abbildung 42 demonstriert das allgemeine Konzept der Einbettung eines reversiblen Wasserzeichens am Beispiel eines Videoframes F. Im Frame F wird ein Einbettungskanal B ausgewählt, in dem später die Wasserzeicheninformation eingebettet wird. Um das ursprüngliche Video wiederherstellen zu können, wird der Inhalt des Einbettungskanals B komprimiert. Die komprimierten Daten C des Einbettungskanals B werden als Teil der Wasserzeicheninformation in das Video eingebettet. Der durch die Komprimierung geschaffene Platz wird zum Einfügen der Nachricht M genutzt. Die Nachricht M beinhaltet die Informationen, die zur Erfüllung der an das Wasserzeichen gestellten Anforderungen benötigt werden.

8.1.2 Verifizierung

Die Verifizierung des Videos ist ein Vergleich zwischen der eingebetteten Nachricht M und den Videodaten, aus denen während der Einbettung die Nachricht M gebildet wurde. Je nachdem wie die Nachricht M aufgebaut ist, müssen einzelne Schritte der Einbettung wiederholt werden. So müssen möglicherweise Hashwerte wieder gebildet oder Entschlüsselungen durchgeführt werden. Der Vergleich führt zu einer eindeutigen Aussage, ob die Videodaten, an denen der Vergleich durchgeführt wurde, authentisch sind.

8.1.3 Wiederherstellung des Originals

Während der Wiederherstellung des Originals werden die komprimierten Daten C wieder dekomprimiert und dann an den ursprünglichen Positionen im Einbettungskanal B eingefügt. Zusätzlich werden die dekomprimierten Daten auf ihre Integrität hin überprüft. Nach dem Rückschreiben der Daten entspricht das Video wieder seinem Originalzustand.

8.1.4 Kryptographische Mechanismen

Mit Hilfe verschiedener kryptographischer Methoden ist es möglich, die Sicherheitsanforderungen zu realisieren, die vom Anwendungsfall in Kapitel 3.2 gefordert werden. Der Fokus liegt dabei auf den folgenden Sicherheitsanforderungen:

(1) Verifikation der Authentizität der Videodaten
Um ein Video als authentisch einstufen zu können, müssen einerseits die Daten integer sein und andererseits die Echtheit des Videos gewahrt werden.

Um eine Aussage über die Integrität von Multimediadaten zu treffen, werden Hashfunktionen angewandt. Die allgemeine Hashfunktion bildet Daten beliebiger Länge auf einen Datensatz fester Länge ab. Zudem ist die Funktion injektiv. Eine kryptographische Hashfunktion ist eine Einwegfunktion, d. h. für einen Hashwert gibt es keine inverse Berechnung. Zudem ist die Hashfunktion schwach kollisionsresistent. Es ist praktisch unmöglich, eine Kollision herzustellen, d. h. dass aus zwei verschiedenen Datensätze identische Hashwerte ermitteln werden [Ec2001].

Mit Hilfe einer digitalen Signatur kann ein digitaler Inhalt signiert werden. Dadurch kann die Authentizität des Videos nachgewiesen werden. Da Videos einfach kopiert werden können, muss durch den Einsatz der Signatur die Echtheit des Videos nachgewiesen werden können. Durch den Einsatz von asymmetrischen Signaturen besteht die Möglichkeit der öffentlichen Verifizierung der Echtheit des Videos. Während der Bildung der Signatur wird von dem Video der Hashwert gebildet und mit Hilfe der Signatur verschlüsselt. Wird eine öffentliche Verifizierung angestrebt, wird der Hashwert mit dem geheimen Schlüssel verschlüsselt. Durch Entschlüsselung des Video mit dem öffentlichen Schlüssel und der nachfolgenden Überprüfung der Hashwerte kann die Echtheit nachgewiesen werden.

(2) Reproduktion des Originals
Die Reproduktion des Originals erfordert eine sichere Einbettung der ursprünglichen Daten des Einbettungskanals in das Video, um sie vor unerlaubten Zugriffen zu schützen. Dies erfordert wiederum eine sichere Verschlüsselung der Daten. Die Verwendung einer symmetrischen Verschlüsselung garantiert den Zugriffsschutz bei Sicherung des Zugriffes auf den geheimen Schlüssel.

8.2 Anforderungen an reversible Videowasserzeichen

Das Ziel möglicher Angreifer ist die Veränderung der Echtheit des Videos. Wie bei Einzelbildern können auch an Videodaten Manipulationen durchgeführt werden. Bei dem hier vorgestellten Wasserzeichen liegt der Fokus auf Angriffen, die einerseits die Integrität und andererseits die Sicherheit des Wasserzeichens verändern. Deshalb konzentrieren wir uns zuerst auf die Angriffsmöglichkeiten die die Integrität und dann die Sicherheit betreffen.

Die Möglichkeiten der Veränderung der Integrität an Videodaten lassen sich für diesen Anwendungsfall in zwei Rubriken unterteilen:

(1) Lokale Manipulationen
Zu dieser Rubrik gehören Manipulationen wie das Entfernen oder Überschreiben von Frameinhalten. Manipulationen dieser Art verändern die Aussage des Videoframes. Die Veränderungen müssen aber auch in den Nachbarframes durchgeführt werden, da ohne diese Anpassung die Veränderung sichtbar werden könnte. Mit entsprechender professioneller Bearbeitungssoftware kann der Bearbeitungsaufwand aber gering gehalten werden.

(2) Zeitliche Manipulationen
In der zweiten Rubrik wird die zeitliche Aussage des Videos manipuliert. Da ein Video aus einer Framesequenz besteht, kann der Angreifer die Sequenz gezielt manipulieren. Dabei können folgende Manipulationen durchgeführt werden:

(a) Austausch von Nachbarframes oder -szenen
(b) Hinzufügen oder Entfernen von Einzelframes bis zu kompletten Szenen

Wenn zusätzlich der Aufwand einer erneuten Enkodierung des Videos so gering wie möglich gehalten wird, führen Manipulationen dieser Art zu keiner deutlichen Verschlechterung der visuellen Qualität.

Die Möglichkeit der Wiederherstellung des Originalvideos erfordert eine zusätzliche Überprüfung der eingebetteten Originaldaten. Um das Original wiederherzustellen, müssen sie während der Einbettung als Bestandteil der Wasserzeicheninformation mit in das Video eingebettet werden. Dies erfordert eine hohe Sicherheit, da dem Angreifer der Einbettungsalgorithmus bekannt sein kann; die eingebetteten Originaldaten müssen also besonders geschützt werden. Durch Verschlüsselung der einzubetteten Originaldaten kann ein direkter Zugriff verhindert werden. Die Integrität der verschlüsselten Originaldaten kann mit

Hilfe eines Hashs überprüft werden, wodurch eine gezielte Fälschung der Originaldaten nachgewiesen werden kann.

Die Sicherheit des Wasserzeichens wird zusätzlich von den folgenden Aspekten bestimmt:

(1) Veränderung der gesicherten Bildinformationen
Eine gezielte Manipulation der Frameinformationen erfordert eine erneute Bildung des Hashwerts, der die Integrität der Frameinformationen überprüfen soll. Der Angreifer sieht sich nach der Modifikation am Video gezwungen, den Hashwert neu zu bilden und den ursprünglichen Hashwert damit zu überschreiben. Durch eine zusätzliche Überprüfung der Integrität des ursprünglich eingebetteten Hashwertes können aber Manipulationen am selbigen nachgewiesen werden. Dies kann durch den Einsatz einer Signatur realisiert werden, die den ursprünglichen Hashwert als Bestandteil beinhaltet. Zudem muss die Einzelbildanordnung im Video mittels eines Index abgesichert werden, der ebenfalls in die Signatur integriert werden muss.

(2) Man-in-the-middle-Angriff
Die Originaldaten des Einbettungskanals werden symmetrisch verschlüsselt. Wenn der geheime Schlüssel bekannt ist, kann ein Angreifer das ursprüngliche Wasserzeichen zerstören und ein eigenes Wasserzeichen in Abhängigkeit an seinen eigenen Schlüssel in das Video einbetten. Um die Authentizität des Videos zu sichern, ist es notwendig, die verschlüsselten Originaldaten als Bestandteil in die digitale Signatur zu integrieren.

(3) Analyse der Ursprungsinformation
Bei diesem Angriff werden die komprimierten Originaldaten dahingegen analysiert, ob anhand des Kompressionsergebnisses Eigenschaften der Originaldaten erkannt werden können. Besteht z. B. der Einbettungskanal des Wasserzeichens aus hochfrequente DCT-Koeffizienten, dann haben diese Koeffizienten mit einer hohen Wahrscheinlichkeit den Wert null. Aufgrund dieser Eigenschaft lassen sie sich häufig sehr gut komprimieren. Um eine Analyse der komprimierten Daten zu erschweren, sollte der Wasserzeichenalgorithmus die übrigen Bits bis zur kompletten Blocklänge der Verschlüsselung mit pseudo-zufällig erstellten Werten komplettieren. Dadurch kann der Angreifer nicht erkennen, wieviel Bits der gesamten Blocklänge zu den eigentlichen Originaldaten gehören.

Wenn die komplette Wasserzeicheninformation mit null überschrieben wird, ist es nicht mehr möglich, das ursprünglich eingebettete Wasserzeichen zu detektieren. Zudem kann das Originalvideo nicht mehr wiederhergestellt werdem. Dieser Angriff kann zudem nicht nachgewiesen werden, wenn sämtliche Werte des Einbettungskanals den Wert null hatten. Ihm kann nur entgegen gewirkt werden, wenn nicht sämtliche Werte den Wert null hatten.

Sämtliche ermittelten Anforderungen an die Integrität und Sicherheit erfordern eine gesonderte Analyse der Wasserzeicheninformation. Sie wird im folgenden Kapitel beschrieben.

8.3 Zusammensetzung der Wasserzeicheninformation

Das generelle Konzept wurde in [DiSt2002], [StDi2003] und [KaDi2004] vorgestellt. In diesem Kapitel wird das allgemeine Konzept für Mediendaten für Videodaten beschrieben. Vor dem Einbetten des Wasserzeichens wird ein Videosignal V in zwei Teilsignale aufgeteilt, die im Folgenden als Teil A und Teil B bezeichnet werden. Der Einbettungskanal ist das Teilsignal B. Die anderen Mediendaten werden dem Teil A zugewiesen:

$$V = A \cup B \qquad (19)$$

Für Videodaten lässt sich Formel 19 wie folgt spezifizieren: Teil A stellt die remainingData $V_{remaining}$ dar und Teil B beinhaltet die Wasserzeicheninformation inklusive zusätzlicher zu füllende Daten. Dadurch lässt sich Formel 19 wie folgt beschreiben:

$$V_W = V_{remaining} + Data_{fill} + W \qquad (20)$$

$Data_{fill}$ sind sogenannte Füllbits, die zur Anwendung kommen, wenn die Kompression der Daten des Einbettungskanals (selectedData) mehr Kapazität zur Verfügung stellt, als von der gesamten Wasserzeicheninformation benötigt.

Die Wasserzeicheninformation W besteht aus den Daten, die zur Wiederherstellung des Originals $W_{\mathrm{Re}c}$ und zur Überprüfung der Authentizität des Videos W_{Aut} notwendig sind. Zur Wiederherstellung des Originals werden die selectedData komprimiert. Zusätzlich wird eine Integritätskontrolle der selectedData hinzugefügt, um garantieren zu können, dass die Daten nach der Entschlüsselung noch integer sind. Die Authentizitätsinformation weist die Echtheit Videos nach und wird ebenfalls mit in die Wasserzeicheninformation integriert. Zur vereinfachten Darstellung wird die vollständige Wasserzeicheninformation wie folgt dargestellt:

$$W = W_{\mathrm{Re}c} + W_{Aut} \qquad (21)$$

Die Wasserzeicheninformation $W_{\mathrm{Re}c}$ beinhaltet die Videodaten, die zur Wiederherstellung des Originals benötigt werden. In diesem Fall ist es die Daten des vorausgewählten Einbettungskanals: selectedData $V_{selected}$. Zur Schaffung der benötigten Kapazität für die Daten zur Überprüfung der Authentizitäts werden die ausgewählten Videodaten $V_{selected}$ zu $C_{V_{selected}}$ komprimiert. Anschließend wird $C_{V_{selected}}$ zweifach verschlüsselt. Während der ersten Verschlüsselung wird $C_{V_{selected}}$ mit dem symmetrischen Schlüssel $K_{\mathrm{sec}ret}$ verschlüsselt. Nur bei Kenntnis des Schlüssels $K_{\mathrm{sec}ret}$ ist die Wiederherstellung des Originalvideos möglich. Eine Verbreitung des Schlüssels kann damit gezielt kontrolliert werden. Für die anschließende Verschlüsselung wird der Schlüssel $K_{H_{remaining}}$ benutzt. Er ist der Hashwert von $V_{remaining}$, entsprechend den Daten des Video, die nicht zum Einbettungskanal gehören. Mit der Verwendung von $K_{H_{remaining}}$ ist die Wiederherstellung des Originals nur möglich, wenn $V_{remaining}$ nicht verändert wurde.

Nach der Entschlüsselung und Dekomprimierung von $V_{selected}$ bedarf es einer zusätzlichen Kontrolle der Integrität von $V_{selected}$. Augenblicklich wäre es nicht möglich nachzuweisen, ob $V_{selected}$ verändert wurde. Deshalb wird zu $W_{\mathrm{Re}c}$ ein verschlüsselter Hash $MAC(V_{selected}, K_{\mathrm{sec}ret})$ hinzugefügt. Er wird mit $K_{\mathrm{sec}ret}$ verschlüsselt, sodass nur die Person, die das Originalvideo wiederherstellen kann, auch in der Lage ist, die Integrität von $V_{selected}$ zu überprüfen. So wird eine Manipulation von $V_{selected}$ ausgeschlossen.

Die Wasserzeicheninformation zur Wiederherstellung des Originalvideos setzt sich wiefolgt zusammen:

$$W_{\mathrm{Re}c} = E\left(E\left(C_{V_{selected}}, K_{\mathrm{sec}ret}\right), K_{H(V_{remaining})}\right) + MAC(V_{selected}, K_{\mathrm{sec}ret}) \qquad (22)$$

Eine Überprüfung der Authentizität wird mit Hilfe der Verwendung einer digitalen Signatur durchgeführt. Die öffentliche Verifikation erfordert die Anwendung einer asymmetrischen Signatur, welche sich aus dem Hashwert von $V_{remaining}$ und $W_{\mathrm{Re}c}$ zusammensetzt und mit dem privaten Schlüssel $K_{private}$ verschlüsselt wird. Die Verwendung des Hashs weist eine Änderung der Daten nach. $W_{\mathrm{Re}c}$ liegt weiter in verschlüsselter Form vor, da auch bei öffentlicher Verifikation ein Zugriff auf die Originaldaten verhindert werden soll.

So setzt sich die Wasserzeichennachricht für die Authentizitätsüberprüfung wie folgt zusammen:

$$W_{Aut} = S(H(V_{remaining} + E\left(E\left(C_{V_{selected}}, K_{\mathrm{sec}ret}\right), K_{H(V_{remaining})}\right) + MAC(V_{selected}, K_{\mathrm{sec}ret})), K_{private}) \qquad (23)$$

Basierend auf Formel (21) kann die Wasserzeicheninformation wie folgt zusammengeführt werden:

$$W = E\left(E\left(C_{V_{selected}}, K_{\mathrm{sec}ret}\right), K_{H(V_{remaining})}\right) + MAC(V_{selected}, K_{\mathrm{sec}ret}) + S(H(V_{remaining} + E\ (E\ (C_{V_{selected}}, K_{\mathrm{sec}ret}), K(_{H(V_{remaining})}) + MAC(V_{selected}, K_{\mathrm{sec}ret})), K_{private}) \qquad (24)$$

Die zusammengesetzte Wasserzeicheninformation gilt für das gesamte Video. Um die Möglichkeit der framegenauen Identifikation zu ermöglichen, wird die Wasserzeicheninformation für jedes Frame berechnet und in das Frame eingebettet. Anders als bei Bilddaten ist bei Videodaten die Möglichkeit einer Veränderung der Frameanordnung zu beachten. Zur Kontrolle der Frameanordnung wird ein Index eingeführt, der eine

framegenaue Identifikation ermöglicht. Dadurch erweitert sich Formel 24 wie folgt, wobei jetzt die Wasserzeicheninformation für ein Frame F berechnet wird:

$$\begin{aligned} W_F = & E\left(E\left(C_{F_{selected}}, K_{\sec ret}\right), K_{H\left(F_{remaining}\right)}\right) \\ & + Index(F) \\ & + MAC((F_{selected} + Index(F)), K_{\sec ret}) \\ & + S(H(F_{remaining} + Index(F) \\ & \quad + E\ (E\ (C_{F_{selected}}, K_{\sec ret}), K(_{H(F_{remaining})}) \\ & \quad + MAC((F_{selected} + Index(F)), K_{\sec ret})), K_{private}) \end{aligned} \tag{25}$$

Nach der Zusammenstellung der Wasserzeicheninformation muß analysiert werden, welche Sicherheitsalgorithmen für die Verschlüsselung E, den Hashwert $H_{remaining}$, dem verwendeten MAC Verfahren und der Signatur S verwendet werden. Die Länge der Hashwerte und der Signatur entscheidet über den Grad der Sicherheit.

Für die Verschlüsselung E wird der augenblicklich als sicher eingestufte Standard AES verwendet [Ec2001]. Der geheime Schlüssel hat eine Länge von 256 Bits. Für den Hashwert $H_{remaining}$ besteht eine Auswahl der Hashalgorithmen MD5 und SHA-1, wobei der MD5 eine Länge von 128 Bit und der SHA-1 von 160 Bit hat. Das verwendete MAC-Verfahren ist das HMAC-Verfahren. Für die Signatur S wird die asymmetrische RSA-Signatur genutzt [Ec2001].

Basierend auf der Auswahl der Kryptographiealgorithmen spezifiziert sich Formel 25 wie folgt:

$$\begin{aligned} W_F = & E_{AES}\left(E_{AES}\left(C_{F_{selected}}, K_{\sec ret}\right), K_{H\left(F_{remaining}\right)}\right) \\ & + Index(F) \\ & + MAC_{HMAC}((F_{selected} + Index(F)), K_{\sec ret}) \\ & + S_{RSA}(H(F_{remaining} + Index(F) \\ & \quad + E_{AES}(E_{AES}(C_{F_{selected}}, K_{\sec ret}), K(_{H(F_{remaining})}) \\ & \quad + MAC_{HMAC}((F_{selected} + Index(F)), K_{\sec ret})), K_{private}) \end{aligned} \tag{26}$$

8.4 Reversibles MPEG-Videowasserzeichen

In diesem Kapitel wird ein reversibles Wasserzeichen für MPEG-Videos vorgestellt [HaDi2005]. Zuerst werden weitere Wasserzeichenlösungen aufgeführt und erläutert. Anschließend folgt die Beschreibung der Einbettung, Verifizierung und Wiederherstellung des Originals. Nach der Evaluierung der Testergebnisse werden Optimierungsvorschläge diskutiert.

8.4.1 Existierende Wasserzeichenlösungen

Um die Authentizität von Videodaten nachzuweisen, wurden auch schon in vorherigen Publikationen verschiedene kryptographische Methoden, wie digitale Signaturen [DiSt1999], [LiCh2001], [ChLi2003] eingesetzt. In [DiSt1999] wird ein inhaltsfragiles Wasserzeichen vorgestellt, dass eine digitale Signatur als Wasserzeicheninformation in das Video einbettet. Mit Hilfe eines Kantendetektors, dem canny edge Detector [Fi1997], werden die Informationen aus dem Video extrahiert und anschließend wird basierend auf den Ergebnissen die Signatur gebildet. Die Kantendetektion generiert ein binäres Bild, das anschließend in eine Binärfolge umgewandelt und mit Hilfe eines VLC-Codes verlustfrei komprimiert wird. Danach wird die Signatur mit dem geheimen Schlüssel gebildet. Die Signatur wird als inhaltsfragiles Wasserzeichen eingebettet.

[FrGo2000], [DuFr2002] und [CeSh2006] präsentieren Ansätze, die das eingebettete Wasserzeichen wieder entfernen können. Dabei wird oft ein LSB (Least Signifikant Bit) Wasserzeichen genutzt. Es bietet den Vorteil, dass die Daten ohne Einbußen der visuellen Qualität des Videos eingebettet werden können.

In [FrGo2000] und [DuFr2002] werden Lösungen präsentiert, die Bild- und Videodaten mit Hilfe von Hashwerten verifizieren. Dabei werden von JPEG-Bild- oder Video-I-Framedaten 128 Bit Hashwerte gebildet und an vorausgewählten LSB-Bits einzelner DCT-Koeffizienten eingebettet. Die Auswahl der DCT-Koeffizienten ist entscheidend für den Kompressionsgrad, da vor dem Einbetten des Wasserzeichens die Originaldaten komprimiert werden müssen. Die komprimierten Daten sind ebenfalls Bestandteil der einzubettenden Wasserzeicheninformation.

Der Einbettungskanal ist das LSB-Bit der Position P(5,5) eines quantisierten blauen Chrominanz-DCT-Blockes mit der Breite und Höhe von je 8 Bildpixeln. Der Wert an der Position P hat mit einer hohen Wahrscheinlichkeit den Wert null und lässt sich effizient verlustfrei komprimieren. Der C_B Farbkanal bietet die beste visuelle Qualität im Vergleich zu den anderen beiden Farbkanälen Y und C_R in einem JPEG-Bild. In [DuFr2002] wird die gleiche Strategie als Erweiterung für MPEG-2-Videodaten beschrieben. Als Erweiterung wird bei Videos die Kontrolle der Bildsequenz als Bestandteil der Verifikationskontrolle hinzugefügt. Durch die Integration eines Bildindex können Angriffe wie Austausch, Entfernen oder Hinzufügen von Frames nachgewiesen werden.

In [CeSh2006] wird ein Framework beschrieben, dass es ermöglicht, die Authentizität der Framedaten vor der Rekonstruktion durchzuführen. Bei vorherigen Überprüfungsschemen ist es immer notwendig gewesen, das markierte Bild zuerst wiederherzustellen und dann eine Authentizitätsüberprüfung durchzuführen. Der Ansatz teilt die Bilddaten in zwei Kanäle P_A und P_I, wobei P_A die LSB-Bits eines vorausgewählten Einbettungskanals und P_I sämtliche andere Daten des Bildes beinhaltet. Vor der Einbettung werden die Daten von P_A zuerst in P_I eingefügt. Im darauffolgenden Schritt werden die Authentifikationsinformationen von P_I gebildet und in P_A eingefügt. Dadurch befinden sich die kompletten Bilddaten in P_I. Folglich kann die Verifizierung der Echtheit vor der eigentlichen Rekonstruktion des Bild durchgeführt werden.

Während die vorherigen Ansätze sich auf MPEG-2 Videos beschränken, wurden in [PrRi2005] und [PrSc2006] h.264 Video als Einbettungsquelle genutzt. So wurden in [PrRi2005] sogenannte skipped Makroblöcke als Einbettungsquelle genutzt. Da diese Makroblöcke keine Daten besitzen, sind sie als idealer Einbettungskanal geeignet. Um visuelle Störungen zu vermeiden, wurden nur die skipped Makroblöcke genutzt, die am wenigsten als Referenzblöcke für Makroblöcke in den nachfolgenden Frames genutzt werden. Zudem werden verschiedene Einbettungsmethoden diskutiert und auf ihre Anwendbarkeit hin analysiert. In [PrSc2006] werden die Objektgrenzen minimal verschoben. Die Verschiebung um einen Pixel ist visuell nicht wahrnehmbar. Als Markierungsmethode wird der QIM Ansatz genutzt.

8.4.2 Einbettung des Wasserzeichens

Das vorgestellte Wasserzeichen für MPEG-Videos basiert auf dem Ansatz von [FRGo2002]. Es bettet die Authentizitätsinformationen in die LSB-Bits der Position P(5,5) des C_B Blockes ein. Da aber möglicherweise die Kapazität pro Frame zu gering sein kann, wird eine Voruntersuchung durchgeführt und bei Bedarf die Kapazität im Frame erhöht, sodass als sicher eingestufte digitale Signaturen und Hashwerte eingebettet werden können.

8.4.2.1 Erweiterung der Kapazität in MPEG-Frames

Da die Wasserzeicheninformation in jedes Frame eingebettet werden, ist eine Untersuchung der maximal zur Verfügung stehenden Kapazität sehr wichtig. Sie entscheidet möglicherweise über eine anzupassende Einbettungsstrategie.

Entscheidende Faktoren sind:

(1) die Anzahl der C_B Blöcke pro Makroblock,
(2) die Bitrate des Videomaterials und
(3) der Aufbau der GoP.

Die Chrominanzwerte eines Makroblocks werden sehr oft nicht komplett wie im ursprünglichen Ausgangsbild abgespeichert. Häufig wird die Auflösung der Chrominanzblöcke verringert, in dem nur jede zweite Zeile und Spalte des Blocks abgetastet wird. Diese Verringerung der Datenmenge ist möglich, da das menschliche Augen nicht so sensibel auf Veränderungen dieser Art reagiert [ISO1995]. Folglich werden nur 25 % der gesamten Chrominanzdaten kodiert. Dadurch ist die Kapazität des Wasserzeichens eingeschränkt.

Die Bitrate des Videos bestimmt die Stärke der Quantisierung und die Abtastrate der blauen Chrominanzblöcke. Eine niedrige Bitrate führt zu einer starken Quantisierung, wodurch die DCT-Koeffizienten, die durch das Wasserzeichen verändert werden, sehr oft den Wert null haben. Dadurch ist die Kompression sehr effizient. Eine hohe Bitrate führt aber gleichzeitig zu einer genaueren Abtastrate der blauen Chrominanzdaten. Es ist möglich, dass die Chrominanzdaten komplett abgespeichert werden. In diesem Fall kann aufgrund der höheren Menge zur Verfügung stehender Chominazdaten gleichzeitig eine bessere Kapazität angeboten werden.

Die GoP besteht häufig aus einer variablen Struktur von beliebig vielen I-, P- und B-Frames. Während bei I-Frames jeder Makroblock kodiert und abgespeichert wird, können basierend auf dem Bewegungsgrad bei P- und B-Frames einzelne Makroblöcke übersprungen werden. Makroblöcke dieser Art verhindern die Einbettung des Wasserzeichens, was je nach Menge dieser Markoblöcke zu einer geringeren Kapazität führt.

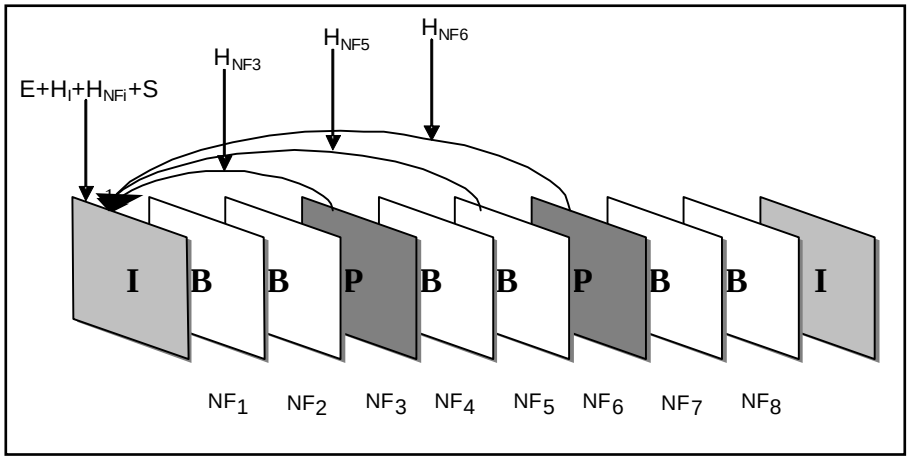

Abbildung 43: Anpassung von reversiblem Wasserzeichen für MPEG-Videos

Die drei zu beachtenden Aspekte erfordern eine Anpassung der Einbettungsmethode. Das Wasserzeichen wird ausschließlich in I-Frames eingebettet [HaDi2005]. Damit kann fast immer garantiert werden, dass ausreichend Kapazität vorhanden ist, um die komplette Wasserzeicheninformation im Frame unterzubringen. Folglich müssen die remainingData so angepasst werden, dass zusätzlich zu den remainingData des I-Frames auch noch die kompletten Daten der P- und B-Frames mit einfließen. Das ermöglicht eine Verifikation der Integrität der gesamten GoP, auch wenn nur das I-Frame als Einbettungsquelle genutzt wird. Abbildung 43 demonstriert den Einbettungsvorgang. Sie zeigt, dass der Hashwert der remainingData sich aus den remainingData des I-Frames H_I sowie der nachfolgenden P- und B-Frames H_{NFI} zusammensetzt. Der Frameindex $Index(F)$ wird in einem Index für die gesamte GoP umgewandelt. Es kann gesagt werden, dass dadurch die kleinste zu analysierende Einheit eine GoP ist. Der Nachteil ist aber, dass bei Austausch von Nachbarbildern in der jeweiligen GoP die Veränderung zwar erkannt werden kann, eine framegenaue Analyse aber nicht möglich ist.

Es werden DCT-Koeffizienten genutzt, die sich effizient komprimieren lassen. Dabei wird auf die Strategie von [FrGo2001] zurückgegriffen. Diese Strategie nutzt die Eigenschaft, dass der verwendete DCT-Koeffizient $P_Q(5,5)$ an der Stelle (5,5) in quantisierter Form häufig den Wert null hat. Dadurch ist eine effiziente Komprimierung möglich. Die Kapazität kann aber durch Hinzunahme weiterer hochfrequenter DCT-Koeffizienten zusätzlich erweitert werden. Der Einbettungskanal für einen einzelnen Block wird durch die Auswahl der Koeffizienten $P_Q(5,5)$ bis $P_Q(7,7)$ vergrößert. Dadurch liegen neun DCT-Koeffizienten pro Block vor.

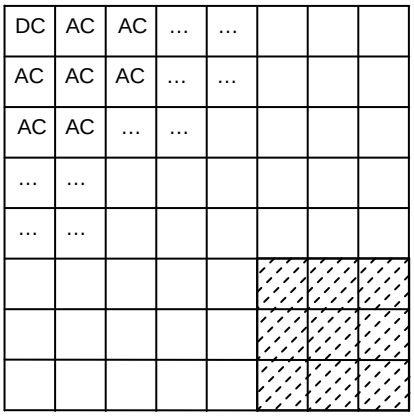

Abbildung 44: Erweiterter Einbettungskanal

Abbildung 44 zeigt die Position des Einbettungsquadrats. Die im unteren rechten Teil des Blocks befindlichen AC-Koeffizienten bilden den erweiterten Einbettungskanal.

Video	Kompressionsrate in %
bbc3	99,0
cact	99,4
flower	98,2
mobil	99,2
mulb	99,3
pulb	99,3
susi	97,2
tens	97,9
time	99,3
v700	99,3

Tabelle 9: Durchschnittliche Kompressionsrate pro GoP

Tabelle 9 stellt die Kompressionsrate dar, die bei verschiedenen MPEG-2-Videos erreicht werden kann, wenn die LSB-Bits des in Abbildung 44 dargestellten Bereichs komprimiert werden. Die Videos beinhalten bei einer Framegröße von 704x576 Bildpixeln maximal 14256 LSB-Bits pro Frame. Die Differenzen zwischen den Video deuten auf Unterschiede im Inhalt der Frames hin, die sich folglich unterschiedlich stark komprimieren lassen. Ein inhaltsreiches Frame führt zu einer schlechteren Kompression und zu einer etwas geringeren Kapazität.

Die drei diskutierten Faktoren führen zu einer ausreichenden Kapazität, um als sicher eingestufte Signaturen und Hashwerte [Bu2007] erfolgreich in das Video einbetten zu können.

8.4.2.2 Einbettungsstrategie

Die Wasserzeicheninformation werden in das LSB Bit der quantisierten AC-Koeffizienten an den Positionen $P_Q(5,5)$ bis $P_Q(7,7)$ des blauen Chrominanz Farbkanals C_B eingebettet. Abbildung 45 stellt den Einbettungsprozess dar.

Nach dem Demultiplexing des Videos in die Video- und Audioströme werden die Videoframes dekodiert (Schritt 1). Das Ziel ist es, auf die LSB-Bits zugreifen zu können. Der Dekodierprozess wird vor der inversen Quantisierung der DCT-Blöcke des C_B Farbkanals gestoppt (Schritt 2), da sich eine Einbettung des Wasserzeichen in komplett dekodierte LSB-Bits als ineffizient erwiesen hat. Die LSB-Bits der Positionen $P_Q(5,5)$ bis $P_Q(7,7)$ werden aus dem DCT-Block extrahiert (Schritt 3) und komprimiert (Schritt 4). Nach der Kompression und der Verschlüsselung der LSB-Bits bilden die komprimierten Daten den ersten Bestandteil der Wasserzeicheninformation (4a). Anschließend werden die Sicherheitsinformationen (4b) und bei noch frei verfügbaten Platz pseudozufällig generierte Füllbits (4c) in die Wasserzeichenzeicheninformation eingefügt. Die Länge des HAMC $L(HMAC)$ und der RSA-Signatur $L(RSA)$ ist dabei in Abhängigkeit zur Anwendung frei wählbar, wobei auch übliche Signaturlängen von 1024 oder 2048 Bit sowie MAC-Längen von 160 Bit ausgewählt werden können. Abschließend werden die Wasserzeicheninformationen an den LSB Bitpositionen $P_Q(5,5)$ bis $P_Q(7,7)$ eingebettet (Schritt 5) und in den Videostrom wieder eingefügt.

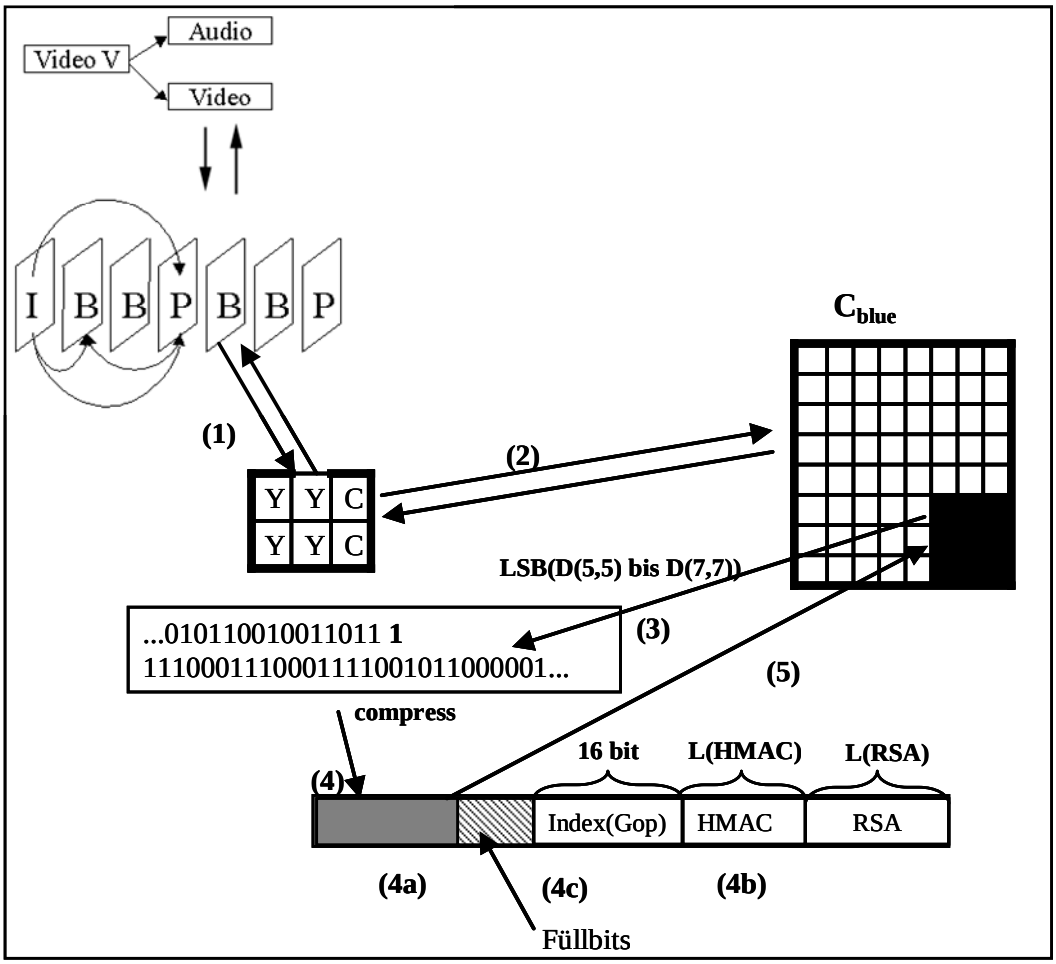

Abbildung 45: Einbetten eines reversiblen MPEG-Wasserzeichen

8.4.3 Verifizierung der Authentizität

Die Integrität und Authentizität des Videos wird während des Auslesemodus des Wasserzeichens überprüft. Die digitale Signatur wird mittels des öffentlichen Schlüssels des RSA-Schlüsselpaars entschlüsselt. Anschließend wird die Integrität der GoP dahingegen überprüft, in dem die verschlüsselten selectedData und remainingData, der HMAC und der Index überprüft werden. Da diese Daten Bestandteil der restlichen Wasserzeicheninformation sind, kann ein Hashwert über diese Daten gebildet werden und der mit dem in der Signatur befindlichen Hashwert verglichen werden. Sind beide Hashwerte identisch, wird diese GoP als integer eingestuft. Diese Integritätskontrolle wird für jede GoP durchgeführt; wurden sämtliche GoPs des Videos als integer eingestuft, kann das Video auch als authentisch eingestuft werden.

8.4.4 Wiederherstellung des Originals

Während des Prozesses der Wiederherstellung des Videos werden beide AES-Verschlüsselungen wieder entschlüsselt. Aufgrund der Schlüsselvergabe ist dies nur möglich, wenn die jeweiligen remainingData nicht verändert wurden und der secretKey bekannt ist. Anschließend werden die komprimierten selectedData $C_{F_{selected}}$ wieder dekomprimiert. Sie ergeben in Kombination mit den remainingData $F_{remaining}$ das komplette Videoframe. Vor der Wiederherstellung der Videoframes wird die Integrität der dekomprimierten Daten mit Hilfe des HMAC durchgeführt, der ebenfalls Teil der Wasserzeicheninformation ist. Dafür wird ein erneuter HMAC mit den dekomprimierten Framedaten gebildet und mit dem eingebetteten HMAC verglichen. Wenn beide HMACs identisch sind, werden die selectedData als integer eingestuft und können in die entsprechenden Positionen im Video gespeichert werden (Formel 27). Die Entfernung des Wasserzeichens wird für sämtliche Videoframes durchgeführt, in denen ein Wasserzeichen eingebettet wurde.

$$\begin{aligned} F_{Inv} &= F_{remaining} + D\left(C_{F_{selected}}\right), wenn \\ & MAC_{HMAC}((F_{selected} + Index(F)), K_{\sec ret}) = \\ & MAC_{HMAC}((D\left(C_{F_{selected}}\right) + Index(F)), K_{\sec ret}) \end{aligned} \qquad (27)$$

8.5 Evaluation der Testergebnisse

Während der Tests soll untersucht werden, in wieweit das Wasserzeichen das Video nach verschiedenen Modifikationen noch als authentisch einstuft. Dabei werden sowohl video- als auch Wasserzeichen-verändernde Modifikationen durchgeführt.

8.5.1 Testsuite

Für die Testumgebung wurde ein Referenzdatensatz von zehn Videos genutzt [Be2007]. Der Datensatz deckt verschiedenste Eigenschaften von Videomaterial ab. Dazu gehören Standbilder, Zoomeffekte, Rotationen und gleichmäßige Kameraschwenks [Su2006]. Sie besitzen eine Bildgröße von 720x576 Bildpixel und eine Datenrate von 4 MBit/s, was einer Qualität von DVD-Videos entspricht.

Es werden folgende Modifikationen betrachtet:

(1) Videomodifikationen:
(2) Re-Enkodierung
    (a) Schnitt innerhalb einer GoP, nach einer GoP, Entfernen von GoPs
    (b) Austausch oder Entfernen von Slices
    (c) Austausch von zwei Szenen
    (d) Entfernen von I-, P- oder B-Frames
    (e) Vertausch und Hinzufügen von Frames innerhalb und außerhalb von GoP
    (f) Farbkorrekturen an vereinzelten Frames oder am gesamten Video
    (g) verschiedene Filteroperationen

(3) Gezielte Angriffe auf das Wasserzeichen:
    (a) Zero Assumption (Überschreiben des kompletten Wasserzeichens mit null),
    (b) Meet in the middle Attack

8.5.2 Evaluation der Testergebnisse

Im folgenden Kapitel wird kurz aufgeführt, nach welchen Angriffen das Wasserzeichen als nicht authentisch eingestuft wurde. Danach wird der Grund für die nicht erfolgreiche Detektion diskutiert. Sind einzelne Angriffe nicht erkannt wurden, ist mit Hilfe eines Hexeditors ein bitgenauer Vergleich des markierten Videos vor und nach dem Angriff durchgeführt wurden. Damit wird untersucht, warum der Angriff nicht erkannt wurde.

Modifikation	**Erkennungsquote in %**
Neu kodiert	70
Schnitt nach einer GoP (Erstellen von zwei Szenen)	60
Einbauen von Übergängen zwischen zwei Szenen	100
Schnitt innerhalb einer GoP (Erstellen von zwei Szenen)	70
drei Schnitte jeweils nach GoPs (Erstellen von vier Szenen)	80
Vertauschen von zwei Szenen	100
GoP innerhalb des Videos entfernen	100
I-Frame entfernen	100
P-Frame oder B-Frame entfernen	100
Vertauschen einzelner Frames innerhalb einer GoP	100
Vertauschen einzelner Frames zwischen GoPs	100
Einfügen von Frame am Ende des Videos	100
Einfügen von Frame innerhalb des Videos	100
Farbkorrektur (Helligkeitskorrektur, dunkler)	100
Farbkorrektur (Änderung der Farbtönung)	100
Farbkorrektur in einem Frame (Änderung der Farbtönung)	100
Veränderung der Geschwindigkeit (Erhöhung)	100
Integration von Text in die Einzelbilder	100
Einarbeitung verschiedener Videoeffekte in das Video: - Mosaik - Wassertropfen - alter Film - Weichzeichnen - Blendeffekt	 100 100 100 100 100
GoP entfernt	100
GoPs vertauschen	100
Frames entfernen	100
Frames vertauschen	100
Slice entfernt	100
Slices vertauschen	80

Tabelle 10: Untersuchte Modifikationen

Tabelle 10 listet die Erkennungsquote der Modifikationen auf. Ein Wert von 100% sagt aus, daß nach Durchführung der Modifikation die Veränderungen in allen Videos erkannt und die Videos als nicht authentisch eingestuft worden sind. Ein Wert unter 100% deutet darauf hin, dass einzelne Videos auch nach Durchführung der Modifikation weiter als authentisch eingestuft werden. Der Grund dafür, warum nicht immer alle Modifikationen komplett erkannt wurden, ist auf den Videoinhalt zurückzuführen. Einzelne Videos bestehen aus einer Sequenz von Standbildern.

Bei einer genaueren Untersuchung dieser Videos hat sich herausgestellt, dass die Modifikationen außerhalb der gespeicherten Framedaten liegen. Entweder wurden Informationen im Frameheader einzelner Frames verändert oder die Menge der Füllbits am Ende von kodierten Frames hat sich verändert. Füllbits werden immer dann eingesetzt, wenn die Menge der kodierten Framedaten nicht ausreichend ist, um die festgelegte Bitrate zu erreichen. Dass auch der Austausch von Slices nicht immer erkannt wurde, lässt darauf schließen, dass der Inhalt der ausgetauschten Slices identisch ist.

Beim Zero-Assumption-Angriff wurde der gesamte Einbettungskanal überschrieben, was zur Folge hat, dass das gesamte Wasserzeichen verloren geht. Das Wasserzeichen konnte die Modifikation an jedem Testvideo zwar komplett nachweisen, Ziel des Angriffes ist es aber, das Original wiederherzustellen. Deshalb wurde auch das modifizierte Video mit dem Original verglichen.

Modifiziertes Video	Vergleich mit Original
100b	unterschiedlich
Bbc3	identisch
cact	identisch
flwr	unterschiedlich
mobl	identisch
mulb	unterschiedlich
pulb	unterschiedlich
susi	unterschiedlich
tens	identisch
time	identisch

Tabelle 11: Zero-Assumption-Angriff

Wie aus Tabelle 11 zu entnehmen ist, sind einzelne Videos nach Durchführung des Angriffs identisch mit ihrem Original. Das lässt sich nur damit begründen, dass die Werte an den Einbettungspositionen schon vor dem Einbetten den Wert null hatten. Gründe sind einerseits eine starke Quantisierung (z. B. bbc3) und andererseits homogene Blockflächen (time), wodurch hohe Frequenzen nicht vorhanden sind.

Der Meet-in-the-middle-Angriff [Ec2001] kann nur theoretisch betrachtet werden. Da eine 128-Bit AES-Verschlüsselung angewandt wird, ist der Aufwand als sehr hoch einzuschätzen, die angewandten Schlüssel zu finden. Der Angriff kann nur durchgeführt werden, wenn die Originaldaten und verschlüsselten Daten von $F_{remaining}$ vorliegen. Das Ziel des Angriffs ist es, die beiden Schlüssel der doppelten AES-Verschlüsselung zu finden. Dabei wird die innere AES-Verschlüsselung mit allen möglichen Schlüsseln durchgeführt und die Ergebnisse abgespeichert. Danach wird die äußere Verschlüsselung mit allen möglichen Schlüsseln entschlüsselt und diese Ergebnisse werden mit den zuvor gespeicherten verschlüsselten Daten verglichen. Ist ein identisches Ergebnis zu finden, sind beide Schlüssel gefunden. Dieser Angriff kann soweit wie möglich erschwert werden, wenn der Aufwand, um die angewandte AES-Verschlüsselung zu brechen, so umfangreich wie möglich ist.

8.5.3 Optimierungsvorschläge

Die Evaluation der Testergebnisse zeigt, dass das Verfahren nach einer Vielzahl von relevanten Modifikationen das Video als nicht mehr authentisch einstuft. Nur bei einzelnen Spezialvideos (Standbildvideo bzw. hoher Kompressiongrad) liegen Schwächen vor. Dabei haben sich besonders zwei Schwachpunkte herausgestellt:

(1) Das Video muss einen hochfrequenten Inhalt in den Frames aufweisen und
(2) Modifikationen außerhalb der Frames werden nicht erkannt.

Bei den Standbildvideos hat sich gezeigt, dass bei Nichtvorhandensein von hochfrequentem Inhalt der Einbettungskanal sehr oft keine Daten beinhaltet. Dadurch kann zwar effizient komprimiert werden, aber durch das Fehlen von Daten ist das Wasserzeichen nicht so effizient gegenüber einzelnen Modifikationen. Deshalb ist es ratsam, den Einbettungskanal so zu erweitern, dass man einen variableren Datenstrom erhält. Als Möglichkeiten bieten sich die Hinzunahme von Koeffizienten aus dem niederfrequenten Bereich, wobei aber der Aspekt der Transparenz immer überprüft werden muss. Eine Beispiel stellt Abbildung 46 dar, bei dem durch Hinzunahme des LSB-Bits von einem der markierten AC-Koeffizienten oben links im DCT-Block die zu komprimierenden Daten des Einbettungskanals nicht komplett null sind.

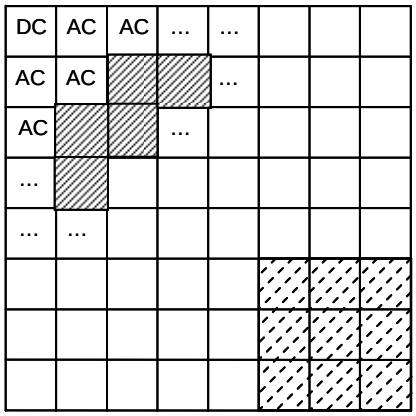

Abbildung 46: Optimierung des Einbettungskanals

Der zweite Schwachpunkt des Wasserzeichens ist die Fokussierung auf die Kontrolle der Daten innerhalb des Frames. Werden Daten außerhalb des Frames, wie zum Beispiel Strukturdaten, verändert, können sie nicht vom Wasserzeichen kontrolliert werden. Eine mögliche Lösung wäre die Integration der Strukturdaten in die remaining Daten, so dass eine Integritätskontrolle möglich ist.

8.6 Zusammenfassung

Es wurde ein Wasserzeichen zur Authentifizierung von Videos vorgestellt. Gleichzeitig besteht die Möglichkeit der Wiederherstellung des Originalvideos. Besonders bei sensiblen Anwendungsbereichen, bei denen kleinste Änderungen durch das Wasserzeichen die Bedeutung des Inhalts verändern, muss es möglich sein, es wieder zu entfernen.

Speziell dafür wurde ein Sicherheitsprotokoll von [KaDi2004] entsprechend den Anforderungen für Videodaten angepasst. Aspekte wie Sequenzanordnung treffen genauso entscheidende Aussagen wie der Inhalt der Frames. Die öffentliche Verifizierung des Wasserzeichens erfordert die Anwendung einer PKI Struktur.

Zudem wurde ein Wasserzeichenalgorithmus für Bilddaten [FrGo2001] den Anforderungen von Videodaten [HaDi2005] angepasst. Dabei ist es notwendig, die kleinste zu verifizierende Einheit von einem Frame auf eine GoP zu erweitern, da die notwendigen Kapazitätsanforderungen in P- und B-Frames nicht erfüllt werden können. Die hohen Anforderungen lassen sich mit der notwendigen Mindestlänge für Signatur und Hashwerte begründen. Dafür wurde der Einbettungskanal in den I-Frames erweitert.

Die Ergebnisse spiegeln die hohe Detektionsrate von Veränderungen im Video wieder. Dafür wurde eine Reihe von Manipulationen am Video durchgeführt. Einzig bei einer geringen Anzahl von Videos wurden einzelne Modifikationen in den Videos nicht erkannt. Ein weiteres Problem kann bei einzelnen Videos auftreten, wenn nach Überschreibung des Einbettungskanals das Original hergestellt wird. Diese einzelnen Spezialfälle zeigen noch vereinzelten Forschungsbedarf, um auch diese Schwachstellen zu beseitigen.

Trotzdem zeigt das Wasserzeichen eine hohe Detektionsrate bei relevanten Modifikationen, wie z. B. das Hinzufügen oder Entfernen vom Inhalt des Videos. Dabei ist es irrelevant, ob der Inhalt Teil des jeweiligen Frames, von GoPs oder Szenen ist. Zusätzlich spricht für das Wasserzeichen, dass es auch in Zukunft weiter eingesetzt werden kann, obwohl die Anforderungen für Hashverfahren und digitale Signaturen ständig angepasst werden müssen. Die durchschnittliche Kapazität von 14 kBits pro GoP zeigt, dass Signaturen der Länge von 8192 Bits und Hashwerte der Länge von 256 Bits ohne Einschränkungen in das Video eingebettet werden können.

9

Effizientes Watermarking für Transaktionswasserzeichen

Da die Verbreitung der Medien immer häufiger über das Internet durchgeführt wird, müssen digitale Wasserzeichen besondere Anforderungen in Bezug auf die Komplexität erfüllen. Wichtig ist es dabei, dass es durch das Einbetten des Wasserzeichens zu keiner Verzögerung kommt. Sie müssen schnell in die Medien eingebettet werden, aber gleichzeitig dürfen andere Eigenschaften wie Robustheit und Sicherheit nicht vernachlässigt werden. In diesem Kapitel werden verschiedene Strategien diskutiert, wie Wasserzeichen schnell in Videodaten eingebettet werden können.

Das Kapitel beinhaltet die Definition des Transaktionswasserzeichens und die technischen Möglichkeiten, die angewandt werden können, um möglichst schnell ein Wasserzeichen in Videodaten einzubetten. Dabei werden verschiedene Ansätze präsentiert, die sich je nach den Anwendungsszenarien aus Kapitel 3.1 richten. In diesem Kapitel werden drei Szenarien präsentiert, wo ein robustes Wasserzeichen angewendet wird. Der Fokus liegt aber eher auf der Echtzeitfähigkeit des Wasserzeichens.

9.1 Konzept eines Transaktionswasserzeichens

Da ein einzelner Verkauf eines digitalen Mediums auch als Transaktion bezeichnet wird, können dafür verwendete digitale Wasserzeichen auch als Transaktionswasserzeichen bezeichnet werden. Mit dem zusätzlichen Einbringen eines Wasserzeichens in die verkaufte Kopie des Mediums soll ein Nachweis einer möglichen unerlaubten Verbreitung nachgewiesen werden.

Transaktionswasserzeichen müssen aber spezielle Eigenschaften erfüllen, damit sie in die Anwendungen integriert werden können. Die erste wichtige Eigenschaft ist die Integration des Wasserzeichens in partiell dekodierte Videodaten. Bei Videodaten ist der De- und Enkodiervorgang der mit Abstand zeitaufwändigste Prozess während der Einbettung. Wenn der Einbettungsvorgang so aufgebaut ist, dass das Video nicht komplett dekodiert werden muss, kann damit Zeit eingespart werden [HaTh2003].

In Data-on-demand-Systemen werden die Daten in komprimierter Form bereitgestellt. Das bringt die Vorteile einer kürzeren Übertragungszeit zum Empfänger und einer geringeren Menge von benötigtem Speicherplatz für den Anbieter mit sich.

Die Kompression der Daten, der damit verbundene geringere Speicherplatzbedarf und die geringere Übertragungszeit erleichtern aber auch die Verteilung von Videodaten über das Internet. Mit Hilfe des Einsatzes von personifizierten digitalen Transaktionswasserzeichen werden notwendige Nachweismöglichkeiten geschaffen, die den Ursprung von möglicher unerlaubter Vervielfältigung nachzuweisen [StDi2002].

9.1.1 Anforderungen

Transaktionswasserzeichen müssen die folgende Anforderungen erfüllen:

(1) Hoher Robustheitsgrad
Das Wasserzeichen muss resistent gegenüber absichtlichen oder unabsichtlichen Manipulationen an den Videodaten sein. Wichtig sind dabei besonders Re-Enkodierung und DA/AD-Wandlung. Für Bilddaten können dafür anerkannte Werkzeuge für Robustheitstests, wie die StirMark Benchmark Suite [PeSt2001], genutzt werden.

(2) Hohe Sicherheit
Da kundenspezifische Informationen eingebettet werden, können beim Vergleich der Kopien die Unterschiede zwischen den Informationen identifiziert werden. Das Wasserzeichen ist dann an diesen Positionen angreifbar. Um eine hohe Sicherheit zu gewährleisten, muss die Wasserzeichennachricht so aufgebaut werden, dass sie resistent gegen sogenannte Koalitionsangriffe ist. Gegen Koalitionsangriffe resistente Wasserzeichennachrichten werden unter anderem in [BoSh1995] und [DiBe1999] vorgestellt.

(3) Hohe Datenrate
Zur Erkennung von Kunden, die an einem Koalitionsangriff teilgenommen haben, benötigen die Wasserzeicheninformationen einen speziellen Aufbau. Um eine genaue Identifizierung zu gewährleisten, wird eine ausreichend hohe Kapazität benötigt.

(4) Hohe Transparenz
Ein weiteres Kriterium ist die Transparenz des Wasserzeichens. Die Qualität des Datenmaterials sollte durch die Einbettung nicht vermindert werden.

(5) Echtzeit-Einbettung
Der Download-Vorgang darf durch die Einbettung des Wasserzeichens nicht signifikant verlängert werden. Um die Echtzeit-Einbettung zu ermöglichen, muss einerseits der Aufwand für den Dekodier- und Enkodiervorgang so gering wie möglich gehalten und andererseits die Komplexität des Wasserzeichens niedrig sein.

9.1.2 Bewertung existierender Wasserzeichenlösungen

Es existieren bereits Ansätze für Echtzeit-Wasserzeichen, die speziell für komprimierte Videodaten entwickelt wurden. Dabei bietet der Videostandard mehrere Möglichkeiten zur Markierung der Videos. In diesem Kapitel werden existierende Wasserzeichenlösungen und digitale Fingerabdrücke vorgestellt. Existierende Lösungen für digitale Fingerabdrücke werden in bezug ihrer Anwendbarkeit für die vorgestellten Anwendungsszenarien hin untersucht.

In [JoKu1998] wird durch eine leichte Modifizierung der Bewegungsvektoren in einem MPEG-4-Video das Wasserzeichen eingebettet. Die Wasserzeichenstärke bestimmt die Amplitude zur Modifikation der Bewegungsvektoren. Beim Auslesen des Wasserzeichens wird die Information aus dem ausgelesenen Bewegungsvektor, der Amplitude, und dem vorgegebenen Suchfenster bestimmt. In [Li1998] dagegen wird der Aufbau der GoP

verändert. Durch Veränderungen der Referenzen, was ein gleichzeitiges Ändern der Frametypen mit sich zieht, kann das Wasserzeichen in das Video eingebettet werden. Das Verfahren bettet eine kleine zusätzliche Sequenz in die GoP ein, ohne dass dabei Störungen während des Abspielvorganges entstehen.

Da Strukturinformationen aber sehr einfach veränderbar sind, bieten diese Wasserzeichen folglich nur eine geringe Robustheit. Ein einfacher Re-Enkodiervorgang kann sowohl die Bewegungsvektoren als auch die GoP-Struktur verändern. Die Kapazität des Wasserzeichens hängt von der Menge der Strukturinformationen ab. So können bei [Li1998] 6 Bits pro GoP eingebettet werden, d.h. also 12 Bits pro Sekunde. Die Strukturinformationen können unterschiedlich markiert werden und so eine Resistenz gegenüber Koalitionsangriffen aufbauen. Beide Verfahren sind echtzeitfähig, da der Dekodiervorgang des Videos nicht komplett durchgeführt werden muss. Man kann aber sagen, dass Verfahren dieses Typs, d.h. solche die MPEG-Strukturinformationen verändern, nicht für Transaktionswasserzeichen genutzt werden können. Sie unterstützen zwar Echtzeiteinbettung und besitzen eine gute Transparenz, haben aber eindeutige Mängel in der Robustheit.

In [DeCs1999] wurde ein Wasserzeichen entwickelt, das eine hohe Robustheit gegenüber geometrischen Operationen am Videomaterial aufweist. Speziell dafür wurde neben dem Wasserzeichen ein spezielles Muster, ein Template, in das Video eingebettet. Das Template dient zur Rückführung von Veränderungen, wie Wechsel der Framerate oder Skalierung der Frames. Nach dem Auslesen des Templates wird das Video in den ursprünglichen Zustand zurückgeführt und das Wasserzeichen ausgelesen. Das Video wird als dreidimensionales Objekt beschrieben. Das eigentliche Wasserzeichen wird auf einen 3D-Block über der Länge von 16 oder 32 Frames aufgetragen. Mit der Einteilung des zu markierenden Frequenzbereiches wird die Qualität der Transparenz bestimmt. Es kann zwischen statischen und bewegten Objekten unterschieden werden. Zwischen den beiden Objekttypen und innerhalb des Frequenzbereiches wird das Wasserzeichen aufgetragen.

Andere Ansätze, wie [SeLa2001] oder [LiLu2004], dekodieren das Frame nur teilweise. [SeLa2001] stoppt die Dekodierung nach der inversen Quantisierung und [LiLu2004] sogar vor der Quantisierung. Beide Verfahren betten das Wasserzeichenbit so ein, dass sie das Energieverhältnis von zwei pseudozufällig ausgewählten Regionen so manipulieren, in dem sie hochfrequente Bereiche manipulieren. Das Energieverhältnis beider Regionen bestimmt das Wasserzeichenbit. In [LuCh2005] wird ein anderes Konzept vorgestellt, bei dem das Wasserzeichen in Abhängigkeit vom jeweiligen Frame gebildet wird. Dabei wird zuerst ein Hash gebildet, der aus Eigenschaften der VLC-Werte des Frames besteht, extrahiert und mit den Wasserzeicheninformationen vermischt. Dadurch soll es für einen Angreifer erschwert werden, das Wasserzeichen durch verschiedene statistische Angriffe aufzuspüren. Das Wasserzeichen verändert die VLC-Werte des Makroblockes durch Addition oder Subtraktion eines konstanten Wertes.

Durch die Integration der Wasserzeichen in die Bilddaten ist es möglich, die Robustheit signifikant zu verbessern. [DeCs1999] ist zusätzlich robust bei einem Wechsel der Framerate und gegenüber zeitlichen Veränderungen. In [LuCh2005] wurde die Datenrate während einer Re-Enkodierung deutlich verringert und trotzdem konnte das Wasserzeichen weiter ausgelesen werden. Die Kapazität stellt kein signifikantes Problem dar. Bei [DeCs1999] und [LuCh2005] können sämtliche Frames markiert werden, wodurch eine hohe Kapazität zur Verfügung gestellt wird. Der PSNR-Wert für die Transparenz verschlechtert sich in

[LuCh2005] um 1 dB. Alle Verfahren dekodieren das Video nicht komplett. Zwar nimmt die Komplexität zu, aber durch den Ausgleich der jeweiligen Rechnerleistung ist es weiter möglich, das Wasserzeichen in Echtzeit einzubetten.

Die vorgeführten Verfahren eignen sich durchaus zum Einsatz als Transaktionswasserzeichen. Dabei ist besonders die verbesserte Robustheit, aber auch die gesteigerte Kapazität hervorzuheben. Dies ist auf die angewandte Markierungstechnologie zurückzuführen. Da deutlich mehr Bild- als Strukturinformationen vorliegen, bietet sie auch mehr Kapazität. Da für diese Markierungsmethode immer bessere Ansätze entwickelt werden, verbessert sich auch zunehmend die Robustheit der Wasserzeichen. Nur der Aspekt der Echtzeitfähigkeit muss durch optimierte Einbettungsstrategien realisiert werden.

Zusammenfassend kann man sagen, dass die Anforderungen an Transaktionswasserzeichen nur erfüllt werden können, wenn Bilddaten als Einbettungsquelle genutzt werden. Der Entwicklungsstand zeigt auch, dass die anderen Eigenschaften der Wasserzeichen gut ausgeprägt sind.

9.1.3 Anwendbarkeit existierender Fingerabdrücke

Da die Echtzeitfähigkeit bei der Erstellung von individuellen Markierungen ein wichtiger Aspekt ist, müssen auch vorhandene Fingerabdrücke als Basis für die Wasserzeicheninformation untersucht werden.

In Untersuchungen zum Thema Urheberrechtsschutz [PfFe2002] hat sich gezeigt, dass zum Schutz vor illegalen Kopien eine Identifizierung der Partei bzw. des Kunden von Vorteil ist, der das Material gekauft hat. Dafür können digitale Fingerabdrücke zur Identifizierung der Kunden eingebettet werden.

Ein digitaler Fingerabdruck ist eine binäre Sequenz mit speziellen mathematischen Eigenschaften, die als Wasserzeichen eingebettet wird. Mit Hilfe von digitalen Fingerabdrücken können Kunden sowohl zurückverfolgt werden, als auch Koalitionsangriffe mehrerer Kunden verifiziert werden. Ein Koalitionsangriff ist ein gezielter Angriff auf das Wasserzeichen, in dem mehrere Kunden durch Vergleich ihrer Kopien die Unterschiede zwischen den einzelnen Fingerabdrücken erkennen und fälschen. Dadurch ist es möglich, das Ausleseergebnis effizient zu verfälschen [St2003]. In [DiHa2001a] werden verschiedene Möglichkeiten beschrieben, wie Koalitionsangriffe Fingerabdrücke verfälschen oder zerstören können.

Zwei Beispielverfahren zur Erzeugung digitaler Fingerabdrücke sind:

(1) Der Schwenk-Algorithmus wird in [DiBe1999] ausführlich dargestellt. Der Algorithmus basiert auf in [BeRo1998] beschriebenen mathematischen Grundlagen. Die grundsätzliche Idee ist dabei, einen Bitvektor zu erzeugen, der nach einer Attacke sämtliche Beteiligten identifiziert. Unschuldige können hierbei durch Koalitionsangriffe nicht verdächtigt werden.

(2) Das Boneh-Shaw-Verfahren wird in [BoSh1995] vorgestellt. Auch hier wird ein Bitvektor erzeugt, der aber nur einen Angreifer identifiziert. Weiterhin besteht bei diesem Verfahren eine grundsätzliche Gefahr, Unschuldige zu verdächtigen. Diese kann aber durch längere Vektoren verringert werden.

Um eine Sicherheit gegenüber Koalitionsangriffen zu garantieren, müssen die Verfahren eine Vielzahl von Informationen in den Bitvektor integrieren. So verfügt das Schwenk-Verfahren über eine Teilinformation im Bitvektor, die Angreifer nicht aufspüren können, da sie in den verschiedenen Bitvektoren identisch ist. Diese Teilinformation identifiziert dann die Angreifergruppe. Da es aber eine Vielzahl von verschiedenen Angreiferkombinationen geben kann und diese auch eindeutig identifiziert werden sollen, hat das Verfahren den allgemeinen Nachteil, dass durch die Integration der Informationen in den Bitvektor die Länge exponentiell anwächst.

Im Boneh-Shaw-Verfahren wird die Menge der zu identifizierenden Angreifer auf genau eine Person eingeschränkt. Diese Vorgehensweise hat bei der Identifikation den Vorteil, dass eine geringere Kapazität benötigt wird.

In [TrWu2003], [WaWu2004] hat man erkannt, dass es sich als wenig sinnvoll erweist, einen Koalitionsangriff von jedem Angreifer mit jedem anderen Angreifer zu identifizieren. So wären bei zwei Angreifern und n Fingerabdrücken genau 2^n mögliche Angriffe zu untersuchen. Deshalb grenzt man die Angreifer erst in vordefinierte Gruppen ein und identifiziert dann die Angreifer in der Gruppe. Das kann darauf zurückgeführt werden, dass bei Koalitionsangriffen doch eher Angreifer aus den gleichen geographischen Gebieten oder der gleichen sozialen Ebene zusammenarbeiten, als per Zufall ausgewählte Angreifer.

Im Allgemeinen ist zu erkennen, dass es besser ist, die Menge der potentiellen Angreifer zuerst einzugrenzen und darauf basierend die Identifikation durchzuführen. Das bietet den eindeutigen Vorteil der Verringerung der einzubettenden Informationsmenge bei gleichbleibender Genauigkeit der Angreiferindentifikation. Durch die geringere Menge an einzubettenden Informationen pro Fingerabdruck kann die Detektionsrate des Wasserzeichens verbessert werden, da jeder Fingerabdruck redundant eingebettet werden kann. Aufgrund dieser signifikanten Verbesserung in Bezug zur Kapazität ist es möglich, digitale Fingerabdrücke auch für eine Vielzahl von Transaktionen anzuwenden, wie z. B. in Online-Shop-Szenarien.

9.2 Bewertung von Alternativen

Neben dem Einsatz von digitalen Wasserzeichen für Online-Transaktionen gibt es aber auch andere Schutzmechanismen, die zum Einsatz kommen.

Eine dieser Möglichkeiten ist die transparente Verschlüsselung. Bei dieser Methode werden Teile der Videodaten so verschlüsselt, dass nach der Verschlüsselung der Videoinhalt erkennbar ist. Wenn der Kunde das Video gekauft hat, wird ihm der Schlüssel für die Entschlüsselung zugesendet. Nach der Entschlüsselung liegt ihm das Video in Originalqualität vor. Danach ist aber die Kontrolle über das Video verloren gegangen. Es kann nicht mehr kontrolliert werden, ob der Kunde auch weiter die während des Verkaufs anerkannten Vertragsbedingungen einhält [ChLi2000].

Eine andere weit verbreitete und anerkannte Methode ist das Digital-Rights-Management (DRM). Eine Vielzahl von aktiven Onlineshops benutzt ein DRM-System, wie z. B. von Microsoft. Das DRM regelt die Zugriffsmöglichkeiten auf das digitale Medium. Es bestimmt die Möglichkeiten des Abspielens und der Verbreitung. Für den Abspielvorgang wird eine spezielle Software benötigt, die die DRM-Bestimmungen erkennen und verwalten kann. Ein DRM-System ist eine anerkannte Methode zum Schutz des Urheberrechts. Der Benutzer muss die Bestimmungen des Systems akzeptieren, um Zugriff auf die Medien zu erlangen. Ein DRM-Schutz kann nur innerhalb des DRM-Systems gewahrt werden. So kann am Ausgang der Sound- oder Videokarte das Audio- oder Videosignal ungeschützt abgefangen und separat aufgenommen werden.

Eine andere Anwendung, die zum Schutz vor unerlaubter Vervielfältigung genutzt wird, ist der Kopierschutz. Er wird neben digitalen Medien auch für Software eingesetzt. Hierbei ist das Ziel, das ein Kopieren der Daten von der CD nicht möglich ist. Aber auch diese Anwendung hat ihre Schwachstellen. Generell entsteht ein Problem, wenn es zu einer Beschädigung des Trägermediums kommt. Da es nicht möglich war, eine Sicherungskopie anzulegen, ist es nur möglich, eine neue CD vom Verkäufer anzufordern. Bei digitalen Medien entsteht zudem der Nachteil, dass an entsprechenden Peripheriekarten das Signal abgefangen und aufgenommen werden kann. Dadurch geht der Kopierschutz komplett verloren.

9.3 Lösungungsvorschläge für die Anwendungsszenarien

Die drei in Kapitel 3.1 vorgestellten Szenarien erfordern effiziente technische Lösungen, die im folgenden Kapitel vorgestellt werden. Die Echtzeit ist dabei der wichtigste Aspekt während der Einbettung des Wasserzeichens. Bei Videowasserzeichen ist der Aspekt des De- und Enkodierens des Videos der mit Abstand zeitaufwändigste Prozess des gesamten Vorganges. Deshalb erweist es sich als nützlich, diese Vorgänge so weit wie möglich zu verringern oder die Performance von mehreren Computern gleichzeitig zu nutzen [StHa2007].

Der Wasserzeichenkontainer ist eine effiziente Lösung für das in Kapitel 3.1.1 vorgstellte Anwendungsszenario des Onlineshops. Das Client-Server-Modell ist eine Lösung für das in

Kapitel 3.1.2 vorgestellte spezifische Onlineshop-Konzept mit der Vorverteilung von Videomaterial. Das Konzept des „Distributed Watermarking“ kommt für das Anwendungszenario für verteilte Bibliotheken (3.1.3) zum Einsatz.

9.3.1 Wasserzeichenkontainer für Videowasserzeichen

Die Idee des Wasserzeichenkontainers für Videowasserzeichen kann im Prinzip für jedes Wasserzeichen verwendet werden [Kl2006], [HaSt2008]. Das Verfahren ist besonders beim klassischen Onlineshop von Vorteil, da von einem Video eine Vielzahl von Kopien erstellt werden, die sich nur in der eingebetteten Wasserzeicheninformation unterscheiden. Bei jedem Wasserzeichen gibt es für den Einbettungsvorgang je eine Definition für das Einbetten des jeweiligen Wasserzeichenbits null und eins. Die Modifikation im Video ist vom Wasserzeichenalgorithmus vorgeschrieben; es können Pixelwerte, DCT-Blöcke oder auch DWT-Frequenzbereiche sein. Sämtliche andere Schritte sind während des Einbettungsprozesses identisch. Bei Videodaten wäre das z. B. das De- und Enkodieren der Videoframes.

9.3.1.1 Konzept des Wasserzeichenkontainers

Bei Verwendung des gleichen geheimen Schlüssels ist somit einzig die Wasserzeicheninformation selbst der Unterschied beim Einbetten von individuellen Wasserzeichen. Somit ist es bei gleichen Schlüssel möglich, ein universell vormarkiertes Video zu erstellen, aus dem ein markiertes Video erstellt werden kann. Dies ist dann für jede mögliche Wasserzeicheninformation durchführbar. Wird ein anderer geheimer Schlüssel verwendet, erfordert dies die Erstellung eines weiteren Wasserzeichenkontainers, da sich die Markierungspositionen zwischen verschiedenen Schlüsseln unterscheiden.

Für die Erstellung des Kontainers gibt es somit zwei Vorgänge, die Erstellung des Kontainers (Preparation Stage) und die Extraktion des markierten Videos (Finalization Stage). Die Erstellung des Kontainers wird nur einmal durchgeführt und hat keine spezifischen Echtzeitanforderungen. Die Echtzeit ist während der Extraktion wichtig. Da aber nur die notwendigen Daten aus dem Kontainer extrahiert werden, und deshalb nur Lese- und Schreiboperationen durchgeführt werden, kann eine vielfache Echtzeit erreicht werden.

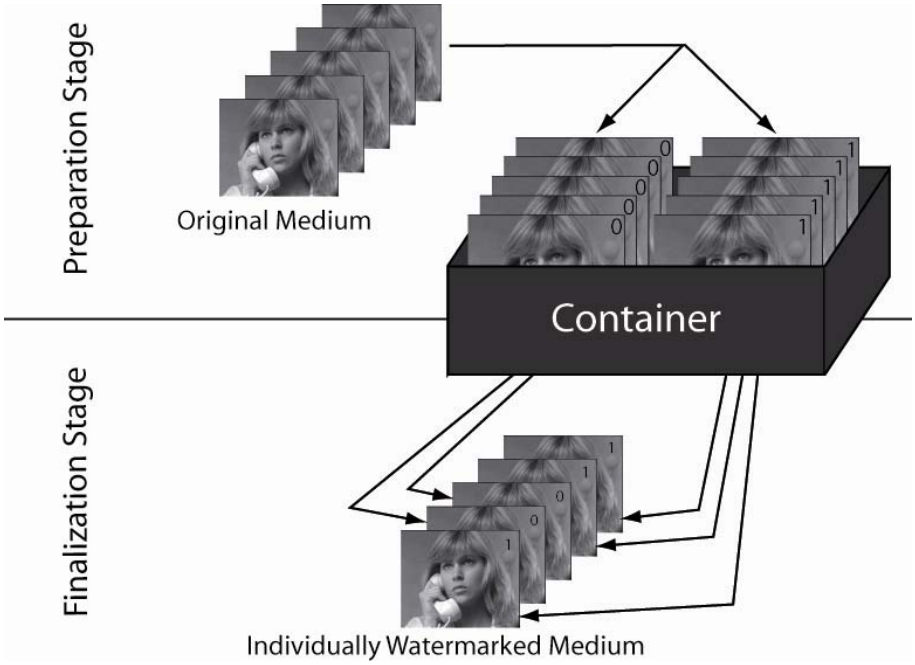

Abbildung 47: Erstellung des Kontainers und Extraktion eines Beispielvideos

Abbildung 47 stellt die beiden Teilprozesse noch einmal beispielhaft dar. Während des ersten Prozesses, der Erstellung des Kontainers, wird jedes Frame des Videos einmal mit einer null und einmal mit einer eins vormarkiert und im Kontainer abgespeichert. Bei diesem Beispiel wird zur vereinfachten Darstellung auf ein 1-Bit-Wasserzeichen pro Frame zurückgegriffen. Im zweiten Prozess, der Extraktion des Videos, werden basierend auf der Wasserzeicheninformation „10011“ die jeweiligen Frames aus dem Kontainer kopiert und zu einem neuen Video zusammengefasst.

Während der Erstellung des Kontainers K wird ein Video V in die Frames F zerlegt. Das Frame F_i wird daraufhin in die Markierungseinheiten M zerlegt. Die Markierungseinheiten entsprechen den Bereichen im Frame, an denen ein Wasserzeichenbit eingebettet wird. In MPEG-Videos sind es sehr oft DCT-Blöcke. Die Markierungseinheiten werden jeweils mit den Wasserzeichen W_0 und W_1 markiert. Die beiden Wasserzeichen präsentieren jeweils die Veränderungen, die durchgeführt werden, um an einer Markierungseinheit M die Wasserzeichenbits einzubetten. Zusammenfassend kann gesagt werden, dass der Kontainer K sich wie folgt zusammensetzt:

$$K = \left\{ F_i, F_i(M_{W_o}), F_i(M_{W_1}) \mid F_i \in V \right\} \qquad (28)$$

Da sich der Kontainer jetzt aber bis auf eine dreifache Größe gegenüber dem Ursprungsvideo vergrößern könnte, sollte die Menge soweit wie möglich verringert werden. Es kann damit erreicht werden, dass nicht die kompletten markierten Markierungseinheiten M, sondern nur die Differenzen zu den ursprünglichen Daten abgespeichert werden. Da Wasserzeichen nur sehr geringe Veränderungen am Trägermedium durchführen, würde somit eine deutliche Ersparnis beim Platzbedarf erreicht werden. Basierend auf den Erkenntnissen modifiziert sich der Kontainerinhalt wie folgt:

$$K = \left\{ F_i, D(F_i(M_{W_o}) - F_i), D(F_i(M_{W_1}) - F_i) \mid F_i \in V \right\} \quad (29)$$

Mit D ist die jeweilige Differenz zwischen den markierten und den ursprünglichen Markierungseinheiten gemeint.

Nachdem die Kontainerdaten erstellt wurden, ist es notwendig, sie mit Hilfe eines definierten Speicherungsprotokolls in den Kontainer abzulegen. Ziel muss es sein, während der Extraktion die notwendigen Daten zur Erstellung des markierten Videos schnell zu finden und zu lesen. Als sehr effizient erweist sich eine vordefinierte Tabelle, die für jeden Datensatz Startposition und Datenlänge der Markierungseinheit im Datenstrom des Kontainers und die kodierten Markierungsdaten für beide Wasserzeichenbits beinhaltet. Aufgrund der möglichen großen Datenmenge bei Videodaten sollte die Tabelle separat für jedes einzelne zu markierende Frame abgespeichert werden. Um ein zeitverzögerndes Parsen durch den Videostrom bis zu einer neuen Frametabelle zu verhindern, beinhaltet die videobasierte Tabelle auch die Starteinträge für die Frametabellen. Basierend auf dem Tabellenaufbau ist es sehr einfach, durch die Startadresse und Datenlänge einen Austausch der Originaldaten und der jeweiligen Markierungseinheit durchzuführen.

Während der Extraktion wird das markierte Video aus dem Kontainer K erstellt. Dazu ist nur die Kontainerdatei und die Wasserzeicheninformation W_I notwendig. Die Wasserzeicheninformation wird in die Einzelbits zerlegt. Für die Erstellung des markierten Videos werden für jede Markierungseinheit die notwendigen Daten aus der Kontainerdatei kopiert und in das Video geschrieben. Bleibt die Markierungseinheit unmarkiert, werden die Originaldaten kopiert, andernfalls die markierten Daten des augenblicklichen Wasserzeichenbits. Da es nur Lese- und Schreiboperationen sind, wird die Echtzeitfähigkeit durch die Effizienz des Kopiervorganges bestimmt.

9.3.1.2 Besonderheiten für MPEG-Videos

Bei verschiedenen Videoanwendungen wird jeweils ein spezielles MPEG-Profil angewandt. So wird z. B. bei der DVD das Main-Profil verwendet [ISO1995]. MPEG-Videos, die nach einem Profil kodiert werden, bestehen aus einem festgelegten Parametersatz. So sind z.B. die Frame- und Datenrate festgelegt. Das Videomaterial ist nur abspielbar, wenn die Parameter, mit denen das Video erstellt wurde, den Bestimmungen des Profils entsprechen. So darf die Datenrate durch das Einbetten eines Wasserzeichens nicht verändert werden. Wenn die Datenrate in den markierten Frames kontinuierlich höher dem des unmarkierten Frame ist, kann dies die Synchronisierung zwischen Bild und Ton beinflussen. Deshalb müssen die Frames der markierten DVD die gleiche Datenmenge wie die Frames der unmarkierten DVD ausweisen.

Das für die Erstellung des Kontainers angewandte Videowasserzeichen muss demnach eine Datenratenkontrolle beinhalten. Eine Möglichkeit zur Datenratenkontrolle ist die Anwendung eines Tiefpassfilters, der hochfrequente Bereiche im Frame entfernt. Aufgrund der geringen Auswirkungen auf die visuelle Qualität werden Manipulationen im hohen Frequenzbereich nur sehr gering wahrgenommen. Die Verringerung der Datenrate des Frames schafft Platz für die Einbettung des Wasserzeichens. Der ursprüngliche Ansatz wurde schon zum Einbetten

von Wasserzeichen genutzt, in dem durch Entfernen von hohen Frequenzbereichen die gesamte Energie der DCT-Blöcke verändert wurde [SeLa2001].

Um aber die Qualitätsminderung so gering wie möglich zu halten, werden so wenige hohe Frequenzen wie möglich entfernt. Es wird untersucht, ab wann das markierte Frame mindestens die gleiche oder eine niedrigere Bitrate als das Originalframe besitzt. Dafür werden die hohen Frequenzen in inverser Richtung des zig-zag-Scans entfernt und danach das Wasserzeichen auf den Block aufgetragen. Nach Durchführung des Prozesses für sämtliche zu markierende DCT-Böcke wird die Bitrate des markierten Frames ermittelt und sie mit der Originalrate verglichen. Ist sie zu hoch, wird im nächsten Durchlauf der Schnittpunkt um eine Position im inversen Scanmodus vorverlegt, d.h. wenn z.B. im letzten Durchlauf ab Position 60 die Frequenzen weggeschnitten wurden, würde dann im nächsten Durchlauf ab Position 59 der Schnitt durchgeführt werden. Ist sie hingegen gleich oder kleiner als die Originalrate, ist der ideale Schnittpunkt für das Frame erreicht. Auch das Einbetten des Wasserzeichens in das Frame führt dann nicht zu einer Erhöhung der Datenrate.

Zusätzlich wird aber auch ein minimaler Schnittpunkt festgelegt, der aussagt, ab welchem Frequenzwert der Schnitt durchgeführt werden kann, ohne zu deutliche Qualitätseinbußen hervorzurufen. In [SeLa2001] liegt der Wert z. B. bei Position 40. Tritt hingegen der Fall auf, dass auch nach Entfernen der Frequenzwerte ab dem minimalen Schnittpunkt die Datenrate immer noch zu hoch ist, wird das Frame als nicht markierungsfähig eingestuft und nur die Originaldaten in den Kontainer geschrieben.

Ein weiterer zu beachtender Punkt sind die DC-Werte der DCT-Blöcke. Sie werden nicht als absoluter Wert, sondern als relative Differenz zum vorherigen DC-Wert abgespeichert. Wenn durch das Einbetten des Wasserzeichens in einen Makroblock die DC-Werte verändert werden, hat das auch direkte Auswirkungen auf die DC-Werte des nachfolgenden Makroblockes. Die abzuspeichernden Differenzwerte müssen dementsprechend angepasst werden. Wenn keine Korrektur durchgeführt wird, besteht die Möglichkeit, dass die einem markierten Block nachfolgenden Blöcke leichte Helligkeitsveränderungen aufweisen. Der Kontainer wird deshalb so aufgebaut, dass neben dem ursprünglichen Makroblock auch zwei zusätzliche Blöcke abgespeichert werden. Sie beinhalten die Daten des Originalblockes, aber mit der Besonderheit des angepassten DC-Werts, wenn im vorherigen Makroblock ein Wasserzeichen eingebettet wird. Dadurch kann die relative Differenz des DC-Werts zum Vorgänger wieder angepasst werden, wenn dieser durch das Wasserzeichen verändert wird.

9.3.1.3 Performance des Videocontainers

Um die Echtzeitfähigkeit des Videowasserzeichenkontainers zu evaluieren, wurden je 160 Markierungsdurchläufe auf zwei verschiedenen Computern durchgeführt. Ziel ist es, nachzuvollziehen, ob eine durchschnittlich gleichschnelle Markierung der Videos möglich ist.

Abbildung 48 stellt die Ergebnisse der Evaluierung dar. Dabei ist deutlich zu erkennen, das der Echtzeitfaktor im Durchschnitt bei 100 bis 120 fach Echtzeit liegt. Die Ergbnisse zeigen aber auch, dass eine hohe Standardabweichung vorherrscht, d.h. das die Geschwindigkeit des Einbettens des Wasserzeichens teils sehr unterschiedlich ist. Die Standardabweichung liegt

bei den durchgeführten Tests bei 30, wobei der Durchschnitt den Wert 115 hat. Um eine sichere und ausagekräftige Testergebnisse zu bekommen, haben die Wasserzeichennachrichten keine signifikanten Unterschiede und es wurde nur ein Wasserzeichenalgorithmus genutzt.

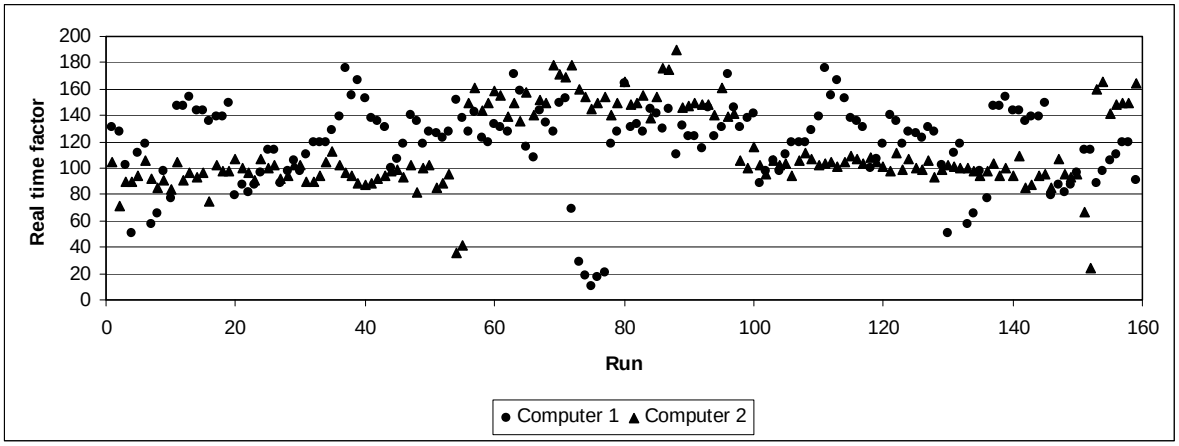

Abbildung 48: Evaluierung der Echtzeitfähigkeit des Wasserzeichenkontainers

Die Ergbnisse zeigen auch neben der großen Standardabweichung eindeutig, dass das Video mit einem Vielfachen der Echtzeit markiert werden kann. Aufgrund dieser hohen Einbettungsgeschwindigkeit eignet sich das Konzept des Wasserzeichenkontainers ideal für einen Onlineshop, wo bei einer Vielzahl von gestarteten Downloads das Wasserzeichen ohne Verzögerung markiert werden muss.

9.3.2 Client-Server-Modell für Videodaten

Mit dem Client-Server-Modell ist es möglich, dass Wasserzeichen ohne großen Aufwand in die Videodaten einzubetten, ohne dass dabei das Einbettungsprogramm verteilt werden muss. Dabei ist es besonders wichtig, nur die wasserzeichenrelevanten Operationen auf dem Server auszuführen. Sämtliche anderen Operationen werden auf dem Client ausgeführt. Da das Einbetten des Wasserzeichens nur ein kleiner Teil des gesamten Einbettungsprozess' darstellt, können die durchzuführenden Operationen auf dem Server in Echtzeit durchgeführt werden. Das kann besonders in Szenarien wichtig sein, in denen eine Vielzahl von Anfragen zum Einbetten von Wasserzeichen auftreten.

Abbildung 49 stellt eine spezielle technische Umsetzung des Client-Server Modells vor [HaSt2005a]. Das in Abbildung 49 vorgeführte Beispielvideo ist absichtlich visuell verschlechtert worden, um es als Vorführvideo ausgeben zu können. Erst mit der Übermittlung der fehlenden Videodaten vom Server zum Client wird das Video vollständig. Im Vorführfilm (linke Seite) werden in jedem I-Frame (schwarz dargestellt) vereinzelte DCT-Blöcke gezielt verändert, sodass die visuelle Qualität des Frames verringert wird. Das kann zum Beispiel durch den Austausch der Originaldaten durch Rauschsignale erreicht werden. Da das I-Frame die Referenz für die restlichen Frames der GoP ist, werden während des Abspielens die eingebrachten Rauschmuster in die nachfolgenden Frames automatisch übertragen.

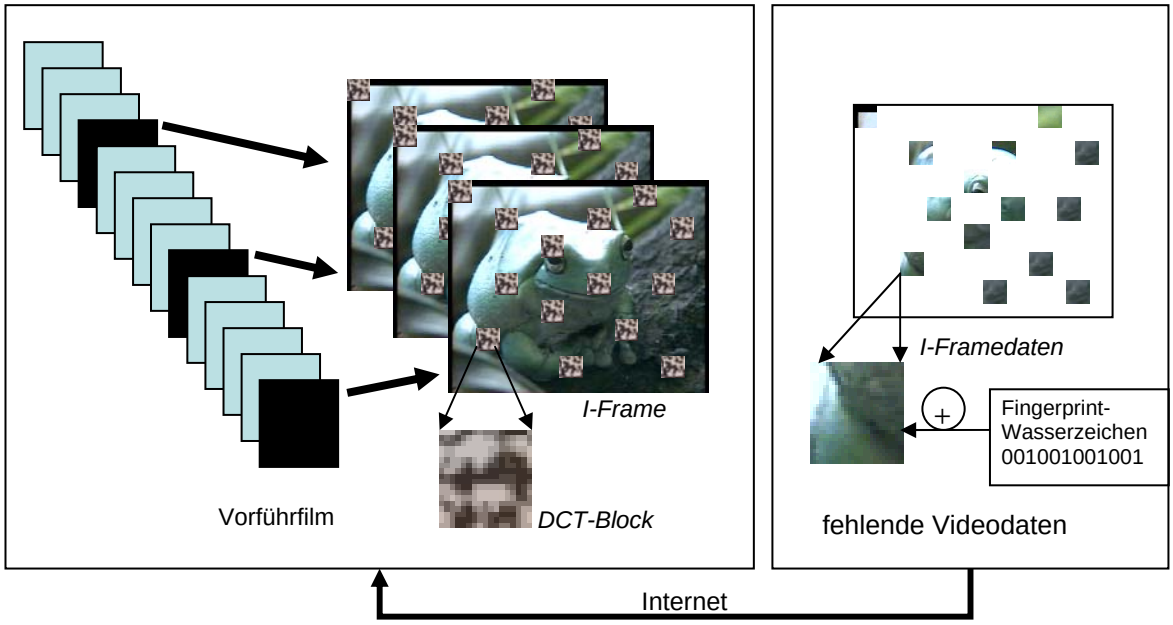

Abbildung 49: Integration der fehlenden Videodaten mit Wasserzeichen in das Video

Auf der Serverseite befinden sich die im Vorführfilm fehlenden Videodaten (rechte Seite). Der Wasserzeichenalgorithmus bettet den Fingerabdruck in die ausgewählten DCT-Blöcke der I-Frames ein. Die markierten Daten werden sicher über das Internet zu dem Kunden übertragen und ersetzen im Vorführvideo die Rauschdaten. Speziell für das MPEG-Datenformat existieren eine Vielzahl von Wasserzeichenansätzen für die Luminanz-DCT-Blöcke. Neben den kodierten Blockdaten muss nur die Position des DCT-Blockes innerhalb des Frames mit übertragen werden. Auf der Clientseite werden die übertragenen Blöcke direkt in das Video kopiert. Zudem weisen kodierte DCT-Blöcke häufig eine geringe Datenmenge hin. Dadurch ist die Menge der übertragenen Daten gering, so dass auch die Echtzeitfähigkeit erreicht werden kann.

9.3.3 Distributed Watermarking in Grid-Netzwerken

Bei Anwendungen dieser Art wird der Einbettungsvorgang an andere Computerressourcen weitergegeben. Dabei kann der Client als Initiator verstanden werden und die relevante Integration der Wasserzeicheninformation wird auf dem Server durchgeführt. Das führt zu einer Ressourcenteilung. Die Resourcen sind die Computer in Netzwerken, die ihre Rechenstärke zur Verfügung stellen.

Eine Form von verteilten Netzwerken ist das sogenannte Grid-Computing. In [FoKe1998] wird ein Grid-Netzwerk wie folgt beschrieben: Bei einem Grid handelt es sich um eine Infrastruktur, die eine integrierte, gemeinschaftliche Verwendung von meist geographisch auseinander liegenden, autonomen Ressourcen erlaubt. Ein Betriebssystem ermöglicht die Kommunikation zwischen den einzelnen Ressourcen des Netzwerks und managt den Informationsfluss zwischen den Komponenten.

Die Idee, wie man die zur Verfügung stehenden Grid-Netzwerke nutzen kann, besteht darin, dass die Einbettung des Wasserzeichens in einem Client-User-Interface als Job definiert wird

und zum Grid-Netzwerk gesendet wird. Die dafür vorgesehene Komponente, das Workload-Management-System (WMS), empfängt die Jobanfragen. Parallel verwaltet es die zur Verfügung stehenden Computerressourcen zur Durchführung eingereichter Jobs. Die dabei relevanten Ressourcen zur Jobausführung sind die Computing-Elemente. Sie nehmen die Anfragen vom WMS entgegen und leiten sie an die ihnen zur Verfügung stehenden Workernodes weiter. Auf den Workernodes selbst wird der Job ausgeführt. Abbildung 50 zeigt einen Ausschnitt des gLite [Gl2007] Grid-Netzwerkes. Dabei werden die Ressourcen dargestellt, die für die Einbettung des Wasserzeichens genutzt werden.

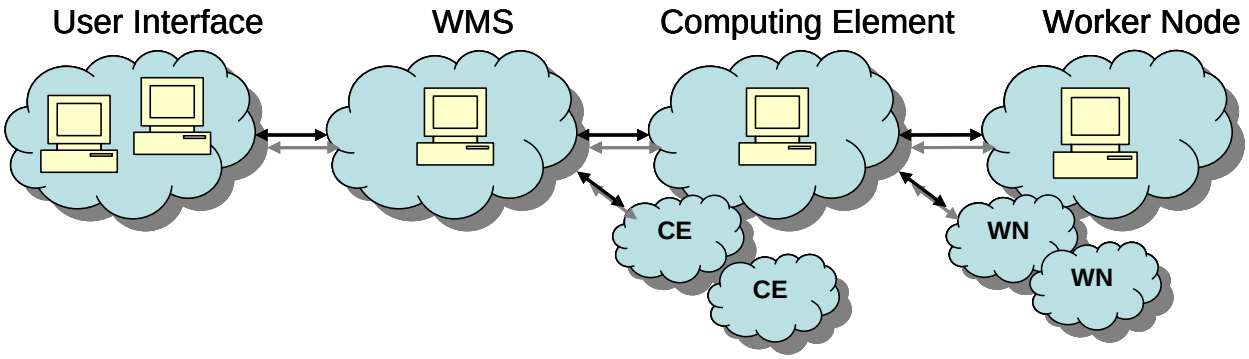

Abbildung 50: gLite Netzwerkelemente für Distributed Watermarking [Gl2007]

Diese Strategie kann als „Distributed Watermarking“ bezeichnet werden, da der Arbeitsaufwand an ein verteiltes Netz weitergegeben wird.

Neben dem klassischen Einbetten des Wasserzeichens kann die Strategie auch dahingegen erweitert werden, dass die Videos zuerst aufgeteilt und die Teilvideos dann als separate Jobs zum Grid-Netzwerk gesendet werden (Abbildung 51). Dadurch ergeben sich folgende drei Teilprozesse:

(1) Splitting bzw. Aufteilen des Videos V in n Teilvideos (Job A)
(2) Watermarking der n Teilvideos (Job B1 bis Bn)
(3) Merging bzw. Zusammenfügen der n Teilvideos zu einem Video V’ (Job C)

Während der Aufteilung wird das Video (Job A) in n Teilvideos aufgeteilt. Wie das Video in einzelne Teilvideo aufgeteilt wird basiert auf dem angewandten Wasserzeichen. Wenn z. B. das in Kapitel 7 vorgestellte Wasserzeichen genutzt wird, erweist es sich als sinnvoll, das Video jeweils am Beginn einer GoP zu teilen. Nach der Zerlegung des Videos in n Teilvideos wird für jedes Teilvideo ein Job (B1 bis Bn) generiert und zum Netzwerk gesendet. Das WMS-System leitet die erhaltenen Jobanfragen an frei zur Verfügung stehende Computing-Elemente weiter, auf denen dann das Wasserzeichen in den Teilvideos eingebettet wird. Wenn alle n Teilvideos markiert wurden und die jeweiligen Jobs B1 bis Bn abgeschlossen ist, wird mit Hilfe eines weiteren generierten Jobs C das markierte Video aus den n Teilvideos zusammengesetzt.

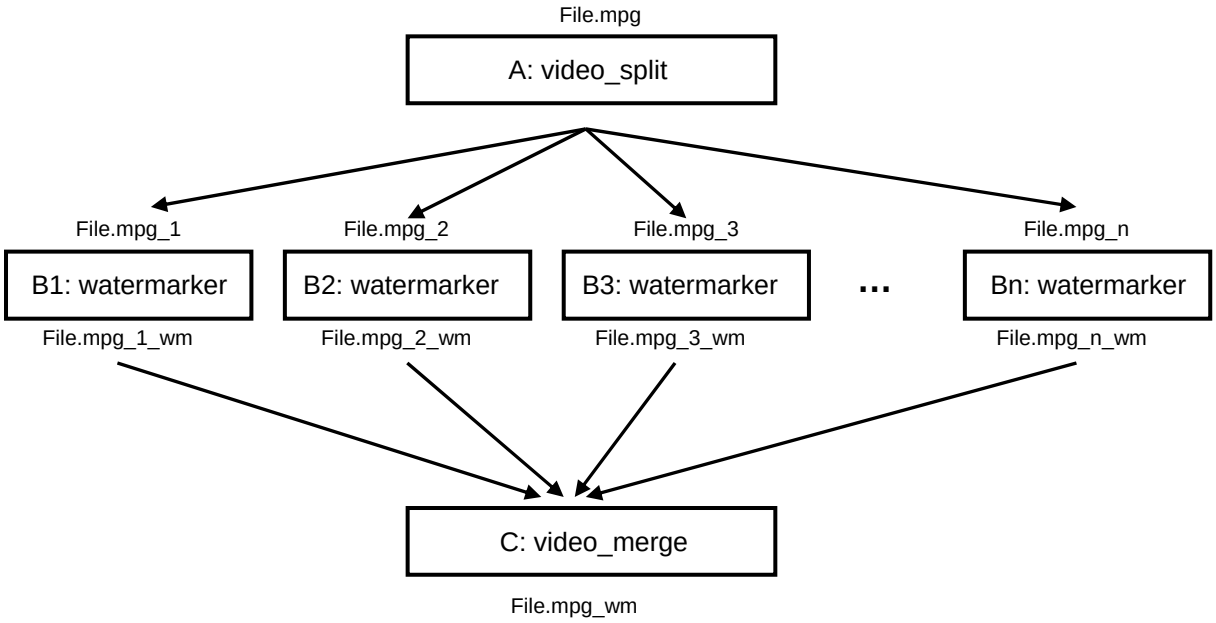

Abbildung 51: Strategie verteiltes Markieren

9.3.3.1 Performance des Distributed Watermarking

Die Analyse setzt sich aus zwei Testschritten zusammen. Zuerst wird untersucht, ob die Einbettung des Wasserzeichens in mehrere Teilvideos Vorteile gegenüber der Einbettung in das Komplettvideo bringt. Aus den gewonnen Kenntnissen wird eine komplette Markierung einer Videosammlung durchgeführt, um einen Vergleich zu einer auf einem Computer lokal durchgeführten Einbettung zu ermöglichen.

Die dafür verwendete Videosammlung besteht aus 535 ASF-kodierten Videos. Der erste Teiljob, das Aufteilen, erweitert sich aufgrund des unterschiedlichen Videokodex um den Prozess einer Konvertierung von ASF nach MPEG. Da dieser Prozess aber für sämtliche Videos durchgeführt werden muss, sind die Testergebnisse weiter vergleichbar. Das Videomaterial wurde innerhalb des Projekts DILIGENT [Di2007] zur Verfügung gestellt. Die Abspiellänge liegt zwischen 0:53 bis 37:35 Minuten. Um die Performance der Markierung eines Komplettvideos gegenüber der von mehreren Teilvideos zu messen, wurde ein Beispielvideo mit der Abspiellänge von 23 Minuten herangezogen. Das Grid-Netzwerk wurde auf drei Computing-Elemente begrenzt und basiert auf dem gLite System [Gl2007]. Die drei Computing-Elemente haben dabei jeweils fünf, sieben und 24 Workernodes. Die Zuteilung der Jobs an das jeweilige Computing-Element bestimmt aber das WMS-System.

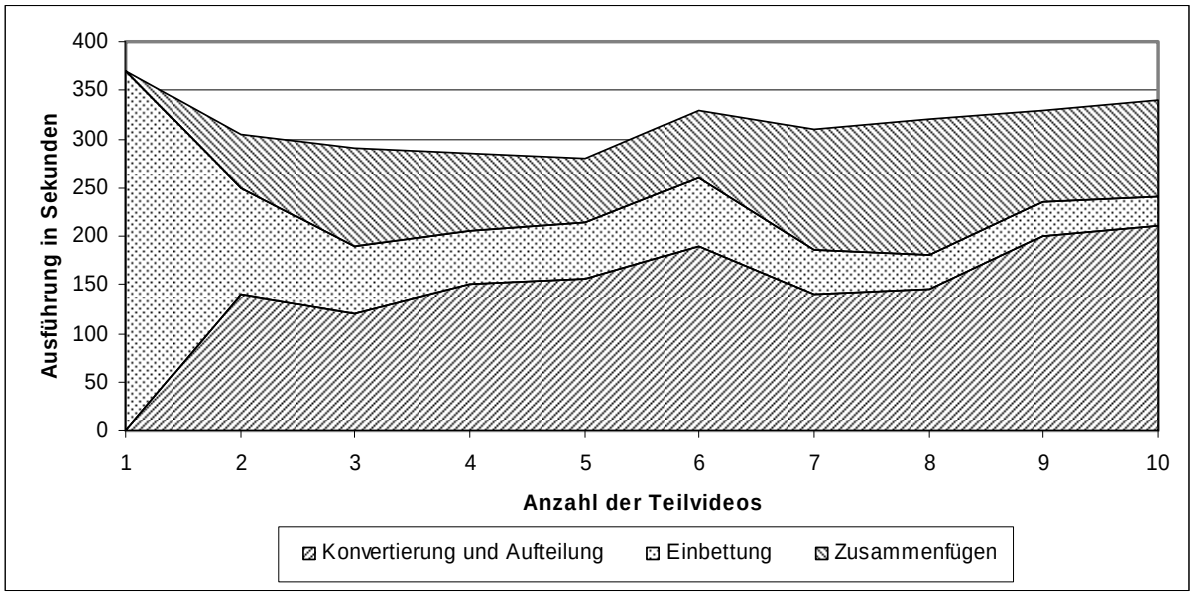

Abbildung 52: Zeitverteilung der Teilprozesse

Wie aus Abbildung 52 zu entnehmen, ist durch eine Aufteilung des Videos in n Teilvideos eine leichte Verbesserung der Performance zu erkennen. Die besten Ergebnisse zeigen sich bei einer Aufteilung von drei bis fünf Teilvideos. Die Teilvideos haben eine Größe von acht bis zehn MB. Das durchschnittliche Einbetten des Wasserzeichens in die Teilvideos wird immer schneller, je mehr Teile vorliegen. Der Vorteil wird aber wieder aufgebraucht, da durch das Aufteilen des Videos in eine hohe Anzahl an Teilen der Teilprozess der Konvertierung und des Aufteilens mehr Zeit benötigt. Die Aufteilung und das Zusammenfügen verbrauchen also die Zeit, die durch das effiziente Einbetten eingespart wurde.

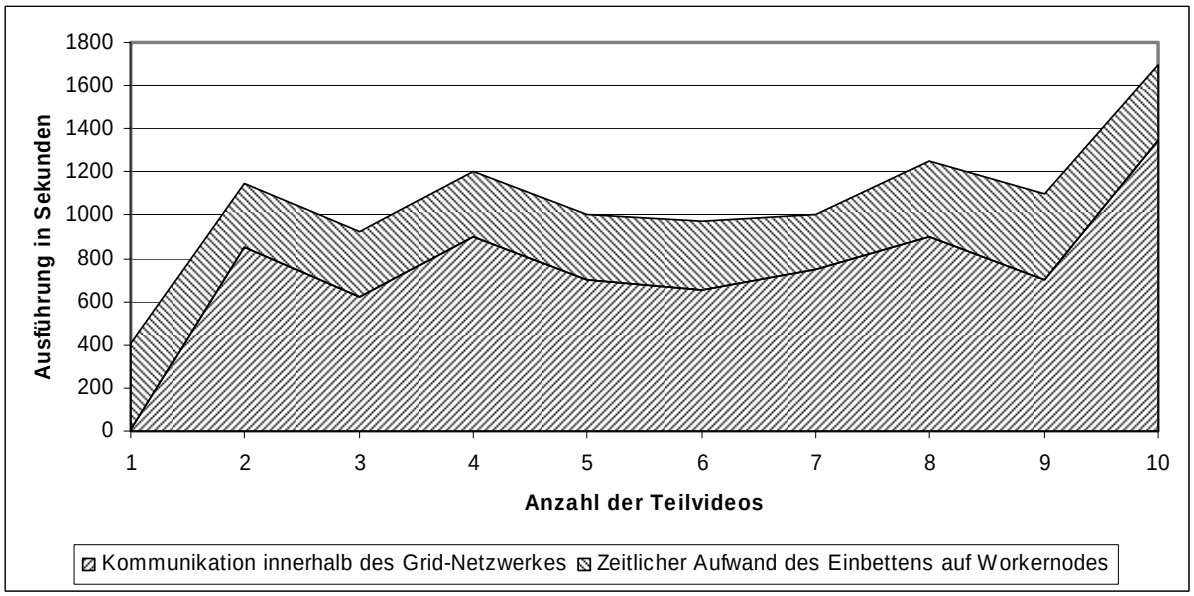

Abbildung 53: Zeitlicher Aufwand der Teilprozesse

Abbildung 53 zeigt hingegen den Nachteil, der durch das Aufteilen und Zusammenfügen der Teilvideos auftritt. Während die Bearbeitungszeit für die beiden Videoprozesse und das Watermarking konstant bleibt, nimmt die Kommunikation innerhalb des Grid-Netzwerks deutlich zu. Die Veränderungen sind teilweise so gravierend, dass ein Aufteilen des Videos zu deutlichen Performanceeinbußen führt. Zusätzlich verschlechtern sich die Ergebnisse ab einer Menge von 10 Teilvideos noch einmal. Es muß aber gesagt werden, dass die Verschlechterung nur auf die Kommunikation innerhalb des Grid-Netzwerks zurückzuführen ist.

In einer weiteren Untersuchung wurde evaluiert, ob es ab einer bestimmten Dateigröße vorteilhafter ist, das Video in Teilvideos aufzuteilen. Dafür wurde in ein Video mit unterschiedlichen Größen das gleiche Wasserzeichen eingebettet. Die Ergebnisse zeigten, dass es sich erst ab einer Dateigröße von 900 bis 1000 MB lohnt, das Video in zwei Teilvideos aufzuteilen.

Aus den gewonnenen Erkenntnissen wurde jedes Video der Kollektion als gesamtes Video markiert. Das Ergebnis wird mit einem lokalen Einbettungsprozess verglichen. Tabelle 12 stellt die Ergebnisse des Vergleichs dar. Dabei wurde die Videokollektion einmal komplett lokal markiert. Für das Distributed Watermarking waren drei Durchläufe (run) notwendig, da zwei der drei Durchläufe nicht komplett abgearbeitet wurden, was man jeweils an den 96 bzw. 82 Fehlversuchen (abort) erkennt. Zudem fällt auch noch die relativ hohe Anzahl von Resubmissions bzw. Wiederaufnahmen auf. Der Vorgang Resubmission wird immer dann vom WMS-System gestartet, wenn der vorherige Einbettungsversuch nicht vollständig abgeschlossen wurde. Folglich führt es zu einem erneuten Einbettungsversuch.

lokal	run 1	run 2	run3
535 videos	535 videos	535 videos	535 videos
0 abort	96 abort	0 abort	82 abort
535 done	439 done	535 done	455 done
0 resubmit	120 resubmit	86 resubmit	120 resubmit
103927 Sek	8918 Sek	9235 Sek	19640 Sek

Tabelle 12: Vergleich lokales und distributed Watermarking

Zusammenfassend ist durch den dennoch möglichen Vergleich zu erkennen, dass aufgrund der Nutzung des Grid-Netzwerks eine Vielzahl von Videos schnell und effizient markiert werden können. Während die sequentielle Abarbeitung der Einbettung des Wasserzeichens in 535 Videos auf dem lokalen Computer ungefähr 29 Stunden benötigte, waren durch Nutzung des Grid-Netzwerks alle Videos schon nach ungefähr zwei bis drei Stunden markiert. Selbst bei Wiederaufnahme von abgebrochenen Jobs liegt die Gesamtzeit immer noch unter der Zeit, die für das lokale Einbetten notwendig ist.

9.4 Zusammenfassung

In diesem Kapitel wurde das Transaktionswasserzeichen vorgestellt und verschiedene technische Lösungen präsentiert.

Die Ausprägungen der Eigenschaften eines Transaktionswasserzeichens werden besonders von Anwendungsszenarien wie Onlineshops oder digitalen Bibliotheken bestimmt. Besonders wichtig ist dabei die Echtzeit-Einbettung, da die hohe Anzahl von gleichzeitigen Einbettungsvorgängen entsprechend hohe Anforderungen an die Computerressourcen stellt.

Bei den präsentierten Lösungen wurden drei Ansätze vorgestellt und auch teils evaluiert. Sie sind keine wasserzeichenspezifische Lösungen, sondern sie setzen die in den Szenarien erkannten Anforderungen um. Der Wasserzeichenkontainer ist für eine Anwendung geeignet, wo eine beliebige Anzahl von Kopien von einem Video erstellt und markiert werden soll, wie z. B. im klassischen Onlineshop (Kapitel 3.1.1). Da die einzigen Unterschiede die Wasserzeicheninformationen sind, kann durch einen Vorarbeitungsprozess ein Video erstellt werden, das sowohl die Original, als auch die beiden markierten Kopien enthält, die je mit dem Bit null und eins markiert sind. So müssen während der Extraktion des markierten Videos nur noch die entsprechenden markierten Daten mit den Originaldaten für das augenblickliche Bit zusammenkopiert werden. Da jegliche Videobearbeitung nicht mehr notwendig ist, wird der Markierungsvorgang sehr schnell abgeschlossen.

Das Client-Server-Modell kann für das Szenario von Kapitel 3.1.2 angewandt werden. Die relevanten Videodaten werden nach Kaufbestätigung mit dem individuellen Wasserzeichen markiert und zum Client transferiert. Dort werden sie in das markierte Video integriert. Dieses Modell ist nur für das spezielle Onlineshop-Szenario sinnvoll, dass in dieser Arbeit vorgestellt wurde. Der Vorteil des Modells liegt darin, dass sämtliche sicherheitsrelevanten Verarbeitungsprozesse auf dem Server durchgeführt werden.

Mit Hilfe von Grid-Netzwerken kann das Szenario von digitalen Bibliotheken aus Kapitel 3.1.3 umgesetzt werden. Die gewünschte skalare Markierung der Medien kann durch die verfügbaren Ressourcen des Grid-Netzwerks realisiert werden. Die Einbettungsvorgänge werden als Jobanfrage an das Grid-Netzwerk geschickt und dort durchgeführt. Die Methode kann mit einem verteilten Einbetten auf mehreren zur Verfügung stehenden Ressourcen verglichen werden. Die Tests haben gezeigt, dass eine Vielzahl von Videos im Grid-Netzwerk schneller als mit lokalen Ressourcen markiert wird. Der einzige Nachteil liegt im Kommunikationsaufwand innerhalb des Netzwerks, der einzelne Einbettungsvorgänge ineffizient macht.

10 Bewertung der präsentierten Ansätze

Diese Arbeit beinhaltet Weiterentwicklungen von robusten und reversiblen Videowasserzeichen. Eine zuerst durchgeführte Evaluierung hat Schwachstellen an existierenden Wasserzeichen erkannt, die dann als Basis für die Ansätze der Arbeit dienen. Die Grundideen für beiden präsentierten Wasserzeichen stammen von Bildwasserzeichen und wurden für MPEG-2 Video optimiert. Die Konzentration auf den MPEG-2 Standard beruht auf der augenblicklichen hohen Akzeptanz von Videos dieser Art. Es kann inzwischen behauptet werden, dass Bildwasserzeichen einen hohen Entwicklungsstand haben. Zusätzlich erweist es sich als sinnvoll, die Ähnlichkeit zwischen JPEG und MPEG zu nutzen und den Fokus auf Wasserzeichen für JPEG-Bilder zu legen.

In diesem Kapitel wird eine Bewertung der entwickelten Ansätze dieser Arbeit durchgeführt. Ziel ist es zu untersuchen, in wieweit die Ansätze der durch die Evaluierung erkannten Schwachstellen gelöst hat. Dabei soll nicht nur der Erfolg sondern auch die nicht gelösten Aspekte bzw. dazugekommenen Probleme aufgeführt werden. Zur Vereinfachung wird die Bewertung anhand der Auflistung der Schwachstellen aus der Evaluierung durchgeführt.

(1) Kapazitätsmangel

Eine Vielzahl von Wasserzeichen markieren nur I-Frames. Da sie nur einen geringen prozentualen Anteil im gesamten Video darstellen, ist die Kapazität begrenzt bzw. das Wasserzeichen kann nicht redundant eingebettet werden. Eine hohe Redundanz verbessert zudem die Robustheit des Wasserzeichens, wodurch das Ausleseergebnis als sicherer eingestuft werden kann. Beim Einsatz von digitalen Fingerabdrücken kann eine begrenzte Kapazität zu einer mangelhaften Sicherheit führen. So verlängert sich z. B. in [DiBe1999] die Länge der Wasserzeichennachricht exponentiell bei gleichzeitiger Erhöhung der Sicherheitsstufe des Fingerabdrucks. Eine höhere Sicherheitsstufe ermöglicht aber den Nachweis der Angreifer bei Koalitionsangriffen mit einer höheren Anzahl von Beteiligten.

Die Arbeit präsentiert eine Lösung, die die Markierung von I-Frames auf weitere Frames im Video erweitert. Durch die erweiterte Markierung wird die Kapazität des Verfahrens verbessert. Die Lösung ist für P-Frames anwendbar. In den P-Frames weisen die korrekt ausgelesen Wasserzeichenbits eine höhere Erfolgsquote als die falsch ausgelesen Bits auf. Es muss aber gesagt werden, dass die Ausleseergebnisse sich verschlechtern, wenn spätere P-Frames der GoP markiert werden.

Es ist aber auch ratsam, die Markierung auf B-Frames zu erweitern, was in dieser Arbeit noch nicht betrachtet wurde. Schließlich bilden sie die größte Anzahl von Frames in einer GoP. Augenblicklich kann nicht gesagt werden, welche möglichen Anpassungen am Wasserzeichenalgorithmus durchgeführt werden müssen. Üblicherweise besitzen B-Frames nur eine sehr geringe Menge an Inhalt, die zusätzlich effizient komprimiert wird.

Aufgrund der unterschiedlich starken MPEG-Kompression erweist es sich als notwendig, den Skalierungsfaktor der MPEG-Kompression in den

Einbettungsprozess mit zu integrieren. Die Ursprungsidee des in dieser Arbeit präsentierten Wasserzeichens beruht auf einem JPEG-Bildwasserzeichen, wo die Kompression nicht so variabel ist. So kann es durchaus möglich sein, dass bei einem geringen MPEG-Bitrate und der daraus resultierenden starken Kompression das Wasserzeichen nicht mehr korrekt ausgelesen werden kann, da die Einbettung schon ineffizient war.

In anderen Lösungen wird auch eine andere Strategie angewandt. In [KaDe1999] wird das Video komplett in die Einzelbilder dekodiert und jedes Einzelbild separat markiert. Für diese Methode können zwar auch JPEG-basierte-Wasserzeichen genutzt werden, für das in dieser Arbeit vorgestellte Anwendungsszenario des Onlineshops ist sie nur bedingt einsetzbar, da die Zeitverzögerungen durch die komplette Kodierung des Videos zu groß sind.

(2) Veränderung der Frametypen

Wenn bei der Bearbeitung einzelne Frames oder Sequenzen hinzugefügt oder entfernt werden oder für eine Re-Enkodierung ein anderes Enkodierschema als vor der Bearbeitung des Videos benutzt wird, kann es sein, dass sich der Frametyp des entsprechenden Frames oder der Sequenzaufbau ändert. So besteht die Möglichkeit, dass die Detektion des Wasserzeichens aus ursprünglich nicht markierten Frames erfolgt.

In der Arbeit wird eine Lösung präsentiert, die auf dem Ansatz der zeitlichen Synchronisierung [HaKu2005] basiert. Ziel dieser Lösung ist es, die ursprünglich markierten Frames wieder aufzuspüren und anschließend die eingebetteten Wasserzeicheninformationen wieder auszulesen. Mit Hilfe des robusten Hashwertes können die ursprünglich markierten Frames detektiert und das Wasserzeichen aus diesen Frames ausgelesen werden. Die Testergebnisse deuten einerseits auf eine hohe Robustheit des Hashs und andererseits ist die Wahrscheinlichkeit einer falschen Identifikation von nicht-markierten Frames sehr klein.

Das einzige Problem des Ansatzes ist die Interpolation zwischen zwei benachbarten Frames zu einem Frame. Eine Verringerung der Framerate erfordert den Prozess der Interpolation. Dadurch kann sich der Frameinhalt so deutlich verändern, dass die extrahierten Merkmale zu einem unterschiedlichen Hashwert führen. Wenn die Framerate im gesamten Video verringert wird, kann es folglich sehr schwierig werden, die ursprünglich markierten Frames zu detektieren. Somit ist ein reines Ausschlusskriterium, dass die interpolierten Frames nicht untersucht werden, nicht anwendbar, da dadurch das gesamte Video nicht untersucht werden könnte. Deshalb muss der Schritt der Merkmalsextraktion robust gegenüber dieser Veränderung sein.

Andere Lösungen konzentrieren sich nicht nur auf einzelne Frames. In [HaKu2005] wird die zeitliche Synchronisation dazu genutzt, eine markierte Framesequenz zu finden.

(3) Referenzen innerhalb des Videos

Die Bewegung im Video kann sich auf das Einbetten von Wasserzeichen auswirken. Wenn in ein Frame ein Wasserzeichen eingebettet wird, muss damit gerechnet werden, dass es aufgrund der Referenzen innerhalb der GoP auch in Nachbarbilder übertragen wird.

Vor dem Einbetten von Wasserzeichen in P- und B-Frames werden die übertragenen Wasserzeichensignale aus den Referenzframes mit in die Berechnung einfließen. In [Ha2000] wurde das Prinzip des Drift-Compensation-Signals entwickelt.

In dieser Arbeit werden vor dem Einbetten des Wasserzeichens in P-Frames die übertragenen Wasserzeichensignale zuerst kompensiert, bevor das für das augenblickliche Frame berechnete Wasserzeichensignal eingebettet wird. Erst nachdem sämtliche übertragenen Wasserzeichensignale entfernt wurden, kann mit der Einbettung der Wasserzeichen in das jeweiligen P-Frame begonnen werden. Untersuchungen haben gezeigt, dass die übertragenen Wasserzeichensignale sich auch um einzelne Pixel im nachfolgenden Frame verschieben können. Deshalb wird in dieser Arbeit die Strategie verfolgt, in dem sämtliche übertragenen Wasserzeichensignale entfernt werden. So ist das Wasserzeichen, das in den P-Frame komplett unabhängig von anderen Signalen. Durch die Bewegungskompensation hat das Wasserzeichen eine gute Transparenz und es können visuelle Artefakte vermieden werden.

(4) Echtzeitfähigkeit

Aufgrund der Komplexität des MPEG-Videostandards erfordert die Re-Enkodierung eines MPEG-Videos erfordert ein entsprechend hohes Maß an verfügbaren Ressourcen.

Um den Einbettungsvorgang in Echtzeit zu realisieren, sollten nur jene Videoteile dekodiert werden, wo auch das Wasserzeichen eingebettet wird [Di2007], [HaSt2008]. So müssen z. B. I-Frames nur dann dekodiert werden, wenn das Wasserzeichen auch nur in I-Frames eingebettet wird. Sämtliche andere Daten können in das markierte Video kopiert werden.

Mit dem in dieser Arbeit vorgestellten Kontainermodell kann das Video für die spätere Erstellung von markierten Kopien vorbereitet werden. Dabei wird der zeitintensive Aufwand des Wasserzeichenalgorithmus während der Kontainererstellung verlegt. Häufig ist der zeitintensive Anteil die Zerlegung des Videos in die Frames und die abschließende Zusammenfügung der markierten Frames in ein Video. Diese Operationen sind zudem noch für jeden Einbettungsvorgang identisch, wodurch es sich als sinnvoll erweist, sie nur einmal für sämtliche Markierungen durchzuführen. Während der Erstellung der markierten Kopie müssen nur noch die relevanten Daten aus dem Kontainer extrahiert und zu einem Video zusammengestellt werden. Damit wurde eine bis zu hundertfache Echtzeit erreicht.

Augenblick ist das Kontainermodell mit einem Wasserzeichen getestet wurden, das die Wasserzeichenbits auch in Makroblöcke einbettet. So kann der Kontainer sehr einfach erstellt werden, da das MPEG-Video auch aus Makroblöcken besteht. Wird hingegen ein anderes Wasserzeichen genutzt, das die Wasserzeicheninformationen nicht direkt auf Makroblockbasis in das Video einbettet, muss das Speicherprotokoll dementsprechend angepasst werden, damit während der Extraktion der markierten Kopie keine Kodieroperationen notwendig sind. Somit kann trotzdem die hohe Echtzeitrate während der Erstellung der markierten Kopie eingehalten werden

Ein weiterer Lösungsvorschlag ist die Zerlegung des Videos in mehrere Teilvideos und die Markierung der Teilvideos auf mehreren zur Verfügung stehenden Computerressourcen. Durch das parallele Einbetten des Wasserzeichens besteht die Möglichkeit, auch Wasserzeichen mit einer hohen Komplexität in Echtzeit einzubetten. Besonders bei großen Videos und bei einer hohen Videoanzahl ist die Verteilung der Einbettungsvorgänge sehr effizient. Der Forschungsbedarf besteht in der Verringerung des Zeitverlusts innerhalb des verteilten Netzwerks. Das Wasserzeichen kann zwar eine sehr gute Echzeitfähigkeit vorweisen, da aber der Einbettungsvorgang auf anderen zur Verfügung stehenden Computer verteilt wird, muss mit Einbußen der Echtzeit gerechnet werden. Videodaten können aus immensen Datenmengen bestehen und diese über das Internet zu versenden bzw. zu empfangen kann zu deutlichen zeitlichen Einbußen führen. Deshalb sollte die Markierung auf einen regional kleineren Kreis oder auf Computer mit sehr guten Datenratenübertragungen begrenzt werden. Als ideal bietet sich das Intranet an, da zudem häufig eine Vielzahl von Computer zur Verfügung steht, wo deren Ressourcen nicht komplett genutzt werden.

Andere Wasserzeichen erreichen eine Echtzeit mit Hilfe einer geringen Komplexität des Algorithmus. So wird der Dekodiervorgang nicht komplett durchgeführt und die Markierung durch Veränderung der VLC Werte durchgeführt [LuCh2005]. Diese Verfahren weisen aber eine mangelhafte Robustheit auf.

(5) Videoeditierung

Durch Veränderungen des Videoinhaltes besteht auch die Möglichkeit, die Bedeutung des Inhaltes zu verändern. Gezielte Modifikation verändern den Inhalt des Videos, in dem sie Elemente in den Frames modifizieren. Besonders in Anwendungsszenarien, wo Überwachungsvideos zum Einsatz kommen, erweist es sich als notwendig, einen Nachweis der Authentizität des Videos durchzuführen.

In der Arbeit wird eine Lösung präsentiert, die nachweisen kann, ob das Video authentisch ist. Die Genauigkeit des Nachweises der Authentizität kann bis auf eine GoP genau nachgewiesen werden. Da aber Wasserzeichen auch ungewünschte Veränderungen sind, die bei sehr sensiblen Anwendungen zu einer Veränderung der Aussage führen, wurde die Lösung dahingegen erweitert, das Wasserzeichen zu entfernen. Mit Hilfe eines Sicherheitsprotokolls wird ein unerlaubtes Entfernen des Wasserzeichens verhindert. Nach einer Vielzahl von Veränderungen am Video wurde es als nicht authentisch eingestuft. Ein Schwachpunkt ist der Aufbau des Einbettungskanals. Es ist möglich, durch Überschreiben des Einbettungskanals das Original unerlaubt wiederherzustellen. Die Lösung muss Inhalte des Frames mit in

den Einbettungskanal integrieren, der ungleich null ist. Da das möglicherweise die Folge hätte, das niederfrequente Inhalte markiert werden, muss ein Konsens zwischen Sicherheit und Transparenz gefunden werden.

11 Zukünftige Arbeit

In dieser Arbeit wurden Lösungen für MPEG-Videowasserzeichen diskutiert. Da aber die Entwicklung der Videostandstandards fortfahren wird und die Verbreitung von Videos tendenziell immer stärker zunimmt, ist in den folgenden Bereichen weiterer Forschungsbedarf notwendig:

(1) Markierung des gesamten MPEG-Videos,
(2) Anpassung der Wasserzeichen an Weiterentwicklungen von Videostandards,
(3) Suche von markierten Videos in Peer-to-Peer-Netzwerken und
(4) Unterstützung im Fernsehbereich durch Videowasserzeichen.

11.1 Markierung des gesamten MPEG-Videos

In der Arbeit wurde die Lösung für P-Frames von MPEG-Videos vorgestellt. Um aber robust gegenüber einer Vielzahl von Modifikationen am Videomaterial zu sein, ist es notwendig, das gesamte Video zu markieren oder eine robuste Methode zu entwickeln, die die markierten Frames vor der Detektion des Wasserzeichens im Video auffindet.

Damit das komplette Video markiert werden kann, muss auf die Besonderheiten der verschiedenen Kompressionsstufen geachtet oder auf ein sehr robustes Bildwasserzeichen zurückgegriffen werden, das für JPEG-Bilder entwickelt wurde und gegenüber sehr hohen JPEG-Kompressionsstufen resistent ist.

Ein mögliches Konzept wäre die Anpassung des in Kapitel 7.4.2 vorgestellten robusten Videowasserzeichens an die höhere Kompression in B-Frames. Da die Kompression zwischen P- und B-Frames auf dem gleichen Algorithmus beruht, wären nur die unterschiedlichen Quantisierungsfaktoren der entscheidende Punkt für die Einbettung des Wasserzeichens in B-Frames. Die Wasserzeichenmuster und Markierungsstellen werden gegenüber dem vorliegenden P-Frame identisch gehalten, um mögliche visuelle Artefakte zu vermeiden [SuKu2002]. Die Wasserzeicheninformation wird mit Beginn einer neuen Sequenz geändert, da sich häufig nach einem Szenenschnitt der Frameinhalt deutlich ändert.

11.2 Anpassung von Wasserzeichen an neue Videostandards

Die in der Arbeit betrachteten Videostandards wurden zwischen 1995 und 1998 veröffentlicht. Parallel dazu wurden weitere Videostandards wie Apple Quicktime und Microsoft Windows Media veröffentlicht. Sie werden häufig für andere Einsatzszenarien als hochqualitatives Videomaterial eingesetzt.

Durch die Verbreitung und Weiterentwicklung von digitalem Fernsehen wird der Videostandard h.264 immer wichtiger. Er kann mit MPEG-4 verglichen werden, was bei

gleicher visueller Qualität eine stärkere Kompression ermöglicht. Einzelne Veröffentlichungen [WuWa2005], [ZhHo2007] markieren h.264 mit einfachen Modellen. Dabei wird besonders auf den 40:1 Kompressionsfaktor von h.264 gegenüber MPEG I-Frames hingewiesen.

11.3 Suche von markierten Videos in Peer-to-Peer Netzwerken

Durch die Verbesserung der Netzbandbreite bei DSL Anschlüssen rückt auch die illegale Verbreitung von Videos über verschiedene P2P (Peer-to-Peer) Netzwerke immer stärker in den Focus. Ein P2P-Netzwerk ist eine Netzwerkform, wo alle Netzwerkteilnehmer die gleiche Priorität besitzen. Es können sowohl Daten zwischen den Teilnehmern ausgetauscht als auch Dienste auf anderen Teilnehmern genutzt werden. Besonders die Verbreitung von DVD-Videos verstärkt sich stetig. Dabei werden zwei Strategien angewandt:

(1) Konvertierung von DVD in DivX in CD-Qualität
Der DVD-Film wird von der DVD kopiert. Die kopierten Dateien des Hauptfilms werden zu einem kompletten MPEG-Video zusammengefügt. Abschließend wird das MPEG-Video in DivX konvertiert und die Kompressionsrate so gewählt, dass er auf eine CD kopiert werden kann. Zusätzlich besteht auch die Möglichkeit, dass die Seitenverhältnisse von 16:9 auf 4:3 geändert werden, in dem der obere und untere schwarze Streifen entfernt wird.

(2) Umwandlung von DVD-9 auf DVD-5 Format
Die Original-DVD wird meistens im DVD-9-Format. Da aber DVD-Rohlinge nur auf dem DVD-5 Format basieren, wird das komplette Video stärker komprimiert und die Videodateien neu erstellt. Die neuen Dateien können dann mit einem üblichen Kopierprogramm auf eine DVD-5 kopiert werden. Die durch die Kompression zugeführten Qualitätsunterschiede sind nur schwer erkennbar und deshalb vernachlässigbar.

Da mit Hilfe der P2P-Netzwerktechnologie sehr häufig Videodaten angeboten werden, kann es sich als sinnvoll erweisen, sie nach Wasserzeichen zu untersuchen. Dafür wird ein Analyseprogramm genutzt, das die Videos von den am Netzwerk angeschlossenen Computern herunterlädt und anschließend den Wasserzeichendedektor startet. Die Strategie kann dabei wie folgt sein:

(1) Suche nach aktuellen Videos mit Hilfe einer Datenbank
Eine regelmäßig aktualisierte Datenbank beinhaltet die Namen und Hashwerte von bekannten Videos. Verschiedene P2P Netzwerke benutzen Hashalgorithmen zur Identifizierung der Dateien. Dadurch wird ein mögliches Herunterladen von falschen Dateien verhindert.

(2) Start von Download und Unterbrechung bei ausreichend vorhandenen Daten des Videos
Der Client startet den Downloadvorgang, analysiert die heruntergeladenen Teilvideos und hält den Downloadvorgang an, sobald die Teildatei groß genug ist, um das Wasserzeichen komplett auslesen zu können. Konnte das Wasserzeichen

mehrmals erfolgreich ausgelesen werden, dann wird der Downloadvorgang protokolliert und abgebrochen. Wurde es nicht ausgelesen, werden weitere Teile heruntergeladen und untersucht. Bei Videodaten ist es aufgrund der Datenmenge ratsam, das Wasserzeichen auch schon aus Teilvideos auszulesen.

(3) Weiterleitung der protokollierten Ausleseergebnisse an die Eigentümer oder Verkäufer der Videos
Die protokollierten Daten werden an die Verkäufer der Videos weitergeleitet. Die Eigentümer können dann mögliche rechtliche Schritte einleiten, um die illegale Verbreitung zu stoppen.

Das Konzept erfordert einen robusten Synchronisationsmechanismus, da es notwendig ist, den Beginn eines eingebetteten Wasserzeichens im Teilvideo zu finden. Wenn die Videodaten zusätzlich mit Datenkompressionsalgorithmen wie RAR oder ZIP komprimiert wurden, ist es notwendig, die gesamte Datei herunterzuladen und dann zu dekomprimieren.

11.4 Unterstützung durch Videowasserzeichen im Fernsehumfeld

Videowasserzeichen können innerhalb des Fernsehumfelds in verschiedensten Bereichen eingesetzt werden [HaSt2007]. Neben dem klassischen Urheberszenario können sie auch unterstützend im Kreislauf der Videoerstellung von der Aufnahme bis zur Ausstrahlung des Materials genutzt werden. Die Unterstützung lässt sich mit Hilfe von Metawasserzeichen durchführen. Die Wasserzeichen enthalten z. B. Informationen über Erstellung und Weiterleitung des Materials von einer Instanz zu einer anderen Instanz. Abbildung 54 zeigt die Möglichkeiten bei der Integration von Wasserzeichen während des Verarbeitungsprozesses. So können sie z. B. während der Erstellung Informationen über Autoren, Aufnahmegeräte und Zeitpunkt der Aufnahme in die Videos eingebettet werden. Im Archiv können Identifikationsnummern in das Video integriert werden. In Produktion und Broadcasting können Zeitstempel als Wasserzeicheninformation Aufschluss über den Weiterleitungszeitpunkt geben.

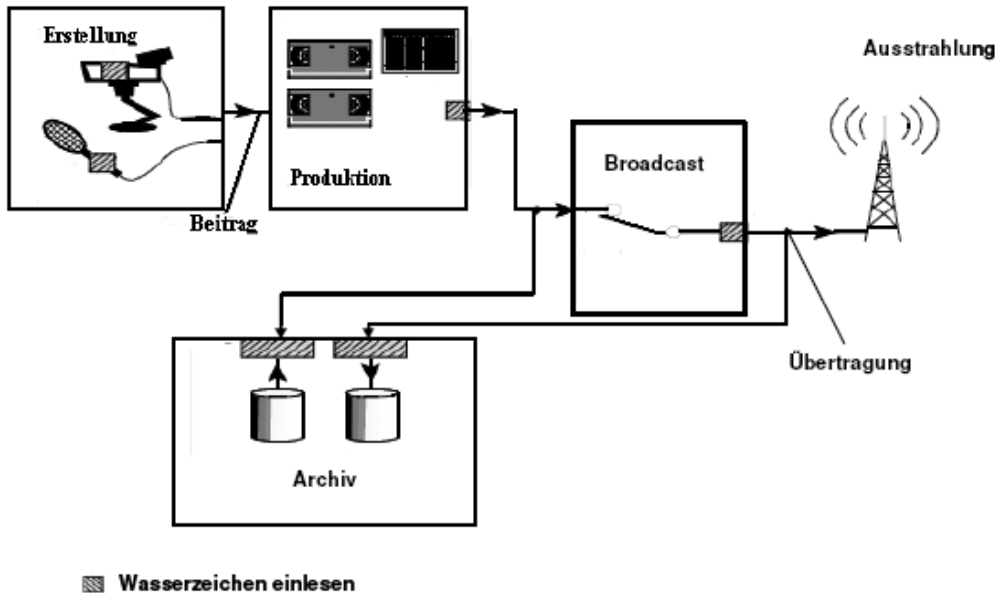

Abbildung 54: Einbettung von Fernsehwasserzeichen

Die Detektion des Wasserzeichens (Abbildung 55) dient insbesondere der Überprüfung des eingegangen Videomaterials. In der Produktion können die während der Erstellung eingebetteten Informationen automatisch ausgelesen werden. Im Archiv geben die Wasserzeicheninformationen Aufschluss über Bearbeitungsschritt und Videokategorie. Auf der Empfangsseite können die Wasserzeichen für das Monitoring genutzt werden.

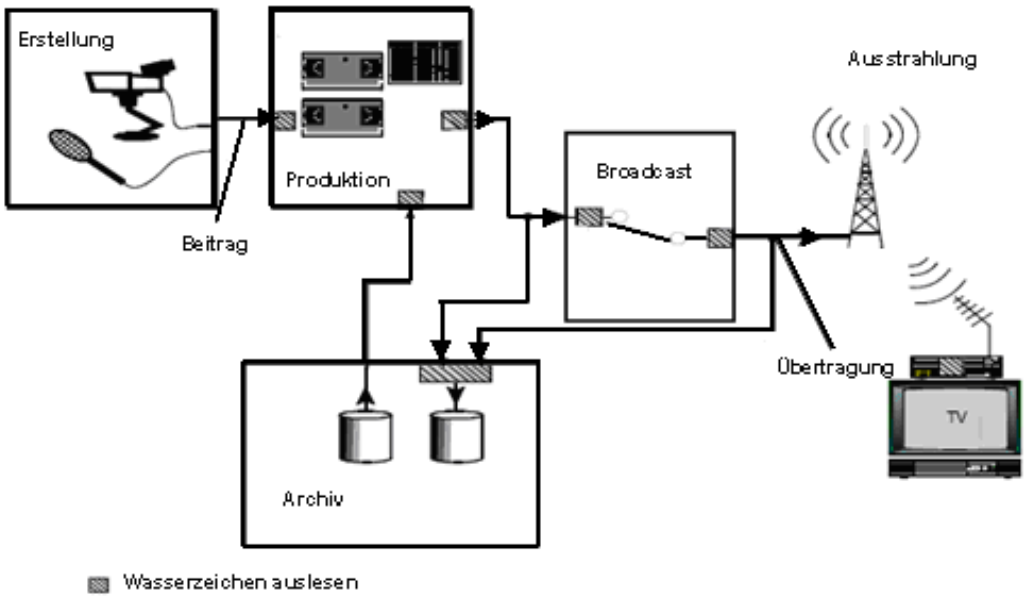

Abbildung 55: Detektion von Fernsehwasserzeichen

Während des Einsatzes von Metawasserzeichen ist die Robustheit kein wichtiger Aspekt. Sie dienen lediglich dazu, nützliche Informationen mit in das Video einzubetten. So ist es

unproblematisch, wenn sie nach dem Auslesen durch eine anschließende Modifikation am Video zerstört werden.

Bevor ein Wasserzeichen im Fernsehbereich eingesetzt werden soll, müssen noch basierend auf dem jeweiligen Szenario die Parameter des Wasserzeichenalgorithmus festgelegt werden. So kann die Kapazität unterschiedlich ausgeprägt sein. Es entscheidet sich daran, ob Informationen über die Aufnahme oder eine Archividentifikationsnummer in das Video eingebettet werden soll.

12 Zusammenfassung

Diese Arbeit beschäftigt sich mit dem Thema digitale Wasserzeichen für MPEG-Videodaten. Dabei liegt der Fokus auf Anwendungsbereichen, bei denen die Authentifizierung des Eigentümers bzw. Käufers oder der Videodaten wichtig ist. Speziell die Authentifizierung des Eigentümers ist bei der heutigen Verbreitung von digitalen Videos ein zentraler Aspekt. Auf der einen Seite können durch den Vertrieb von Videos über Netzwerke zwar Kosten eingespart werden, auf der anderen Seite besteht aber immer die Möglichkeit, dass die Kontrolle über die Verbreitung der Videos relativ schnell verloren gehen kann. Der Vorteil von digitalen Medien liegt in der einfache und verlustfreien Vervielfältigung, ohne dabei wie bei analogen Medien Qualitätsverluste zu erleiden.

Natürlich wurden schon entsprechende Mechanismen entwickelt, die eine unerlaubte Vervielfältigung verhindern oder wenigstens eindämmen sollen. Ein sehr bekannter und verbreiteter Mechanismus ist das Digital-Right-Management (DRM). Mit Hilfe der Zuweisung von speziellen Nutzungsrechten kann dem Käufer des Medium die Nutzung vorgeschrieben werden. So kann festgelegt werden, unter welchen Bedingungen das Medium abgespielt bzw. kopiert werden darf.

Digitale Wasserzeichen sind eine Alternative gegenüber einem DRM-System. Der Käufer hat keine Einschränkung bei der Benutzung des Videos. Es werden während des Verkaufs nur Nachweisinformationen im Video eingebettet. Die Veränderungen sind so minimal, dass keine visuellen Einschränkungen auftreten. Anschließend hat der Käufer das vollständige Zugriffsrecht auf das von ihm gekaufte Video.

Die in dieser Arbeit durchgeführte Evaluation wird auf Basis schon existierender Wasserzeichenalgorithmen durchgeführt. Da die Verfahren schon ein gewisse Qualität erreicht haben, fokussiert sich die Arbeit auf bestimmte Schwachpunkte:

- das Wiederfinden des Wasserzeichens im Video,
- die Markierung des gesamten Videos,
- die schnellen Markierung des Videos und
- die Authentifizierung des Videos.

Für jedes dieser Probleme werden entsprechende Ansätze vorgestellt und abschließend bewertet.

Bevor diese Arbeit die Wasserzeichenalgorithmen betrachtet, werden zuerst die Anwendungsszenarien vorgestellt. Mit Hilfe der Anwendungen werden die Parameter für die späteren vorgestellten Wasserzeichen festgelegt. Im Bereich der Authentifizierung des Urhebers bzw. Käufers wird die Verbreitung über das Internet und digitale Bibliotheken betrachtet. Über das Internet werden Videos mit Hilfe von Onlineshops verbreitet. Da mit einer Vielzahl von gleichzeitig parallel stattfindenden Verkäufen eines Videos gerechnet werden muss, sollte das Wasserzeichen sehr schnell eingebettet werden. Dafür wird in der Arbeit der Wasserzeichenkontainer vorgestellt. Die Idee liegt in der Aufteilung des gesamten Einbettungsprozesses in zwei Schritte. Der erste Schritt umfasst die zeitintensiven

Operationen und wird nur einmal durchgeführt. Das begründet sich durch die fast identische Markierung jeder Kopie des Videos. Der einzige Unterschied ist die Wasserzeicheninformationen; sämtliche andere Prozesse, die während der Einbettung vorgenommen werden, sind identisch.

Dagegen liegt bei der Markierung von digitalen Bibliotheken eine Vielzahl von verschiedenen Videos vor. Dadurch ist eine Vorverarbeitung nicht möglich. Die Videos können nur effizient und schnell markiert werden, wenn entsprechende Computerrressourcen genutzt werden. Eine in dieser Arbeit erläuterte Möglichkeit ist die Nutzung von Grid-Netzwerken. Ein Grid-Netzwerk ist ein eigenständiges über verschiedene geographische Standorte verteiltes Netzwerk, dass verschiedene Ressourcen zur Verfügung stellt. Der in dieser Arbeit vorgestellte Ansatz verteilt die Markierungsvorgänge über die Standorte und ermöglicht so ein verteiltes paralleles Markieren der Videobibliothek. Die zeitliche Untersuchung hat gezeigt, dass die Markierung der Bibliothek schneller abgeschlossen werden kann, als wenn eine lokale Ressource genutzt wird.

Ein Schwachpunkt vorhandener Verfahren ist das Wiederfinden eines eingebetteten robusten Wasserzeichens. Da der in dieser Arbeit betrachtete Videostandard MPEG die Einzelbilder unterschiedlich stark komprimiert, markiert eine Vielzahl von Wasserzeichen nur vorausgewählte Einzelbilder im Video. Mit der Anwendung der Technologie des robusten Hashs ist es möglich, diese markierten Bilder wiederzufinden und anschließend das Wasserzeichen erfolgreich auszulesen. Der Vorteil des robusten Hashs liegt darin, daß er für eine Vielzahl von Wasserzeichen genutzt werden kann und zwar genau immer dann, wenn vorausgewählte Einzelbilder markiert werden. Die anschließende Untersuchung hat gezeigt, daß der robuste Hash gegenüber einer Vielzahl von Veränderungen am Video robust ist.

Des weiteren bietet sich auch die Möglichkeit, das Video vollständig mit einem robusten Wasserzeichen zu markieren. Da MPEG aber die Einzelbilder unterschiedlich stark komprimiert, wird das Wasserzeichen an die verschiedenen Komprimierungsstufen angepasst. Zusätzlich wird noch die Methode der Bewegungskompensation durchgeführt, da die Videos auch zeitlich komprimiert werden, d.h. identischer Bildinhalt, der in mehreren benachbarten Einzelbilder auftritt, wird nur einmal komprimiert. Auf die anderen folgenden Einzelbilder wird er übertragen. Diese Besonderheiten müssen während der Markierung sämtlicher Einzelbilder betrachtet werden, um eine Mehrfachmarkierung zu vermeiden. Die Ergebnisse der Untersuchung stellen ein positives Gesamtbild dar. Nach der Untersuchung der P-Frames muss der in dieser Arbeit entwickelte Ansatz noch auf B-Frames erweitert werden.

In einem weiteren Anwendungsszenario wird die Authentizität der Videodaten betrachtet. Dabei kommen Sicherheitstechnologien wie der fragile Hash und die digitale Signatur zum Einsatz. Der Hash dient der Überprüfung der Integrität und die digitale Signatur überprüft die Authentizität. Da sich das Anwendungsszenario auf sehr sensible Daten beschränkt, wird es dahingegen erweitert, dass die Möglichkeit bestehen muss, das Wasserzeichen wieder aus dem Video zu entfernen, nachdem es als authentisch eingestuft wurde. Deshalb wird ein reversibles Wasserzeichen entwickelt. Da die Möglichkeit der Entfernung des Wasserzeichens Sicherheitsbedenken aufruft, wird die Wasserzeicheninformation in zwei Teile getrennt:

- Wiederherstellung des Originals
- Überprüfung der Authentizität

Bei der Wiederherstellung des Originals werden die während der Markierung des Video symmetrisch verschlüsselten Originaldaten wieder entschlüsselt und an ihre ursprünglichen Positionen im Video geschrieben. Durch die Anwendung einer symmtrischen Verschlüsselung ist nur den Personen vorbehalten, das Video zu entschlüsseln, die auch den geheimen Schlüssel besitzen.

Die Überprüfung der Authentizität ist hingegen öffentlich durchführbar. Ist der öffentliche Schlüssel des asymmetrischen Schlüsselpaares bekannt, kann die Authentizität überprüft werden. Gleichzeitig ist es aber nicht möglich, das Originalvideo wiederherzustellen.

Da aber das symmetrische AES und das asymmetrische RSA Verfahren inzwischen eine gewisse Länge ihrer Schlüssel aufweisen müssen, kann mit dem Wasserzeichenverfahren nicht mehr Einzelbild-genau sondern nur noch GoP-genau die Integrität des Videoinhalts überprüft werden. Die notwendige Kapazität für die Wasserzeicheninformationen stellt nicht jedes Einzelbild zur Verfügung. Es ist nur möglich, die Wasserzeichninformation in ein I-Frame einzubetten. Die anderen Einzelbilder der GoP können nur auf ihre Integrität überprüft werden, in dem der Hashwert dieser Bilder mit in das I-Frame eingebettet wird.

Die Untersuchung zeigt, dass nach einer Vielzahl von durchgeführten Modifikationen die Überprüfung der Authentizität ein verändertes Video aufweist. Der einzige Schwachpunkt ist die Möglichkeit, einzelne Originalvideos wiederherzustellen, wenn die markierten Positionen im Video mit null überschrieben werden.

Sämtliche in der Arbeit präsentierten Ansätze können die Schwachstellen kompensieren, die sich während der Evaluation herausgestellt haben. So ist es mit Hilfe des robusten Hashs möglich, die ursprünglich markierten Frames wiederzufinden. Das Markierungsschema des robusten Wasserzeichens konnte von intrakodierte auf interkodierte Frames erweitert werden. Mit Hilfe des Containermodells können eine Vielzahl von unterschiedlich markierten Kopien eines Videos erstellt werden. Komplette digitale Bibliotheken werden durch Nutzung verteilter Netzwerke in Echtzeit markiert. Auch ist es möglich, mit Hilfe des reversiblen Wasserzeichens die Authentizität des Video zu untersuchen und zusätzlich das Originalvideo wiederherzustellen. Dabei ist die Sicherheit des Wasserzeichens garantiert.

Gleichzeitig haben sich aber noch offene Punkte herausgestellt, die für sehr gute Wasserzeichenalgorithmen noch optimiert werden müssen.

Der Ausblick zeigt eindeutig, dass die Entwicklungen dieser Arbeit sich zukünftige Weiterentwicklungen anpassen müssen. So werden augenblicklich Optimierungen an den Videostandards vorgenommen, die Auswirkungen auf die Wasserzeichen haben können. Durch die zunehmende Akzeptanz von Breitbandnetzwerken wie DSL muss auch für Videodaten nach effzienten Suchmechanismen geforscht werden. Da im Allgemeinen das Fernsehen in den nächsten Jahren generell digital wird, bietet sich im Fernsehumfeld ein weiteres großes Forschungsfeld an. Dabei können eine Vielzahl von Szenarien auftreten, bei denen digitale Videowasserzeichen zum Einsatz kommen.

Anhang A

Dieser Anhang beinhaltet die Ergebnistabellen aus der Evaluation von Kapitel 5.1.

Tabelle 13 beinhaltet die Ausleseraten des Wasserzeichens (A-Rate) und die False-Bit-Rate (FBR) der ausgelesenen Fingerabdrücke. Es werden die drei Wasserzeichenverfahren separat evaluiert, um die bestmöglichen Werte für die Wasserzeichenparameter und basierend darauf die bestmöglichen Detektionsergebnisse zu finden.

	WS = 1		**WS = 1,5**		**WS = 2**		**WS = 4**	
	A-Rate	**FBR**	**A-Rate**	**FBR**	**A-Rate**	**FBR**	**A-Rate**	**FBR**
[DiHa2001a]	88,24%	57,36%	31,70%	82,87%	31,70%	82,87%	0,45%	99,59%
[DiSt1998]	97,00%	2,07%	98,98,00%	0,07%	100,00%	0,07%		
	$R_1 = 0,25$		$R_1 = 0,5$		$R_1 = 0,75$		$R_1 = 1,00$	
	A-Rate	FBR	A-Rate	FBR	A-Rate	FBR	A-Rate	FBR
[DiHa2001a]	28,87%	65,40%	29,61%	64,50%	30,51%	64,18%	31,40%	61,86%
	$R_2 = 5$		$R_2 = 10$		$R_2 = 15$			
	A-Rate	FBR	A-Rate	FBR	A-Rate	FBR		
[DiHa2001a]	30,80%	65,40%	34,39%	64,50%	25,00%	61,22%		
	$R_3 = 3$		$R_3 = 5$		$R_3 = 9$			
	A-Rate	FBR	A-Rate	FBR	A-Rate	FBR		
[DiHa2001a]	33,33%	64,08%	33,33%	64,08%	30,34%	65,10%		
	Range = 5		Range = 10					
	A-Rate	FBR	A-Rate	FBR				
[DiHa2001a]	29,02%	65,42%	31,18%	62,56%				
	Cut-Off-Point = 3		Cut-Off-Point = 10		Cut-Off-Point = 20			
	A-Rate	FBR	A-Rate	FBR	A-Rate	FBR		
[SeLa2001]	100,00%	0,68%	100,00%	3,52%	86,67%	7,49%		
	Difference = 5000		Difference = 10000		Difference = 20000			
	A-Rate	FBR	A-Rate	FBR	A-Rate	FBR		
[SeLa2001]	100,00%	2,07%	93,33%	1,99%	93,33%	6,17%		

Tabelle 13: Auslesen der Wasserzeicheninformation

Tabelle 14 erläutert die Ergebnisse der Robustheit der Verfahren nach dem Auftragen eines additiven Wertes.

	3 addiert		5 addiert		3 subtrahiert		5 subtrahiert	
	A-Rate	FBR	A-Rate	FBR	A-Rate	FBR	A-Rate	FBR
[DiHa2001a]	0,00%	-	0,00%	-	0,00%	-	0,00%	-
[DiSt1998]	100,00%	0,83%	100,00%	0,83%	100,00%	0,83%	100,00%	83,00%
[SeLa2001]	100,00%	3,27%	100,00%	4,03%	100,00%	3,17%	97,67%	3,98%

Tabelle 14: Auslesen der Wasserzeicheninformation nach dem Auftragen eines zusätzlichen Wertes [Th2002]

Bei dem Skalierungsangriff (Tabelle 15) wird zuerst das markierte Video auf eine bestimmte Größe skaliert. Im zweiten Fall wird das markierte Video wieder auf Originalgröße re-skaliert.

	auf 480x352 skaliert		reskaliert	
	A-Rate	FBR	A-Rate	FBR
[DiHa2001a]	0,00%	-	0,00%	-
[DiSt1998]	30,00%	9,25%	80,00%	2,96%
[SeLa2001]	6,98%	38,00%	88,37%	8,32%

Tabelle 15: Auslesen der Wasserzeicheninformation nach Videoskalierung [Th2002]

Tabelle 16 beinhaltet die Detektionsergebnisse der eingebetteten Fingerabdrücke. Dabei wurden die beiden Fingerabdruckverfahren [DiBe1999] (Fingerabdruck 1) und [BoSh1995] (Fingerabdruck 2) angewendet. Für eine Analyse der möglichen Abhängigkeit der Fingerabdrücke von den Videodaten, wurden zwei verschieden aufgebaute Videos untersucht. Video1 ist ein Zeichentrickvideo und Video2 ein Werbevideo.

		Testvideo	Kundenliste	einzelne Kunden	falsche Kunden
[DiHa2001a]	[DiBe1999]	Video 1	44,44%	0,00%	0,00%
[DiHa2001a]	[DiBe1999]	Video 2	77,78%	22,22%	0,00%
[DiSt1998]	[DiBe1999]	Video 1	33,33%	33,33%	0,00%
[DiSt1998]	[DiBe1999]	Video 2	33,33%	33,33%	0,00%
[DiHa2001a]	[BoSh1995]	Video 1	0,00%	88,89%	11,11%
[DiHa2001a]	[BoSh1995]	Video 2	0,00%	100,00%	0,00%
[DiSt1998]	[BoSh1995]	Video 1	0,00%	94,44%	5,56%
[DiSt1998]	[BoSh1995]	Video 2	0,00%	100,00%	0,00%

Tabelle 16: Auslesen der Kundeninformationen [Th2002]

Anhang B

Die folgende Tabelle stellt die prozentualen Detektionsraten des robusten Videohash für Videoset 1 dar. Dabei haben die in den einzelnen Spalten der Tabelle aufgeführten Angriffe die folgende Bedeutung:

- GoP9 – Verkleinerung der GoP Größe von 12 auf 9 Frames pro GoP
- GoP15 – Vergrößerung der GoP Größe von 12 auf 15 Frames pro GoP
- Bitrate 1 MB/s – Re-Enkodierung des Videos mit einer Bitrate von 1 MB/s
- Bitrate 750 kB/s – Re-Enkodierung des Videos mit einer Bitrate von 750 kBit/s
- FPS 24 – Verringerung der Framerate von 25 auf 24 fps
- FPS 30 – Erhöhung der Framerate von 25 auf 30 fps

VideoNr	GOP9	GOP15	Bitrate 1 MB/s	Bitrate 750 kBit	FPS 24	FPS 30
1	93,12	92,96	92,96	92,96	51,84	100,00
2	91,37	92,01	92,01	91,85	43,77	100,00
3	98,08	98,08	98,08	98,08	66,99	100,00
4	95,36	95,36	95,52	95,36	48,96	100,00
5	97,60	97,76	97,76	97,60	60,58	100,00
6	89,92	89,92	89,76	89,92	39,20	99,52
7	75,84	75,84	76,00	76,00	36,32	89,92
8	93,28	93,28	93,12	93,12	38,56	100,00
9	90,88	90,88	90,88	90,88	40,32	100,00
10	88,00	88,32	88,16	87,84	38,72	100,00
11	95,83	95,83	95,99	95,83	68,43	98,88
12	96,96	96,80	96,80	96,80	52,64	100,00
13	83,09	83,09	83,09	82,93	47,53	91,23
14	97,44	97,60	97,60	97,76	68,64	98,56
15	98,08	98,08	98,24	98,24	70,03	100,00
16	94,87	94,87	95,03	94,87	52,56	98,40
17	92,80	92,80	92,80	92,96	47,04	100,00
18	93,28	93,28	93,44	93,12	61,76	100,00
19	95,19	95,19	95,19	95,19	47,76	100,00
20	95,69	95,69	95,85	95,69	57,35	100,00
21	75,40	75,24	75,40	75,40	22,36	99,68
22	84,29	84,29	84,29	84,46	41,67	100,00
23	84,32	84,48	84,32	84,32	30,40	100,00
24	94,72	94,72	94,72	94,72	53,92	100,00
25	85,44	85,28	85,44	85,28	24,96	100,00
26	83,68	83,84	84,16	84,00	32,16	100,00
27	87,20	87,20	87,20	87,20	32,80	98,72
28	85,94	85,94	85,94	85,94	31,95	100,00
29	91,68	91,84	92,00	92,16	39,68	100,00
30	67,52	67,36	67,68	67,20	22,40	100,00
31	92,16	92,16	92,00	92,16	46,88	100,00
32	74,72	74,88	74,40	74,72	22,08	99,84
33	87,20	87,20	87,36	87,36	42,08	100,00
34	88,32	88,48	88,16	88,16	41,76	100,00
35	81,44	81,28	81,28	81,60	26,72	99,36
36	75,84	75,84	76,00	76,00	34,88	97,44
37	85,12	85,12	85,44	85,28	34,24	100,00
38	94,24	93,92	94,24	94,24	70,72	100,00
39	96,49	96,49	96,65	96,49	76,36	99,84
40	98,40	98,40	98,40	98,56	78,72	100,00

Tabelle 17: Robuster Videohash Detektionsrate Videoset 1

Die folgende Tabelle stellt die Detektionsraten des robusten Videohash für Videoset 2 dar. Dabei haben die in den einzelnen Spalten der Tabelle aufgeführten Angriffe die folgende Bedeutung:

- GoP9 – Verkleinerung der GoP Größe von 12 auf 9 Frames pro GoP
- GoP15 – Vergrößerung der GoP Größe von 12 auf 15 Frames pro GoP
- Bitrate 1 MBit/s – Re-Enkodierung des Videos mit einer Bitrate von 1 MBit/s
- Bitrate 2 MBit/s – Re-Enkodierung des Videos mit einer Bitrate von 2 MBit/s
- FPS 24 – Verringerung der Framerate von 25 auf 24 fps
- FPS 30 – Erhöhung der Framerate von 25 auf 30 fps

VideoNr	GOP9	GOP15	Bitrate 1 Mbit/s	Bitrate 2 Mbit/s	FPS 24	FPS 30
1	100	100	100	100	100	100
2	100	100	100	100	100	100
3	100	100	100	100	100	100
4	100	100	100	100	100	100
5	100	100	100	100	100	100
6	100	100	100	100	100	100
7	100	100	100	100	100	100
8	100	100	100	100	100	100
9	100	100	100	100	100	100
10	100	100	100	100	100	100
11	98,14	98,14	98,14	98,14	100	100
12	99,07	99,07	99,07	99,07	100	100
13	95,33	95,33	95,33	95,33	100	100
14	99,07	99,07	99,07	99,07	100	100
15	97,21	96,75	96,75	96,75	100,48	99,54
16	100	100	100	100	100	100
17	100	100	100	100	100	100
18	96,28	96,28	96,28	96,28	100	99,54
19	99,54	99,54	99,54	99,54	100	100
20	98,6	98,6	98,6	98,6	100	100

Tabelle 18: Robuster Videohash Detektionsrate Videoset 2

Referenzen

[Av2007] Audio-Video-Interleave, http://de.wikipedia.org/wiki/Audio_Video_Interleave, abgerufen am 14.10.2007

[BaCh2001] P. Bas, J.-M. Chassery, B. Macq: *Geometrically invariant watermarking using feature points*, In IEEE TRANSACTIONS ON IMAGE PROCESSING, VOL. 11, NO. 9, SEPTEMBER 2002

[Be2007] BeBerkeley Multimedia Research Center, http://bmrc.berkeley.edu/ftp/-pub/multimedia/mpeg/mpeg2, abgerufen 07.02.2007

[BeRo1998] Beutelspacher, Rosenbaum: *Projective Geometry*, Cambridge University Press, ISBN: 0-5214-8364-6, 1998

[BiDa2005] S. Biswas, S. R. Das, E. M. Petriu: *An adaptive compressed MPEG-2 video watermarking scheme*, In IEEE Transactions on Instrumentation and Measurement, Vol. 54, No. 5, October 2005

[Bo2005] M. E. Borda: *Digital rights protection - a great challenge of the new millennium*, In 7th International Conference on Telecommunications in Modern Satellite, Cable and Broadcasting Services, pp 207 - 214 vol. 1, 28-30 September 2005

[BoSh1995] D. Boneh, J. Shaw: *Collusion-Secure Fingerprinting for Digital Data*, In Proc of CRYPTO ´95, Springer LNCS 963, S. 452-465, 1995

[Br1992] L. Brown: *A survey of image registration techniques*, In Proc. Of ACM Computing Surveys 24, pp. 325-376, Dezember 1992

[Bu2007] Bundenetzagentur: *Geeignete Algorithmen für die elektronische Signatur http://www.bundesnetzagentur.de/enid/11a7b9316dc2e4bdc95eb7be8d24686b ,0/Veroeffentlichungen/Algorithmen_sw.html*, abgerufen am 10.06.2007

[CeSh2006] M. U. Celik, G. Sharma, A. M. Tekalp: *Lossless watermarking for image authentication: a new framework and an implementation*, In IEEE Transactions on Image Processing, Volume 15, Issue 4, April 2006, pp. 1042 – 1049

[ChLi2000] H. Cheng, X. Li: *Partial Encryption of Compressed Images and Videos*, In IEEE Trans. Signal Processing, Volume 48, No. 8, pp. 2439--2451, 2000

[ChLi2003] L. Chun-Shien, H.-Y.M. Liao: *Structural digital signature for image authentication: an incidental distortion resistant scheme*, IEEE Transactions on Multimedia, Volume: 5, Issue: 2, pp. 161- 173, Juni 2003

[ChXu2006] Y. Chung, F. Xu: A Secure Digital Watermarking Scheme for MPEG-2 video Copyright Protection, In IEEE International Conference on Video and Signal Based Surveillance (AVSS'06), p.84 – 89, ISBN 0-7695-2688-8, 2006

[CoMi2002] I. Cox, M. Miller, J. Bloom: *Digital Watermarking*. Academic Press, San Diego, USA, 2002, ISBN 1-55860-714-5

[CoSa2004] B. Coskun, B. Sankur, B. Bogazici: *Robust video hash extraction*, In Proceedings of the IEEE 12th Signal Processing and Communications Applications Conference, pp. 292-295, April 2004, ISBN: 0-7803-8318-4

[DeCs1999] F. Deguillaume, G. Csurka, J. O Ruanaidh, T. Pun: *Robust 3d DFT video watermarking*, Security and Watermarking of Multimedia Contents, Vol. 3657 of SPIE Proceedings, San Jose, USA, Januar 1999

[Di2000] J. Dittmann: *Digitale Wasserzeichen*, Springer Verlag, Berlin, 2000, ISBN 3-540-6661-3

[DiHa2001a] J. Dittmann, E. Hauer, C. Vielhauer, J. Schwenk, E. Saar: *Customer Identification for MPEG-Video based on Digital Fingerprints*, In Proceedings of Advances in Multimedia Information Processing – PCM 2001, The Second IEEE Pacific Rim Conference on Multimedia, Springer Verlag Berlin, Beijing, China, pp. 383-390, ISBN 3-540-42680-9, October 2001

[DiHa2001b] J. Dittmann, E.Hauer, M. Steinebach: *Qualitätssicherung von digitalen Wasserzeichen für Video- und Audiodaten*: Evaluierung der Robustheit und Sicherheit anhand von Angriffsmodellen, Im Tagungsband des 9. Dortmunder Fernsehseminar, Dortmund, Deutschland, pp. 271-275, September 2001

[DiBe1999] J. Dittmann, A. Behr, M. Stabenau, P. Schmitt, J. Schwenk, J. Ueberberg: Combining digital Watermarks and collusion secure Fingerprints for digital Images, In Proceedings of Security and Watermarking of Multimedia Contents, Volume 3657, San Jose, USA, pp. 171-182, January 1999

[DiSt1998] J. Dittmann, M. Stabenau: *Digitale Wasserzeichen für MPEG Video*, In: GMD Report 34, September 1998, S.53-57, ISSN 1435-2702

[DiSt1999] J. Dittmann, A. Steinmetz, R. Steinmetz: *Content-based Digital Signature for Motion Pictures Authentication and Content-Fragile Watermarking*, In: Proceedings of IEEE International Conference on Multimedia Computing and Systems, Volume 2, Florenz, Italien, pp. 209-213, Juni 1999

[DiSt2000] J. Dittmann, M. Steinebach, T. Kunkelmann, L. Stoffels: H204M - Watermarking for Media: Classification, Quality Evaluation, Design Improvements, In Proceedings of the ACM Multimedia 2000 Workshops, Los Angeles, USA, pp. 107-110, Okober 2000

[DiSt2002] J. Dittmann, M. Steinebach, L. Ferri: *Watermarking protocols for authentication and ownership protection based on timestamps and holograms*, In Proceedings of IEEE Security and Watermarking of Multimedia Contents IV, Volume 4675, San Jose, USA, pp. 240 - 251, Januar 2002

[DiSt2004] J. Dittman, M. Steinebach, A. Lang, S. Zmudizinski: *Advanced audio watermarking benchmarking*, In Proceddings of IEEE Security, Steganography, and Watermarking of Multimedia Contents VI, Volume 5306, pp.224-235, Januar 2004

[Di2007] DILIGENT - A testbed DIgital Library Infrastructure on Grid ENabled Technology, EU Projekt, http://www.diligentproject.org/index.php, abgerufen am 22.02.2007

[DRA2006] Deutsches Rundfunkarchiv DRA, www.dra.de, zuletzt abgerufen am 20.02.2006

[DuFr2002] R. Du, J. Fridrich: *Lossless Authentication of MPEG-2 Video*, In Proceedings of ICIP 2002, Rochester, USA, pp.893-896, September 2002

[EBU2001] L. Cheveau, E. Goray, R. Salmon: *Watermarking – Summary results of EBU tests*, EBU Technical Review, Grand-Saconnex, Switzerland, March 2001

[Ec2001] C. Eckert: *IT Sicherheit: Konzepte – Verfahren – Protokolle*, Oldenburg Verlag, München, 2001, ISBN 3-486-25298-4

[Ec2005] Ecrypt – European Network of Excellence in Cryptology: *D.WVL.1 First Summary Report on Fundamentals*, EU Projekt IST-2002-507932, Januar 2005

[Fi1997] S. Fischer: Indikatorenkombination zur Inhaltsanalyse digitaler Filme, D 180 (Diss. Universitat Mannheim), Shaker Verlag Aachen., 1997, ISBN 3-826529-7-66

[FoKe1998] I. Foster, C. Kesselman: *The Grid: Blueprint for a New Computing Infrastructure*, Morgan Kaufmann Publishers, 1998, ISBN 978-1558604759

[Fr2000] J. Fridrich: *Visual Hash for Oblivious Watermarking*, in Proc. Of SPIE Photonic West Electronic Imaging 2000, Security and Watermarking of Multimedia Contents, San Jose, USA, pp. 286–294, January 2000

[FrGo2000] J. Fridrich, M. Goljan: *Robust Hash Functions for Digital Watermarking*, In Proc. of ITCC 2000, Las Vegas, USA, pp. 173–178, March 2000

[FrGo2001] J. Fridrich, M. Golian: *Invertible authentication for JPEG images*, In Proc. of ICIP 2001, Las Vegas, USA, pp. 223-228, Oktober 2001

[Fu2007] Y. Fu: Novel object based robust vieo watermarking scheme, Novel Object Based Robust Video Watermarking Scheme, In Proceedings of 2007 International Conference on Machine Learning and Cybernetics, Volume 3, pp.1705 – 1710, Honglong, China, 2007

[FuSh2007] Y. Fu, R. Shen: *An object based robust watermarking scheme*, In Proceedings of IEEE International Workshop on Anti-counterfeiting, Security, Identification, pp. 188 – 193, April 2007

[Gl2007] gLite – Lightweight Middleware for Grid Computing, http://glite.web.cern.ch/glite/, abgerufen am 22.02.2007

[Ha2000] F. Hartung: *Digital Watermarking and Fingerprinting of Uncompressed and Compressed Video*, Shaker Verlag, Aachen 2000, ISBN 3-8265-7052-9

[Hau2000] E. Hauer: Optimierung eines Wasserzeichenverfahrens zum Einbringen von kundenspezifischen Kundendaten in digitales Datenmaterial, Im Tagungsband des Workshops Sicherheit in Netzen und Medienströmen, Berlin, Deutschland, pp. 169-180, September 2000

[HaDi2005] E. Hauer, J. Dittmann, M. Steinebach: *Digital Signatures Based on Invertible Watermarks for Video Authentication*, In Proceedings of Communications and Multimedia Security, 9th IFIP TC-6 TC-11 International Conference, CMS 2005, Salzburg, Austria, pp. 271-273, September, 2005

[HaKu2005] O. Harmanci, M. Kucukgoz, M. Ki. Mihcak: *Temporal Synchronization of Watermarked Video Using Image Hashing*, In IEEE Security and Watermarking of Multimedia Contents VII, 5681, San Jose, USA, pp. 370-380, Januar 2005

[HaSt2005a] E. Hauer, M. Steinebach: Digitale Fingerabdrücke in Video-on-demand-Szenarien, Im Tagungsband von DACH Security 2005, Darmstadt, Deutschland, März 2005, pp. 398 – 407

[HaSt2005b] E. Hauer, M. Steinebach: Robust digital watermark solution for intercoded frames of MPEG video data, In IEEE Security and Watermarking of Multimedia Contents VII, 5681, San Jose, USA, pp. 381-390, Januar 2005

[HaSt2006] E. Hauer, M. Steinebach: *Temporal synchronization of marked MPEG video frames based on image hash system*, In Proc of IEEE Security and Watermarking of Multimedia Contents VIII, Volume 6072, San Jose, USA, pp. 480-488, Januar 2006

[HaBo2007] E. Hauer, T. Bölke, M. Steinebach: *Framework for combined video frame synchronization and watermark detection*, In Proceedings of SPIE Security, Steganography, and Watermarking of Multimedia Contents IX, -- Volume 6505, San Jose, USA, Januar 2007, pp 650ff

[HaSt2007] E. Hauer, M.Steinebach: *Digitale Bildwasserzeichen im Fernsehumfeld*, 12. Dortmunder Fernsehseminar, Dortmund, Deutschland, März 2007

[HaSt2008] E. Hauer, M. Steinebach, P.Wolf: *The video watermarking container: efficient real-time transaction watermarking*, In Proceedings of IEEE Security and Watermarking of Multimedia Contents IX, Volume 6819, San Jose, USA, Januar 2008, pp. 681 - 690

[HaTh2003] E. Hauer, S. Thiemert, M. Steinebach, J. Dittmann, A. Lang: *Transaktionswasserzeichen für Online-Shops*, 2. Thüringer Medienseminar der FKTG, Erfurt, Germany, 2003

[HaTh2004] E. Hauer, S. Thiemert: *Synchronization techniques to detect MPEG video frames for watermark retrieval*, IEEE Security and Watermarking of Multimedia Contents III, Volume 5306, San Jose, USA, January 2004

[ISO1993] ISO/IEC 11172: *Information Technology - Coding of moving pictures and associated audio for digital storage media up to about 1,5 Mbit/s (MPEG-1) – Part 2: Video*, Genf, Schweiz, 1993

[ISO1995] ISO/IEC 13818-2: *Information Technology - Generic coding of moving pictures and associated audio (MPEG-2) – Part 2: Video*, Genf, Schweiz, 1995

[ISO2001] ISO/IEC 14496-2: *Information Technology – Coding of Audiovisual Objects – Part 2: Visual*, Genf, Schweiz, 2001

[ISO2003a] ISO/IEC 15938: *Multimedia content description interface*, Genf, Schweiz, 2003

[ISO2003b] ISO/IEC 21000: *Multimedia framework (MPEG 21)*, Genf, Schweiz, 2003

[JiYo2000] J. Jing, G. Yong Liang: *Robust Compressed Video Watermarking*, In Proc of IEEE Security and Watermarking of Multimedia Contents II, Volume 3971, San Jose, USA, pp. 477 - 485, January 2000

[JoKu1998] F. Jordan, M. Kutter, T. Ebrahimi: *Proposal of a watermarking technique to hide/retrieve copyright data in video*, ISO/IEC document JTC1/SC29/WG11 MPEG97/M2281, Mai 1998

[KaDe1999] T. Kalker, G. Depovere, J. Haitsma, M. Maes: *A Video Watermarking System for Broadcast System*, In Proceedings of the IEEE International Conference of Image Processing (ICIP `99), Kobe, Japan, pp. 103-112, October 1999

[KaDi2004] S. Katzenbeisser, J. Dittmann: *Malicious Attacks on Media Authentication Schemes Based on Invertible Watermarks*, In IEEE Security, Steganography, and Watermarking of Multimedia Contents VI, Volume 5306, San Jose, USA, pp. 838 – 847, January 2004

[KoYa2007] C.-C. Ko, B.-Z. Yang: An integrated technique for video watermarking, An Integrated Technique for Video Watermarking, In 6th IEEE/ACIS International Conference on Computer and Information Science, ICIS 2007, pp. 37 – 42, July 2007

[Kl2006] C. Kloos: *Evaluierung und Optimierungsvorschläge eines digitalen Videowasserzeichens*, Bachelorarbeit, Fachhochschule Darmstadt, Fachbereich Informatik, Darmstadt, Mai 2006

[Kr2008] Dr. H. J. Krieger: *Das deutsche Urheberrechtsgesetz - UrhG – Stand September 2008*, http://transpatent.com/gesetze/urhg.html, abgerufen am 15.10.2007

[KuJo2002] M. Kutter, F. Jordan: *Digital Watermarking Technology*, abgerufen von der Internetseite http://www.alpvision.com/DOWNLOAD/watermarking.pdf

[LaLa1997] G. C. Langelaar, R.L. Lagendijk, J. Biemond: *Real-time labeling methods for MPEG compressed video*, Proceedings 18th Symposium on Information Theory in the Benelux, Veldhofen, Niederlande, Mai 1997

[LaTh2003] A. Lang, S. Thiemert, E. Hauer, H. Liu, F. A. P. Petitcolas: *Authentication of MPEG-4 data: risks and solutions*, , In IEEE Security, Steganography, and Watermarking of Multimedia Contents V, Volume 5020, Santa Clara, USA, pp. 452 – 461, January 2003

[LeCz2003] F. Lefebvre, J. Czyz, B. Macq: *A robust soft hash algorithm for digital image signature*, In Proceedings ICIP 2003 International Conference on Image Processing, Volume 2, pp. II-495-8 vol.3, September 2003, ISBN: 0-7803-7750-8

[LiCh2001] C.-Y. Lin, S.-F. Chang: *A robust image authentication method distinguishing JPEG compression from malicious manipulation*, IEEE Trans. Circuits Syst. Video Technol., vol. 11, pp. 153–168, Februar 2001

[Li1998] J.-P. Linnartz: MPEG PTY Marking. http://diva.eecs.berkeley.edu/~linnartz/pty.html, 1998, abgerufen Februar 2003

[LiDe2002] E.T. Lin, E.J. Delp: *Temporal Synchronization in Video Watermarking*, In Proceedings of SPIE, Security and Watermarking of Multimedia Contents IV, 21-24 January 2002, San Jose, USA, pp. 478-490

[LiDe2003] E.T. Lin, E.J. Delp: *Temporal Synchronization in Video Watermarking - Further Studies*, In Proceedings of SPIE, Security and Watermarking of Multimedia Contents V, Januar 2003, Santa Clara, USA, pp. 493-504

[LiLu2004] H. Li, Z. Lu, F. Zou: *New Real-time Watermarking Algorithm for compressed video in vlc Domain*, In International Conference on Image Processing ICIP '04, Volume 4, pp. 2171-2174, ISBN: 0-7803-8554-3

[LiHu2006] Y.-R. Lin, H.-Yi. Huang, W.-H. Hsu: *An embedded watermark technique in video for copyright protection*, In 18th International Conference on Pattern Recognition, Volume 4, pp. 795 – 798, 2006

[LiZo2004] H. Ling, Z. Lu, F. Zou: *Improved Differential Energy Watermarking (IDEW) Algorithm for DCT Encoded Images and Videos*, In 7th International Conference on Signal Processing (ICSP '04), Volume 3, pp. 2326 – 2329, September 2004

[LuCh2005] C. S. Lu, J. R. Cheny, K. C. Fany: *Real Time Frame Dependent Watermarking in MPEG Video*, In the Proc of IEEE Int. Conference on Communications, Paris, France, Juni 2004

[MaEl2006] S. Mabtoul, E. Ibn-Elhaj, D. Aboutajdine: A blind Image Watermarking Algorithm Based on Dual Tree Complex Wavelet Transform, In Proceedings of ISCCSP 2006. Marokko, März, 2006.

[Mil1995] T. Milde: *Videokompressionsverfahren im Vergleich*, dpunkt – Verlag für digitale Technologie GmbH, Heidelberg, 1995, ISBN 3-920993-23-3

[MiVe2001] K. Mihcak, R. Venkatesan: *New Iterative Geometric Methods for Robust Perceptual Image Hashing*, In Proceedings of ACM Workshop on Security and Privacy in Digital Rights Management, Springer Verlag London, pp 13 – 21, 2001, ISBN:3-540-43677-4

[Ne2005] S. Neichtadt: *Robust Audio-Hash als Synchronisationsmechanismus für Audiowasserzeichenverfahren*, Diplomarbeit, Techische Universität Darmstadt, Fachbereich Informatik, Darmstadt, 2005

[NiSc2002] X. Niu, M. Schmucker, C. Busch: *Video Watermarking Resisting to Rotation, Scaling, and Translation*, In Proceedings of SPIE, Security and Watermarking of Multimedia Contents IV, pp. 512-519, Januar 2002, San Jose, USA

[NoMe2006] M. Noorkami, R. M. Mersereau: Improving Perceptual Quality in video watermarking using motion estimation, 2006 IEEE International Conference on Image Processing, pp. 1389-1392, Oktober 2006, ISBN: 1-4244-0481-9

[OoKa2001] J.C. Oostveen, A.A.C. Kalker, J.A. Haitsma: *Visual hashing of digital video: applications and techniques*, In Proc of SPIE Applications of Digital Image Processing XXIV, Volume 4472, San Diego, USA, pp. 121-131, July 2001

[PfFe2002] Pfitzmann, Federrath, Kuhn: *DRM-Studie dmmv*, technischer Teil, Download unter http://www.dmmv.de/shared/data/zip/2501_006_019_druck-version020904.zip, September 2002

[PeSt2001] F. A. P. Petitcolas, M. Steinebach, F. Raynal, J. Dittmann, C. Fontaine, N. Fates: *Public automated web-based evaluation service for watermarking schemes: Stirmark Benchmark*, Security and Watermarking of Multimedia Contents III, Ping Wah Wong, Edward J. Delp III, Editors, Proceedings of SPIE Vol. 4314, Bellingham WA, USA, pp. 575 - 584, ISBN 0-8194-3992-4, 2001

[PrRi2005] D. Pröfrock, H. Richter, M. Schlauweg, E. Müller: *H.264/AVC video authentication using skipped macroblocks for an erasable watermark*, In Proceedings of Conference Visual Communications and Image Processing (VCIP 2005). July 2005, Beijing, China

[PrSc2006] D. Pröfrock, M. Schlauweg, E. Müller: *A new uncompressed-domain video watermarking approach robust to H.264/AVC compression*, In Proceedings of Conference Signal Processing, Pattern Recognition and Applications (SPPRA 2006), Februar 2006, Innsbruck, Austria

[RoVl2005] C. De Roover, C. De Vleeschouwer, F. Lefebvre, B. Macq: Key-Frame Radial Projection for Robust Video Hashing, In Proceedings of IEEE Transactions on Signal Processing, [see also IEEE Transactions on Acoustics, Speech, and Signal Processing,], Volume: 53, Issue: 10, Part 2, Oct. 2005, Montreux

[SeKa2002] I. Setyawan, G. Kakes, R.L. Lagendijk: *Synchronization-insensitive Video Watermarking using Structured Noise Pattern*, In Proceedings of SPIE, Security and Watermarking of Multimedia Contents IV, pp. 520-530, San Jose, USA, Januar 2002

[SeLa2001] I. Setyawan, R.L. Lagendijk: *Low bit-rate video watermarking using temporally extended Differential Energy Watermarking (DEW) algorithm*, In Proc of IEEE Security and Watermarking of Multimedia Contents III, Volume 4314, San Jose, USA, pp. 73-84, January 2001

[SkUh2003] C. Skrepth, A. Uhl: *Robust Hash Functions for Visual Data: An Experimental Comparison*, In: Pattern Recognition and Image Analysis, Volume 2652/2003, Springer Verlag Berlin/Heidelberg, ISBN: 978-3-540-40217-6

[St2003] M. Steinebach: *Digitale Wasserzeichen für Audiodaten*, Shaker Verlag, Aachen, 2003, ISBN 3-8322-2507-2

[StDi2002] M. Steinebach, J. Dittmann; C. Neubauer: Anforderungen an digitale Transaktionswasserzeichen für den Einsatz im e-Commerce, Von e-Learning bis e-Payment - Das Internet als sicherer Marktplatz, Klaus P. Jantke, Wolfgang S. Wittig, Jörg Herrmann (Hrsg.), Tagungsband LIT 2002, September 2002, Leipzig, pp. 209 - 217, ISBN 3-89838-033-5, 2002

[StDi2003] M. Steinebach, J. Dittmann: *Watermarking-based digital audio data authentication*, EURASIP Journal on Applied signal processing, No. 10, September; Hindawi Publishing Corporation, pp. 1001 - 1015, 2003

[StHa2003] M. Steinebach, E. Hauer, J. Dittmann: *Digitale Audiowasserzeichen im Archivbereich – Das H2O4M Projekt* Informatik 2003, Mit Sicherheit Informatik, Beiträge des Schwerpunktes "Sicherheit, Schutz und Zuverlässigkeit", 29.09 – 2.10.2003 in Frankfurt am Main, Rüdiger Grimm, Hubert B. Keller, Kai Rannenberg (Hrsg.), S. 395 - 398, ISBN 3-88579-365-2, 2003

[StHa2007] M. Steinebach, E.Hauer, P.Wolf: *Efficient watermarking strategies*, In Proceedings of Proceedings of the Third International Conference on Automated Production of Cross Media Content for Multi-Channel Distribution (AXMEDIS) 2007 Conference, November 2007, Barcelona, Spanien, pp. 65 - 71

[SuCh2004] S.W. Sun and P.C. Chang: *Video Watermarking Synchronization Based on Profile Statistics*, In IEEE Aerospace and Electronic Systems Magazine, Vol.19, No. 5, pp. 21-25, May 2004

[Su2006] D. Sutlar: Sicherheitsanalyse eines invertierbaren Wasserzeichens, Praktikumsbericht, Technische Universität Darmstadt, Fachbereich Informatik, Darmstadt, November 2006

[SuKu2002] K. Su, D. Kundur, D. Hatzinakos: *A Novel Approach to Collusion-resistant Video Watermarking*, In Proc of IEEE Security and Watermarking of Multimedia Contents IV, Volume 4675, San Jose, USA, pp. 491-502, Januar 2002

[SwZh1998] M. D. Swanson, B. Zhu, und A. H. Tewfik: *Multiresolution Scene-based Video Watermarking Using Perceptual Models*, In IEEE Journal on Selected Areas in Communications, Volume 16, No. 4, pp. 540-550, May 1998

[Th2002] S. Thiemert: *Werkzeuge zur Qualitätsevaluierung und Vorschläge zur Optimierung von MPEG-Video-Wasserzeichen*, Diplomarbeit, Hochschule Anhalt (FH), Köthen, Deutschland, Januar 2002

[ThVo2004] S. Thiemert, T. Vogel, J. Dittmann, M. Steinebach: A High-Capacity Block based Video Watermark, In Procedding of EUROMICRO 2004 (30th Conference), Rennes, France, pp. 457 – 460, August 2004

[TrWu2003] W. Trappe, M. Wu, Z. J. Wang, K. J. R. Liu: Anti-collusion Fingerprinting for Multimedia, In IEEE Transaction Journal on Signal Processing, Volume 51, No. 4, April 2003

[WaPe2006] Y. Wang, A. Pearmain: *Blind MPEG-2 video watermarking in DCT domain robust against scaling*, In IEEE Proceedings - Vision, Image, and Signal Processing, Vol. 153, No. 5, Oktober 2006

[WaWu2004] Z. J. Wang, M. Wu, W. Trappe, K. J. R. Liu: Group-Oriented Fingerprinting for Multimedia Forensics, In EURASIP Journal on Applied Signal Processing, Volume 14, pp. 2142-2162, 2004

[WiMc1999] I. H. Witten, R. J. McNab, S. Jones, M. Apperley, D. Bainbridge, S. J. Cunningham: *Managing complexity in a distributed digital library*, In Computer Volume 32, Issue 2, pp. 74 – 79, Februar 1999

[Wi2007] *Wikipedia DVD Video*, http://de.wikipedia.org/wiki/DVD-Video, abgerufen am 08.04.2007

[Wi2008] *Wikipedia Authentizität*, http://de.wikipedia.org/wiki/Authentizit%C3%A4t, abgerufen am 16.03.2008

[WiSu2003] T. Wiegand, G. J. Sullivan, G. Bjontegaard, A. Luthra: *Overview of the H.264 / AVC Video Coding Standard*, in IEEE Transactions on circuits and systems for video technology, Volume 13, Issue 7, pp. 560 - 576, July 2003

[WuTr2004] M. Wu, W. Trappe, Z. Wang, K. J. R. Liu: *Collusion Resistant Fingerprinting for Multimedia*, in Magazine of IEEE Signal Processing, Special Issue on Digital Rights Management, March 2004, pp.15-27

[WuWa2005] G.-Z. Wu, Y-J. Wang, W-H. Hsu: *Robust watermark embedding/detection algorithm for H.264 video*, In Journal of Electronic Imaging, Volume 14, Issue 1, März 2005, Issue 1

[ZhCh2001] Y. J. Zhang, T. Chen, J. Li: *Embedding Watermarks into Both DC and AC Components od DCT*, In Proc of IEEE Security and Watermarking of Multimedia Contents III, Volume 4314, San Jose, USA, pp. 424-435, Januar 2001

[ZhHo2007] J. Zhang, A. T. S. Ho, G. Qiu, P. Marziliano: *Robust Video Watermarking of H.264/AVC*, In IEEE Transactions on Circuits and Systems II: Analog and Digital Signal Processing, Volume 54, Issue 2, Februar 2007, pp. 205-209, ISSN: 1057-7130

[ZhKo1995] J. Zhao, E. Koch: *Embedding Robust Labels into Images for Copyright Protection*, In Proc. International Congress on Intellectual Property Rights for Specialized Information, Knowledge and New Technologies, Vienna, Austria, August 1995

[ZhSc2006] X. Zhou, M. Schmucker, C. L. Brown: *Video Perceptual Hashing Using Interframe Similarity*. In Proceedings of: Sicherheit 2006: Sicherheit - Schutz und Zuverlaessigkeit, Beitraege der 3. Jahrestagung des Fachbereichs Sicherheit der Gesellschaft fuer Informatik e.v. (GI), Magdeburg, pp. 107 – 110, Februar 2006