

TECHNISCHE
UNIVERSITÄT
DRESDEN

Fakultät Informatik

Technische Berichte
Technical Reports
ISSN 1430-211X

TUD-FI05-03 Januar 2005

Sven Buchholz, Martin Heidel

Institut für Systemarchitektur

**Modellierung der Encoding-Kosten und
der Qualität kodierter Videoobjekte
am Beispiel von WMV-9 und MPEG-1**



Technische Universität Dresden
Fakultät Informatik
D-01062 Dresden
Germany
URL: <http://www.inf.tu-dresden.de/>

Modellierung der Encoding-Kosten und der Qualität kodierter Videoobjekte am Beispiel von WMV-9 und MPEG-1

Sven Buchholz und Martin Heidel

Technische Universität Dresden, Fakultät Informatik

Institut für Systemarchitektur, Professur Rechnernetze

{*Sven.Buchholz,Martin.Heidel*}@inf.tu-dresden.de

TUD-FI05-03 Januar 2005

Zusammenfassung

In Systemen, die Videoobjekte zur Laufzeit dynamisch an die Eigenschaften und Fähigkeiten der Clients und deren Netzwerkanbindung anpassen, werden Modelle benötigt, die eine Vorhersage der Adaptionkosten und der Qualität der adaptierten Videoobjekte erlauben. Darüber hinaus werden solche Modelle der Adaptionoperationen auch benötigt, um Systeme zur adaptiven Auslieferung von Videos analytisch oder durch Simulation zu evaluieren.

Die Qualität der kodierten Videoobjekte wird maßgeblich durch den Encoding-Prozess bestimmt. Außerdem spielt der kostenintensive Encoding-Prozess auch für die Gesamtkosten der Adaption eine entscheidende Rolle.

Ziel der vorliegenden Arbeit ist es, am Beispiel der Kodierung adaptierter Videoobjekte mittels des WMV-9-Codecs und eines MPEG-1-Codecs Modelle zur Vorhersage der Kosten zur Kodierung von Videoobjekten und der Qualität der kodierten Videoobjekte bei dynamischer Inhaltsadaption zu entwickeln.

Dazu wird (1) eine Menge von Testvideos mit unterschiedlichen Auflösungen und Bitraten kodiert, (2) dabei die Encoding-Kosten in Form der Encoding-Zeit sowie die Qualität der kodierten Videos in Form des *Peak-Signal-to-Noise-Ratio* (PSNR) ermittelt und (3) Modelle zur Beschreibung des empirisch ermittelten Zusammenhangs zwischen den Eigenschaften der Videosequenzen, den Zielparametern der Adaptionoperationen, der Encoding-Zeit und der Qualität der kodierten Videos abgeleitet.

1 Einleitung

Bei der Auslieferung von digitalen Videoströmen im Internet an Clients mit heterogenen Fähigkeiten zur Videodarstellung (z. B. Displayauflösung, verfügbare Dekoder usw.), die über Netz-

werkanbindungen mit heterogenen Übertragungseigenschaften (z. B. Datenrate) an das Internet angebunden sind, muss

- die Darstellung (räumliche und zeitliche Auflösung, Farbauflösung, Bildausschnitt) und
- die Kodierung (Kodierverfahren, Komprimierung)

der Videos an die Eigenschaften und Fähigkeiten der Endgeräte und Netzwerkanbindungen angepasst werden. Diese Anpassung wird als Inhaltsadaption bezeichnet. Inhaltsadaption kann sowohl statisch als auch dynamisch ausgeführt werden. Statische Inhaltsadaption bedeutet, dass bereits vor der Anforderung durch einen Client (oft bereits zum Zeitpunkt der Inhaltserstellung) verschiedene Repräsentationen der Inhalte erzeugt werden. Bei einer Anfrage eines Clients zur Ausführungszeit wird dann nur noch eine voradaptierte Repräsentation ausgewählt und an den Client ausgeliefert. Bei dynamischer Inhaltsadaption hingegen werden die adaptierten Repräsentationen erst zum Zeitpunkt der Anforderung durch den Client erzeugt [1].

Um zur dynamischen Inhaltsadaption geeignete Adaptionoperationen auswählen und parametrisieren zu können, ist es notwendig, die Kosten für die Ausführung der Adaptionoperationen sowie die Eigenschaften und Qualität der adaptierten Videoobjekte im Voraus abschätzen zu können. Während die Eigenschaften der adaptierten Objekte im allgemeinen als Parameter vorgegeben werden, erfordert eine Vorhersage der Adaptionkosten und der Qualität der adaptierten Videoobjekte geeignete Modelle der Adaptionoperationen. Darüber hinaus werden solche Modelle der Adaptionoperationen auch benötigt, um Systeme zur adaptiven Auslieferung von Videos analytisch oder durch Simulation zu evaluieren.

Die Qualität der kodierten Videoobjekte wird maßgeblich durch den Encoding-Prozess bestimmt. Außerdem spielt der kostenintensive

Encoding-Prozess auch für die Gesamtkosten der Adaption eine entscheidende Rolle.

Ziel der vorliegenden Arbeit ist es, am Beispiel der Kodierung adaptierter Videoobjekte mittels

1. des Windows Media Video 9 Codecs und
2. des MainConcept MPEG-1-Encoders

Modelle zur Vorhersage der Kosten zur Kodierung von Videoobjekten und der Qualität der kodierten Videoobjekte bei dynamischer Inhaltsadaption zu entwickeln. Kosten und Qualität sollen durch die Modelle als Funktion der räumlichen Auflösung, der Bitrate und weiterer objektspezifischer Parameter, die invariante Eigenschaften der Videos repräsentieren, beschrieben werden.

Dazu wird (1) eine Menge von repräsentativen Videosequenzen zu verschiedenen Auflösungen skaliert, (2) die skalierten Videos mittels der genannten Video-Codecs mit unterschiedlichen Bitraten kodiert, (3) dabei die Encoding-Zeit als Maß der Kosten gemessen und (4) die Qualität der kodierten Videoobjekte ermittelt. Aus den empirisch ermittelten Zusammenhängen werden schließlich mathematische Modelle zu deren Beschreibung abgeleitet.

Die weiteren Teile dieses Berichtes gliedern sich folgendermaßen: Im nächsten Abschnitt wird die grundlegende Vorgehensweise zur Analyse und Modellierung der Encoding-Kosten und der Qualität der adaptierten und kodierten Medienobjekte dargestellt. Der Aufbau und die Durchführung der Experimente, die die Basis für die Entwicklung der Modelle bilden, sind in Abschnitt 3 beschrieben. In Abschnitt 4 werden die Ergebnisse der Experimente gezeigt und Modelle der Encoding-Zeiten und der Qualität der kodierten Videoobjekte abgeleitet. Eine Zusammenfassung der Arbeit wird in Abschnitt 5 gegeben.

2 Grundlegende Vorgehensweise

Abbildung 1 zeigt die Vorgehensweise zur Modellierung der Encoding-Kosten und der Qualität der kodierten Videoobjekte, die in der vorliegenden Arbeit verfolgt wird. Dabei wird eine Menge von Beispielvideos adaptiert, dabei neu kodiert und anschließend die Qualität der kodierten Videos bestimmt. Bei der Kodierung wird die Encoding-Zeit als Maß der Encoding-Kosten gemessen.

Als Beispielvideos werden 20 Testsequenzen von der *Video Quality Experts Group (VQEG)* [2] verwendet. Diese repräsentieren ein breites Spektrum von Videoinhalten.

Bei den durchgeführten Experimenten werden folgende Adaptionsoperationen betrachtet:

1. Skalierung der räumlichen Auflösung der Videoobjekte und
2. Reduktion der Bitrate

in deren Abhängigkeit Vorhersagemodelle für die Encoding-Zeit und Qualität der kodierten Videoobjekte bestimmt werden.

Die Skalierung der einzelnen Videoobjekte wird unabhängig von den eingesetzten Kodierverfahren im Voraus der Kodierung vorgenommen. Als Zielauflösungen für die Skalierungsoperation werden die folgenden Auflösungen betrachtet: 96×72 , 128×96 , 160×120 , 192×144 , 240×180 , 320×240 , 480×360 , 640×480 und 704×528 . Die geringeren Auflösungen sind beispielsweise an Displaygrößen von Mobiltelefonen bzw. Handhelds orientiert. Dagegen reichen die höheren Auflösungen bis hin zu Endgeräten mit fernsehtypischen Formaten. Die Vielzahl an betrachteten Auflösungen gewährleistet eine hinreichende Zahl von Stützstellen für die Modellbildung. Die genannten Zielauflösungen gehen davon aus, dass das Seitenverhältnis von 4:3 bei allen Skalierungen der räumlichen Auflösung beibehalten wird.

Die Reduktion der Bitrate wird erreicht, indem das Video mit einer geringeren Bitrate als das Ausgangsvideo neu kodiert wird. Dazu wird die gewünschte Bitrate bei der Kodierung – mittels des WMV-9- und des MPEG-1-Codecs – dem Codec als Zielbitrate vorgegeben. Im vorliegenden Bericht wird ausschließlich die Kodierung mit konstanter Bitrate betrachtet.

Da der WMV-9-Codec für den Einsatz in Umgebungen mit geringen Bandbreiten entworfen wurde, werden hier typische Bitraten für die Modem- oder Mobilfunkübertragung von 9.6 kbit/s, 28.8 kbit/s und 56 kbit/s sowie auch Bitraten von 200 kbit/s, 350 kbit/s und 700 kbit/s betrachtet. Der MPEG-1-Codec ist im Gegensatz zu WMV-9 nicht in der Lage die geringen Bitraten von ≤ 700 kbit/s zu kodieren. Deshalb erfolgt die Kodierung der Videoobjekte mittels MPEG-1 in folgenden Bitraten: 750 kbit/s, 950 kbit/s, 1150 kbit/s, 1350 kbit/s, 1500 kbit/s und 1700 kbit/s.

Die gewählte Bitrate des Videostroms beeinflusst unmittelbar die Qualität des kodierten Videos. Bei geringen Bitraten werden größere Quantisierungsfaktoren gewählt, d. h. es sind weniger Bilddetails reproduzierbar. Bei höheren Bitraten können im Encoding-Prozess kleinere Quantisierungsfaktoren verwendet werden, was sich in einer detaillierteren Bildreproduktion widerspiegelt und einer höheren Qualität entspricht.

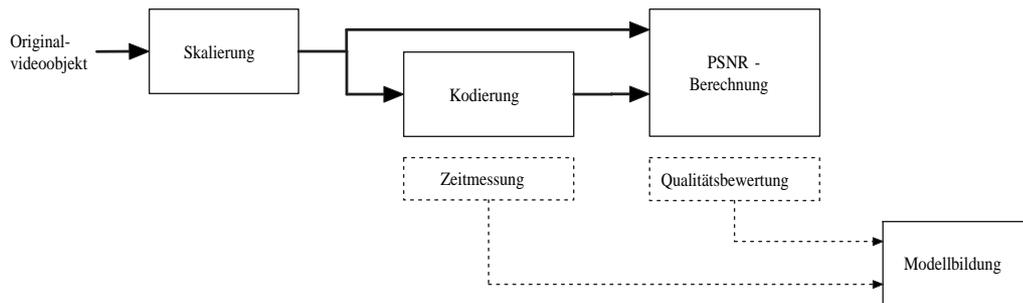


Abbildung 1: Schema zur allgemeinen Vorgehensweise der Modellbildung

Die Beurteilung der Qualität eines adaptierten und kodierten Videobjektes basiert in der vorliegenden Arbeit auf dem objektiven Bewertungsverfahren *Peak-Signal-to-Noise-Ratio* (PSNR). PSNR beschreibt die Qualität eines kodierten Videos anhand der punkweisen Abweichung vom Originalvideo. Originalvideo und kodiertes Video werden dabei allein auf Basis der Helligkeitskomponente jedes Bildpunktes miteinander verglichen. Das PSNR-Verfahren wird in dieser Arbeit aufgrund seiner Einfachheit und schnellen Berechenbarkeit gegenüber kostenintensiven und zeitaufwendigen subjektiven Verfahren (vgl. [3, 2, 4]) zur Qualitätsbewertung vorgezogen. Andere objektive Bewertungsverfahren, die teilweise den subjektiven Qualitätseindruck besser wiedergeben als PSNR, sind zum einen wegen ihrer aufwendigen Parametrisierung (z. B. Video Quality Metric, siehe [5]) und zum anderen aufgrund von Lizenzgebühren (z. B. JNDmetrix [6]) nicht zum Einsatz gekommen.

3 Aufbau und Durchführung der Experimente

3.1 Beispielvideos in den Analysen

Als Beispielvideos werden die 20 Testsequenzen der *Video Quality Experts Group (VQEG) FRTV Phase I* [2] verwendet. Diese Videobjekte wurden von der VQEG zur Evaluation von subjektiven und objektiven Verfahren zur Bewertung der Qualität von Videobjekten zusammengestellt. Sie bildeten darüber hinaus auch die Basis für die Evaluation von Video-Codern in [7, 8].

Alle Videobjekte liegen in einem der Fernsehformate PAL oder NTSC mit den entsprechenden Auflösungen und Bildwiederholraten vor. Sie sind verlustfrei im YUV-Format kodiert. Die hohe Auflösung und verlustfreie Kodierung stellt sicher, dass keines der Ausgangsvideos bereits Skalierungs- oder Komprimierungsartefakte aufweist.

3.1.1 Charakteristika der Beispielvideos

In Tabelle 1 sind die 20 verwendeten Videosequenzen aufgelistet und mit einer Inhaltsangabe sowie einer Kurzcharakteristik versehen.

Das Videomaterial enthält sowohl TV-Aufnahmen als auch Filmmaterial, bis hin zu synthetisch erzeugten Clips und Computeranimierten Szenen. Es wird dabei ein breites Spektrum an Szeneninhalten bereitgestellt, welches die Kodierverfahren in den unterschiedlichsten Bereichen fordert – so gilt es u. a. filigrane Details ebenso wie schnelle Bewegungsabläufe korrekt zu kodieren.

Jedes der 20 Videos hat eine Länge von 8 Sekunden plus 20 zusätzliche Frames. Die 20 zusätzlichen Frames sind für die hier durchgeführten Experimente nicht notwendig und können daher vom Ende einer jeden Videosequenz entfernt werden, so dass für alle Videos in den beiden Fernsehformaten PAL und NTSC die Laufzeit exakt 8 Sekunden beträgt – unabhängig von den voneinander verschiedenen Bildwiederholraten beider Fernsehformate.

Die Fernsehformate PAL und NTSC beinhalten 720 Pixel pro Zeile. Die ersten 10 Videos umfassen 576 Zeilen bei 50 Hz (Format PAL), wogegen die restlichen 10 Videos 486 Zeilen bei 60 Hz (Format NTSC) enthalten. Sie sind *interlaced* abgetastet. Interlaced-Abtastung ist ein Verfahren, das ursprünglich aus der analogen Videotechnik stammt. Dabei bilden zwei aufeinander folgende Halbbilder das darzustellende Bild. Das erste Halbbild enthält nur die ungeradzahligen Zeilen und das zweite Halbbild enthält die entsprechenden geradzahligen Zeilen. Somit sind beim PAL-Format nur 25 Vollbilder pro Sekunde und beim NTSC-Format 29.97 Vollbilder pro Sekunde kodiert.

Alle Videos der VQEG sind als verlustfrei komprimierter UYVY-Binärdatenstrom bereitgestellt. UYVY gehört zur Klasse der YUV-4:2:2-Formate. Das bedeutete, dass die Luminanz (Y) in voller Auflösung gespeichert wird und die Chrominanz-

Video	Inhalt	Charakteristik	Format
1	Baum	Standbild, verschiedene Ausrichtungen	PAL
2	Eröffnungsfeier	kräftige Farben, Verdeckunseffekte	PAL
3	Harfe	kräftige Hintergrundfarbe, Zoom, filigrane Details	PAL
4	Grafiken (Desktop)	bewegter Text, bewegte Objekte, scharfe Kanten, synthetisch	PAL
5	Kanu	Wasseroberfläche, Wasserbewegung, Bewegung in verschiedene Richtungen, detaillierter Hintergrund	PAL
6	Formel 1 Auto	schnelle Bewegungen, kräftige Farben verschiedene Strukturen	PAL
7	Fast-Food- Restaurant	rasche Objektverfolgung, bewegte vertikale Streifen	PAL
8	Schriftband	horizontales Scrollen von Texten	PAL
9	Rugby	viele und schnelle Bewegungen, Verwendung von Komplementärfarben	PAL
10	Modelleisenbahn	Bewegung in Vorder- und Hintergrund, intensive Farben, filigrane Details	PAL
11	Luftballons	kräftige Farben, scharfer Umschnitt, Bewegung	NTSC
12	New York	Kameraschwenk, detaillierte Objekte, strenge, schmale Linienverläufe	NTSC
13	Modelleisenbahn	wie Video 10	NTSC
14	Animationsfilm	synthetische Objekte, Szenenschnitte, Bewegung, flächige Farbverläufe	NTSC
15	kreisende Buchstaben	Bewegung in sämtliche Richtungen, Transparenz, scharfe Kanten	NTSC
16	Herbstbäume	Landschaft, Wasserfallbewegungen, Zoom, ähnliche Objekte und Farben	NTSC
17	Football	viel Bewegung, schneller Kameraschwenk	NTSC
18	Segelboot	kaum Bewegung, Helligkeitskontrast	NTSC
19	Susie	Hautfarbe, Haare, kaum Bewegung	NTSC
20	Pflanze	filigrane Details, kräftige Farben, Zoom, kleine „Stör-“ Objekte	NTSC

Tabelle 1: Charakteristik der Videoobjekte

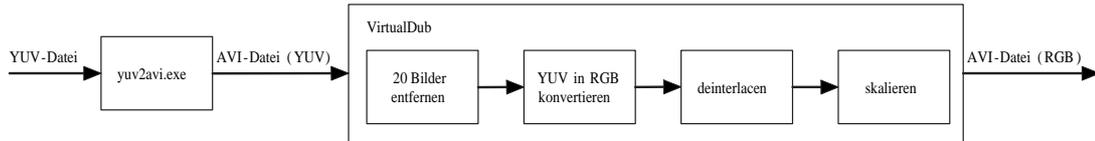


Abbildung 2: Ablauf der Videovorverarbeitung

werte U und V lediglich in der halben Auflösung. Dadurch kann jeder Bildpunkt des Bildes durch 16 Bit anstatt mit 24 Bit im unkomprimierten YUV 4:4:4 Format repräsentiert werden. Je ein Byte belegt dabei der Y-Wert und alternierend wird das andere Byte vom U- bzw. V-Wert besetzt.

3.1.2 Vorbereitung der Beispielveideos

Der vorliegende Rohdatenstrom in der UYVY-Formatierung ist nicht direkt für typische Videoverarbeitungsprogramme, wie sie z. B. im Windows-Betriebssystem zur Verfügung stehen, nutzbar. Für die folgenden Untersuchungen ist es notwendig diese Daten in einem AVI-Container zu speichern. Zu diesem Zweck ist von der Firma Streamcrest (www.streamcrest.com) das Kommandozeilen-Tool 'yuv2avi.exe' entwickelt worden. Damit können die binären Videodaten verlustfrei in das AVI-Dateiformat überführt werden und stehen für die Weiterverarbeitung durch verschiedene Videoverarbeitungs-Werkzeuge in der Windows-Umgebung zur Verfügung.

Für die weitere Vorverarbeitung des Videomaterials wurde ausschließlich das flexible OpenSource-Tool VirtualDub (in Version 1.5.8) eingesetzt. In VirtualDub sind alle für die Vorverarbeitung notwendigen – und im Folgenden beschriebenen – Operationen durch Filter implementiert. Die Filterkette ist in Abbildung 2 illustriert.

Da die 20 zusätzlichen Frames eines jeden Videos (vgl. Abschnitt 3.1.1) für die hier durchgeführten Experimente nicht notwendig sind, werden in Vorbereitung der Experimente zunächst diese 20 Extraframes jeweils vom Ende des Videostroms entfernt. Damit entstehen Videosequenzen mit einer Laufzeit von exakt 8 Sekunden. Somit enthalten die untersuchten Videosequenzen im PAL-Format insgesamt $8\text{ s} \cdot 25\text{ fps} = 200$ und im NTSC-Format $8\text{ s} \cdot 29.97\text{ fps} \approx 240$ Vollbilder.

Anschließend müssen die interlaced abgetasteten Videos in eine progressive Abtastung umgewandelt werden. Dieser Vorgang wird auch als *Deinterlacing* bezeichnet. Beim *Deinterlacing* werden je zwei aufeinander folgende Halbbilder unter Berücksichtigung der Bewegung zu einem Vollbild zusammengefügt. Dieser Schritt ist not-

wendig, da digitale Kodierverfahren, wie WMV-9 und MPEG-1, für eine effektive Bewegungserkennung und Komprimierung progressiv abgetastete Videoströme voraussetzen. Zum *Deinterlacing* wird der *deinterlace*-Filter in VirtualDub verwendet. Bei der Vorbereitung der Videos für die Untersuchungen wurde die höchste Qualitätsstufe *blend fields together* des *deinterlace*-Filters verwendet.

Da der *deinterlace*-Filter in VirtualDub Ausgangsvideos erfordert, die im RGB-Farbsystem kodiert sind, müssen die Ausgangsvideos zuvor vom YUV- in das RGB-Farbsystem überführt werden. Auch hierfür wird ein spezieller Filter in VirtualDub verwendet.

Für die Skalierung der Videos in die in Abschnitt 2 genannten Zielauflösungen wird der *resize*-Filter in VirtualDub eingesetzt. Als Interpolationsverfahren wird bei der Skalierung bikubische Interpolation verwendet. Bikubische Interpolation ermöglicht die Erzeugung scharfer Bilder und verhindert unregelmäßige Kantenverläufe bei unterschiedlichen Skalierungen.

Die so erzeugten Videos können nun direkt für die PSNR-Berechnung und die Kodierung verwendet werden. Die beiden zur Kodierung eingesetzten Codecs (s. Abschnitt 3.2.1 und 3.2.2) setzen zusätzlich noch spezielle Optionen, die die Qualität und die Bitrate des Ausgabevideos beeinflussen.

3.2 Verwendete Codecs

Bevor die verwendeten Codecs genauer betrachtet werden, sollen zunächst Anforderungen an die für die Untersuchung eingesetzten Codecs aufgezeigt werden.

Voraussetzung für den Einsatz eines Codecs ist, dass dieser Videos in verschiedenen räumlichen Auflösungen (vgl. Abschnitt 2) kodieren kann. Darüber hinaus soll die Kodierung mit unterschiedlichen Bitraten möglich sein. Zusätzlich muss der Codec die Kodierung im Stapelverarbeitungsbetrieb unterstützen. Das kann entweder über die Einbindung in die *Video for Windows*-Umgebung (was die Nutzung der Stapelverarbeitungsmöglichkeiten von Videobearbeitungssoftware wie VirtualDub oder Adobe Premiere er-

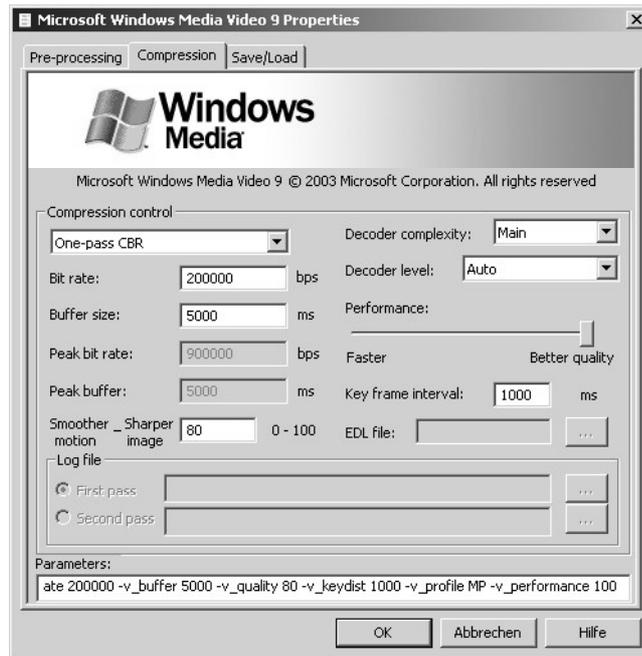


Abbildung 3: Einstellungsfenster für *Windows Media Video*

möglichst) oder durch ein eigenes Konvertierungsprogramm mit entsprechenden Fähigkeiten erfolgen. Codecs, die nur eine interaktive Schnittstelle zur Verfügung stellen, sind aufgrund des Aufwands für die Durchführung einer großen Zahl von Experimenten nicht geeignet.

Als proprietäres Kodierverfahren wurde der frei verfügbare *Windows Media Video* (WMV) Codec in der Version 9 eingesetzt. Dieser Codec wird in die *Video for Windows*-Umgebung eingebunden und ist somit in Adobe Premiere zugänglich, welches die kodierten Videodaten in einer AVI-Datei ablegt.

Für die *MPEG-1*-Kodierung wurde der MPEG Encoder der Firma MainConcept ausgewählt. Das zugehörige Konvertierungsprogramm verfügt über eine Stapelverarbeitung und Videos in den verschiedenen räumlichen Auflösungen und mit den Bildraten von 25 fps und 29.97 fps werden verarbeitet.

3.2.1 Windows Media Video

Alle für die Untersuchung notwendigen Einstellungen des WMV-9-Codec sind anhand des in Abbildung 3 dargestellten Eigenschaften-Fensters vorgenommen und werden im Folgenden erläutert.

Als erstes wird der Encoding-Modus *One-Pass CBR* gewählt, dieser hat den größten Einfluss auf die Qualität der Ausgabe. WMV bietet zwar die Möglichkeit qualitativ hochwertigere Video-

sequenzen durch den Encoding-Modus *Two-Pass CBR* zu generieren, dieser führt jedoch bei der Anfragebeantwortung zu unakzeptabel langen Wartezeiten. Bei *Two-Pass*-Encoding-Verfahren kann erst im zweiten Durchlauf mit der Übertragung des kodierten Videostroms begonnen werden, d. h. während des ersten Durchlaufs – in dem das komplette u. U. sehr lange Video einmal verarbeitet werden muss – muss der Anfragende warten, ohne ein Bild angezeigt zu bekommen. Bei *One-Pass CBR* wird dagegen nach einer viel kürzeren Startverzögerung der Anfragende mit dem kodierten Videostrom versorgt.

Zusätzlich zur Wahl des Encoding-Modus muss nun eine beliebige aber feste *Bit Rate* zusammen mit einer *Buffer Size* eingestellt werden. Die Größe des zur Verfügung gestellten Buffers hat einen entscheidenden Einfluss auf die erzeugbare Bildqualität. Je größer die Zeitspanne für die *Buffer Size* gewählt wird, umso größer ist der Spielraum des Codecs eine höhere, als die angestrebte *Bit Rate* kurzer Videoabschnitte mit benachbarten Videoabschnitten geringerer Bitraten auszugleichen. Können diese kurzfristig auftretenden höheren Bitraten durch eine zu klein gewählte Zeitspanne nicht ausgeglichen werden, so müssen Daten und damit Bildinhalte bei der Kodierung verworfen werden. Bei der Zuteilung der *Buffer Size* von 1 s (1000 ms) lassen sich nur stark ruckelnde Videosequenzen erzeugen, dagegen liefert die Standardeinstellung mit 5 s angenehm weiche

Bewegungsübergänge. Erst durch eine enorme Erhöhung der Bitrate kann dieses Manko wieder ausgeglichen werden. In den Untersuchungen wird deshalb eine *Buffer Size* von 5000 ms gewählt.

Die Bitraten werden entsprechend Abschnitt 2 zwischen 9 600 bps, 28 800 bps, 56 000 bps, 200 000 bps, 350 000 bps und 700 000 bps variiert. Das *Key Frame Interval* wird auf 1000 ms eingestellt, damit sich Bildausfälle infolge verlorengegangener I- und P-Frames auf möglichst kurze Zeitabschnitte beschränken lassen. Der Regler *Performance* wird zwischen der ganz linken und der ganz rechten Position variiert. Diese beiden Optionen werden im folgenden als Low Quality (LQ) und High Quality (HQ) bezeichnet.

Konkret werden 7 Auflösungen (96×72, 128×96, 160×120, 240×180, 320×240, 640×480 und 704×528) aus Abschnitt 2 gewählt und es werden die 6 niedrigeren Bitraten (9.6, 28.8, 56, 200, 350 und 700 kbit/s) für die Messungen in WMV in Kombination mit den beiden Qualitätsstufen (LQ und HQ) eingestellt. D. h., es ergeben sich für ein zu untersuchendes Videoobjekt 84 verschiedene Testfälle ($7 * 6 * 2 = 84$).

Alle restlichen Einstellungsmöglichkeiten werden entsprechend den in Abbildung 3 gezeigten Standardeinstellungen belassen.

Da der Codec in die *Video for Windows*-Umgebung eingebunden wird, kann er in Adobe Premiere (hier in der Version 6.0) genutzt werden. Die Kodierung in WMV-9 ist über sämtliche Videosequenzen, alle Auflösungen und alle Bitraten problemlos durchführbar. Allerdings enthalten insbesondere Videos, die mit der kleinsten Bitrate von 9.6 kbit/s kodiert werden, kaum Inhalte des Quellvideos. Die kodierten Videosequenzen haben häufig überhaupt keinen Bildinhalt bzw. nur ein oder zwei Bilder von 200 bzw. 240 Bildern des gesamten Videostroms enthalten.

Adobe Premiere bestimmt parallel zur Kodierung die Encoding-Zeit. Die Weiterverarbeitung dieser Messgröße ist in Abschnitt 3.3.1 beschrieben.

Das Kommandozeilenskript `wmcmd.vbs` von Microsoft, welches eine Konvertierung in beliebige Windows Media Formate mit benutzerdefinierten Einstellungen erlaubt, konnte nicht verwendet werden, da der Ausgabestrom immer in das proprietäre WMV-Dateiformat eingebunden wird. Das WMV-Dateiformat konnte im Gegensatz zu einem AVI-Container, wie er durch Adobe Premiere erzeugt wird, nicht für die PSNR-Berechnung (Abschnitt 3.3.2) eingelesen werden. Auch das frei verfügbare Videobearbeitungswerkzeug VirtualDub war nicht für die WMV-9-Kodierung nutzbar, da die Kodierung bei Auflösungen über 352×288 Bildpunkten (*CIF*) regelmäßig mit nicht

nachvollziehbaren Fehlermeldungen abgebrochen wurde.

3.2.2 MainConcept MPEG-1-Encoder

Zur Erzeugung von MPEG-1-kodierten Videoobjekten wird eine Evaluationsversion des MPEG-1-Encoder der Firma MainConcept in der Version 1.4.2 gewählt. Diese Evaluationsversion bringt bei der Kodierung ein Logo einer Größe von etwa 80 × 60 Bildpunkten in die linke obere Ecke jedes Einzelbildes des kodierten Videos auf (vgl. Abbildung 4). Um daraus resultierende Verfälschungen bei der Bewertung der Qualität des kodierten Videos zu vermeiden, ist es notwendig bei der Ermittlung des PSNR-Wertes den Bildausschnitt mit dem Logo unberücksichtigt zu lassen. Um trotzdem einen hinreichenden Bildausschnitt für die PSNR-Bewertung zur Verfügung zu haben, sollte die Größe der untersuchten Videos eine Auflösung von 160 × 120 Bildpunkten nicht unterschreiten. Deshalb werden in den Experimenten mit dem MPEG-1-Encoder die folgenden Auflösungen berücksichtigt: 160 × 120, 192 × 144, 240 × 180, 320 × 240, 480 × 360, 640 × 480 und 704 × 528.



Abbildung 4: Logo in Videos, die mittels des MainConcept MPEG-1-Encoders kodiert worden sind

Im Konvertierungsprogramm von MainConcept existiert ein spezifisches Eigenschaftensfenster für den MPEG-1-Codec (Abb. 5), um die notwendigen Einstellungen für die unterschiedlichen Adaptionsoperationen vornehmen zu können.

Die *Breite* und *Höhe* sind entsprechend der räumlichen Auflösung des vorab skalierten Videos anzugeben. Die *Framerate* wird entsprechend der Bildwiederholrate der Videoobjekte auf 25 fps bzw. 29.97 fps eingestellt. Die angestrebten *Bitraten* können in der Eingabemaske, die über die Schaltfläche „Erweitert...“ erreicht wird, angegeben werden. Sie werden gemäß Abschnitt 2



Abbildung 5: Einstellungsfenster für MPEG-1

zwischen 750 kbit/s, 950 kbit/s, 1150 kbit/s, 1350 kbit/s, 1500 kbit/s und 1700 kbit/s variiert.

Die Schieberegler *Suchmethode* und *Suchbereich* sind laut Dokumentation für das Ausbalancieren zwischen Encoding-Zeit und Qualität des kodierten Videos vorgesehen. Bei den Experimenten konnte jedoch kein Einfluss dieser Einstellungen auf Encoding-Zeit oder PSNR-Wert des kodierten Videos festgestellt werden. Daher wird für alle Experimente sowohl *Suchmethode* als auch *Suchbereich* auf den Wert 8 eingestellt. Somit entfällt eine Unterscheidung zwischen Low Quality und High Quality, wie sie beim WMV-9-Codec vorgenommen wird.

Alle weiteren Einstellungen werden entsprechend den in Abbildung 5 gezeigten Standardwerten belassen.

Während der Kodierung wird vom MainConcept MPEG-1-Encoder die benötigte Encoding-Zeit automatisch erfasst und steht für die Auswertung in einer Logdatei zur Verfügung.

3.3 Bestimmung der Messgrößen

In den Experimenten wird die Encoding-Zeit als Maß der Encoding-Kosten und der PSNR-Wert des kodierten Videos als Qualitätsmetrik gemessen.

3.3.1 Encoding-Zeit

Als Encoding-Zeit wird die Gesamtausführungszeit des Encoding-Prozesses bei Verwendung der verschiedenen Codecs gemessen.

Entsprechend Abschnitt 3.2.1 erfolgt die Kodierung mittels des WMV-9-Codecs innerhalb von Adobe Premiere. Folglich kann auch die Zeitmessung nur innerhalb dieses Programms durchge-

führt werden. Adobe Premiere gibt im Anschluss an die Stapelverarbeitung eine Logdatei aus, in der alle Anfangs- und Endzeiten der ausgeführten Videobearbeitungsoperationen festgehalten sind. Die Differenzen der Anfangs- und Endzeiten ergeben die Gesamtausführungszeiten der einzelnen Encoding-Prozesse. Die kleinste Zeiteinheit entspricht dabei einer Sekunde. Genauere Zeitmessungen sind mittels Adobe Premiere nicht möglich. Da Adobe Premiere nicht im Quellcode vorliegt, sind auch keine entsprechenden Erweiterungen realisierbar.

Der MainConcept MPEG-1-Encoder legt detaillierte Informationen über den Encoding-Prozess in Logdateien ab. Diese enthalten unter anderem die Bitrate und die Encoding-Zeit des Videoobjektes. Auch hier ist die Zeitmessung lediglich sekunden-genau verfügbar.

Um den Einfluss zufälliger Messfehler auf die Ergebnisse der Zeitmessung zu minimieren, wird jede Konvertierung, und damit auch die Ermittlung der Encoding-Zeit, zehnmal nacheinander ausgeführt und ein Mittelwert der Encoding-Zeiten bestimmt. Dabei werden für jede Konvertierung die zwei kleinsten und die zwei größten Messwerte unberücksichtigt gelassen. Für die Mittelwertbildung wurden also lediglich 6 Messwerte herangezogen. So ist sichergestellt, dass Ausreißer den Mittelwert nicht signifikant verfälschen.

3.3.2 PSNR-Wert

Die Qualitätsbeurteilung für die erzeugten Videos erfolgt mittels des objektiven Bewertungsverfahrens *Peak Signal-to-Noise Ratio (PSNR)*. PSNR beschreibt anhand der Differenz der Helligkeit, der *Luminanz*, der einzelnen Bildpunkte den Unterschied zwischen dem Original und dem komprimierten Videoobjekt.

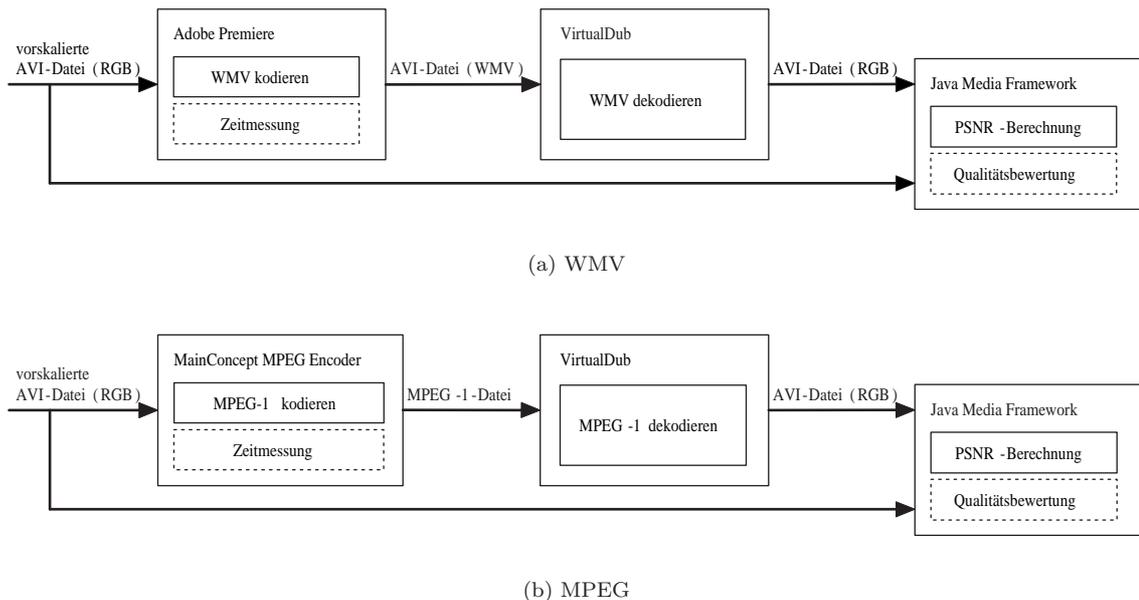


Abbildung 6: Zeitmessung und Qualitätsbewertung

In den Analysen im Zusammenhang mit dem vorliegenden Bericht wird eine eigene Implementierung des PSNR-Verfahrens verwendet. Die Implementierung basiert auf dem *Java Media Framework (JMF)* [9]. JMF ist eine Java-Klassenbibliothek, die eine Vielzahl von Funktionalitäten zur Bearbeitung und Wiedergabe von Videostreamen zur Verfügung stellt.

Für die Berechnung des PSNR müssen zwei Videostreamen, das Original und das kodierte Video, parallel eingelesen und Bild für Bild miteinander verglichen werden. Da JMF für die Videobearbeitung unter Echtzeitbedingungen entwickelt wurde, stellt die parallele Auswertung der zwei Videosequenzen vor allem bei hohen Auflösungen (ab ca. 320×240) enorme Anforderungen an das System. Deshalb kann es vorkommen, dass beim Einlesen und Dekodieren mittels JMF einzelne Frames nicht erfasst werden, wodurch die PSNR-Auswertung unmöglich wird.

Um dieses Problem zu umgehen werden die kodierte Videos bereits vor dem Einlesen in Java in das unkomprimierte RGB-Format überführt (vgl. Abbildung 6). Damit entfällt das ressourcenintensive Dekodieren in Java. Bei der Umwandlung in das RGB-Format bleiben die Fehler und Verfälschungen des kodierte Videos erhalten. Die Kompressionsergebnisse stehen also nach wie vor zur Verfügung. Die RGB-Konvertierung wird, wie bereits die Vorbereitung der Beispielveideos (Abschnitt 3.1.2), mittels VirtualDub vorgenommen.

Die unkomprimierten RGB-Videos werden in einem AVI-Container gespeichert.

Um zwei Videos anhand der Luminanz vergleichen zu können, müssen zunächst die Luminanzwerte $Y_{h,i,j}$ aller Bildpunkte (i, j) aller Einzelbilder h der beiden RGB-kodierten Videos bestimmt werden. Dazu wird die in der ITU-R BT.601 [10] empfohlene Umrechnungsvorschrift verwendet:

$$Y_{h,i,j} = 0.299R_{h,i,j} + 0.587G_{h,i,j} + 0.114B_{h,i,j} \quad (1)$$

Wie in Formel 1 zu sehen ist, fließen dabei die einzelnen Farbanteile Rot $R_{h,i,j}$, Grün $G_{h,i,j}$ und Blau $B_{h,i,j}$ der Bildpunkte mit unterschiedlichen Gewichtungen in den Luminanzwert $Y_{h,i,j}$ ein. Aus den Luminanzwerten $Y_{h,i,j}$ des Originals und $Y'_{h,i,j}$ des kodierte Videos wird der mittlere quadratische Abstand MSE ermittelt:

$$MSE = \frac{\sum_{h=1}^H \sum_{i=1}^I \sum_{j=1}^J (Y_{h,i,j} - Y'_{h,i,j})^2}{H \cdot I \cdot J} \quad (2)$$

Wie in Abschnitt 3.2.2 dargelegt, bringt der MainConcept MPEG Encoder auf jedes Bild des Videostroms ein Logo mit den Abmessungen von ca. 80×60 Bildpunkten auf. Dies kann insbesondere bei der Bewertung von Videoobjekten in kleineren Auflösungen ($< 160 \times 120$) zu signifikanten Verfälschungen des mittleren quadratischen Fehlers, der durch den Encoding-Prozess verursacht wird, führen. Um dadurch bedingte Fehler bei der Qualitätsbewertung mittels PSNR zu vermeiden, sollen bei der Bestimmung des mittleren quadratischen



Abbildung 7: Bildausschnitt von $\frac{3}{4}$ der Bildfläche, der zur PSNR-Berechnung herangezogen wird

Fehlers die Teile der Bildfläche mit Logo nicht berücksichtigt werden. Um trotzdem in allen Experimenten den gleichen Bildausschnitt zur Qualitätsbewertung heranzuziehen, werden grundsätzlich die Fehler im linken oberen Viertel der Bildfläche ignoriert (auch bei hohen Auflösungen, bei denen das Logo nur einen geringen Teil der Bildfläche belegt).

Der mittlere quadratische Fehler wird demzufolge nur über den verbleibenden $\frac{3}{4}$ der Bildfläche (vgl. Abbildung 7) berechnet:

$$MSE = \frac{\sum_{h=1}^H (\sum_{i=1}^I \sum_{j=1}^J (Y_{h,i,j} - Y'_{h,i,j})^2 - \sum_{i=1}^{I/2} \sum_{j=1}^{J/2} (Y_{h,i,j} - Y'_{h,i,j})^2)}{\frac{3}{4} \cdot H \cdot I \cdot J} \quad (3)$$

Aus dem Verhältnis des Quadrates der maximalen Signalstärke (bei 8-bit-Kodierung der Luminanz ist das $2^8 - 1 = 255$) und dem mittleren quadratischen Fehler ergibt sich durch logarithmieren der PSNR-Wert:

$$PSNR = 10 \log_{10} \frac{255^2}{MSE} \quad (4)$$

Für den Zusammenhang zwischen dem PSNR-Wert eines Videos und dem subjektiven Qualitätseindruck, wie er zum Beispiel durch eine Gruppe von Versuchspersonen beurteilt würde und der entsprechend ITU-R BT.500-11 [11] als *Mean Opinion Score (MOS)* beschrieben werden kann, gilt laut [12] näherungsweise die Abhängigkeit gemäß Tabelle 3. Diese Zuordnung ist jedoch keinesfalls als absolut anzusehen, sondern soll lediglich einen Anhaltspunkt bieten, auf welche visuell wahrnehmbare Qualitätsstufe ein ermittelter PSNR-Wert hindeutet.

3.4 Testumgebung

Die wichtigsten Eigenschaften der Testumgebung sind in Tabelle 2 aufgeführt.

Alle generierten Videos werden auf einer neu angelegten, 24 GB großen Partition der genannten Festplatte abgelegt. Um störende Einflüsse auf den Kodiervorgang zu minimieren, werden nicht benötigte Prozesse, wie Antivirensoftware und andere Überwachungsprogramme, soweit wie möglich deaktiviert.

Prozessor:	AthlonXP 1800+ (1.53 GHz), 128/256 KB Cache
Hauptspeicher:	512 MB, DDR-PC2100
Festplatte:	WD 120 GB, eIDE UDMA-100, 2 MB Cache, 7200 rpm
Betriebssystem:	Windows XP Professional

Tabelle 2: Charakteristik des Testsystems

4 Ergebnisse und Modellbildung

In diesem Abschnitt werden die Ergebnisse der Experimente ausgewertet. Zunächst wird ein Überblick über die Messergebnisse gegeben (Abschnitt 4.1). Auf Basis der Ergebnisse der Experimente werden in Abschnitt 4.2 die Zusammenhänge zwischen den Zielparametern der Adaptionoperationen (räumliche Auflösung, Bitrate) und den Encoding-Zeiten (Encoding-Kosten) bzw. den PSNR-Werten (Qualität) der kodierten Videos beschrieben und mathematisch modelliert.

4.1 Messergebnisse

Die Experimente haben gezeigt, dass es dem WMV-9-Codec im Allgemeinen nicht gelingt, Videos mit einer räumlichen Auflösung von 640×480 oder mehr Bildpunkten mit weniger als 56 kbit/s zu kodieren. Videos mit Auflösungen von 640×480 und 704×528 Bildpunkten, die mit einer geringeren Zielbitrate kodiert worden sind, enthalten oft überhaupt keine Bildinformationen oder es wurden nur 1 bis 2 Einzelbilder der Videosequenz kodiert. Das spiegelt sich auch in PSNR-Werten von unter 10 dB wider. Die Encoding-Zeiten zur Erzeugung dieser unbrauchbaren Ergebnisse lagen jedoch unverhältnismäßig hoch (vgl. Tabelle 5). Die hohen Encoding-Zeiten in Verbindung mit der

	subjektive Kriterien	PSNR-Werte (dB) bei		
		Detailgehalt		
		gering	mittel	hoch
<i>hervorragend</i> (<i>excellent</i>)	keine sichtbare Verzerrung zum Original	> 40	> 37	> 35
<i>gut</i> (<i>good</i>)	gering sichtbare Verzerrungen	34 – 40	31 – 37	29 – 35
<i>zufriedenstellend</i> (<i>fair</i>)	mittlere bis starke Verzerrungen	28 – 34	25 – 31	24 – 29
<i>gering</i> (<i>poor</i>)	erkennbarer Bildinhalt, aber schlechte Qualität	22 – 28	20 – 25	20 – 24
<i>schlecht</i> (<i>bad</i>)	kaum erkennbarer Bildinhalt	< 22	< 20	< 20

Tabelle 3: MOS-Skala und Approximation der vergleichbaren PSNR-Werte, nach [12] S. 188

ungenügenden Bildqualität deuten auf eine erfolglose Suche des Encoders nach geeigneten Quantisierungsfaktoren zur Kodierung des Videos hin – ohne die Anforderungen nach geringen Bitraten bei hohen Auflösungen befriedigen zu können. Aus diesem Grund bleiben die Ergebnisse der Experimente, die als Zielparame- ter der Adaption Bitraten von 9.6 kbit/s oder 28.8 kbit/s in Verbindung mit den Auflösungen 640×480 bzw. 704×528 verwenden, für alle weiteren Untersuchungen mit dem WMV-9-Codec unberücksichtigt.

Darüber hinaus weichen beim WMV-9-Codec die tatsächlich erzielten Bitraten teilweise erheblich von den angestrebten Zielbitraten (9.6 kbit/s, 28.8 kbit/s, 56 kbit/s, 200 kbit/s, 350 kbit/s und 700 kbit/s) ab (Tab. 4). Dies deutet auf eine Schwäche des Codecs bei der Abschätzung geeigneter Quantisierungsfaktoren hin. Aus diesem Grund werden bei der Modellbildung nicht die angestrebten Zielbitraten, sondern die tatsächlich erreichten Bitraten als Parameter herangezogen.

Die Mittelwerte der ermittelten Encoding-Zeiten und PSNR-Werte sind in den Tabellen 5, 6, 7 und 8 zusammengefasst. Zur leichteren Einordnung sind die PSNR-Werte in den Tabellen 7 und 8 mit verschiedenen Grautönen unterlegt, welche den entsprechenden subjektiven Qualitätseindruck repräsentieren (vgl. Tabelle 3). Ein dunkler Hintergrund steht dabei für eine schlechte Bildqualität, wohingegen ein weißer Hintergrund eine sehr gute Qualität verdeutlichen soll.

Wie in Abschnitt 3.2.2 dargelegt, bringt der MainConcept MPEG Encoder auf jedes Bild des Videostroms ein Logo auf. Um dadurch

bedingte Fehler bei der Qualitätsbewertung mittels PSNR zu vermeiden, wird bei der Berechnung des PSNR-Wertes der Fehler im linken oberen Viertel der Bildfläche ignoriert (vgl. Abschnitt 3.3.2). Der Einfluss der Wahl des Bildausschnitts auf den ermittelten PSNR-Wert lässt sich aus den Unterschieden zwischen den PSNR-Werten die auf Grundlage der gesamten Bildfläche in den Tabellen 7(a) und 8(a) und auf Grundlage des Ausschnitts von $\frac{3}{4}$ der Bildfläche in den Tabellen 7(b) und 8(b) ersehen. Beim MPEG-1-Encoder ergeben sich bei Berücksichtigung aller Bildpunkte signifikant geringere PSNR-Werte, wohingegen beim WMV-9-Codec die Qualität nur marginal durch die Wahl des Bildausschnitts beeinflusst wird. Das zeigt, dass die Beschränkung auf $\frac{3}{4}$ der Bildfläche zur Bewertung der Qualität der MPEG-1-kodierten Videos sinnvoll und notwendig ist um Verfälschungen des Ergebnisses infolge des Logos zu vermeiden, und darüber hinaus, dass der gewählte Bildausschnitt bzgl. der PSNR-Bewertung repräsentativ für das Gesamtbild ist.

Für den WMV-9-Codec (Tabelle 5 und 7) sind die mittleren Messwerte nach den beiden Kodieroptionen *Low Quality (LQ)* und *High Quality (HQ)* gegenübergestellt. Es fällt auf, dass die Encoding-Zeiten für HQ bei gleicher räumlicher Auflösung des Videos und gleicher Bitrate immer signifikant höher sind als bei LQ-Kodierung. Dagegen liegen die mittleren PSNR-Werte für HQ und LQ sehr nah beieinander. Die Option HQ erzielt meist nur einen geringfügig höheren PSNR-Wert als LQ. Teilweise kehrt sich dieser Effekt sogar um,

Bitrate	9.6 kbit/s		28.8 kbit/s		56 kbit/s		200 kbit/s		350 kbit/s		700 kbit/s	
	LQ	HQ	LQ	HQ	LQ	HQ	LQ	HQ	LQ	HQ	LQ	HQ
96×72	21.0	20.5	49.5	48.0	86.8	82.0	231.8	215.5	350.4	337.0	591.2	564.0
128×96	20.9	20.8	50.6	49.4	89.8	87.1	253.0	234.8	388.9	368.7	661.5	637.1
160×120	20.9	20.8	51.1	50.7	91.9	91.0	270.5	252.0	418.0	389.1	717.6	683.5
240×180	20.7	18.5	50.9	50.6	94.2	93.5	293.9	281.5	466.6	438.0	804.4	757.8
320×240	18.0	16.6	50.2	50.7	94.6	93.8	308.1	302.5	497.2	474.6	858.6	809.2
640×480	10.2	9.0	43.9	39.9	90.7	90.0	325.0	322.4	548.3	539.2	1009.0	985.9
704×528	8.8	7.6	40.0	36.4	90.9	87.0	321.8	320.4	542.0	538.1	1019.0	991.6

Tabelle 4: Mittlere tatsächliche Bitraten (in kbit/s) über alle Videos bei Kodierung mittels des WMV-9-Codexs mit den Kodieroptionen Low Quality (LQ) und High Quality (HQ)

Bitrate	9.6 kbit/s		28.8 kbit/s		56 kbit/s		200 kbit/s		350 kbit/s		700 kbit/s	
	LQ	HQ	LQ	HQ	LQ	HQ	LQ	HQ	LQ	HQ	LQ	HQ
96×72	0.64	2.4	0.53	2.63	0.53	2.64	0.58	2.65	0.63	2.78	0.74	2.89
128×96	0.98	3.66	0.86	3.93	0.83	4.07	0.86	4.09	0.88	4.17	0.98	4.43
160×120	1.41	5.68	1.14	6.27	1.08	6.45	1.0	6.41	1.0	6.53	1.03	6.8
240×180	3.71	11.6	2.85	13.4	2.43	13.8	1.97	14.1	1.93	14.1	1.98	14.3
320×240	7.66	18.8	5.44	21.4	4.51	22.9	3.42	23.5	3.2	23.5	3.22	23.5
640×480	45.0	71.2	28.8	78.8	21.5	81.4	15.7	90.0	13.5	89.8	11.7	90.6
704×528	43.8	65.0	36.0	70.5	25.2	75.7	20.5	83.2	16.9	84.0	15.3	85.4

Tabelle 5: Mittlere Encoding-Zeiten (in s) über alle Videos bei Kodierung mittels des WMV-9-Codexs mit den Kodieroptionen Low Quality (LQ) und High Quality (HQ)

Bitrate in kbit/s	750	950	1150	1350	1500	1700
Auflösung 160×120	0.9	0.88	0.97	0.98	0.9	0.95
Auflösung 192×144	1.03	1.06	1.07	1.08	1.13	1.04
Auflösung 240×180	1.5	1.65	1.75	1.68	1.73	1.79
Auflösung 320×240	2.53	2.57	2.67	2.73	2.69	2.84
Auflösung 480×360	5.18	5.48	5.59	5.61	5.66	5.63
Auflösung 640×480	9.02	9.1	9.05	9.17	9.18	9.2
Auflösung 704×528	10.8	10.8	10.8	10.8	10.8	10.8

Tabelle 6: Mittlere Encoding-Zeiten (in s) über alle Videos bei Kodierung mittels des MPEG-1-Codexs

Bitrate	9.6 kbit/s		28.8 kbit/s		56 kbit/s		200 kbit/s		350 kbit/s		700 kbit/s	
	LQ	HQ	LQ	HQ	LQ	HQ	LQ	HQ	LQ	HQ	LQ	HQ
96 × 72	18.3	18.7	22.1	22.8	25.4	25.1	30.6	28.9	33.2	31.7	36.8	36.0
128 × 96	17.7	17.7	20.8	21.6	23.9	24.6	29.4	28.3	31.6	30.3	35.0	33.6
160 × 120	17.1	17.1	19.7	20.5	22.5	23.6	28.4	27.9	30.5	29.4	33.1	32.1
240 × 180	16.1	16.7	18.4	18.5	20.5	21.3	26.9	27.1	29.0	28.8	31.4	30.7
320 × 240	14.2	12.5	17.5	17.4	19.4	20.4	25.8	26.7	29.1	28.8	31.4	31.0
640 × 480	9.37	8.77	14.7	13.1	17.6	16.6	21.5	22.4	24.9	26.2	30.0	31.4
704 × 528	8.79	8.05	13.9	13.2	16.4	16.2	20.7	21.5	23.0	24.4	27.3	28.4

(a) Betrachtung der gesamten Bildfläche

Bitrate	9.6 kbit/s		28.8 kbit/s		56 kbit/s		200 kbit/s		350 kbit/s		700 kbit/s	
	LQ	HQ	LQ	HQ	LQ	HQ	LQ	HQ	LQ	HQ	LQ	HQ
96 × 72	18.1	18.6	22.0	22.7	25.3	25.0	30.6	28.8	33.2	31.7	36.8	36.0
128 × 96	17.6	17.6	20.7	21.5	23.8	24.5	29.4	28.2	31.6	30.3	34.8	33.6
160 × 120	17.0	17.0	19.6	20.3	22.4	23.5	28.4	27.8	30.4	29.3	33.1	32.1
240 × 180	16.1	16.6	18.3	18.4	20.4	21.2	27.1	27.3	29.3	29.1	31.8	31.0
320 × 240	14.2	12.6	17.4	17.3	19.2	19.8	25.7	26.7	28.7	29.0	31.7	31.3
640 × 480	9.34	8.81	14.7	13.1	16.8	16.5	21.4	22.4	25.0	26.4	30.2	31.6
704 × 528	8.85	8.18	13.4	13.1	16.3	16.2	20.6	21.3	22.9	24.4	27.4	28.4

(b) Betrachtung eines Bildausschnitts von $\frac{3}{4}$ der gesamten Bildfläche

Tabelle 7: Mittlere PSNR-Werte (in dB) über alle Videos bei Kodierung mittels des WMV-9-Codecs mit den Kodieroptionen Low Quality (LQ) und High Quality (HQ)

Bitrate in kbit/s	750	950	1150	1350	1500	1700
Auflösung 160 × 120	21.6	21.7	21.8	21.9	21.9	21.9
Auflösung 192 × 144	22.8	23.0	23.2	23.3	23.3	23.4
Auflösung 240 × 180	24.2	24.5	24.7	24.8	24.9	25.0
Auflösung 320 × 240	25.6	26.1	26.4	26.6	26.8	27.0
Auflösung 480 × 360	25.8	27.2	28.0	28.5	28.7	28.9
Auflösung 640 × 480	24.6	25.7	27.0	27.8	28.9	29.6
Auflösung 704 × 528	24.5	25.2	26.2	27.2	27.9	29.1

(a) Betrachtung der gesamten Bildfläche

Bitrate in kbit/s	750	950	1150	1350	1500	1700
Auflösung 160 × 120	32.8	34.0	34.9	35.7	36.2	36.8
Auflösung 192 × 144	32.0	33.2	34.2	34.9	35.4	36.0
Auflösung 240 × 180	31.0	32.3	33.3	34.1	34.6	34.9
Auflösung 320 × 240	29.9	31.2	32.2	33.0	33.6	34.3
Auflösung 480 × 360	27.8	29.8	31.1	32.0	32.6	33.1
Auflösung 640 × 480	25.5	27.0	28.6	29.7	31.0	32.2
Auflösung 704 × 528	25.2	26.1	27.5	28.7	29.5	31.1

(b) Betrachtung eines Bildausschnitts von $\frac{3}{4}$ der gesamten Bildfläche

Tabelle 8: Mittlere PSNR-Werte (in dB) über alle Videos bei Kodierung mittels des MPEG-1-Codecs

so dass bei LQ ein höherer PSNR-Wert als bei HQ zu beobachten ist. Aus diesem Grund werden im Folgenden nur noch die Ergebnisse für die Kodieroption LQ betrachtet. Im praktischen Einsatz ist kaum zu erwarten, dass ein Kodierverfahren Verwendung finden wird, welches bei vergleichbarer Qualität deutlich mehr Zeit und Ressourcen zum Kodieren des Videos benötigt als ein vergleichbares Verfahren.

4.2 Modellbildung

Aus den empirisch ermittelten Gesetzmäßigkeiten sollen mathematische Modelle zu Beschreibung der Zusammenhänge zwischen den Zielparametern der Adaption (Auflösung und Bitrate) und den Messwerten (Encoding-Zeit und PSNR) gebildet werden, die eine Vorhersage der Kosten zur Kodierung von Videoobjekten und der Qualität der kodierten Videoobjekte erlauben. Im folgenden Abschnitt wird die Vorgehensweise bei der Modellbildung vorgestellt. In Abschnitt 4.2.2 wird dieses Vorgehen an einem Beispiel demonstriert. Die Ergebnisse der Modellbildung werden in Abschnitt 4.2.3 zusammengefasst und bewertet. Für die Messwertanalyse und Modellbildung wird *MATHEMATICA* [13] von der Firma Wolfram Research eingesetzt.

4.2.1 Eine allgemeine Vorgehensweise zur Modellbildung

Ziel der Modellbildung ist es, die Abhängigkeit der Messgrößen Encoding-Zeit t_{enc} und PSNR-Wert $PSNR$ von den Zielparametern der Adaption Auflösung $aufl$ und Bitrate $bitr$ als Funktionen

$$t_{enc} = f_{\mathbf{p}}(aufl, bitr) \quad (5)$$

$$PSNR = g_{\mathbf{p}}(aufl, bitr) \quad (6)$$

zu beschreiben. Da sowohl die Encoding-Zeiten als auch die PSNR-Werte der kodierten Videos datenabhängig sind, sind die Funktionen $f_{\mathbf{p}}$ und $g_{\mathbf{p}}$ jeweils spezifisch für ein konkretes Videoobjekt. Dies wird in den Formeln 5 und 6 durch den Parametervektor $\mathbf{p} = (p_1, p_2, \dots, p_m)$ ausgedrückt, der die invarianten Eigenschaften des konkreten Videos ausdrückt. Die Funktionen $f_{\mathbf{p}}$ und $g_{\mathbf{p}}$ sind so zu bestimmen, dass der mittlere quadratische Fehler minimal wird.

Die Abbildungen $f_{\mathbf{p}}$ und $g_{\mathbf{p}}$ sind Funktionen von zwei Veränderlichen ($aufl, bitr$). Somit beschreiben beide eine Fläche im dreidimensionalen Raum (vgl. Abbildung 8(a)). Die Wahl geeigneter Prototypen für die Funktionen $f_{\mathbf{p}}$ und $g_{\mathbf{p}}$ zur Beschreibung der Fläche im dreidimensionalen Raum ist nicht trivial. Aufgrund dessen wird ein mehrstufiges Verfahren verwendet (Abbildung 8(b) zeigt

die ersten beiden Stufen der Modellbildung), um geeignete Prototypen zu finden.

In der **1. Stufe** wird nur die Abhängigkeit der Messgrößen von einer der beiden Veränderlichen, zum Beispiel von der Auflösung $aufl$ untersucht:

$$t_{enc} = f_{\mathbf{p}'}(aufl) \quad (7)$$

$$PSNR = g_{\mathbf{p}'}(aufl) \quad (8)$$

Der jeweils andere Parameter (hier: $bitr$) wird entsprechend konstant gehalten. Die Parameter $\mathbf{p}' = (p'_1, p'_2, \dots, p'_n)$ der Funktionen $f_{\mathbf{p}'}$ und $g_{\mathbf{p}'}$ drücken die invarianten Eigenschaften der Videoobjekte bei konstanter Bitrate aus. Da $f_{\mathbf{p}'}$ und $g_{\mathbf{p}'}$ Funktionen nur einer Veränderlichen sind, können geeignete Prototypen aus der graphischen Darstellung der Funktionen erraten werden. Die Parameter \mathbf{p}' lassen sich dann leicht mittels einfacher Regression bestimmen.

In der **2. Stufe** wird dann die Abhängigkeit der Parameter $\mathbf{p}' = (p'_1, p'_1, \dots, p'_n)$ von der anderen Veränderlichen (hier: von der Bitrate $bitr$) untersucht:

$$p'_i = h_{\mathbf{p}'}^{(i)}(bitr) \quad (9)$$

Auch die Funktionen $h_{\mathbf{p}'}^{(i)}$ sind Funktionen nur einer Veränderlichen, deren Parameter \mathbf{p} objekt-spezifische Invarianten sind. Diese lassen sich ebenfalls leicht durch einfache Regression bestimmen.

In der **3. Stufe** werden schließlich die Parameter $\mathbf{p}' = (p'_1, p'_2, \dots, p'_n)$ in den Formeln 7 und 8 durch die Zusammenhänge $p'_i = h_{\mathbf{p}'}^{(i)}(bitr)$ entsprechend Formel 9 ersetzt. Damit erhält man jeweils ein Gesamtmodell für t_{enc} und $PSNR$ gemäß den Formeln 5 und 6, welches den dreidimensionalen Zusammenhang entsprechend Abbildung 8(a) beschreibt. Die Parameter \mathbf{p} werden mittels multipler Regression nochmals angepasst, um den quadratischen Fehler im dreidimensionalen Modell zu minimieren.

Da das beste Modell in Stufe 1 nicht unbedingt die beste Modellierung der Parameter in Stufe 2 erlaubt, werden auf jeder Stufe mehrere geeignete Modelle bestimmt und zu mehreren Gesamtmodellen kombiniert. Die endgültige Auswahl eines Gesamtmodells erfolgt im Anschluss an Stufe 3 auf Grundlage der Residuen zwischen Messwerten und Vorhersagewerten im dreidimensionalen Modell. Die Residuen beschreiben den Vorhersagefehler des jeweiligen Modells in Form der absoluten Abweichung des Modells von den Messwerten. Neben den absoluten Residuen sollen auch die relativen Residuen für die Bewertung mit herangezogen werden. Ein relatives Residuum ergibt sich aus

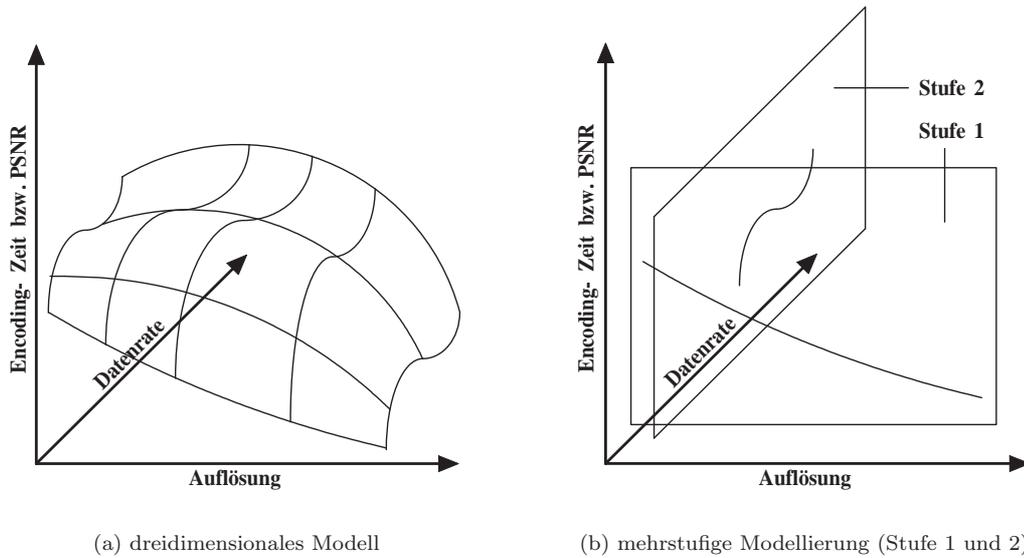


Abbildung 8: Modellierung

dem Verhältnis zwischen dem absoluten Residuum und dem Messwert:

$$\text{relatives Residuum} = \frac{\text{absolutes Residuum}}{\text{Messwert}} \quad (10)$$

Mit Hilfe der relativen Residuen wird berücksichtigt, dass sich gleichgroße Vorhersagefehler bei kleinen Werten gravierender auswirken als bei großen Werten. So entspricht eine Abweichung von 1 s bei einem Messwert von 2 s einem relativen Residuum von 50%, während die selbe Abweichung bei einem Messwert von 60 s einem relativen Residuum von lediglich 1.67% entspricht. Als Gütemetriken für die Auswahl eines geeigneten Gesamtmodells werden die Quantile der Ordnung 0.9 der relativen Residuen herangezogen.

4.2.2 Modellbildung am Beispiel

In diesem Abschnitt wird die dreistufige Modellbildung anhand eines konkreten Beispiels gezeigt. Als Beispiel wird das Vorhersagemodell für die Encoding-Zeit t_{enc} , welche mittels des WMV-9-Codex bei Low Quality ermittelt wurde, gewählt. Für die durchzuführenden Regressionsanalysen werden die Messwerte der Encoding-Zeit für alle 20 Videoobjekte zugrunde gelegt. In jeder Stufe werden für alle 20 Videos die gleichen Prototypen bestimmt. Die Datenabhängigkeit der Zusammenhänge findet ausschließlich in den Parametern \mathbf{p}' bzw. \mathbf{p} Ausdruck.

In der **1. Stufe** wird zunächst die Bitrate $bitr$ als konstant betrachtet und nur die Abhängigkeit von der veränderlichen Auflösung $aufl$ erfasst. Die Abbildung 9 zeigt beispielhaft die Kurvenverläufe,

die sich unter Verwendung von Video 10 (vgl. Tabelle 1) bei jeweils konstanter Bitrate über die betrachteten Auflösungen ergeben haben. Für alle Videoobjekte lassen sich ähnliche empirische Kurvenverläufe für die Encoding-Zeiten beobachten. Diese können durch die Regressionsfunktion

$$t_{enc} = p'_1 \cdot aufl^{\left(\frac{1}{p'_2}\right)} - p'_3 \quad (11)$$

näherungsweise beschrieben werden. Abbildung 10 zeigt die angenäherte Modellkurve für Video 10 bei einer konstanten Bitrate von 350 kbit/s. In der 1. Stufe werden für jedes Videoobjekt sechs verschiedene Parametersätze (p'_1, p'_2, p'_3) jeweils einer für jede der sechs verschiedenen Bitraten bestimmt.

In der **2. Stufe** werden die Parameter p'_1 , p'_2 und p'_3 jeweils den sechs erzielten Bitraten $bitr$ für jedes Videoobjekt gegenübergestellt (Abbildung 11). An dieser Stelle ist es wichtig, die tatsächlich erzielten Bitraten zu verwenden, da diese von den angestrebten Zielbitraten (9.6 kbit/s, 28.8 kbit/s, 56 kbit/s, 200 kbit/s, 350 kbit/s und 700 kbit/s) teilweise erheblich abweichen (vgl. Abschnitt 4.1). Die Zusammenhänge zwischen der tatsächlich erzielten Bitrate $bitr$ und den Parametern (p'_1, p'_2, p'_3) lassen sich näherungsweise durch die folgenden Funktionen beschreiben:

$$p'_1 = p_1 \left(\frac{1}{bitr}\right)^{p_2} \quad (12)$$

$$p'_2 = 1.15 \left(\frac{1}{bitr}\right)^{-p_5} \quad (13)$$

$$p'_3 = p_3 e^{\left(\frac{p_4}{bitr}\right)} \quad (14)$$

Zur Bildung des Gesamtmodells werden in der **3. Stufe** die in der 2. Stufe ermittelten Modelle

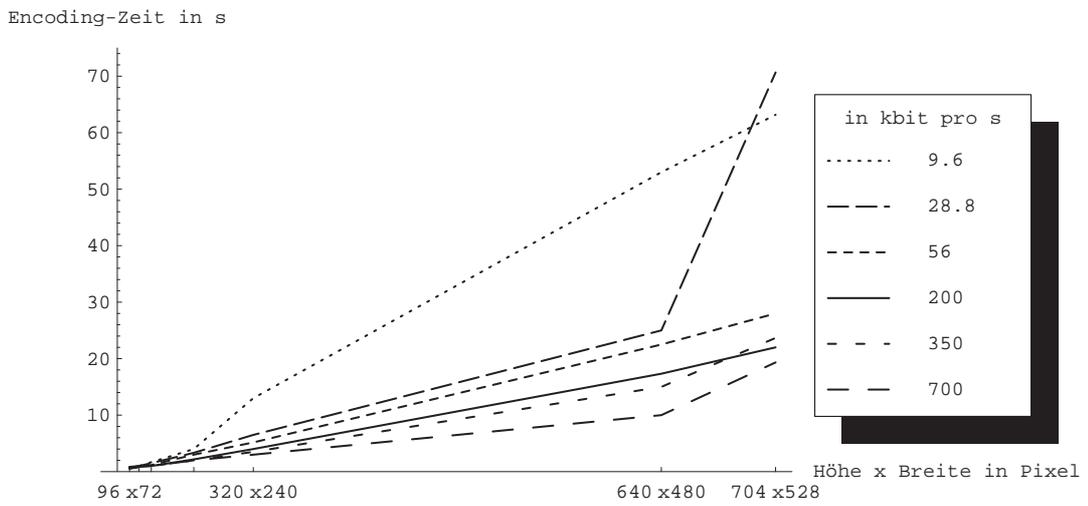


Abbildung 9: Messwerte der Encoding-Zeit für Video 10 bei der WMV-9-Kodierung mit der Kodieroption Low Quality

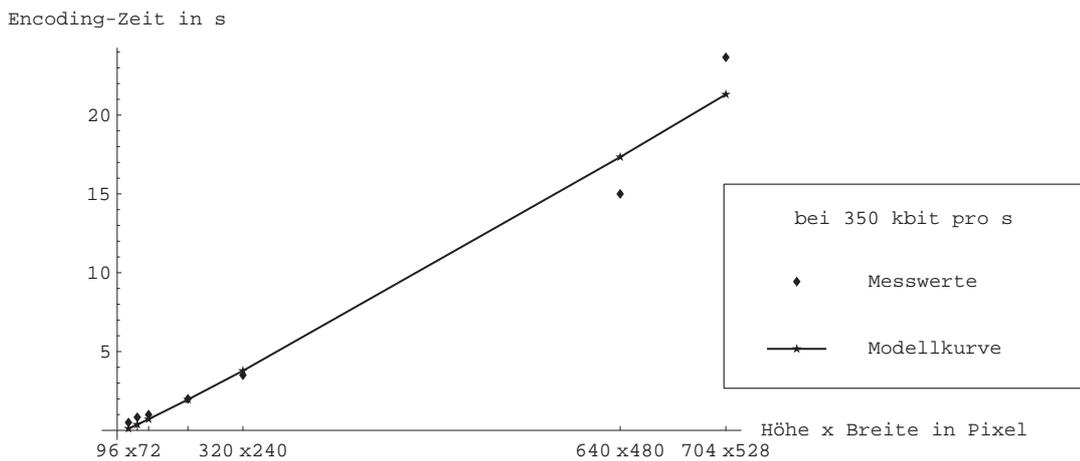


Abbildung 10: Modellkurve der Encoding-Zeit für Video 10 bei der WMV-9-Kodierung mit einer Zielbitrate von 350 kbit/s und der Kodieroption Low Quality

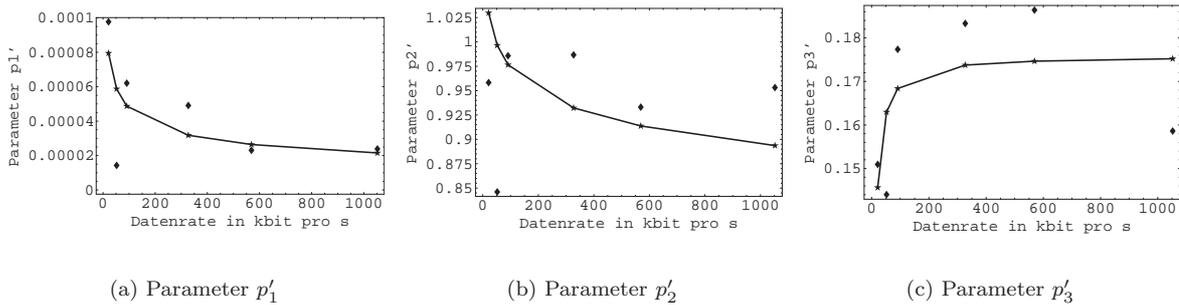


Abbildung 11: Abhängigkeit der Parameter p'_1 , p'_2 und p'_3 von der tatsächlich erzielten Bitrate $bitr$ für Video 10 bei WMV-9-Kodierung mit der Kodieroption Low Quality

der Parameter p'_1 , p'_2 und p'_3 in Formel 11 (aus Stufe 1) eingesetzt:

$$t_{enc} = p_1 \left(\frac{1}{bitr}\right)^{p_2} \cdot aufl^{1.15} \left(\frac{1}{bitr}\right)^{p_5} - p_3 e^{\left(\frac{p_4}{bitr}\right)} \quad (15)$$

Die Parameter $\mathbf{p} = (p_1, p_2, \dots, p_m)$, die sich durch multiple Regression in der 3. Stufe ergeben, beschreiben die objektspezifischen, invarianten Eigenschaften der Videos. Die ermittelten Parameterwerte zum Modell entsprechend Formel 15 sind in Tabelle 14 für den WMV-Codec in Anhang A zusammengefasst.

Eine Interpretation, welche Parameter welche invarianten Objekteigenschaften repräsentieren, geht über den Fokus des vorliegenden Berichtes hinaus.

Die Quantile der absoluten und relativen Residuen zur Bewertung der Güte des Modells gemäß Formel 15 sind in Tabelle 9 gezeigt. Eine differenzierte Diskussion der Modellgüte folgt im Abschnitt 4.2.3.

Quantile der Ordnungen	0.75	0.90	0.95	0.99
relative Residuen	25.5%	44.6%	64.9%	90.7%
absolute Residuen	0.664 s	1.90 s	3.15 s	4.11 s

Tabelle 9: Residuen der Encoding-Zeiten (Modell nach Formel 15) bei LQ-Kodierung in WMV

4.2.3 Ergebnisse der Modellbildung und Bewertung der Modelle

Auf Grundlage der Vorgehensweise, die in Abschnitt 4.2.1 eingeführt und in Abschnitt 4.2.2 beispielhaft demonstriert wurde, werden Modelle der Encoding-Zeiten und PSNR-Werte der kodierten

Videoobjekte für alle 20 Beispieldemos bestimmt. Diese Modelle werden anhand der absoluten und relativen Residuen bewertet und diskutiert.

Encoding-Zeit

Für den Zusammenhang der Encoding-Zeit t_{enc} zur Kodierung mittels des WMV-9-Codescs bei der Kodieroption Low Quality (LQ)¹ von der Auflösung $aufl$ und der Bitrate $bitr$ des kodierten Videoobjektes wurde in Abschnitt 4.2.2 das Modell gemäß Formel 15 ermittelt. Die objektspezifischen, invarianten Parameter für die 20 untersuchten Beispieldemos sind in Anhang A.1 zusammengefasst.

Auch die Encoding-Zeiten bei Verwendung des MPEG-1-Encoders von MainConcept lassen sich durch den Zusammenhang in Formel 15 modellieren. Jedoch unterscheiden sich die objektspezifischen Parameter (s. Anhang A.2 für Parameter des Encoding-Zeit-Modells für MPEG-1-Kodierung) bei den verschiedenen Codescs deutlich voneinander, so dass die Encoding-Zeiten bei gleicher Auflösung und Bitrate unterschiedlich sind.

Die Vorhersagegüte der Modelle ist in Tabelle 10 und 11 anhand der Quantile der Ordnungen 0.75, 0.90, 0.95 und 0.99 der absoluten und relativen Residuen bewertet. Neben der Bewertung der Modelle für den gesamten betrachteten Definitionsbereich wird die Auswertung zusätzlich nach verschiedenen Auflösungsstufen vorgenommen.

Tabelle 10 zeigt, dass bei der WMV-Kodierung einige Vorhersagewerte bis zu einigen Sekunden von den gemessenen Encoding-Zeiten abweichen. Bei ca. 1% weicht die vorhergesagte Encoding-Zeit über 4s ab (vgl. Quantil der Ordnung 0.99). In 90% der Fälle liegt der Fehler der Vorhersage jedoch unter 2s und in 75% unabhängig von der Auflösung der Videos unter 1s (vgl. Quantile der Ordnung 0.90 und 0.75).

¹Die Kodieroption High Quality (HQ) wird bei der Modellierung nicht betrachtet (vgl. Abschnitt 4.1).

Quantile der Ordnung	rel. Residuen				abs. Residuen			
	0.75	0.90	0.95	0.99	0.75	0.90	0.95	0.99
gesamter Definitionsbereich	25.5%	44.6%	64.9%	90.7%	0.664 s	1.90 s	3.15 s	4.11 s
Aufgeschlüsselt nach Auflösung								
96 × 72, 128 × 96, 160 × 120	21.2%	41.3%	60.1%	83.7%	0.439 s	1.26 s	2.46 s	3.79 s
240 × 180, 320 × 240	26.5%	44.2%	53.7%	98.6%	0.799 s	2.14 s	3.22 s	4.54 s
640 × 480, 704 × 528	33.2%	55.5%	70.2%	146.0%	0.955 s	2.18 s	3.24 s	4.11 s

Tabelle 10: Abweichungen der Encoding-Zeiten bei WMV-Kodierung, Modell 15

Quantile der Ordnung	rel. Residuen				abs. Residuen			
	0.75	0.90	0.95	0.99	0.75	0.90	0.95	0.99
gesamter Definitionsbereich	11.5%	21.1%	28.0%	40.4%	0.251 s	0.360 s	0.428 s	0.585 s
Aufgeschlüsselt nach Auflösung								
160 × 120, 192 × 144	15.7%	24.9%	30.8%	51.2%	0.286 s	0.399 s	0.462 s	0.633 s
240 × 180, 320 × 240, 480 × 360	11.7%	19.7%	27.0%	44.1%	0.230 s	0.340 s	0.427 s	0.540 s
640 × 480, 704 × 528	7.36%	14.9%	23.6%	35.8%	0.245 s	0.343 s	0.401 s	0.568 s

Tabelle 11: Abweichungen der Encoding-Zeiten bei MPEG-Kodierung, Modell 15

Die relativen Residuen zeigen diese Abweichungen noch drastischer. So muss bei 10% der vorhergesagten Werte eine Abweichung zur tatsächlichen Encoding-Zeit von über 40% in Kauf genommen werden, für die hohen Auflösungen 640x480 und 704x528 liegt dieser Wert sogar bei 55,5%. Damit wird deutlich, dass lange nicht für alle, noch nicht einmal für 90% der untersuchten Adaptionsoperationen und deren Kombinationen, eine sichere Vorhersage der Encoding-Zeit getroffen werden kann.

Die Schwierigkeiten einer genauen Vorhersage werden bereits in den Abbildungen 10 und 11 deutlich. Diese zeigen, dass die Kurvenverläufe teilweise nur sehr schlecht an alle Messwerte angenähert werden können. Besonders groß fällt die Abweichung der Vorhersagewerte von den Messwerten bei hohen Auflösungen aus, da bei diesen die Suche nach geeigneten Quantisierern und Bewegungsvektoren gegebenenfalls sehr aufwendig werden kann.

Bei der Vorhersage der MPEG-1-basierten Encoding-Zeiten (Tabelle 11) fallen sowohl die absoluten als auch die relativen Residuen viel geringer aus, als dies bei der WMV-9-Kodierung der Fall ist. Besonders anhand der absoluten Residu-

en zeigt sich, wie gut die Encoding-Zeiten durch das Modell vorhergesagt werden können. 99% der Messwerte werden mit einem Fehler von weniger als 0.6s vorhergesagt. Für 90% der betrachteten Fälle liegt die Abweichung des Vorhersagewertes unabhängig von der Auflösung der Videos deutlich unter 0.5s. Das entspricht für mittlere und hohe Auflösungen relativen Residuen von weniger als 20%. Nur für kleine Auflösungen liegt das Quantil der Ordnung 0.90 der relativen Residuen über 20%. Für alle Quantile der relativen Residuen gilt, dass sie bei größeren Auflösungen geringer sind, als bei niedrigen Auflösungen. Das bedeutet, dass das Modell für Videobjekte in größeren Skalierungen eine genauere Vorhersage der Encoding-Zeiten erlaubt.

Wie in Abschnitt 4.2.1 erläutert, wurden für die Abhängigkeit der Encoding-Zeiten von der Auflösung *aupl* und der Bitrate *bitr* neben dem Modell nach Formel 15 weitere Modelle abgeleitet und bewertet. Die Modelle und deren Bewertung für die WMV-9- und MPEG-1-Kodierung sind in Anhang B.1 zusammengefasst. Die Tabellen 18 bis 25 zeigen, dass kein anderes Modell eine genauere Vorhersage der Encoding-Zeiten erlaubt, als das Modell in Formel 15.

Quantile	rel. Residuen				abs. Residuen			
der Ordnung	0.75	0.90	0.95	0.99	0.75	0.90	0.95	0.99
gesamter Wertebereich	9.19%	15.0%	20.5%	37.0%	2.18 dB	3.24 dB	4.19 dB	8.15 dB
Aufgeschlüsselt nach PSNR-Wert								
$30 \leq PSNR < 40$	8.03%	11.4%	17.0%	20.5%	2.54 dB	4.04 dB	6.06 dB	8.14 dB
$20 \leq PSNR < 30$	7.09%	11.9%	14.8%	37.4%	1.85 dB	3.00 dB	3.66 dB	7.81 dB
$10 \leq PSNR < 20$	11.9%	20.3%	24.7%	51.9%	1.97 dB	2.98 dB	3.77 dB	10.3 dB

Tabelle 12: Abweichungen PSNR-Vorhersage bei WMV-Kodierung, Modell 16

Quantile	rel. Residuen				abs. Residuen			
der Ordnung	0.75	0.90	0.95	0.99	0.75	0.90	0.95	0.99
gesamter Wertebereich	2.98%	5.39%	7.07%	15.0%	0.90 dB	1.52 dB	2.10 dB	3.46 dB
Aufgeschlüsselt nach PSNR-Wert								
$40 \leq PSNR$	2.67%	3.60%	4.03%	4.36%	1.11 dB	1.48 dB	1.65 dB	1.80 dB
$30 \leq PSNR < 40$	2.18%	4.02%	5.80%	8.37%	0.78 dB	1.41 dB	2.06 dB	3.15 dB
$20 \leq PSNR < 30$	4.13%	6.31%	9.04%	16.8%	1.06 dB	1.55 dB	2.27 dB	4.32 dB
$10 \leq PSNR < 20$	13.2%	17.8%	26.6%	26.6%	2.56 dB	3.46 dB	5.22 dB	5.22 dB

Tabelle 13: Abweichungen PSNR-Vorhersage bei MPEG-Kodierung, Modell 16

PSNR

Die PSNR-Werte lassen sich sowohl für WMV-9-kodierte als auch für MPEG-1-kodierte Videoströme am besten durch das folgende Modell beschreiben:

$$y = p_1 \text{ aufl} - p_2 + \text{bitr}^{p_3 + e^{-\frac{p_4}{\text{aufl}}}} \quad (16)$$

Die objektspezifischen Parameter (p_1, p_2, p_3, p_4) sind für die beiden untersuchten Codecs in den Anhängen A.3 und A.4 dargestellt.

Die Bewertung der PSNR-Modelle für WMV-9 und MPEG-1 ist in Tabelle 12 bzw. Tabelle 13 zusammengefasst. Um die Ergebnisse detailliert zu diskutieren, wird nicht nur die Vorhersagegüte für den gesamten Wertebereich an PSNR-Werten angegeben, sondern auch für verschiedene Intervalle des Wertebereichs. Die Unterteilung des Wertebereichs orientiert sich an der Zuordnung von PSNR-Werten zum subjektiven Qualitätseindruck (vgl. Tabelle 3). Für WMV-9 (Tabelle 12) ist kein Intervall $40 \leq PSNR$ angegeben, da kein Video mit einer so guten Qualität kodiert werden konnte.

Die Quantile der relativen Residuen in Tabelle 12 und Tabelle 13 zeigen, dass die Güte der

PSNR-Modelle deutlich größer ist, als die der Encoding-Zeit-Modelle. Der relative Fehler der Modelle ist bei Videos hoher Qualität am geringsten und steigt monoton mit abnehmender Qualität. Darüber hinaus zeigen die Quantile der absoluten Residuen, dass das WMV-9-Modell für mindestens 99% der Fälle einen PSNR-Wert vorausagt, der im gleichen oder einem benachbarten Intervall des Wertebereichs liegt wie der empirisch ermittelte – also in der MOS-Skala nach Tabelle 3 mit dem gleichen oder einem ähnlichem Prädikat bewertet werden würde. Für das MPEG-1-Modell ist dieser Anteil sogar noch deutlich größer.

Insgesamt kann für einen Großteil der durchgeführten Messungen der PSNR-Wert sehr gut vorausgesagt werden. Es gibt deutlich weniger und geringere Abweichungen als bei der Vorhersage der Encoding-Zeit. Auch für die Vorhersage der PSNR-Werte der kodierten Videos wurden neben dem Modell entsprechend Formel 16 weitere Modelle abgeleitet und bewertet, die eine weniger genaue Vorhersage der empirisch ermittelten Werte erlauben als Formel 16. Diese Modelle und deren Bewertung sind in Anhang B.2 zusammengefasst.

5 Zusammenfassung

Der vorliegende Bericht beschreibt am Beispiel der Kodierung von Videoobjekten mittels des WMV-9-Codecs und eines MPEG-1-Codecs die Entwicklung von Modellen der Encoding-Kosten und der Qualität kodierter Videoobjekte. Die Videoobjekte werden dabei mit verschiedenen räumlichen Auflösungen und Bitraten kodiert.

Solche Kosten- und Qualitätsmodelle werden benötigt, um in Systemen zur dynamischen Inhaltsadaption geeignete Adaptionoperationen auswählen und parametrisieren zu können. Darüber hinaus werden die vorgestellten Modelle der Adaptionoperationen auch benötigt, um Systeme zur adaptiven Auslieferung von Videos analytisch oder durch Simulation zu evaluieren.

Für die Entwicklung dieser Modelle wurde (1) eine Menge von Testvideos mit unterschiedlichen Auflösungen und Bitraten kodiert, (2) dabei die Encoding-Kosten in Form der Encoding-Zeit sowie die Qualität der kodierten Videos in Form des *Peak-Signal-to-Noise-Ratio* (PSNR) ermittelt und (3) Modelle zur Beschreibung des empirisch ermittelten Zusammenhangs zwischen den Eigenschaften der Videosequenzen, den Zielparametern der Adaptionoperationen, der Encoding-Zeit und der Qualität der kodierten Videos abgeleitet. Die Encoding-Zeit steht dabei als Maß für die Encoding-Kosten.

Sowohl für die Vorhersage der Encoding-Zeit als auch des PSNR-Wertes konnten für beide untersuchten Codecs – WMV-9 und MPEG-1 – geeignete Modelle bestimmt werden. Die Güte der Modelle wurde anhand der absoluten und relativen Residuen untersucht. Dabei hat sich gezeigt, dass der MPEG-1-Encoder von MainConcept eine präzisere Modellierung erlaubt als der WMV-9-Codec. Darüber hinaus ist die Qualität der kodierten Videoobjekte durch die gefundenen Modelle genauer vorherzusagen als die Encoding-Zeiten.

Die Datenabhängigkeit der Encoding-Zeiten und PSNR-Werte wird in den abgeleiteten Modellen durch objektspezifische Parameter erfasst. Die entsprechenden Parameter für die untersuchten Videos sind im Anhang A dieses Berichtes zusammengefasst. Sie wurden durch Regression bestimmt. Diese Parameter beschreiben invariante Eigenschaften der Videoobjekte. Eine Interpretation, welche Eigenschaften durch die einzelnen Parameter beschrieben werden, war jedoch nicht Gegenstand dieser Arbeit. In zukünftigen Arbeiten sollte deshalb untersucht werden, wie die Parameter aus offensichtlichen Eigenschaften der Videos, z.B. aus Farben und Detailgehalt der Szene, abzuleiten sind. Dies würde ermöglichen, die benötigten Parameter für eine große Zahl von

Videoobjekten als Metadaten zur Verfügung zu stellen.

Ein weiterer Ansatzpunkt für weiterführende Arbeiten ist die Beziehung der verschiedenen Codecs untereinander genauer zu untersuchen. In dieser Arbeit konnten sowohl für die Encoding-Zeit als auch für die PSNR-Werte der kodierten Videos für beide Codecs ähnliche Modelle abgeleitet werden. Während der Funktionszusammenhang bei den beiden betrachteten Codecs gleich ist unterscheiden sich die objektspezifischen Parameter. Deshalb ist die Annahme naheliegend, dass sich die gefundenen Parametersätze aus objektspezifischen und codecspezifischen Faktoren zusammensetzen.

Aufgrund des Aufwandes zur Bestimmung der objektspezifischen Parameter sind die gefundenen Modelle bisher nur eingeschränkt für den Einsatz zur Vorhersage der Adaptionkosten und der Qualität der adaptierten Videoobjekte in Systemen zur dynamischen Inhaltsadaption geeignet. Jedoch ermöglichen die Ergebnisse die Modellierung der Videokodierung in simulativen oder analytischen Untersuchungen von Systemen zur adaptiven Auslieferung von Videos. Die hier gezeigten Ergebnisse für den WMV-9-Codec wurden zu diesem Zweck bereits in [14] genutzt.

LITERATUR

- [1] BUCHHOLZ, Sven: *Adaptivitätssensitive Platzierung von Replikaten in Adaptiven Content Distribution Networks*. Fakultät Informatik, Technische Universität Dresden, Dissertation, 2005
- [2] Final Report from the Video Quality Experts Group on the Validation of Objective Models of Video Quality Assessment / Video Quality Experts Group. 2000. – Final Report
- [3] RRNR-TV Group Test Plan / Video Quality Experts Group. 2000. – Draft Version 1.4
- [4] Final Report from the Video Quality Experts Group on the Validation of Objective Models of Video Quality Assessment, Phase II / Video Quality Experts Group. 2003. – Final Report
- [5] WOLF, Stephen ; PINSON, Margaret H.: Video Quality Measurement Techniques. In: *NTIA Report Series* Bd. 02-392, 2002
- [6] SARNOFF CORPORATION: *JNDmetrix Technology*. <http://www.jndmetrix.com>,

- [7] DIPERT, Brian: Video characterization creates hands-on headaches - Part 1. In: *EDN* (2002), July 25, S. 53–58
- [8] DIPERT, Brian: Video characterization creates hands-on headaches - Part 2. In: *EDN* (2002), August 8, S. 81–88
- [9] SUN MICROSYSTEMS, INC.: *Java Media APIs*. <http://java.sun.com/products/java-media/>,
- [10] ITU-R BT.601 / International Telecommunication Union. 1995. – Recommendation. „Studio encoding parameters of digital television for standard 4:3 and wide-screen 16:9 aspect ratios“
- [11] ITU-R BT.500-11 / International Telecommunication Union. 2002. – Recommendation. „Methodology for the subjective assesement of the quality of television pictures“
- [12] OHM, Jens-Rainer: *Digitale Bildcodierung - Repräsentation, Kompression und Übertragung von Bildsignalen*. Springer-Verlag, 1995
- [13] WOLFRAM RESEARCH, Inc.: *Mathematica*. <http://www.wolfram.com/products/mathematica/index.html>,
- [14] UMBACH, Matthias: *Simulationsstudie zu Replica Placement in A-CDNs am Beispiel eines A-CDN zur Verteilung von Videoströmen in Pervasive Computing Umgebungen*. Fakultät Informatik, Technische Universität Dresden, Großer Beleg, 2005

A Parameterwerte

Hier folgen die Angaben der invarianten, objektspezifischen Parameter, welche für die in Abschnitt 4.2.3 detailliert bewerteten Modelle für jedes Videoobjekt ermittelt wurden und für eine Simulation anhand der Modelle verwendet werden können.

A.1 Parameter des Encoding-Zeit-Modells für WMV bei Low Quality

$$t_{enc} = p_1 \left(\frac{1}{bitr} \right)^{p_2} a_{ufl} \frac{1}{1.15} \left(\frac{1}{bitr} \right)^{p_5} - p_3 e^{p_4/bitr} \quad (17)$$

Parameter	p_1	p_2	p_3	p_4	p_5
Video 1	0.00612785	1.61264	0.154175	26.7487	-0.0782218
Video 2	0.00132296	0.927693	-0.372942	-42.2338	-0.0520167
Video 3	0.00359786	1.78607	-0.963686	-37.8249	-0.0900674
Video 4	0.0182383	2.65366	-1.02076	-45.1034	-0.117485
Video 5	0.00177174	1.21472	-0.707057	-75.7398	-0.0663707
Video 6	0.00353729	1.74584	-0.952533	-61.2082	-0.0890959
Video 7	0.0019898	1.36692	-0.76487	-57.8628	-0.0738138
Video 8	0.00576942	2.0355	-1.04831	-47.3362	-0.0987174
Video 9	0.00146818	1.19464	-0.640533	-54.5583	-0.0681624
Video 10	0.00167593	1.02309	-0.569991	-51.5306	-0.0554232
Video 11	0.00112465	0.313105	0.228005	36.753	-0.00727556
Video 12	0.00174868	0.958307	-0.969391	-120.288	-0.0459866
Video 13	0.00154191	0.232595	0.579238	3.68082	0.00406115
Video 14	0.00215906	1.01351	-0.855565	-105.971	-0.0477242
Video 15	0.00127407	0.267249	0.307271	31.7903	-0.00206502
Video 16	0.00341801	1.31165	-1.13101	-129.242	-0.0616416
Video 17	0.00118692	0.430772	0.196533	42.2192	-0.016829
Video 18	0.00259586	0.793448	0.0730391	57.1157	-0.0309402
Video 19	0.00206076	1.14843	-1.09698	-118.858	-0.0564066
Video 20	0.00138335	0.38642	0.150299	46.1974	-0.00876349

Tabelle 14: objektspezifische Parameterwerte, Modell 17

A.2 Parameter des Encoding-Zeit-Modells für MPEG

$$t_{enc} = p_1 \left(\frac{1}{bitr} \right)^{p_2} \text{aufl}^{\frac{1}{1.15}} \left(\frac{1}{bitr} \right)^{p_5} - p_3 e^{p_4/bitr} \quad (18)$$

Parameter	p_1	p_2	p_3	p_4	p_5
Video 1	0.000117845	0.125169	-35.9741	-9068.55	-0.0134799
Video 2	0.000105232	-0.158226	0.047208	0.978218	0.0120069
Video 3	0.000171178	0.208232	-1.86928	-2337.61	-0.0150218
Video 4	0.000269934	0.548173	-1.27975	-1620.4	-0.0367299
Video 5	0.000163593	0.225436	-1.70282	-1648.25	-0.0173983
Video 6	0.000106211	-0.157814	0.155391	-0.832717	0.0119096
Video 7	0.000107075	0.140726	-0.620265	-1333.82	-0.0155246
Video 8	0.000143822	-0.176455	0.141381	834.844	0.017846
Video 9	0.000137847	0.0832429	-0.217602	-136.369	-0.00777559
Video 10	0.000179675	-0.000047412	$-2.07 \cdot 10^{12}$	-50543.2	0.00377486
Video 11	0.000104581	-0.298369	0.586025	-153.465	0.0220147
Video 12	0.000176824	-0.141958	0.0999364	1103.12	0.014724
Video 13	0.000261434	0.0922935	-1.09082	-2294.09	-0.00168156
Video 14	0.000170346	0.0964163	-0.654781	-1605.72	-0.0074479
Video 15	0.000325107	-0.0884278	0.165941	957.167	0.0161229
Video 16	0.000243435	0.204302	-1.55506	-2340.	-0.013122
Video 17	0.0000868027	-0.215579	12.3209	-6971.97	0.0111914
Video 18	0.000206896	0.51697	-1.00239	-473.872	-0.0382631
Video 19	0.000229243	0.232818	-0.510163	-726.24	-0.0166738
Video 20	0.000170696	0.0304519	-0.243156	-361.363	-0.00206836

Tabelle 15: objektspezifische Parameterwerte, Modell 18

A.3 Parameter des PSNR-Wert-Modells für WMV bei Low Quality

$$PSNR = p_1 \text{ aufl} - p_2 + \text{bitr}^{p_3 + e^{-\frac{p_4}{\text{aufl}}}} \quad (19)$$

Parameter	p_1	p_2	p_3	p_4
Video 1	-0.0000604827	-32.2354	-0.370799	71848.3
Video 2	$-5.69 \cdot 10^{-6}$	-13.1781	-0.631852	-301.403
Video 3	-0.0000112716	-12.5552	-0.565554	-507.207
Video 4	-0.0000105253	-20.0191	-0.572945	-748.729
Video 5	-0.0000192226	-14.0142	-0.5729	-473.987
Video 6	-0.0000219033	-15.0949	-0.578091	-495.776
Video 7	-0.0000249152	-10.2409	-0.523233	-423.593
Video 8	-0.0000199408	-14.3087	-0.530456	-200.579
Video 9	-0.0000182289	-13.4699	-0.599404	-590.175
Video 10	$-7.38 \cdot 10^{-6}$	-9.53782	-0.597416	-505.86
Video 11	-0.0000180757	-11.6696	-0.58666	-528.182
Video 12	-0.0000177994	-17.0478	-0.536259	-242.085
Video 13	-0.0000102747	-8.58822	-0.605898	-474.68
Video 14	-0.0000212846	-12.9931	-0.51035	-364.707
Video 15	-0.0000114227	-9.34845	-0.620462	-809.439
Video 16	$-3.95 \cdot 10^{-6}$	-19.3605	-0.58751	-251.691
Video 17	-0.0000210823	-14.9382	-0.579046	-458.881
Video 18	$-9.25 \cdot 10^{-6}$	-18.0986	-0.577836	-799.036
Video 19	-0.0000216414	-21.7155	-0.533348	-138.053
Video 20	$-9.47 \cdot 10^{-6}$	-12.5158	-0.590206	-497.215

Tabelle 16: objektspezifische Parameterwerte, Modell 19

A.4 Parameter des PSNR-Wert-Modells für MPEG

$$PSNR = p_1 \text{ aufl} - p_2 + \text{bitr}^{p_3 + e^{-\frac{p_4}{\text{aufl}}}} \quad (20)$$

Parameter	p_1	p_2	p_3	p_4
Video 1	-0.0000247919	-24.6863	-0.676349	-1136.22
Video 2	$4.51 \cdot 10^{-6}$	-13.9507	-0.650108	-1043.78
Video 3	$-1.58 \cdot 10^{-6}$	-14.2264	-0.612675	-854.183
Video 4	-0.0000168947	-28.6585	-0.661983	-459.795
Video 5	-0.0000249751	-16.5281	-0.598925	-301.739
Video 6	-0.0000165235	-17.6685	-0.624777	-588.122
Video 7	-0.0000446643	-18.4661	-0.583567	11.3952
Video 8	-0.0000284649	-21.0313	-0.583261	416.067
Video 9	-0.0000242762	-14.6682	-0.596613	-254.915
Video 10	$8.92 \cdot 10^{-7}$	-9.08765	-0.604977	-509.452
Video 11	-0.0000217327	-17.5061	-0.630123	-247.592
Video 12	$-3.67 \cdot 10^{-6}$	-24.6198	-0.637491	-259.593
Video 13	$-5.30 \cdot 10^{-6}$	-9.38539	-0.605656	455.179
Video 14	$-7.51 \cdot 10^{-6}$	-23.4528	-0.618476	243.728
Video 15	-0.0000257739	-11.1662	-0.588869	-101.663
Video 16	$5.20 \cdot 10^{-6}$	-20.2535	-0.645931	-751.115
Video 17	-0.000032177	-17.6886	-0.592662	91.4092
Video 18	-0.0000167073	-22.4414	-0.633196	-622.365
Video 19	$-7.94 \cdot 10^{-6}$	-29.0566	-0.65708	309.839
Video 20	$1.24 \cdot 10^{-6}$	-12.8905	-0.618067	-705.738

Tabelle 17: objektspezifische Parameterwerte, Modell 20

B Weitere Modelle

B.1 Modelle zur Vorhersage der Encoding-Zeit

Quantile der Ordnung	rel. Residuen				abs. Residuen			
	0.75	0.90	0.95	0.99	0.75	0.90	0.95	0.99
WMV-9	34.8%	51.1%	62.1%	77.0%	0.61 s	2.08 s	3.10 s	4.03 s
MPEG-1	12.0%	22.5%	28.1%	47.5%	0.26 s	0.40 s	0.47 s	0.64 s

Tabelle 18: Güte des Modells $t_{enc} = p_1 e^{\frac{p_2}{bitr}} \left(aufl^{\frac{(\frac{1}{bitr})^{p_4}}{p_3}} - \frac{1}{p_5} \right)$

Quantile der Ordnung	rel. Residuen				abs. Residuen			
	0.75	0.90	0.95	0.99	0.75	0.90	0.95	0.99
WMV-9	36.3%	55.5%	66.9%	80.1%	0.63 s	1.84 s	2.89 s	4.21 s
MPEG-1	12.2%	22.9%	28.5%	47.7%	0.26 s	0.40 s	0.48 s	0.62 s

Tabelle 19: Güte des Modells $t_{enc} = (p_1 bitr^{p_3} + p_2) aufl^{(-p_4 + e^{p_5/bitr})}$

Quantile der Ordnung	rel. Residuen				abs. Residuen			
	0.75	0.90	0.95	0.99	0.75	0.90	0.95	0.99
WMV-9	36.6%	54.6%	66.0%	80.2%	0.63 s	1.84 s	2.88 s	4.21 s
MPEG-1	11.9%	22.9%	28.4%	46.0%	0.26 s	0.39 s	0.46 s	0.59 s

Tabelle 20: Güte des Modells $t_{enc} = (p_1 bitr^{p_3} + p_2) aufl^{p_4 + \frac{p_5}{bitr}}$

Quantile der Ordnung	rel. Residuen				abs. Residuen			
	0.75	0.90	0.95	0.99	0.75	0.90	0.95	0.99
WMV-9	-	-	-	-	-	-	-	-
MPEG-1	11.0%	21.1%	28.0%	44.3%	0.25 s	0.38 s	0.44 s	0.61 s

Tabelle 21: Güte des Modells $t_{enc} = p_1 e^{-\frac{p_2}{bitr}} aufl^{p_3} + \frac{1}{p_4} + \frac{bitr}{p_5}$

Quantile der Ordnung	rel. Residuen				abs. Residuen			
	0.75	0.90	0.95	0.99	0.75	0.90	0.95	0.99
WMV-9	32.0%	55.8%	77.7%	162%	0.67 s	2.14 s	3.18 s	4.26 s
MPEG-1	10.8%	21.0%	28.5%	45.1%	0.26 s	0.39 s	0.45 s	0.62 s

Tabelle 22: Güte des Modells $t_{enc} = (p_1 \text{ bitr}^{p_3} + p_2) \text{ aufl}^{p_4} + \frac{1}{p_5} + \frac{\text{bitr}}{p_6}$

Quantile der Ordnung	rel. Residuen				abs. Residuen			
	0.75	0.90	0.95	0.99	0.75	0.90	0.95	0.99
WMV-9	31.4%	48.6%	60.2%	109%	0.69 s	2.30 s	3.20 s	4.29 s
MPEG-1	14.2%	22.8%	33.8%	55.4%	0.34 s	0.47 s	0.57 s	0.88 s

Tabelle 23: Güte des Modells $t_{enc} = p_1 \left[p_2 - \left(\frac{1}{\text{bitr}} \right)^{-0.02} \right] \text{ aufl} + p_3 \left(1 - e^{-\frac{p_4}{\text{bitr}}} \right)$

Quantile der Ordnung	rel. Residuen				abs. Residuen			
	0.75	0.90	0.95	0.99	0.75	0.90	0.95	0.99
WMV-9	29.9%	49.2%	65.2%	94.6%	0.62 s	2.33 s	3.27 s	4.23 s
MPEG-1	12.6%	23.5%	32.4%	56.1%	0.30 s	0.42 s	0.52 s	0.83 s

Tabelle 24: Güte des Modells $t_{enc} = p_1 \left[p_2 - \left(\frac{1}{\text{bitr}} \right)^{-0.02} \right] \text{ aufl} + p_3 - e^{-\frac{p_4}{\text{bitr}}}$

Quantile der Ordnung	rel. Residuen				abs. Residuen			
	0.75	0.90	0.95	0.99	0.75	0.90	0.95	0.99
WMV-9	28.4%	49.4%	71.0%	160%	0.74 s	1.90 s	3.18 s	4.17 s
MPEG-1	-	-	-	-	-	-	-	-

Tabelle 25: Güte des Modells $t_{enc} = p_1 \left(\frac{1}{\text{bitr}} \right)^{p_2} \text{ aufl}^{\frac{1}{1.25} \left(\frac{1}{\text{bitr}} \right)^{p_3}} + p_4 - e^{-\frac{p_5}{\text{bitr}}}$

B.2 Modelle zur Vorhersage des PSNR-Wertes

Quantile der Ordnung	rel. Residuen				abs. Residuen			
	0.75	0.90	0.95	0.99	0.75	0.90	0.95	0.99
WMV-9	22.0%	31.1%	39.0%	51.1%	4.8 dB	6.4 dB	7.2 dB	9.1 dB
MPEG-1	3.69%	6.57%	10.1%	15.7%	1.1 dB	2.0 dB	2.5 dB	4.8 dB

Tabelle 26: Güte des Modells $PSNR = -1000 \cdot \left(p_1 - p_2 e^{\frac{1}{bitr}} + e^{\frac{1}{bitr} + \frac{p_3}{aufl}} - e^{\frac{p_4}{aufl}} \right)$

Quantile der Ordnung	rel. Residuen				abs. Residuen			
	0.75	0.90	0.95	0.99	0.75	0.90	0.95	0.99
WMV-9	19.1%	28.7%	35.0%	52.9%	4.2 dB	5.5 dB	6.6 dB	8.7 dB
MPEG-1	4.38%	8.52%	11.1%	21.3%	1.3 dB	2.4 dB	3.0 dB	5.7 dB

Tabelle 27: Güte des Modells $PSNR = 1000 \cdot \left(-p_1 + e^{\frac{p_2}{aufl}} \right) e^{-\frac{p_3 e^{\frac{p_4}{aufl}}}{bitr}}$

Quantile der Ordnung	rel. Residuen				abs. Residuen			
	0.75	0.90	0.95	0.99	0.75	0.90	0.95	0.99
WMV-9	19.1%	28.7%	35.0%	52.9%	4.2 dB	5.5 dB	6.6 dB	8.7 dB
MPEG-1	3.97%	7.09%	9.07%	14.0%	1.2 dB	2.1 dB	2.6 dB	3.6 dB

Tabelle 28: Güte des Modells $PSNR = \left(p_1 + \frac{p_2}{aufl} \right) e^{-\frac{p_3 e^{\frac{p_4}{aufl}}}{bitr}}$

Quantile der Ordnung	rel. Residuen				abs. Residuen			
	0.75	0.90	0.95	0.99	0.75	0.90	0.95	0.99
WMV-9	18.5%	27.0%	33.1%	52.3%	3.7 dB	5.4 dB	6.4 dB	8.9 dB
MPEG-1	3.63%	5.67%	7.65%	14.2%	1.1 dB	1.6 dB	2.3 dB	3.6 dB

Tabelle 29: Güte des Modells $PSNR = (p_1 aufl - p_2) e^{-\frac{p_3 + aufl p_4}{bitr}}$

Quantile der Ordnung	rel. Residuen				abs. Residuen			
	0.75	0.90	0.95	0.99	0.75	0.90	0.95	0.99
WMV-9	21.1%	31.3%	40.0%	55.6%	4.5 dB	6.2 dB	7.4 dB	9.1 dB
MPEG-1	6.87%	12.6%	17.3%	31.7%	2.2 dB	3.7 dB	4.5 dB	7.3 dB

Tabelle 30: Güte des Modells $PSNR = 1000 \cdot \left(-p_1 + e^{\frac{p_2}{aufl}}\right) e^{-\frac{p_3 \left(\frac{p_4}{aufl}\right)^{0.2}}{bitr}}$

Quantile der Ordnung	rel. Residuen				abs. Residuen			
	0.75	0.90	0.95	0.99	0.75	0.90	0.95	0.99
WMV-9	19.2%	29.6%	36.1%	51.0%	4.2 dB	5.5 dB	6.7 dB	8.3 dB
MPEG-1	6.0%	10.7%	14.4%	25.1%	1.9 dB	3.1 dB	3.8 dB	5.6 dB

Tabelle 31: Güte des Modells $PSNR = \frac{1}{aufl} (p_1 aufl + p_2) e^{-\frac{p_3 + e^{\frac{p_4}{aufl}}}{bitr}}$

Quantile der Ordnung	rel. Residuen				abs. Residuen			
	0.75	0.90	0.95	0.99	0.75	0.90	0.95	0.99
WMV-9	17.5%	24.7%	32.3%	48.4%	3.6 dB	5.0 dB	6.2 dB	8.2 dB
MPEG-1	3.57%	5.80%	7.72%	13.1%	1.1 dB	1.6 dB	2.1 dB	3.4 dB

Tabelle 32: Güte des Modells $PSNR = p_1 aufl^{p_2} e^{-\frac{p_3 + aufl^{p_4}}{bitr}}$

Quantile der Ordnung	rel. Residuen				abs. Residuen			
	0.75	0.90	0.95	0.99	0.75	0.90	0.95	0.99
WMV-9	18.1%	25.4%	32.7%	49.1%	3.6 dB	5.0 dB	6.2 dB	8.6 dB
MPEG-1	4.1%	7.3%	10.4%	17.8%	1.3 dB	2.1 dB	2.7 dB	4.9 dB

Tabelle 33: Güte des Modells $PSNR = p_1 aufl^{p_2} e^{-\frac{p_3 aufl + p_4}{aufl bitr}}$