

Universität Leipzig

Fakultät für Mathematik und Informatik

Institut für Informatik

Thema: Qualitätsgerechte Übertragung komprimierter Audio-/ Videoströme in
IP- Netzen im Vergleich verschiedener Kompressionsverfahren.

Diplomarbeit

Am Lehrstuhl für Rechnernetze und Verteilte Systeme

Verantwortlicher HSL: Prof. Dr. Klaus Irmischer

Betreuer: Prof. Dr. Hänßgen

Leipzig, Oktober 2002

vorgelegt von:

Sonne, Dirk

geb. am: 06.10.1972

Studiengang Informatik

Gliederung

1	EINLEITUNG	1-5
2	DIE NOTWENDIGKEIT DER VIDEOKOMPRESSION	2-9
3	DIE VIDEOKOMPRESSION.....	3-15
3.1	DIE VIDEOKOMPRESSIONSVERFAHREN	3-16
3.2	DIE INTRAFRAME-KOMPRESSION	3-20
3.2.1	<i>Die Farbraumkonvertierung.....</i>	<i>3-20</i>
3.2.2	<i>Die Filterung.....</i>	<i>3-21</i>
3.2.3	<i>Die Digitalisierung.....</i>	<i>3-22</i>
3.2.4	<i>Die Skalierung</i>	<i>3-24</i>
3.2.5	<i>Die Transformation</i>	<i>3-24</i>
3.2.6	<i>Das Packen.....</i>	<i>3-34</i>
3.2.7	<i>Die Bildkompression mit Wavelets</i>	<i>3-34</i>
3.2.8	<i>Die Bildkompression mit Fraktalen</i>	<i>3-35</i>
3.3	DIE INTERFRAME-KOMPRESSION	3-36
3.3.1	<i>Die Picture Interpolation (predictive coding)</i>	<i>3-36</i>
3.3.2	<i>Die Kompensation (motion estimation).....</i>	<i>3-36</i>
3.4	DIE BESCHREIBUNG EINIGER VIDEOKOMPRESSIONSVERFAHREN	3-38
3.4.1	<i>Motion JPEG</i>	<i>3-38</i>
3.4.2	<i>MPEG 1.....</i>	<i>3-40</i>
3.4.3	<i>MPEG 2.....</i>	<i>3-47</i>
4	DIE AUDIOKOMPRESSION	4-53
4.1	DIE MPEG 1 AUDIOKOMPRESSION [MPEG].....	4-54
4.2	DIE MPEG 2 AUDIOKOMPRESSION.....	4-58
5	DIE ARBEITSUMGEBUNG.....	5-60
5.1	DIE VERWENDETE SYSTEMHARDWARE	5-60
5.2	DIE EINGESetzten AUDIO-VIDEO KARTEN	5-61
5.3	DER TIME BASE CORRECTOR (TBC).....	5-65
5.4	DIE PERIPHEREN GERÄTE	5-66
5.4.1	<i>Der Videorecorder (VCR) [JVC1]</i>	<i>5-66</i>
5.4.2	<i>Der Monitor [JVC2]</i>	<i>5-66</i>
5.5	DAS ZUSAMMENSPIEL DER EINZELNEN KOMPONENTEN	5-68
5.5.1	<i>Die Aufnahme von Audio- Videostreams.....</i>	<i>5-69</i>
5.5.2	<i>Die Wiedergabe der Audio- Videostreams.....</i>	<i>5-69</i>
5.6	DIE SOFTWARE	5-70
5.6.1	<i>Der Encoder</i>	<i>5-70</i>

5.6.2	<i>Der Player</i>	5-72
5.6.3	<i>Die Fernsteuersoftware</i>	5-72
5.6.4	<i>Der Mainactor Video Editor</i>	5-73
5.6.5	<i>Das Optibase RTE SDK</i>	5-74
5.6.6	<i>Der Borland C++ Builder</i>	5-74
5.6.7	<i>Das Microsoft Visual Studio C++</i>	5-74
5.6.8	<i>Das Time Measure Tool (TMT)</i>	5-75
5.6.9	<i>Das PSNR-Tool</i>	5-80
6	QUALITÄTS- UND KOMPRESSIONSZEITMESSUNGEN	6-82
6.1	DIE MESSUNG DER KOMPRESSIONSZEIT	6-82
6.1.1	<i>Die Vorstellung der Testmethode</i>	6-82
6.1.2	<i>Die Ergebnisse</i>	6-83
6.1.3	<i>Diskussion der Ergebnisse</i>	6-84
6.2	DIE QUALITÄTSMESSUNGEN	6-85
6.2.1	<i>Die Vorstellung der Testmethode</i>	6-85
6.2.2	<i>Die Testfilme im Überblick</i>	6-85
6.2.3	<i>Vergleich der Qualität der Bilder in den Videofilmen</i>	6-86
6.3	AUSWERTUNG DER ERZIELTEN ERGEBNISSE	6-92
6.3.1	<i>Die durchschnittliche Framegröße</i>	6-92
6.3.2	<i>Overhead bei der Codierung</i>	6-94
6.3.3	<i>Entwicklung der Qualitätswerte innerhalb der GOP</i>	6-95
6.3.4	<i>Der Qualitätsvergleich mittels PSNR</i>	6-97
6.3.5	<i>Der Qualitätsvergleich durch Betrachten der Bilder</i>	6-108
7	AUSBLICK AUF KÜNFTIGE STANDARDS	7-109
7.1	JPEG 2000.....	7-109
8	ZUSAMMENFASSUNG	8-113
8.1.1	<i>Zusammenspiel der Hard und Software</i>	Fehler! Textmarke nicht definiert.
9	ANHANG	9-114
9.1	LITERATURVERZEICHNIS	9-114
9.2	VERZEICHNIS DER ABBILDUNGEN, TABELLEN UND GLEICHUNGEN.....	9-117
9.2.1	<i>Abbildungen</i>	9-117
9.2.2	<i>Tabellen</i>	9-120
9.2.3	<i>Gleichungen</i>	9-122
9.3	MESSERGEBNISSE IM TABELLARISCHEM ÜBERBLICK	9-123
9.4	ANLAGEN.....	9-127
10	DANKSAGUNG	10-128
11	ERKLÄRUNG	11-129

1 Einleitung

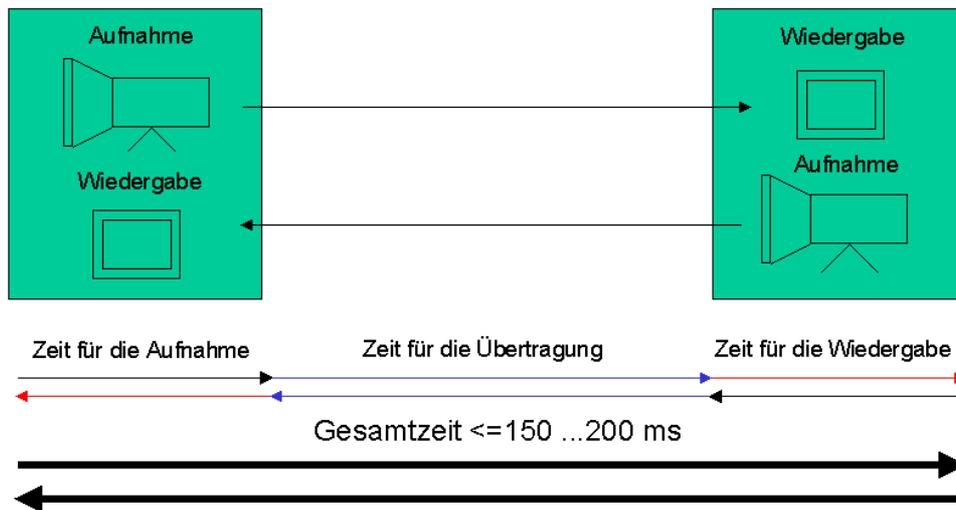
Im Laufe der Jahre hat sich der Computer vom reinen Rechenggerät, welches hauptsächlich zum Aufbereiten und Speichern der Daten verwendet wurde, zum PC entwickelt. Mit den heutigen PC, welche teilweise erstaunliche Multimediafähigkeiten besitzen, können die gesammelten Daten auf vielfältige Art und Weise präsentiert werden. Zum Präsentieren von Daten aller Art gehören auch das Aufnehmen, Übertragen und Abspielen von Videodaten.

Wird eine Videoübertragung für eine Konferenzschaltung genutzt, so muss der Prozess der Videokonferenz „interaktiv“ gestaltet sein. Es soll eine interaktive Wechselbeziehung zwischen den an der Videokonferenz beteiligten Partner entstehen.

Interaktivität im technischem Sinne ist das Reagieren auf Prozesse in Echtzeit oder das Eingreifen von außen in einen gerade laufenden Prozess. Dabei muss das Eingreifen oder das Reagieren auf den Prozess innerhalb 150 ms erfolgen. Einer Dialoganwendung für die Kodierung, Übertragung und Dekodierung ist eine Zeitspanne von 150 bis maximal 200 ms gegeben, denn diese Zeitspanne (verzögerte Reaktion) ist für den Benutzer kaum wahrnehmbar (nach [MIL95]). Ist dies erfüllt, spricht man von einer Echtzeitanwendung oder nennt die Anwendung interaktiv.

In einem Beispielszenario (Abbildung 1-1) ist eine solche Dialoganwendung beschrieben. Das Szenario sei eine Videokonferenz zwischen einem operierenden Ärzteteam und einem Spezialisten, der zur Unterstützung während einer Operation zur Verfügung stehen soll. Im Operationssaal und beim Spezialisten sei je eine Kamera und ein Überwachungsmonitor installiert. Der Spezialist soll während der Operation innerhalb von Bruchteilen einer Sekunde interaktiv, durch Hilfestellungen per Bild und Ton, dem Ärzteteam zur Seite stehen.

Abbildung 1-1 Beispielszenario einer Dialoganwendung



Damit ein assistierender Spezialist rechtzeitig in das Geschehen eingreifen kann, darf die Zeitverzögerung, entstanden durch das Aufnehmen, Übertragen und Wiedergeben des Audio – Video – Stromes, insgesamt nicht länger als 150 bis 200 ms dauern.

Zur weiteren Verdeutlichung soll das Szenario bidirektional aufgebaut werden. Das hieße für das obige Beispiel, dass der Spezialist seine Anweisungen ebenso über eine Kamera und Mikrofon an das Ärzteteam weiter geben könnte. Dann müssten zwei verschiedene Videostreams über das Übertragungsmedium geschickt werden. Dies bedeutet eine Verdopplung der Übertragungslast auf dem LAN¹. Deshalb wird das besondere Augenmerk auf die Art der Kompression der Audio-Videodaten gelegt.

In den folgenden Kapiteln wird herausgefunden, welche Zeitverzögerung durch das Aufnehmen der Videodaten (Kodieren), für das Übertragen des Videostromes und das Wiedergeben der Videodaten (Dekodieren) entsteht. Die Gesamtzeit errechnet

¹ LAN -(Local Area Network) Lokale Netze sind Systeme für den Informationstransfer, die es einer Anzahl Benutzer ermöglichen, auf einem Übertragungsmedium Nachrichten auszutauschen.

sich aus Zeit für die Aufnahme addiert mit der Zeit für die Übertragung in einem LAN und der Zeit welche die Wiedergabe benötigt.

Zu diesem Zwecke wurden im Rahmen dieser Arbeit über 160 unterschiedliche Videos aufgenommen, kategorisiert und miteinander verglichen. Um die Qualität der aufgenommenen Bilder noch besser miteinander vergleichen zu können, wurden aus den Videofilmen über 2400 Bilder extrahiert.

Es wird von konstanten Zeitanteilen für den Zeitteil des Kodierens und Dekodierens der Videodaten ausgegangen. Als variabler Anteil bleibt die Zeit der Videoübertragung.

Des weiteren ist augenscheinlich, dass die Breite des Videostroms, welcher das übertragende Medium belastet und die mögliche Bandbreite der verwendeten Übertragungstechnologie, entscheidend für die Übertragungszeit sind. Der Videostrom ist in seiner Breite abhängig von seiner Art, also vom verwendeten Videokompressionsverfahren und der Kompressionsrate. Deshalb wird im folgenden Kapitel anhand von Rechenbeispielen und praktischen Anschauungen die Notwendigkeit der Kompression von Audio-Videodaten gezeigt.

In dieser Arbeit sollen die Art des Videostroms, dessen Videokompressionsverfahren und Kompressionsrate, sowie spezifische Parameter verändert werden, um für unterschiedliche LAN-Konfigurationen die optimale Wiedergabequalität zu erreichen. Darum wird ein Maß für die optische Qualität der Bilder des übertragenen Videostroms eingeführt und es werden verschiedene Audio-Videoströme miteinander in den Parametern Qualität, Kompressionsrate und Kompressionstechnik verglichen.

Eine Einführung in die zur Verarbeitung von digitalen Videodaten notwendigen Grundlagen, insbesondere Kompressionsverfahren und Formate, bietet das dritte Kapitel.

Das vierte Kapitel beschäftigt sich mit den Grundlagen der MPEG 1 und MPEG 2 Audiokomprimierung.

Um die einzelnen Videoströme miteinander vergleichen zu können, kommt ein Hardwaresystem mit einer Optibase Videokarte, welche im Kapitel sechs näher beschrieben wird, zum Einsatz. Mit deren Hilfe wurden die Audio- /Videoströme aufgenommen.

Des Weiteren werden im fünften Kapitel die verwendeten Softwaretools vorgestellt.

Im Kapitel Sechs werden die Messmethoden vorgestellt, die Ergebnisse ausgewertet.

Im Kapitel Sieben wird ein Ausblick auf künftige Standards gegeben.

Eine Zusammenfassung der Ergebnisse dieser Arbeit zeigt das Kapitel Acht.

Kapitel Neun bildet den Anhang in welchem sich Tabellen und Diagramme finden. Diese Tabellen und Diagramme zeigen die Daten und Parameter der aufgenommenen Filme und daraus extrahierten Bilder.

2 Die Notwendigkeit der Videokompression

Im folgenden sollen die wichtigsten digitalen Bildformate vorgestellt und verglichen werden. Im Anschluss wird durch ein Rechenbeispiel die Notwendigkeit einer Videokompression verdeutlicht.

Zur Zeit existieren mehrere gebräuchliche Bildformate. Tabelle 2-1 zeigt die wichtigsten Eigenschaften.

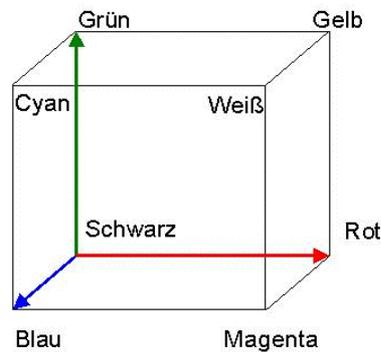
Tabelle 2-1 Eigenschaften digitaler Bildformate (auszugsweise aus [Ohm95])

	QCIF	CIF/ SIF	CCIR- 472	CCIR- 601	EQTV	HD-1440	HD-I	HD-P
Bildpunkte/ Zeile (Y)	176	352	256	720	960	1440	1920	1920
Zeilenanzahl (Y)	144	288	576	576	576	1152	1152	1152
Bildpunkte/ Zeile (U,V)	88	176	128	360	480	720	960	960
Zeilenanzahl (U,V)	72	144	576	576	288	1152	1152	576
Bildseiten- Verhältnis	4:3	4:3	4:3	4:3	16:9	4:3	16:9	16:9
Bildwechsel- Frequenz[Hz]	5-15	10-30	25	25	25	25	25	50
Datenmenge Einzelbild [Kbyte] ²	38,02	152,1	294,9	829,4	829,4	3318	4424	3318
Datenrate für Bildsequenz [Mbit/s]	0,84-3,8	10,1- 30,4	59,0	165,9	165,9	663,5	884,7	1327

Will man einen Farbpunkt auf Computerbildschirmen darstellen, wird das RGB-System zur Kodierung verwendet. Dies ist ein additives Mischfarbsystem, bei welchem die drei Komponenten Rot(R), Grün(G) und Blau(B) zu Mischfarben addiert werden.

² Datenmenge -Die Farbkomponenten werden mit einer Genauigkeit von 8 Bit gespeichert

Abbildung 2-1 Würfeldarstellung des RGB-Farbsystem



Wie in Abbildung 2-1 zu sehen, lässt sich eine Farbe als Punkt in einem dreidimensionalen Würfel darstellen. Es wird üblicherweise jeder Farbwert mit einer Auflösung von 8 Bit gespeichert, was eine Farbvielfalt von $256^3 = 16,7$ Millionen Farben ermöglicht.

Um einen RGB-Wert eines Bildes in ein digitales Bildformat zu konvertieren, werden pro Bild ein Luminanzvektor (Helligkeitsvektor, Y-Vektor) und zwei Farbvektoren (Chrominanzvektor, U -und V-Vektor) gespeichert. Die Farbvektoren bilden einen Differenzialvektor. Die Umwandlung der RGB in die YUV Daten ist eindeutig mit Hilfe der Gleichung 2-1 möglich.

Gleichung 2-1 Umwandlung von RGB-Farbwerten in das innerhalb Europas übliche PAL³ und SECAM⁴ YUV-Komponentenmodell

$$\begin{pmatrix} Y \\ U \\ V \end{pmatrix} = \begin{pmatrix} 0,299 & 0,587 & 0,114 \\ -0,146 & -0,288 & 0,434 \\ 0,617 & -0,517 & -0,1 \end{pmatrix} * \begin{pmatrix} R \\ G \\ B \end{pmatrix}$$

Gleichung 2-2 zeigt, wie sich aus der Auflösung der Anzahl der Bildzeilen und der Farbtiefe die Datenmenge pro Einzelbild errechnet.

Gleichung 2-2 Berechnung der Datenmenge eines Bildes, wenn mittels eines Luminanzvektors und zwei Farbvektoren digital gespeichert wird

$$DMEB = \frac{[(Bpkt / Zeile(Y) * ZA) + 2 * (Bpkt / Zeile(U, V) * ZA(U, V))]}{8000}$$

DMEB = Datenmenge pro Einzelbild in bit
 ZA = Zeilenanzahl
 Bpkt = Bildpunkte
 Y = Helligkeitsvektor
 U,V = Farbvektor

Gleichung 2-3 Datenrate einer Bilderfolge in Mbit/s

$$Datenrate = \left(\frac{DMEB * BFreq}{10000} \right)$$

BFreq = Bildfrequenz in Hz (1/s)

Datenrate in bit/s

U,V = Farbvektor

³ PAL -Phase Alternating Line, in weiten Teilen Westeuropas gültige Fernsehnorm für Farbfernsehen

⁴ SECAM -Sequentiel Couleur Avec Memoire, in Frankreich und Osteuropa gültige Farbf Fernsehnorm

Gleichung 2-3 zeigt, wie sich aus der Datenmenge eines Einzelbildes die Datenrate einer Bildsequenz errechnen lässt.

Die Daten aus der Tabelle 2-1 in die Gleichung 2-3 eingesetzt errechnen eine sehr hohe Datenrate, welche bei der Videobearbeitung entsteht: In einem TV-Studio entstehen Videofilme in einer Auflösung von 720 x 576 Pixel und 25 Bildern in der Sekunde. Wird dieser Videostream abgespeichert, so entsteht gemäß [CCIR601] eine Nettodatenrate von 166 Mbit/s.

Soll dieser Datenstrom aus Video und Audiodaten für eine CD oder DVD produziert und zusammenhängend wiedergegeben werden, dann übersteigt diese Datenrate schnell die maximal mögliche Datenübertragungsrate jedes der zur Zeit gängigen CD ROMs, welche eine maximale Datenübertragungsrate 5,5 Mbyte/s⁵ haben, oder sogar die der Festplatten, mit bis zu 12 Mbyte/s⁶. Sogar die meisten LANs⁷, welche zur Zeit im SOHO⁸ Bereich Verwendung finden, werden mit dieser Übertragungsrate überfordert.

Tabelle 2-2 Vergleich des Datenaufkommens eines Schwarz-Weiß-Standbildes, eines RGB Standbildes und eines 1s Videostreams in gleicher Auflösung (EQTV)

	SW- Bild	RGB – Bild	1s RGB Bilder
MByte	0,55	0,83	20,7

Tabelle 2-2 vergleicht das Datenaufkommen bei der Digitalisierung von Schwarz-Weiß -, RGB Bildern und einer Filmsequenz. In der Spalte SW-Bild wird die Datenmenge für ein Schwarz – Weißbild in EQTV Auflösung (960 x 576) Pixel

⁵ 5,5 Mbyte/s -CD Rom Test in c't 15/99

⁶ 12 Mbyte/s -Festplatten Test in c't 17/99

⁷ LANs -zur Zeit gängige LANs (im SOHO Bereich) haben einen maximalen Datendurchsatz von 5 bis 10 Mbit/s (Stand 1/2000)

⁸ SOHO -Small Office, Home Bereich

gezeigt. Die nächste Spalte zeigt dieses Bild, wenn man eine 8 Bit Farbkodierung zugrunde legt und in der Spalte 1s RGB Bilder wird die Datenmenge für eine Sekunde dieser Bilder (PAL = 25 Bilder/Sekunde) gezeigt.

Tabelle 2-3 Dauer der Übertragung von 2,8Mbyte über verschiedene Netzwerkbandbreiten (abgewandelt übernommen aus [KYA96] Seite 31)

	64KBit	1Mbit	155 Mbit
Sekunden	360	24	0,15

Die Tabelle 2-3 zeigt wie lange die Übertragung eines 2,8 Mbyte großen Bildschirmes auf unterschiedlichen Netzwerkbandbreiten dauert. Die Spalte zwei verdeutlicht die Übertragung auf einer 64kbit/s ISDN Leitung, bei welcher die Bildübertragung 6 min dauert. Die dritte Spalte zeigt die Übertragung auf einem Kanal mit 1Mbit/s Bandbreite und die Spalte Nummer Vier zeigt die Übertragungszeit bei einer 155Mbit/s ATM⁹ Leitung (0,15 Sekunden).

Soll eine Audio-Video-Übertragung in Studioqualität, Fernsehauflösung oder ähnlich, über eine LAN-Strecke in einer bestimmten, so kurz wie möglichen, Zeit übertragen werden, so kann diese nur in einer geeigneten komprimierten Form vonstatten gehen. Die heutzutage verfügbare Übertragungs- und Verarbeitungs-Technik ist nur zum Teil der Übertragung der Bildinformationen in unkomprimierter Form gewachsen. Eine unkomprimierte Videoübertragung ist z.Z. nur mit Hilfe teurerer und

⁹ ATM -(asynchronous transfer mode) Übertragungstechnologie [für die nächste Generation von globalen Kommunikationsnetzen] mit hoher Übertragungsgeschwindigkeit, hoher Skalierbarkeit und Zeittransparenz, wird im WAN und LAN Bereich genutzt.

im Desktopbereich nur wenig verbreiteten Übertragungstechniken wie z.B. ATM oder Gigabit Ethernet¹⁰ möglich.

Ein Beispiel zum Verdeutlichen dieser Aussage soll dem Kapitel zur Kompression vorgreifen: Wird auf einen Audio-Video-Stream mit einer Auflösung von 352 mal 288 Pixel und einer Bildfrequenz von 25 Bildern pro Sekunde der Kompressionsalgorithmus MPEG 1 angewandt, so kann durch diese Kompression ein Audio-Videostrom von lediglich 1,15 Mbit/sek erzeugt werden, was eine Reduzierung um das 26-fache bedeutet. Unkomprimiert wäre dies ein Datenstrom von über 30,4 Mbit/sek .

¹⁰Gigabit Eth. -Netzwerkübertragungstechnoloie mit der eine Verzehnfachung des Datendurchsatzes der Fast Ethernet -Technologie mit Übertragungs von bis zu 1 Gbit/s erreicht wird. Gigabit Ethernet ist Rückwärtskompatibilität mit dem bestehenden 802.3 -Ethernet-Format.

3 Die Videokompression

Alle heutigen Videobearbeitungsumgebungen unterliegen einem Kompromiss.

- Auf der einen Seite steht ein hoher Anspruch an die Qualität (keine Artefakte usw.) sowie eine hohe Auflösung der zu bearbeitenden Bilder,
- auf der anderen die Bandbreitenbeschränkung des Übertragungsmediums.

Um diesem Kompromiss bestmöglich Rechnung zu tragen, wie im Kapitel 2 nachgewiesen, ist bei der Verarbeitung von Audio–Video Strömen eine Kompression der Videodaten notwendig.

Es gibt mehrere Ansätze, wie die Datenmenge einer Videosequenz reduziert werden kann:

- Es kann die Anzahl Pixel pro Bild verringert werden, was deutlich die Qualität der Bilder reduziert. Ein Bild in der Auflösung für EQTV (960 x 576) herunter gerechnet auf CIF/SIF (352 x 288) verringert die Datenmenge nur auf etwa ein Fünftel.
- Auch die Beschränkung auf nur wenige übertragene Bilder pro Sekunde reduziert die Datenmenge. Dies wirkt sich aber auf die Kontinuität der Bilder, den Bilderfluss, aus, es kommt zum „Ruckeln“ des Videostromes. Bei 10 übertragenen Bildern, statt 25 bei PAL, verringert sich die Datenmenge nur um etwa den Faktor 2,5.

All diese Techniken verringern die Datenmenge nur um einen geringen, unwesentlichen Faktor und das nur bei Verschlechterung der Qualität und Kontinuität der Bilder. Es werden effektivere Kompressionsverfahren benötigt.

3.1 Die Videokompressionsverfahren

Es wird grundsätzlich zwischen verlustlosen und verlustbehafteten Kompressionsverfahren unterschieden.

Bei verlustlosen Videokompressionsverfahren kann ein exaktes Abbild der Originaldaten rekonstruiert werden, da bei der Kompression im Idealfall (Verluste entstehen durch Rechenungenauigkeiten der Rechner) keinerlei Informationen verloren gehen. Zu verlustlosen Kompressionsverfahren zählt man z.B. die Arithmetische-Kodierung, die Lauflängen-Kodierung und die Huffman-Kodierung.

Anders als bei den verlustlosen Videokompressionsverfahren können bei den verlustbehafteten die Originaldaten nach einer Kompression nicht wieder vollständig hergestellt werden. Die verlustbehafteten Kompressionsverfahren werden in zwei große Gruppen eingeteilt, in die Intra-Frame-Kompressionsverfahren und die Inter-Frame-Kompressionsverfahren.

Die Intra-Frame-Kompressionsverfahren benutzen die Reduzierung der räumlichen Redundanz zur Kompression eines Bildes. Auch hier sei ein Verfahren genannt.: DPCM ¹¹.

Inter-Frame-Kompressionsverfahren nutzen die zeitliche Redundanz der Bilder einer Bilderfolge aus. Es werden Informationen aus den im Videostrom vorhergehenden und/oder nachfolgenden Bildern mittels Differenzbildern oder Bewegungsvektoren zur Kodierung des Bildes herangezogen. Zu den Inter-Frame-Verfahren zählt man die Bewegungskompensation.

Des weiteren werden die Kodierungsverfahren in Entropie- und Source-Kodierungen unterteilt.

¹¹ DPCM -(Differential Pulse Code Modulation) ist ein verlustbehaftetes Kompressionsverfahren.

Entropiekodierungen sind Verfahren, welche die Ausgangsdaten zeilenweise kodieren und dadurch Kompression erreichen.

Sourcekodierungen trennen relevante und nicht relevante¹² Daten und erreichen eine Kompression durch Entfernen der nicht relevanten Daten.

Es wird in der Praxis auch noch eine dritte Art von Kompressionsverfahren, die Hybrid-Kodierung, welche eine Mischform der beiden erstgenannten ist.

Zu den Hybrid-Kodierungen seien XCCC¹³, JPEG¹⁴ und MPEG¹⁵ genannt. Des weiteren können die Komprimierungsverfahren anhand der Kodier- und Dekodierzeit unterschieden werden.

Es gibt Komprimierverfahren bei denen die Zeiten des Komprimierens und Dekomprimierens etwa gleich sind, diese werden symmetrische Kompressionsverfahren genannt.

Bei unsymmetrischen Komprimierverfahren ist der Aufwand des Komprimierens größer als die des Dekomprimierens.

Für die Kompression von Bildern oder Bilderfolgen (Videos) eignen sich alle angesprochenen Kompressionsverfahren. Zur Kompression von Programmen oder Zahlenfolgen dürfen keine verlustbehafteten Kompressionsverfahren verwendet werden, weil im Gegensatz zu Bildern bei Programmdateien kein Detailverlust hingenommen werden darf.

¹² Daten -(nicht relevante) sind hier Daten, welche vom menschlichem Auge nicht oder nur schlecht wahrzunehmen sind.

¹³ XCCC -(Extended Colour Cell Compression) ist ein nicht lizenziertes Hybridkompressionsverfahren.

¹⁴ JPEG -(Joint Photographic Experts Group) ist ein weitverbreitetes und standardisiertes Einzelbildkompressionsverfahren [ISO10918].

¹⁵ MPEG -(Moving Pictures Experts Group) standardisiertes Kompressionsverfahren zur Kompression von Bewegbildern mit zugehörigem digitalen Audio.

Die nachfolgenden Tabellen, Tabelle 3-1 und Tabelle 3-2 zeigen eine Übersicht über die Einteilung der Kompressionsverfahren und die Kategorisierung einiger Beispiele.

Tabelle 3-1 Gliederung der Kompressionsverfahren, aus [MIL95]

Charakteristik	Erläuterung
Verlustfrei	Ausgangsdaten sind vollständig rekonstruierbar
Intraframe	Zeitliche Redundanzen zwischen zu kodierenden Bildern bleiben unberücksichtigt
Interframe	Zeitliche Redundanzen zwischen zu kodierenden Bildern werden berücksichtigt
Symmetrisch	Kodierung und Dekodierung sind ungefähr gleich schnell durchführbar
Asymmetrisch	Kodierung benötigt wesentlich mehr Aufwand als die Dekodierung

Tabelle 3-2 Kategorisierung von Kodierungstechniken, aus [MIL95]

Kategorie	Beispiel	Verlustfrei
Entropie-Kodierungen	Laufängen-Kodierung	Ja
	Huffman-Kodierung	Ja
	Arithmetische-Kodierung	Ja
Source-Kodierung	DCT	Ja ^a
	DPCM	Ja
	Bewegungskompensation	Ja
	DWT ¹⁶	Ja ^a
	IFS ¹⁷	Nein
Hybrid-Kodierung	XCCC	Nein
	H.261 ¹⁸	Nein
	JPEG	Nein
	MPEG –1	Nein
	Bildkompression mit Wavelets	Nein
	Fraktale Bildkompression	Nein

a Durch die begrenzte Rechengenauigkeit in der Praxis häufig mit Verlusten behaftet.

¹⁶ DWT -Diskrete Wavelet Transformation

¹⁷ IFS -Iterated Function System

¹⁸ H.261 -Symmetrisches Echtzeitkompressionsverfahren zur Bildübertragung (Videokonferenzen) in ISDN Netzen.

3.2 Die Intraframe-Kompression

Die in diesem Kapitel beschriebenen Methoden dienen der Kompression von einzelnen Bildern. Diese Standbildkompressionsverfahren nutzen die Reduzierung der räumlichen Redundanz zur Kompression eines Bildes.

3.2.1 Die Farbraumkonvertierung

In Computersystemen auf Intel Basis (auch in den meisten anderen) wird die Farbe jedes Pixels¹⁹ durch das RGB-Farbmodell beschrieben. Im RGB-Modell ist die Farbe jedes einzelnen Pixels durch die drei Grundfarben Rot, Grün, Blau definiert. Da aber das menschliche Auge wesentlich stärker Helligkeitskontraste als Farbkontraste wahrnimmt (eine Reduzierung der Farbdaten von 24 auf 15 bit wird von dem menschlichem Farbsensorsystem nicht wahrgenommen) und da der Mensch nur 50.000 bis 300.000 Farben unterscheiden kann, ist es sinnvoll diese Eigenschaft durch eine Farbraumkonvertierung auszunutzen.

Die RGB-Farbinformationen werden mit Hilfe der Gleichung 3-1 in das YUV-Farbmodell konvertiert (Abbildung 3-1). Die Farbaufösung wird dadurch auf die Hälfte reduziert. Die Farben eines 2 x 2 Pixel großen Blockes werden gemittelt, während die Luminanzaufösung erhalten bleibt. Nach der Konvertierung setzt sich die Pixelinformation aus der Helligkeit (Luminanz, „Y“) und den Farbwerten Chrominanz, („U“ und „V“) zusammen.

Gleichung 3-1 Umwandlung der RGB Werte ins YUV- Farbmodell

$$\begin{pmatrix} Y \\ U \\ V \end{pmatrix} = \begin{pmatrix} 0,299 & 0,587 & 0,114 \\ -0,146 & -0,288 & 0,434 \\ 0,617 & -0,517 & -0,1 \end{pmatrix} * \begin{pmatrix} R \\ G \\ B \end{pmatrix}$$

¹⁹ Pixel -darzustellender Punkt eines Bildes

Abbildung 3-1 Zuordnung der Farb- und Helligkeitskomponenten vor und nach der Farbraumkonvertierung



Beispiel:

Ausgehend von einem 2 x 2 Pixel großen Block summiert sich bei einer Farbtiefe von 8 Bit pro Grundfarbe der Speicherbedarf auf:

$$4 \text{ (Pixel)} * 3 \text{ (Grundfarben)} * 8 \text{ (Bit Farbtiefe)} = 96 \text{ Bit.}$$

Im YUV-Modell ergeben sich nur:

$$2(U \text{ und } V) * 8(\text{Bit Quantisierungstiefe}) + 4 \text{ (Pixel)} * 8 \text{ (Bit Helligkeitstiefe)} = 48 \text{ Bit.}$$

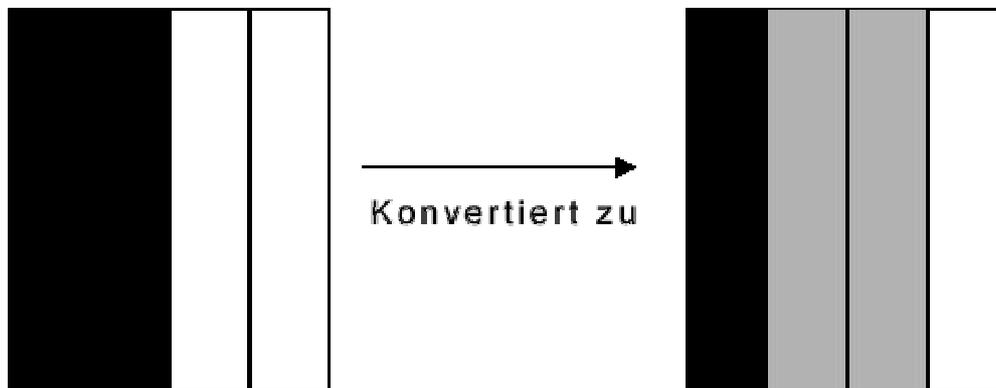
Der Qualitätsunterschied ist nicht sichtbar, trotzdem wurde das Datenaufkommen um die Hälfte reduziert.

3.2.2 Die Filterung

Bei der Filterung von Bildern wird versucht, das Bild auf eine besser zu komprimierende Darstellung zu bringen. Scharfe Kanten, die bei der später beschriebenen Transformation zu schlechten Kompressionsfaktoren führen, werden eliminiert.

Das heißt, dass zum Beispiel bei einem Übergang von einer schwarzen zu einer weißen Fläche anstatt dem sehr scharfen Schwarz – Weiß – Übergang eine graue Linie eingesetzt wird (Abbildung 3-2).

Abbildung 3-2 Kantenglättung bei der Filterung



Einerseits verbessert diese Vorgehensweise den Kompressionsfaktor, andererseits leidet die Qualität des dekodierten Bildes.

Die Balance zwischen Qualität und Datenmengenreduzierung hängt hier im entscheidendem Maße von der Qualität der benutzten Filter ab.

3.2.3 Die Digitalisierung

Auch bei der Digitalisierung der Bildinformationen wird die Eigenschaft des menschlichen Auges, besser auf Helligkeitsveränderungen als auf Farbveränderungen zu reagieren, ausgenutzt. Es wird bei der Digitalisierung die Reduzierung der Farbinformationen mittels einer Abbildung der Farbwerte auf der Zeitachse durchgeführt.

Um ein analoges Signal zu digitalisieren, ist in bestimmten zeitlichen Abständen eine Messung der RGB-Werte durch Abtastung bzw. Sampling erforderlich (Abbildung 3-3).

Abbildung 3-3 Abtastung der RGB-Komponenten bei der Digitalisierung

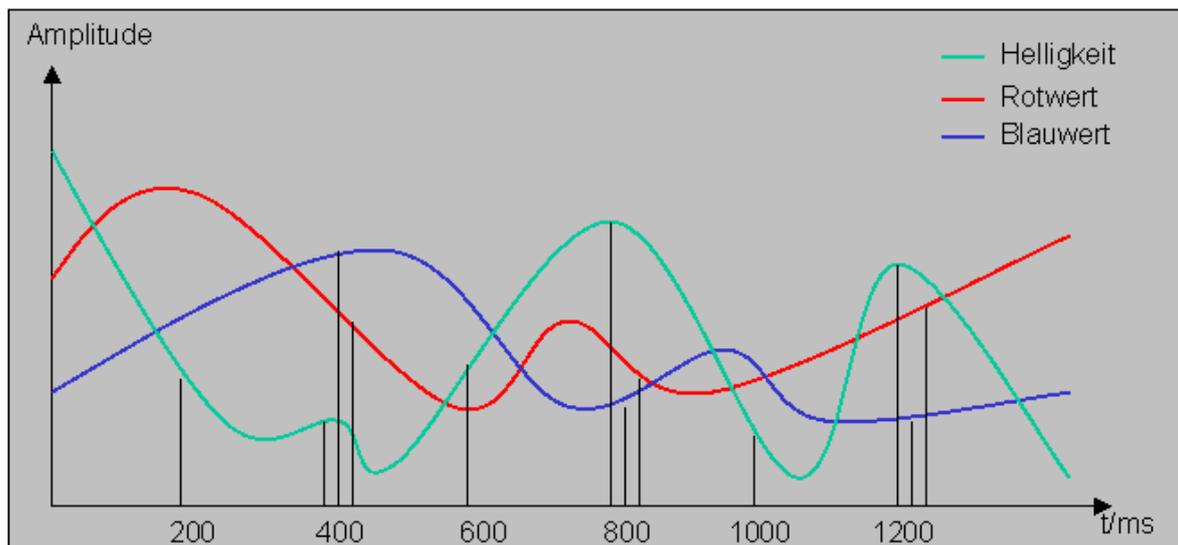
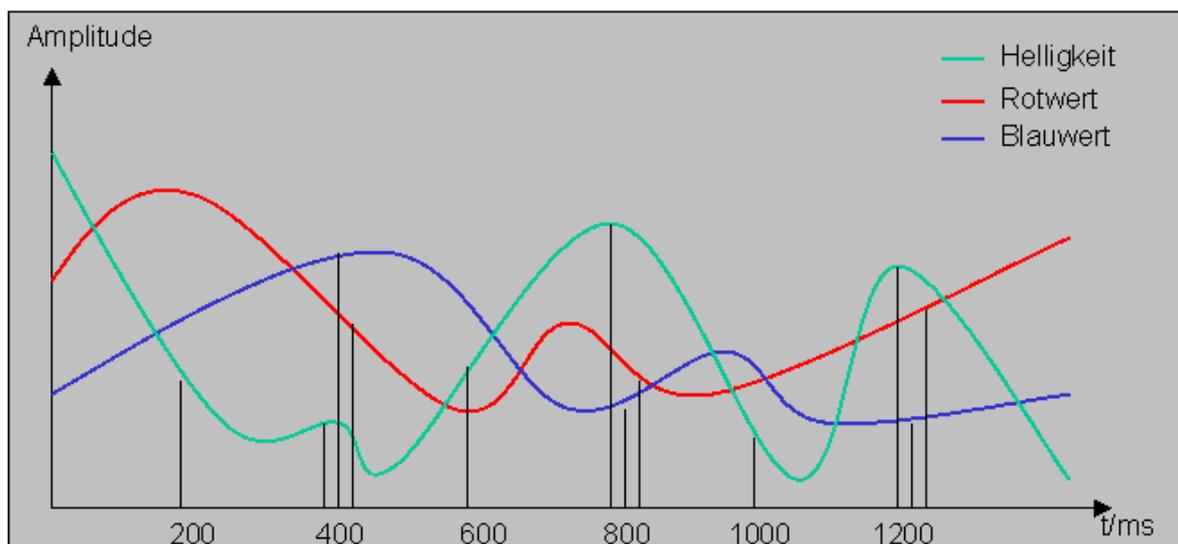


Abbildung 3-4 Abtastung der Helligkeits- und Farbwerte bei der Digitalisierung



Im YUV-Farbraum werden, wie in Kapitel 4.1.2 beschrieben, Luminanz und Chrominanz getrennt gespeichert. Bei dem YUV 4:2:2 Verfahren werden z.B. die Luminanzwerte 4 mal und in der gleichen Periode die Chrominanzwerte nur 2 mal

abgetastet (Abbildung 3-4). Das heißt, die Helligkeit wird mit der doppelten Frequenz wie die Farbwerte gemessen.

Daraus ergibt sich eine Datenmengenreduzierung um den Faktor 1,5, ohne dass optisch sichtbare Qualitätsverluste auftreten. Das digitalisierte Video hat immer noch Studioqualität.

3.2.4 Die Skalierung

Die Skalierung ist das wichtigste Werkzeug der Videokompression, durch sie kann die erforderliche Bandbreite an die verschiedensten Systemvoraussetzungen angepasst werden.

Bei der Skalierung wird Einfluss auf die entscheidenden Videoqualitätsmerkmale, nämlich die Farbtiefe, die Auflösung und die Bildwiederholfrequenz genommen.

Diese Parameter beeinflussen im wesentlichen die benötigte Bandbreite zur Übertragung bzw. Bearbeitung des Videos. Durch Variierung dieser Parameter kann das Video nahezu beliebig komprimiert werden, allerdings kommt es bei zu großer Komprimierung auch zu erheblichen Qualitätseinbußen.

Wenn zum Beispiel die Auflösung von 640 x 480 auf 320 x 240 Bildpunkte reduziert wird, verringert sich die benötigte Bandbreite auf ein Viertel des ursprünglichen Wertes.

3.2.5 Die Transformation

Ziel von Transformationen ist es, die für die Videobearbeitung relevanten Informationen von den weniger wichtigen unterscheiden zu können. Eine Transformation muss umkehrbar sein, um die ursprünglichen Informationen wieder herstellen zu können.

Zur Transformation bei der Videokompression werden zwei verschiedene Verfahren angewandt. Zum einen die diskrete Kosinus-Transformation (DCT) und zum anderen die diskrete Wavelet-Transformation (DWT). Es handelt sich um eine diskrete Transformation, da durch die Digitalisierung diskrete Werte vorliegen.

3.2.5.1 Die diskrete Kosinus- Transformation (DCT)

Um eine DCT durchführen zu können, wird eine Blocking-Strategie angewendet: Das Bild wird in mehrere quadratische Teilblöcke mit jeweils 64 Pixel aufgeteilt.

Die Werte der einzelnen Pixel dieses 8 x 8 Blockes werden den Zellen einer Matrix zugeordnet, die diese Farbinformationen aufnimmt (Abbildung 3-5, hier für leichteres Verständnis nur Grauwerte).

Es ergibt sich eine Matrix in $M_8(\mathbb{R})$. Die 64 Elemente dieser neuen Basis werden Basisbilder genannt.

Diese 8 x 8 Blöcke (Tabelle 3-3, die Grauwerte sind frei gewählt, sind aber durchaus typisch) werden mit Hilfe der Gleichung 3-2 so transferiert, dass sich in der neuen Matrix (Tabelle 3-4) die wichtigsten Informationen oben links (kleine x, y Werte) befinden.

Tabelle 3-3 Blockbildung für die diskrete Kosinus-Transformation

Y\X	0	1	2	3	4	5	6	7
0	140	144	147	140	140	155	179	175
1	144	152	140	147	140	148	167	179
2	152	155	136	167	163	162	152	172
3	168	145	156	160	152	155	136	160
4	162	148	156	148	140	136	147	162
5	147	167	140	155	155	140	136	162
6	136	156	1123	167	162	144	140	147
7	148	155	136	155	152	147	147	136

Nach dieser Transformation wird die DCT angewendet. Es handelt sich um eine 2-dimensionale DCT, welche definiert ist durch die Gleichung 3-2.

Gleichung 3-2 Die zweidimensionale DCT

$$f(k, n) := \frac{C(k)}{2} \frac{C(n)}{2} \sum_{x=0}^7 \sum_{y=0}^7 F(x, y) \cos\left(\frac{p(2x+1)k}{16}\right) \cos\left(\frac{p(2y+1)n}{16}\right)$$

wobei $k, n \in \{0, \dots, 7\}$ und $C(z) := \begin{cases} 1/\sqrt{2}, & \text{für } z = 0 \\ 1, & \text{für } z > 0 \end{cases}$ gilt.

Hieraus entsteht eine Koeffizientenmatrix mit $k+1$ Zeilen und $n+1$ Spalten (Abbildung 3.4).

Tabelle 3-4 Ergebnismatrix nach Anwendung der DCT (Integerwerte)

Y\X	0	1	2	3	4	5	6	7
0	186	-18	15	-9	23	-9	-14	-19
1	21	-34	26	-9	-11	11	14	7
2	-10	-24	-2	6	-18	3	-20	-1
3	-8	-5	14	-15	-8	-3	-3	8
4	-3	10	8	1	-11	18	18	15
5	4	-2	-18	8	8	-4	1	-7
6	9	1	-3	4	-1	-7	-1	-2
7	0	-8	-2	2	1	4	-6	0

Abbildung 3-5 Veranschaulichung der Frequenzzunahme über der Hauptdiagonalen (übernommen aus [MIL95] Seite 20)

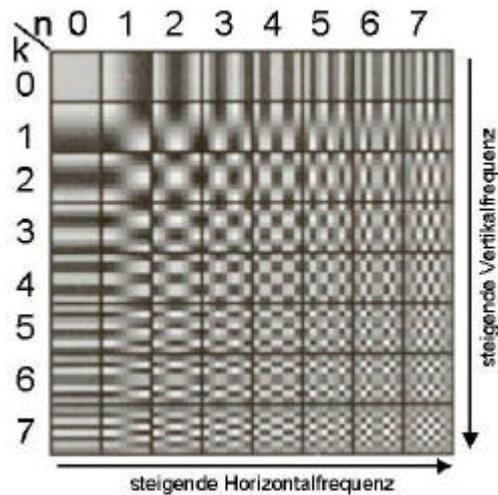


Tabelle 3-4 zeigt, dass die Beträge der Koeffizienten über der Hauptdiagonalen der Matrix in Richtung größerer x-, y-Werte abnehmen. Bei Betrachtung der Abbildung 3-5 kann festgestellt werden, dass die Frequenz von kleinen x-, y-Werten zu großen

x-, y-Werten zunimmt. Es werden also hohen Frequenzen kleine Beträge der Koeffizienten zugeordnet.

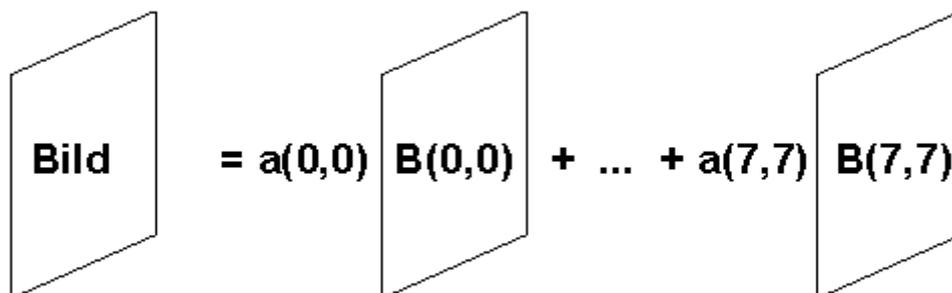
Bei der DCT entsprechen diese Koeffizienten den Amplitudenwerten.

Je mehr dieser Koeffizienten vorhanden sind, desto genauer wird die ursprüngliche Funktion beschrieben. Genau wie bei der Fouriertransformation, wird auch bei der DCT durch eine Abarbeitung von mehr Koeffizienten, eine immer besser werdende Annäherung an die Originalfunktion erreicht.

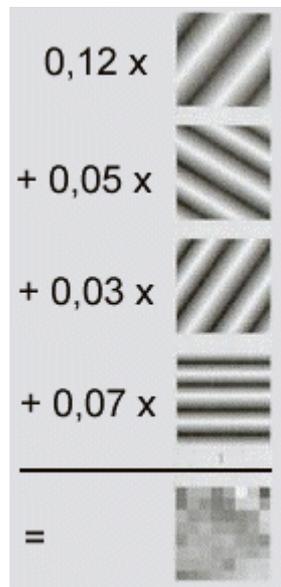
Zu diesem Zeitpunkt hat aber noch keine Kompression stattgefunden. Die Transformation dient lediglich der Umwandlung in eine besser zu komprimierende Darstellung der Bildinformation.

In Abbildung 3-6 und in Abbildung 3-7 wird gezeigt, wie sich ein Bild aus den einzelnen Basisbildern zusammensetzt. Hierbei entspricht $f(k,n)$ aus Gleichung 3-2 dem Basisbild $B(k,n)$ zugehörigen Koeffizienten $a(k,n)$.

Abbildung 3-6 Darstellung eines 8x8 Bildes mit 64 Basisbildern $\{B(k,n)\}$ (nach [MIL95])

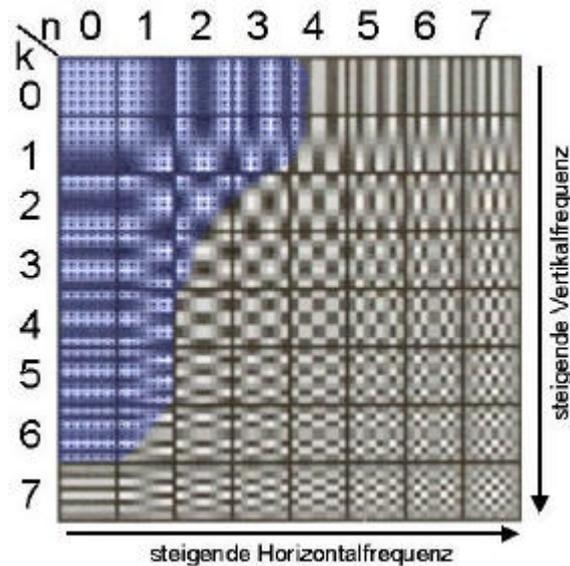


**Abbildung 3-7 Anschauliche Darstellung eines 8x8 Bildes mit 64 Basisbildern
(nach [c't2699])**



Die nach der DCT erhaltenen 8 x 8-Blöcke haben den Vorteil, dass sich die für das menschliche Auge wichtigen Informationen hauptsächlich innerhalb des markierten Bereiches der Abbildung 3-8 befinden. Hochfrequente Anteile, welche durch das Auge schlechter wahrgenommen werden können, beinhalten demnach weniger wichtige Informationen als tieffrequente Anteile, welche das Auge besser verarbeiten kann.

Abbildung 3-8 Bedeutung der DCT-Koeffizienten, Unterteilung in wichtige (blau markiert) und unwichtige Werte (nach [MEI94])



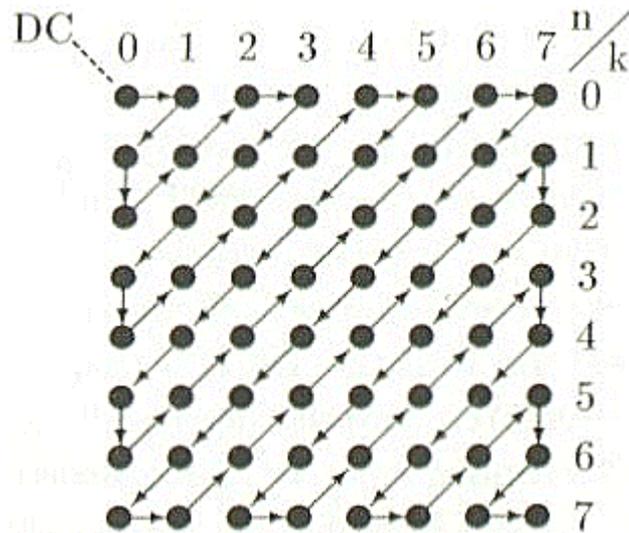
Bei der Abarbeitung der erzeugten Koeffizientenmatrix wäre es naheliegend, zeilen- oder spaltenweise vorzugehen. Um aber eine Reihenfolge zu erzeugen, die der Wichtigkeit der Informationen entspricht, wird eine neue Bearbeitungsreihenfolge festgelegt. Die DCT-Koeffizienten werden in einer speziellen Reihenfolge, von für $(n,k)=0,0$ zu $(n,k)=7,7$ wie in Abbildung 3-9 zu sehen, durchlaufen.

Gleichung 3-3 Bearbeitungsreihenfolge der DCT-Koeffizienten

$$R(n,k)=\{(0,0);(1,0);(0,1);(0,2);(1,1);(2,0) \dots (,7,7)\}$$

Die Reihenfolge entspricht der schlechter werdenden Sichtbarkeit der Basisbilder (Abbildung 3-8).

Abbildung 3-9 Bearbeitungsreihenfolge der DCT-Koeffizienten



Die DCT-Koeffizienten, die in dieser Abfolge durchlaufen, ergeben eine Reihe mit den wesentlichen Informationen des Bildes (in den ersten Werten der Reihe Gleichung 3-3), bis hin zu denen, welche die für das menschliche Auge schlecht sichtbaren Informationen enthalten, welche vernachlässigt werden können.

Die Kompression der Daten erfolgt durch anschließende Quantisierung und der Lauflängenkodierung.

Tabelle 3-5 Quantisierungsmatrix mit dem Qualitätsfaktor 2

Y\X	0	1	2	3	4	5	6	7
0	3	5	7	9	11	13	15	17
1	5	7	9	11	13	15	17	19
2	7	9	11	13	15	17	19	21
3	9	11	13	15	17	19	21	23
4	11	13	15	17	19	21	23	25
5	13	15	17	19	21	23	25	27
6	15	17	19	21	23	25	27	29
7	17	19	21	23	25	27	29	31

Eine Besonderheit dieser Matrix ist die Vergrößerung der Koeffizienten entlang der Hauptdiagonalen (Tabelle 3-5). Dieses hat zu Folge, dass bei der Quantisierung die in der Bearbeitungsreihenfolge weiter hinten platzierten (für große i der Reihe aus Gleichung 3-3) DCT-Koeffizienten eher zu Null werden als weiter vorn platzierte (für kleine i der Reihe aus Gleichung 3-3). Daraus ergibt sich Tabelle 3-6.

Der Vorteil liegt auf der Hand: Durch das Nullsetzen der hochfrequenten Koeffizienten wird eine sehr effiziente Lauflängenkompression ermöglicht.

Quantisiert wird, indem nur der ganzzahlig durch den Quantisierungskoeffizienten (eine natürliche Zahl $Q \in \mathbb{N}$, dem Q-Faktor) teilbare Wert beibehalten wird.

Beispiele zur Berechnung der quantisierten Koeffizienten:

$$\text{Koeffizient (0;0): } 186:3 = 62; \text{ Rest} = 0 \Rightarrow 186-0=186$$

$$\text{Koeffizient (1;1): } (-34):7 = -4; \text{ Rest} = -6 \Rightarrow (-34)-(-6) = -28$$

$$\text{Koeffizient (2;2): } (-2):11 = 0; \text{ Rest} = -2 \Rightarrow (-2)-(-2) = 0$$

Die Kompression kann durch Erhöhen des Q-Faktors gesteigert werden, was allerdings zu einer Verschlechterung der Bildqualität führt.

Theoretisch ist die DCT wegen ihrer Invertierbarkeit (Gleichung 3-4) verlustfrei. Es entstehen aber durch die begrenzte Rechengenauigkeit Verluste.

Tabelle 3-6 DCT-Koeffizienten nach der Quantisierung

Y\X	0	1	2	3	4	5	6	7
0	186	-15	14	-9	22	0	0	-17
1	20	-28	18	0	0	0	0	0
2	-7	-18	0	0	-15	0	-19	0
3	0	0	13	-15	0	0	0	0
4	0	0	0	0	0	0	0	0
5	0	-17	0	0	0	0	0	0
6	0	0	0	0	0	0	0	0
7	0	0	0	0	0	0	0	0

Die inverse DCT ist beschrieben durch Gleichung 3-4.

Gleichung 3-4 Die inverse zweidimensionale DCT

$$F(x, y) = \sum_{k=0}^7 \sum_{n=0}^7 \frac{C(k)}{2} \frac{C(n)}{2} f(k, n) \cos\left(\frac{p(2x+1)k}{16}\right) \cos\left(\frac{p(2y+1)n}{16}\right)$$

Mit $x, y \in \{0, \dots, 7\}$ und C wie in Gleichung 3-2.

3.2.6 Das Packen

Für das Packen von Daten stehen verschiedene Varianten zur Verfügung.

Zum größten Teil wird die Huffman-Kodierung benutzt. Bei der Huffman-Kodierung werden die Zeichen zuerst nach der Auftrittswahrscheinlichkeit sortiert. Anschließend werden den Zeichen mit einer hohen Auftrittswahrscheinlichkeit ein kurzes Kodewort und den Zeichen mit geringer Auftrittswahrscheinlichkeit ein langes Kodewort zugeordnet.

Es gibt auch arithmetische Kodierungen, die aber lizenzrechtlich von IBM, AT&T und Mitsubishi geschützt und damit wegen der Lizenzgebühren unattraktiv für Programmierer sind.

Andere Verfahren verwenden oft auch eine Lauflängenkodierung, die nicht die einzelnen Zeichen kodiert, sondern Art und Anzahl beschreibt. So würde z.B. aus 222244441 die Folge 424411.

Nach der Quantisierung einer DCT-Koeffizientenmatrix beispielsweise, entsteht bedingt durch das Durchlaufen der Matrix (in der Reihenfolge wie in Gleichung 3-3 beschrieben) eine große Anzahl aufeinander folgender Nullen, die bei der Lauflängenkompromierung zu guten Kompressionsfaktoren führen.

3.2.7 Die Bildkompression mit Wavelets

Im Gegensatz zu der DCT wird bei der Wavelet-Transformation keine Blockbildung benötigt. Die Wavelet-Transformation arbeitet wie ein Filter der immer feiner wird.

Die Kompression mittels Wavelets basiert auf einer Zerlegung des Ausgangsbildes in Teilbilder unter Zuhilfenahme von Hoch²⁰- und Tiefpassfiltern²¹. Dieses Filtern wird

²⁰ Hochpassfilter -Frequenzfilter, welcher tiefe Frequenzen herausfiltert

²¹ Tiefpassfilter -Frequenzfilter, welcher hohe Frequenzen herausfiltert

rekursiv auf die entstandenen Bilder angewandt. Es entstehen jeweils 4 Teilbilder mit jeweils der halben Auflösung des Ausgangsbildes.

Zuerst werden die feinen Details von den gröberen (und somit besser sichtbaren) Strukturen des zu komprimierenden Bildes getrennt. Dies wird sukzessive bis zu einer vorgegebenen Anzahl von Schritten, oder immer wieder bis das Bild nur noch ein Pixel groß ist, wiederholt. So wird ständig zwischen wichtigen (für das Auge besser zu sehenden) und weniger wichtigen Bilddaten (welche für das Auge nicht gut wahrnehmbaren Details beinhalten) unterschieden und eine Kompression erreicht.

Die Bildkompression mit Wavelets ist nicht standardisiert.

3.2.8 Die Bildkompression mit Fraktalen

Ähnlich der Bildkompression mit Wavelets gibt es auch keinen Standard, welcher die Kompression durch Fraktalbildung definiert.

Grundlage dieses Verfahrens ist die Fraktalgeometrie. Im Bild werden Selbstähnlichkeiten gesucht und dann dieses oft wiederkehrende Teilbild durch Fraktale beschrieben. Es werden keine Pixel sondern Bildfunktionen definiert. Bei natürlichen Bildvorlagen können so hohe Kompressionsraten erreicht werden. Abstrakte Bilder sollten nicht mit diesem Verfahren verarbeitet werden.

Das komprimierte Bild liegt mit beliebiger Auflösung vor. Es müssen dazu keine speziellen Verfahren, wie Pixelverdopplung oder Interpolation, angewandt werden.

Fraktale-Verfahren sind im Gegensatz zu DCT-basierenden, wegen ihrer sehr langen Codierzeit, nur asymmetrisch durchführbar. Weiterhin ist die fraktale Bildkompression immer mit Verlusten behaftet, weil Bilder in der Regel nicht vollständig zu sich selbst ähnlich sind.

3.3 Die Interframe-Kompression

Bei der Interframe-Kompression werden Redundanzen verschiedener, zeitlich aufeinander folgender Bilder ausgenutzt. Dieses Verfahren eignet sich gut für die Kompression von natürlichen Videos, weil in der Bilderfolge eine große Anzahl zeitlicher Übereinstimmungen vorliegt.

3.3.1 Die Picture Interpolation (predictive coding²²)

Bei der Picture Interpolation wird ein intraframe-komprimiertes Bild als Referenzbild festgelegt. Dieses sogenannte I-Frame (Intra-Frame) oder Key-Frame ist das Schlüsselbild.

Ausgehend vom Schlüsselbild wird nun versucht ein sogenanntes P-Frame (Predictive-Frame) vorherzusagen. Ist dieses geschehen, interpoliert der Codec weitere Bilder zwischen dem I- und dem P-Frame. Diese Bilder heißen B-Frames.

Die Datenmengenreduzierung entsteht dadurch, dass nur die Bildunterschiede und nicht die kompletten Bilder verarbeitet werden. Es wird also nur hin und wieder ein Originalbild verwendet, während die anderen Bilder nur geschätzt bzw. interpoliert sind.

3.3.2 Die Kompensation (motion estimation)

Bei diesem Verfahren werden Bewegungen im Bild abgeschätzt. Viele Bildbereiche verschieben sich quasi nur im Bild. Diese Bildbereiche werden gesucht und anschließend nur durch Bewegungsvektoren beschrieben.

²² Predictive -engl. Vorhersagend

Wenn beispielsweise ein über den Rasen rollender Ball abgebildet wird, so ist schnell klar, dass das gesamte Bild nur einmal gebraucht wird und dann für die folgenden Bilder zum größten Teil nur noch die Positionsveränderungen des Balles auf dem Rasen bestimmt werden muss. Zusätzlich müssen natürlich die vorher vom Ball verdeckten Rasenflächen als Information vorliegen.

3.4 Die Beschreibung einiger Videokompressionsverfahren

Die wichtigsten Kriterien eines Kompressionsverfahrens sind die bildliche Qualität und der Kompressionsfaktor der durch Kompression erzeugten Videodaten. Es ist immer das erste Ziel bei gegebenen Bedingungen, den bestmöglichen Kompromiss zwischen höchst mögliche Darstellungsqualität bei möglichst großer Kompression zu erreichen. Oft werden aber Kompromisse bezüglich der Qualität gemacht werden müssen. Dadurch gewinnen Eigenschaften, wie zum Beispiel die der Skalierbarkeit, an Bedeutung.

3.4.1 Motion JPEG

Der M-JPEG Codec basiert auf dem JPEG-Einzelbildkompressionsverfahren.

JPEG (Joint Photographic Experts Group) ist ein Standard, der Einzelbilder mittels Filterung, Digitalisierung, Skalierung, DCT, Quantisierung, Lauflängenkodierung und abschließender Huffman-Kodierung bis auf ein Hundertstel der Originalgröße komprimiert.

Bis zu einem Kompressionsverhältnis von 15:1 sind kaum sichtbare Qualitätseinbußen zu verzeichnen. Bei höherer Komprimierung (ab ca. 25:1) wird allerdings eine starke Blockbildung (Artefakte), wie in Abbildung 3-10 und Abbildung 3-11 gezeigt, eintreten. Ursache dafür sind die vernachlässigten DCT-Koeffizienten.

Abbildung 3-10 Darstellung der Blockbildung bei geringer Komprimierung



Abbildung 3-11 Darstellung der Blockbildung bei großer Komprimierung, es sind deutlich Artefakte zu sehen



Motion-JPEG ist eine zum Teil nicht standardisierte Weiterentwicklung von JPEG. Bei diesem Verfahren findet eine Aneinanderreihung der JPEG komprimierten Einzelbilder des Videostromes statt.

Das hat den Vorteil, dass jedes Bild einzeln zur Bearbeitung zur Verfügung steht.

Der Nachteil besteht darin, dass keine Inter-Frame Kompression stattfindet und die Redundanzen nicht ausgenutzt werden. Dadurch können die Redundanzen zwischen den Bildern zur besseren Kompression nicht herangezogen werden.

Das Verfahren ist nicht umfassend standardisiert. Deshalb ist nicht gewährleistet, dass die M-JPEG-Hardware verschiedener Hersteller kompatibel ist.

3.4.2 MPEG 1

Die „Motion Picture Expert Group“ ist eine Gruppe der ISO²³, welche sich mit der Standardisierung im Videobereich beschäftigt. Von dieser wurde 1993 MPEG-1, „Kodierung von Bewegtbildern und assoziierten Audio für digitale Speichermedien mit bis zu 1,5 Mbit/s“, als Standard der ISO/IEC (ISO 11172) verabschiedet.

Zielsetzung bei der Entwicklung des MPEG-1 Codecs war es, Videos mit Ton in befriedigender Qualität bei nur geringer verfügbarer Bandbreite (1Mbit/s bis 1,5 Mbit/s), auf einem PC abspielen zu können. Es wurde speziell für CD-ROM-Laufwerke der ersten Generation entwickelt. Der Datendurchsatz solch eines Laufwerkes betrug ca. 150 kByte/s.

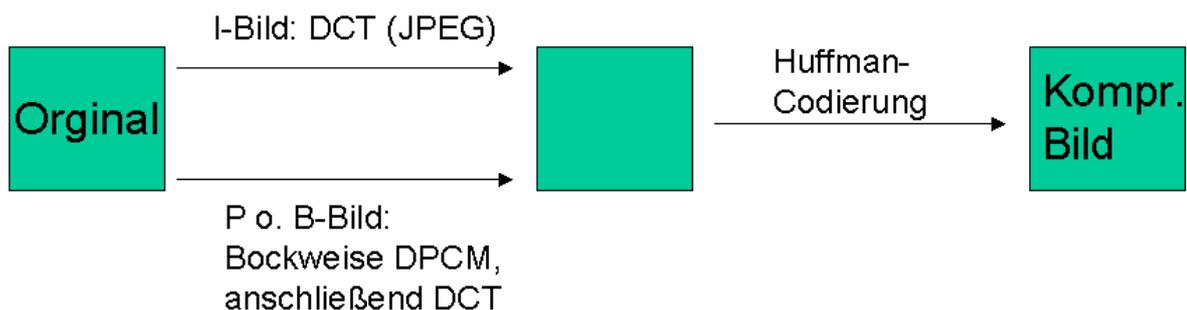
Der internationale Standard ISO 11172, der MPEG-1 Standard, ist in drei Einzelstandards unterteilt:

- Video: dieser Teil beschreibt die Videokodierungs- und Kompressionstechniken
- Audio: dieser Teil beschreibt die Audiokomprimierung mit Hilfe des sogenannten psychoakustischen Modells
- System: dieser Teil behandelt die Synchronisation und das Multiplexing von Video- und Audioströmen

Video- und Audioteil werden hier explizit vorgestellt.

Das MPEG 1 Verfahren ist eine Hybrid-Kodierung und bedient sich JPEG-komprimierter Einzelbilder, nach erfolgter Einzelbildkompression wird noch eine Interframe-Kompression durchgeführt (Abbildung 3-12).

Abbildung 3-12 Schritte bei der MPEG 1 Komprimierung



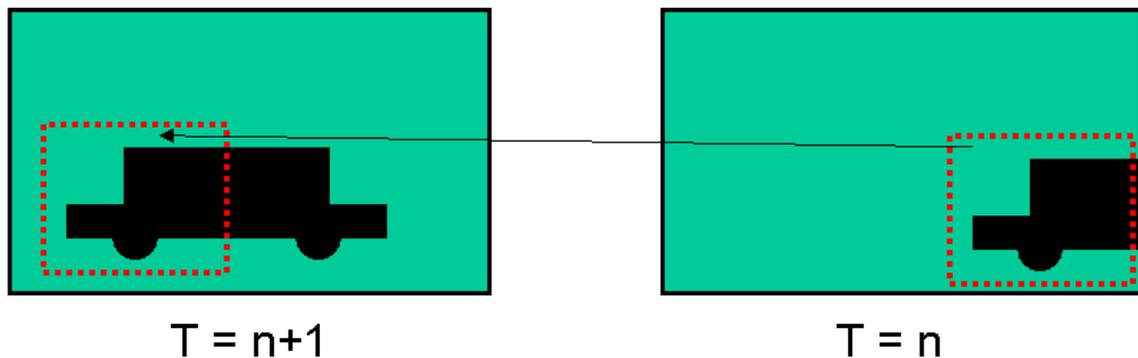
Wie in Abbildung 3-12 angedeutet, unterscheidet MPEG 1 vier verschiedene Arten von Bildern, der I-Bilder, P-Bilder, B-Bilder und der D-Bilder.

Die I-Bilder sind intraframe, mit Hilfe des JPEG-Verfahrens kodierte Bilder. Es werden eventuelle Redundanzen zwischen aufeinanderfolgenden Bildern nicht genutzt. Daraus ergibt sich nur eine begrenzte Kompressionsmöglichkeit.

Die I-Bilder dienen als Vorlage für B- und P-Bilder.

P-Bilder (prädiktiv) werden mit Hilfe von Informationen des vorhergehenden I- oder P-Bildes dekodiert und kodiert (Abbildung 3-13). Durch Ausnutzung zeitlicher und räumlicher Redundanzen können bei P-Bildern wesentlich bessere Kompressionsraten erzielt werden. Auf Grund dieser Eigenschaften kann der Zugriff auf ein P-Bild erst erfolgen, wenn das vorhergehende I- oder P-Bild dekodiert wurde.

Abbildung 3-13 Ausnutzen räumlicher und zeitlicher Redundanzen bei P-Bildern



Auch P-Bilder können als Vorlage zur Komprimierung anderer Bilder benutzt werden.

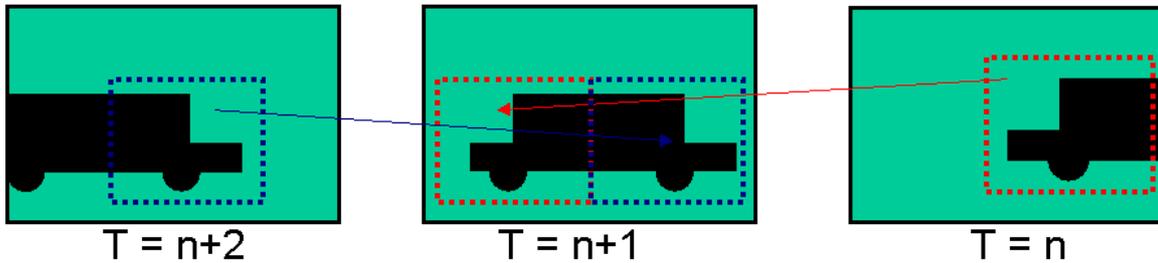
Eine weitere Steigerung der Kompressionsrate ist mit den B-Bildern (bidirektionale) möglich. Wie in Abbildung 3-14 zu sehen werden bei der Komprimierung dieser B-Bilder Bilddetails aus den vorhergehenden und den nachfolgenden I- oder P-Bildern verwandt (Bewegungskompensation).

Weil bei diesen B-Bildern Referenzen auf vorhergehende und nachfolgende Bilder Verwendung finden, ist die Reihenfolge beim Dekodieren (Kodierfolge) unterschiedlich zur logischen Reihenfolge (Anzeigefolge).

Durch die Benutzung von B-Bildern müssen neben einem erhöhten Datenaufkommen, durch die Speicherung der Referenzen aus den Referenzbildern, auch erhebliche Verzögerungen bei der Komprimierung bzw. Dekomprimierung in Kauf genommen werden.

B-Bilder werden nicht als Vorlagen für andere Bilder benutzt.

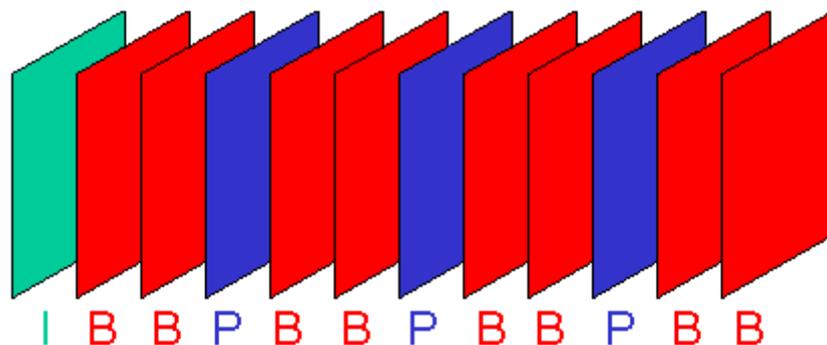
Abbildung 3-14 Ausnutzen zeitlicher und räumlicher Redundanzen der vorhergehenden und nachfolgenden I- oder P-Bilder bei B-Bildern



Die vierte Art von MPEG definierten Bildern ist das D-Bild. Dieses D-Bild darf nicht in einem Datenstrom mit I-, P- oder B-Bildern zusammen genutzt werden. Dies ermöglicht den schnellen Vor- und Rücklauf des komprimierten Videos.

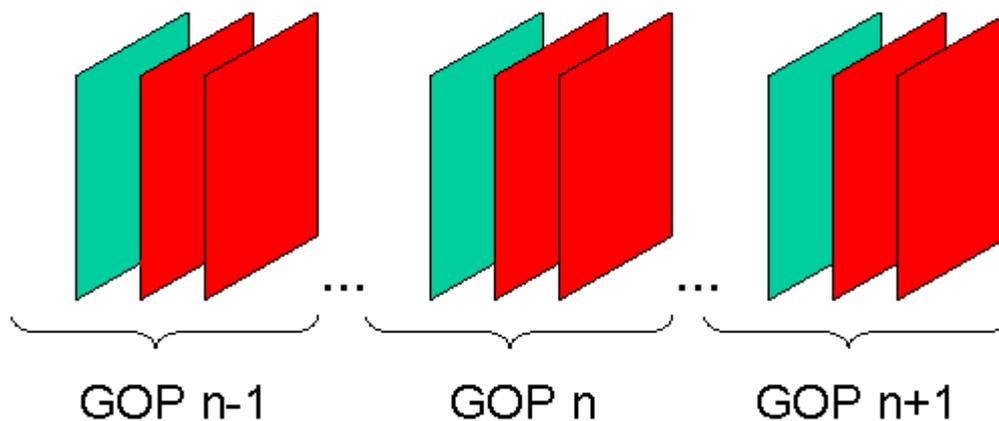
Da die Reihenfolge der Bilder in einem Videostream nicht festgelegt ist, wird eine Datenstruktur, die GOP (Group of Pictures) wie in Abbildung 3-15, verwendet, um die Abfolge zu speichern.

Abbildung 3-15 GOP einer MPEG komprimierten Bilderfolge



Die GOP garantiert eine feste Anzahl und Abfolge der Bilder. Das erste Bild in einer GOP ist ein I-Bild. Innerhalb einer Videobildfolge wiederholt sich die GOP ständig (Abbildung 3-16).

Abbildung 3-16 MPEG kodiertes Video als Abfolge von GOP's dargestellt



In Abbildung 3-15 ist eine komplette GOP der Art IBBPBBPBBPBB zu sehen, sie startet mit einem I-Bild, alle 3 Bilder ein P-Bild und dazwischen jeweils 2 B-Bilder. Diese Abfolge der unterschiedlichen Bilder ist ein Kompromiss zwischen Kompressionsrate und Qualität der Bilder.

Wird die Anzahl der Bilder in der GOP zu klein gewählt, kommen im Videostream zu viele I-Bilder vor und die Kompressionsrate verringert sich, es können zeitliche Redundanzen nicht optimal ausgenutzt werden. Ein besonderer Fall tritt ein, wenn die GOP aus nur einem I-Bild besteht. In diesem Fall wird jedes Bild einer Standbildkompression unterzogen und die Kompressionsrate verschlechtert sich bedeutend.

Andererseits, wird die Anzahl der Bilder in einer GOP zu groß gewählt, so kommen mehr B-Bilder vor und die Qualität des Videostreames verschlechtert sich im Vergleich zu einer optimal gewählten GOP. Weiterhin wird mehr Zeit zur Dekomprimierung der

B-Bilder als zur Dekomprimierung von P-oder I-Bildern benötigt. Es ist also in zeitkritischen Anwendungen auf eine GOP mit vielen B-Bildern zu verzichten.

Der Dekodierer erkennt anhand der GOP die Abfolge der Bilder und kann somit entscheiden, welches Bild als Referenz für ein anderes gebraucht wird.

MPEG 1 ermöglicht langsames Rückwärtsabspielen sowie wahlfreien Bildzugriff. Die maximale Bildgröße beträgt 768 x 576 Pixel bei maximal 25 Hz Bildwiederholfrequenz beziehungsweise 720 x 480 bei 30 Bildern pro Sekunde für die amerikanische Norm.

Es handelt sich bei MPEG nur um eine Definition des Datenstromes, nicht aber um eine Implementierung. Programmierern von Software, die das MPEG-Verfahren benutzen, ist also nur ein Rahmen und ein Format vorgegeben, nicht aber die absolute Größe der verschiedenen Parameter.

So kann z.B. der Suchbereich für Redundanzen zwischen den Bildern über die maximale Größe des Bewegungsvektors von der Software frei definiert werden. Die Vorgabe durch den ISO-Standard besagt lediglich, dass der Bewegungsvektor in einem Bereich von -64 bis +63 Pixeln liegen muss.

MPEG definiert nicht nur den Datenstrom für Video-, sondern auch für Audiodaten. Außerdem ist auch die Synchronisation der Bild- und Tondaten in diesem Standard definiert.

Um Inkompatibilitäten aufgrund der Anwendung verschiedener Parametereinstellungen zu vermeiden, wurde der Constrained Parameter Set (CPS) definiert. Ein MPEG 1-Dekoder erfüllt den CPS, wenn er Videos der Auflösungen 352x240 bei 30 fps²⁴ (USA) und 352x288 (Europa) bei 25 fps dekodieren kann. Bei diesen Formaten handelt es sich um das Standard-Interchange-Format (SIF).

²⁴ fps -frames per second, (Bilder pro Sekunde)

Der MPEG-1-Codec ermöglicht Kompressionsfaktoren bis ca. 125:1. Bis zu einem Kompressionsverhältnis von ca. 15:1 treten keine sichtbaren Qualitätsverluste auf, ab 50:1 nimmt die Bildqualität stark ab.

3.4.3 MPEG 2

Im November 1994 wurde der internationale Standard MPEG 2 (ISO 13818) mit dem übersetzten Titel „Generische Kodierung von Bewegtbildern und synchronisiertem Audio“ vorgelegt.

Bei MPEG 2 handelt es sich um eine Erweiterung des MPEG 1 Standards. Es verfolgt genau wie MPEG 1 das Prinzip der ungleichen Enkodier – und Dekodierverhältnisse, um die Kodierung zu optimieren.

MPEG 2 ist für Datenraten bis 100 Mbit/s ausgelegt und ermöglicht so auch hochauflösende Videos mit Audio in verschiedenen wählbaren Qualitäten wiederzugeben.

MPEG 2 wurde konzipiert um Video per Satellit, Breitbandkabel oder anderer Medien mit hoher Bandbreite zu übertragen. So ermöglicht MPEG 2 zum Beispiel Video on demand.

Seit seiner Einführung ist MPEG 2 der Standard, welcher digitale Fernsehstudioqualität im Bereich von 8 MBit/s liefert und der auch für hochauflösendes Fernsehen (HDTV) einsetzbar ist. Alle weltweit laufenden Projekte zur Einführung neuer digitaler Fernsehstrahlungsnormen setzen, trotz aller anderen Inkompatibilitäten wie Frequenzen und Auflösungen, auf MPEG-2. Auch das Digitalangebot des deutschen Pay-TV-Senders Premiere World arbeitet mit MPEG 2. Damit mehr Kanäle in die vorhandenen Frequenzbänder des Fernseekabels und die Satellitentransponder passen verwendet Premiere zum Teil aber recht niedrige Bitraten.

3.4.3.1 Die Profiles und Levels

MPEG 2 ist abwärtskompatibel und ist in verschiedene Level und Profils, wie in Tabelle 3-7 dargestellt, eingeteilt worden. Die Level und Profils schränken die zur Verfügung stehenden Parameter der Kodierung ein, um diese Einschränkungen

dann in den Kompressionsalgorithmen ausnutzen zu können. Gleichzeitig standardisieren die Profils und Levels auch die Kodierungsparameter.

Tabelle 3-7 Profile und Level Organisation in MPEG 2 Video nach [MIL95]

Max. Anzahl von Pixel/s		max. Bitrate	Levels?					
1920 x 1080 x 30	1920 x 1152 x 25	80 Mbit/s ^a	High	*	MP @HL	*	*	HP @HL
1440 x 1080 x 30	1400 x 1152 x 25	60 Mbit/s ^b	High -1440	*	MP @H14L	*	SSP @H14L	HP @14L
720 x 480 x 30	720 x 576 x 25	15 Mbit/s ^c	Main	SP @ML	MP @ML	SNRP @ML	*	HP @ML
352 x 240 x 30	352 x 288 x 25	4 Mbit/s	Low	*	MP @LL	SNRP @LL	*	*
			Profiles ?	Simple	Main	SNR Scalable	Spatial Scalable	High
				4:2:0, keine bidirektionale Prädiktion	4:2:0, keine Skalierbarkeit	Main + SNR – Skalierbarkeit	Main + Auflösungs – Skalierbarkeit	Gesamte Funktionalität (einschl. 4:2:2)

^a = bei HP: 100 Mbit/s

^b = bei HP: 80 Mbit/s

^c = bei HP: 20 Mbit/s

* = noch nicht definiert

Es existieren also 5 Profiles (High - HP, Spatial Scalable - SSP, SNR Scalable - SNRP, Main - MP, Simple - SP) und 4 Levels (High – HL, High 1440 - H14L, Main - ML, Low - LM). Der grau hervorgehobene Bereich der Tabelle kennzeichnet die möglichen Kombinationen von Profiles und Levels. Der Bereich unter den graugefüllten Feldern die Eigenschaften der einzelnen Profiles, der Bereich links davon, die Eigenschaften der einzelnen Levels.

Wie in dieser Tabelle zu sehen, lässt die Kombination von Profiles (Komplexität der Kompression) und Levels (sample Rate, frame Größe, codierte Bitrate) immer genügend Spielraum für die unterschiedlichsten Anwendungen. MPEG 2 soll laut Anwendungsprofil auf alle Eventualitäten in der digitalen Video-, Übertragungs- und Fernsehtechnik vorbereitet sein.

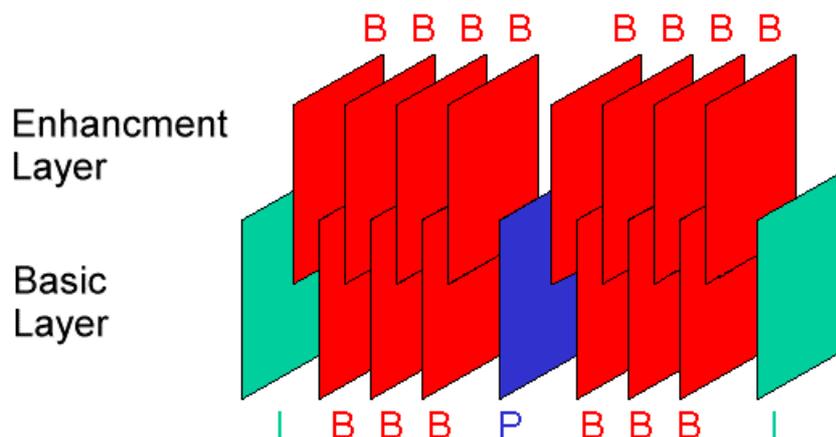
3.4.3.2 Die Scalability

Die Scalability von Video und Audio wurde in MPEG 2 verwirklicht. Sie liefert die Möglichkeit von unterschiedlicher Endnutzung bei ein und derselben feststehenden Kodierung. So kann der Endnutzer entscheiden, welche Teile einer Übertragung er konsumieren möchte.

Dadurch wird MPEG 2 weitgehend speichermedienunabhängig. Innerhalb der Einzelbild-Scalability lassen sich z.B. die beiden folgenden Modi realisieren (andere Kombinationen sind natürlich möglich):

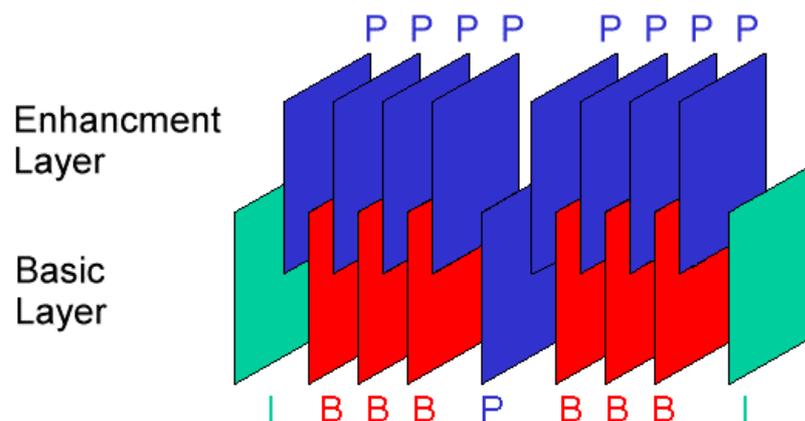
1. Die zeitliche Scalability (Abbildung 3-17), d.h. sollten die Ressourcen des konsumierenden Systems ausreichen, kann die Bildrate erhöht werden, indem in den normalen Ablauf des "Base Layers" zusätzliche B-Bilder des "Enhancement Layers" dekodiert werden.

Abbildung 3-17 Zeitliche-Scalability



2. Die Qualitäts-Scalability (Abbildung 3-18), d.h. anstelle der ungenaueren B-frames des "Base Layers" werden P-frames des "Enhancement Layers" dekodiert und angezeigt. Dies erfordert erheblich mehr Rechenzeit.

Abbildung 3-18 Qualitäts-Scalability



Zusätzlich sind noch die Kodierung von 3-dimensionalen Bildern (stereoscopic view) und die sichere oder geheime Übertragung von Teilströmen mit der Scalability-Funktion möglich.

Weiterhin wird noch die Makroblock-Scalability definiert. Dabei handelt es sich um einen zusätzlichen Eintrag in jedem Makroblock, der auf den nächsten "wichtigen", mit vielen Änderungen gegenüber dem letzten Bild, Makroblock verweist. Sollte also die Dekodierzeit nicht ausreichen, können "unwichtige" Makroblöcke übersprungen werden.

Schließlich gibt es noch die Pan-Scan-Scalability. Diese ermöglicht die Definition von Ausschnitten im aktuellen Bild. Diese Ausschnitte können zwar von Bild zu Bild unterschiedlich eingeteilt werden, die Hauptanwendung ist jedoch die Definition eines 4:3 Fernsehbildes innerhalb eines 16:9 HDTV-Bildes.

3.4.3.3 Die Sicherheit

MPEG 2 war die erste der Multimedia-Video-Definitionen und Standards, in der auch der Bereich Sicherheit (Vertraulichkeit und Integrität) berücksichtigt wurde. MPEG 1 hatte schon im Audiostrom einen rudimentären CRC-Check²⁵ eingeführt. MPEG 2 ist robust gegenüber Bitfehlern und Zellenverlusten. Ein speziell für MPEG 2 optimierter, schneller Algorithmus ist beschrieben, der ein Scrambling²⁶ des kompletten Datenstromes ermöglicht.

²⁵ CRC-Check -Prüfsummen-Check zum Testen der Datenintegrität

²⁶ Scrambling -Beim Scrambling wird ein Datenstrom nach einem festgelegten mathematischen Polynom verwürfelt. Die Verwürfelung wird im Prinzip durch einen Pseudozufallsgenerator mit der Länge 2^n-1 mit n-Rückkopplungsstufen ausgeführt. Diese Technik dient der Vermeidung gleichbleibender Signalmuster und sorgt für eine gleichmäßigere Verteilung der Signalenergie über den gesamten Frequenzbereich .
[Network]

4 Die Audiokompression

Die Qualität einer digitalisierten Audiosequenz hängt von der Auflösung der einzelnen Samples²⁷ und der Anzahl der Abtastungen (Quantisierungsauflösung genannt) in einem Zeitintervall ab.

Je besser die Quantisierungsauflösung gewählt wird, um so geringer wird der Rauschanteil der Sequenz gegenüber dem Nutzsignal sein und um so mehr Daten müssen pro Sample verarbeitet werden. Durch mehr Abtastung pro Zeitintervall nähert sich das digitale Signal dem analogen Original an.

²⁷ Samples -Als Sample wird ein einzelner bei der Digitalisierung anfallender Wert bezeichnet. Ein Sample ist die digitale Entsprechung eines analogen Signals zu einem bestimmten Zeitpunkt.

4.1 Die MPEG 1 Audiokompression [MPEG]

Der Audio-Teil des MPEG 1 Standards beschreibt Mechanismen und Algorithmen, mit denen die digitale Speicherung von Audiosignalen auf kostengünstigen Speichermedien auf der einen Seite und die digitale Übertragung von Audiosignalen auf Kanälen mit begrenzter Kapazität auf der anderen Seite ermöglicht wird.

Bei diesem Zielstreben steht jedoch die Erhaltung der Qualität in einem bestimmten Bereich im Vordergrund. Im Audibereich wird eine der Compact Disc (CD) nahekommende Qualität erreicht.

Bei einem zu erzielenden Datendurchsatz von 1,5 Mbit/s werden maximal 386 kbit/s für den Audibereich bereitgestellt.

Die Darstellung eines stereophonen Audiosignals im Studioformat erfordert eine Abtastfrequenz von 48 kHz und eine gleichförmige Quantisierung von 16 Bit pro Abtastwert. Daraus ergibt sich eine Datenrate von 768 kbit/s für ein Monosignal, als Produkt der Multiplikation der 48 kHz mit den 16 bit pro Abtastwert.

Daraus resultierend ergibt sich für ein Stereosignal eine Datenrate von 2×768 kbit/s, also ca. 1,5 Mbit/s. Als Vergleich dazu wird auf einer CD mit einer Abtastfrequenz von 44,1 kHz bei der gleichen Quantisierung von 16 bit pro Abtastwert gearbeitet, wodurch sich eine Datenrate von ca. 706 kbit/s (Mono) ergibt.

Im MPEG-Audio-Standard werden drei Abtastfrequenzen verwendet, 32, 44,1 und 48 kHz. Aber im Gegensatz zu den oben beschriebenen Fällen ergeben sich hier im Endeffekt Datenraten zwischen 32 kbit/s und 192 kbit/s für ein Monosignal. Für ein Stereosignal liegen sie zwischen 128 kbit/s und 384 kbit/s. Mit einer Datenrate unter 112 kbit/s (mono 56 kbit/s) können leider noch keine zufriedenstellenden Ergebnisse erzielt werden. Diese untere Grenze wird im Bereich der Video-CD (VCD) verwandt.

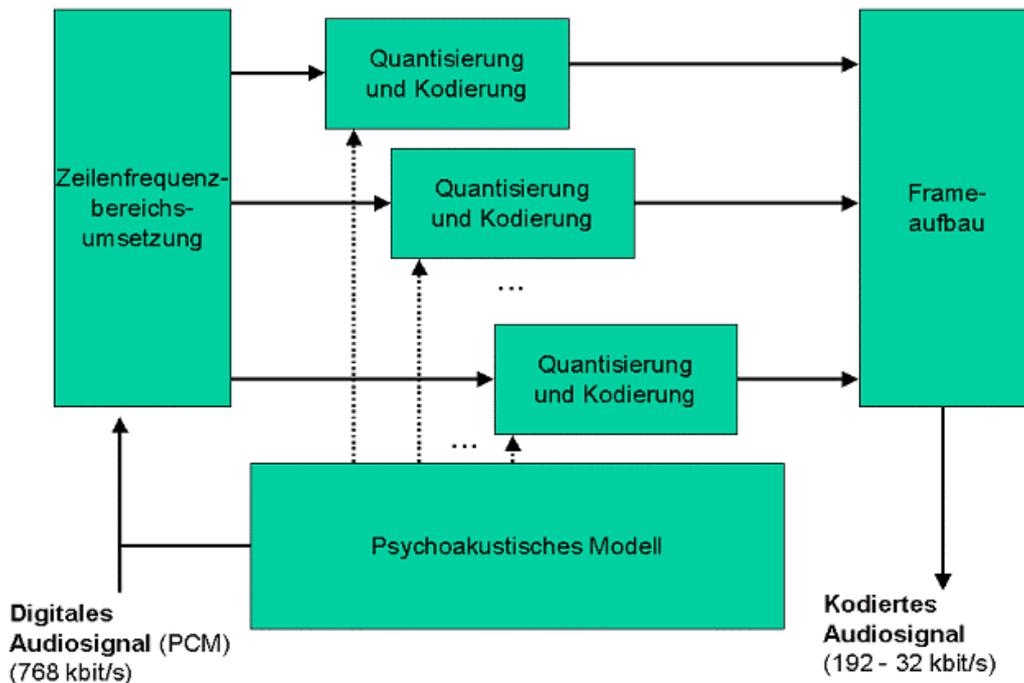
Das Ziel des Standards ist es, mit einer von 1,5 Mbit/s im Studioformat auf 384 kbit/s reduzierten Datenrate eine der Compact Disc ebenbürtige Qualität zu erreichen, wobei auch bei niedrigeren Datenraten wie 192 kbit/s bis hinunter zu 112 kbit/s noch

akzeptable Qualitäten erzielt werden sollen. Das menschliche Gehör ist im allgemeinen bei Störungen im Audiobereich empfindsamer als das Auge im visuellen Bereich, d.h. kurzzeitiges Rauschen und Knacken ist störender als z.B. Flimmern.

Innerhalb der Kodierung sind vier Modi zu unterscheiden. Single Channel Coding für Monosignale, Dual Channel Coding zur Kodierung von z.B. bilingualen Monosignalen (wie z.B. Zweikanalton im Bereich des TV), Stereo Coding zur Kodierung eines Stereosignals, bei dem die beiden Kanäle separat kodiert werden. Zusätzlich ist das Joint Stereo Coding zu nennen, das ebenso wie das Stereo Coding zur Kodierung eines Stereosignals benutzt wird. Bei diesem Verfahren wird die Datenredundanz und -irrelevanz zwischen den beiden Kanälen ausgenutzt und somit eine höhere Kompressionsrate erreicht.

Wie in Abbildung 4-1 zu sehen, wird das digitale Eingangssignal vom Zeitfrequenzbereichsumsetzer in 32 gleichförmige Spektralkomponenten (Frequenzgruppen, Teilbänder) zerlegt. Dieses Grundprinzip entspricht dem Vorgang im menschlichen Gehör (Psychoakustik). Der Vorgang wird als Zeit-Frequenzbereichs-Umsetzung bezeichnet. Die Spektralkomponenten werden dann in Abstimmung auf die Wahrnehmungseigenschaften des menschlichen Gehörs kodiert. Die schlussendliche Kodierung wird von einer der drei definierten Schichten (Layer) durchgeführt.

Abbildung 4-1 Generelles Blockschaltbild eines MPEG 1 Audiokodierers



Sowohl Quantisierung als auch Kodierung werden unter Einbeziehung einer Maskierungsschwelle realisiert. Diese Maskierungsschwelle wird vom psychoakustischen Modell für jede Komponente individuell mit Hilfe der Diskreten-Fourier-Transformation (DFT) berechnet und gibt die maximal erlaubte Quantisierungsfehlerleistung an, mit der noch kodiert werden darf, ohne dass eine Wahrnehmung dieses Fehlers durch das menschliche Gehör befürchtet werden muss.

Die oben erwähnten drei Layer des MPEG-Audio-Standards arbeiten alle nach dem beschriebenen Grundprinzip. Die Zerlegung des Eingangssignals und der Kodierung unterscheidet sich jedoch sowohl in der benötigten Rechenleistung als auch in der erreichten Kompressionsrate. Layer II ist im Verhältnis zum Layer I komplexer aber im Bezug auf die Kodierung auch effizienter.

Der Layer III besitzt die größte Komplexität und zugleich die höchste Effizienz. Er kann auf Computern (Pentium-PC, PowerMac oder Workstations) in Echtzeit in

Software dekodiert werden (mp3). Im Unterschied zu Layer I und II verwendet dieser Layer analog zur Umsetzung im Videoteil eine Modified Discrete Cosine Transformation (MDCT), die es erlaubt, die Anzahl der notwendigen Quantisierungswerte optimal zu errechnen. Laute, niederfrequente Signale überdecken dabei leise, hochfrequente Signale. Die letzteren müssen dann also nicht mit der vollen Bit-Tiefe, sprich Quantisierung, kodiert werden. Die Art der Generierung der Quantisierungswerte mit Hilfe der MDCT wird "Psychoakustisches Modell" genannt und wurde durch langwierige Hörtests von Musikern, Tontechnikern und anderen Spezialisten optimiert.

Wie im Videoteil wird zusätzlich eine Entropie- und Huffmann-Kodierung vorgenommen, die die Werte nach der Umsetzung durch die MDCT optimal komprimieren. Die CD-Qualität wird hierbei mit der selben Datenrate wie in Layer II erreicht (384 kbit/s). Seine Stärke zeigt die Kodierung durch den Layer III erst bei niedrigeren Datenraten, bei denen immer noch eine sehr gute Qualität erzielt werden kann.

Insgesamt erhält man bei der MPEG-Audiokomprimierung als sehr nützlichen Nebeneffekt die Rauschfreiheit, d.h. es sind keine zusätzlichen Verfahren zur Geräuschminderung, wie z.B. das Dolby-System, nötig.

Die einzelnen Layer sind vollständig als integrierte Schaltungen realisiert. Die Layer I Kodierung wurde bereits seit der Einführung der DCC-Recorder (Digital-Compact-Cassette) verwendet. Die Layer II Kodierung wird vom europäischen DAB (Digital Audio Broadcasting) System verwendet.

4.2 Die MPEG 2 Audiokompression

Der Audioteil von MPEG-2 ist wiederum genau wie bei MPEG 1 auf die Kodierung von Audiosignalen hoher Qualität (CD, Studio) zugeschnitten, weist aber eine viel höhere Flexibilität auf. Zusätzlich zu den festgelegten, bereits bei MPEG 1 benutzten Sampling-Frequenzen von 33, 44,1 oder 48 kHz, sind auch solche von 8, 11, 16, 22 oder 24 kHz möglich, dies insbesondere zur Kodierung von Signalen aus digitalen Quellen wie ISDN, G.722 oder Samples aus dem Studiobereich. Ein variables Bit/Sample-Verhältnis wird erlaubt.

Um die Rückwärtskompatibilität zu MPEG 1 und insbesondere die Unterstützung von HDTV und Dolby-Surround zu ermöglichen, kann man in MPEG 2 bis zu 5 Kanäle kodieren, plus einen Kanal mit tiefstfrequenten Anteilen bis zu 100 Hz für Spezialeffekte. Diese 5 Kanäle können in drei verschiedene Kanäle aufgeteilt werden. Links, mitte, rechts sind HDTV-Richtungskanäle, die anderen beiden Kanäle sind Raumkanäle für Dolby-Surround. Speziell über die drei HDTV-Kanäle lassen sich vielfältige Kombinationen zur multilingualen Übertragung wählen.

Durch die Vielfältigkeit in der Kodierung und der hohen Anzahl der Kanäle werden auch neue Möglichkeiten zur Kompression der Audiosignale erschlossen. Zu nennen sind:

- Intensity Stereo Coding (ISC)
- Phantom Coding of Center (PCC)
- Dynamic Transmission Channel Switching
- Dynamic Cross Talk
- Adaptive Multi-Channel Prediction

Eine erweiterte Funktionalität im Bereich Motion Compensation und Motion Estimation wurde definiert. Die Integrität aller Samples (bzw. Offset-Tabellen und Subbandinformationen) kann nun durch zugefügte Prüfsummen geschützt werden.

Von modernen DVD - Playern werden heutzutage verschiedene Tonformate unterstützt. (Tabelle 4-1) Jedes dieser Formate bietet unterschiedliche Möglichkeiten der Kompression und steht für verschiedene wirtschaftliche Interessen.

In der Tabelle 4-1 werden die gebräuchlichsten Tonformate für DVD - Player vorgestellt, die Hersteller, bzw. der Verkäufer, das benutzte Kodierungsverfahren und weitere Eigenschaften, wie Abtastrate und maximale Datenrate genannt.

Tabelle 4-1 Tonformate auf DVD Medien [c't1300]

	PCM	Dolby Digital	MPEG Audio	Dts	SDDS
Herkunft		Dolby	ISO Standard	Digital Theatre Sound	Sony
Verfahren	Lineare Puls-Code-Modulation	AC3 Datenreduktion	MPEG1 Layer 1 und MPEG2 Layer 2	APT	ATRAC
Abtast-Raten	48 und 96 kHz	48 kHz	48 kHz	48 kHz	48 kHz
Max. Datenrate	6,144 Mbit/s	448 kBit/s	640 kbit/s	768 kbit/s	1,280 Mbit/s
Max. Kanalzahl	8	5.1	7.1	7.1	5.1

5 Die Arbeitsumgebung

Entsprechend der Aufgabenstellung wurde eine Versuchsanordnung vorgegeben. Im folgenden Abschnitt wird die Arbeitsumgebung also die verwendete Hard- und Software der Arbeitsumgebung vorgestellt.

5.1 Die verwendete Systemhardware

Die Untersuchungen fanden auf einem Pentium Pro²⁸ System statt. Der Prozessor wurde mit 200 MHz getaktet, der Zugriff auf das Hauptspeicher-Subsystem erfolgte mit 66 MHz und der PCI Bus²⁹ wurde mit 33 MHz betrieben. Das System verfügte über 64 MB Hauptspeicher.

Als Festplatten kamen Ultra Wide SCSI Platten mit einer Gesamtkapazität von 10 GB zum Einsatz. Mit diesen Festplatten sind Dauertransferraten von bis zu 10 Mbyte/s möglich.

Als Grafikkarte wurde eine ELSA Victory 3D benutzt.

In dem Personal Computer kam ein Grafiksystem der Firma Optibase, das MPEG Forge³⁰ Grafiksystem und ein TBC (Time Base Corrector) der Firma FOR-A zum Einsatz.

Diese Komponenten werden im Kapitel 5.2 näher beschrieben.

²⁸ Pentium Pro -die Marke Pentium Pro ist eingetragenes Warenzeichen der Firma INTEL

²⁹ PCI Bus -peripheral component interconnect, Bus für Personalcomputer zum Einstecken von Erweiterungskarten

³⁰ FORGE -die Marke FORGE ist eingetragenes Warenzeichen der Firma Optibase Inc.

5.2 Die eingesetzten Audio-Video Karten

Wie in Abbildung 5-1 zu sehen, besteht das MPEG FORGE System von Optibase aus mehreren Karten.

Abbildung 5-1 MPEG FORGE von Optibase Inc.(entnommen aus [OPT])

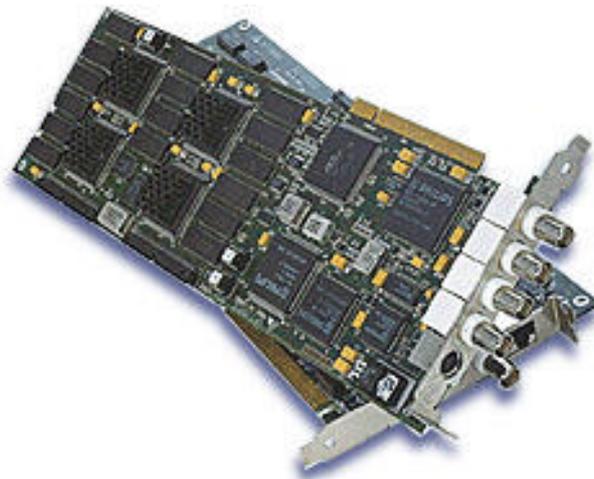
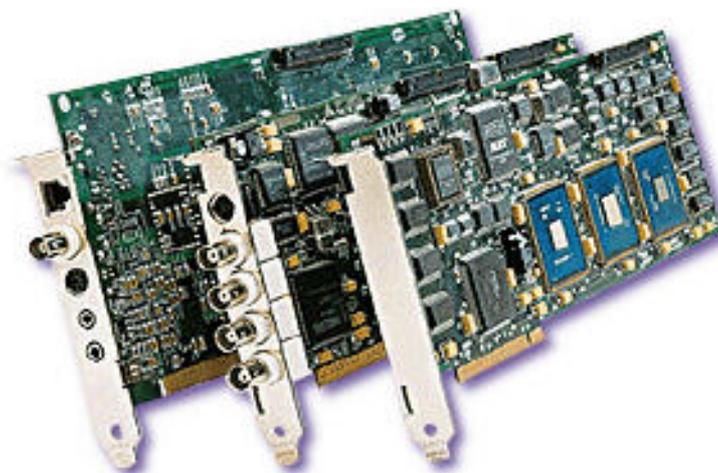


Abbildung 5-2 MPEG Fusion³¹ von Optibase Inc. (entnommen aus [OPT])



³¹ FUSION -die Marke FUSION ist eingetragenes Warenzeichen der Firma Optibase Inc.

Im einzelnen besteht das genutzte System aus drei Karten:

- der MPEG 6000,
- der AV 6000 und
- der VideoPlex³².

Die MPEG 6000 ist das Video-Input-Board welches auch MPEG 1 und Half D-1 kodiert.

Das AV 6000 Board bewerkstelligt das Kodieren der Audioströme und ist auch das Audio-Input-Board.

Soll gleichzeitig zum Kodieren der kodierte Audio-Videostream kontrolliert werden, so ist noch eine dritte Karte, die VideoPlex, notwendig. Diese Karte ist das sogenannte „playback“ Board, welches nicht unbedingt erforderlich wäre. In dieser Arbeit wurde, wie in Abbildung 5-3 zu sehen, eine VideoPlex-Karte verwendet, weil dies das Arbeiten wesentlich vereinfacht.

Wird als kodierter Audio-Videostream ein Full D-1 Stream benötigt, so muss ein MPEG FUSION Board der Firma Optibase eingesetzt werden (in Abbildung 5-2 zu sehen), denn das in dieser Arbeitsumgebung eingesetzte MPEG FORGE Board beherrscht den MPEG 2 Standard nur bis Half D-1.

Einen technischen Vergleich der zwei MPEG Encoder Boards bietet die Tabelle 5-1.

³² VideoPlex -die Marke VideoPlex ist eingetragenes Warenzeichen der Firma Optibase Inc.

Tabelle 5-1 Technical Specifications des MPEG FORGE Audio-Video Systems der Firma Optibase im Vergleich mit dem MPEG FUSION System vom selben Hersteller [OPT]

	MPEG FORGE	MPEG FUSION
Standard Compliance	ISO 11172 (MPEG 1) ISO 13818 (MPEG 2)	
Input/Output Signal		
Video Inputs	Digital Serial Component (BNC) Analog Component (BNC) S- Video (Mini- DIN) Composite (BNC)	
Audio Inputs	Digital AES / EBU (XLR) Balanced Analog Stereo (2xXLR) Unbalanced Analog Stereo (Mini-Jack)	
Output Streams		
NTSC Video	MPEG 1 QSIF 176x120 30fps ³³ MPEG 1 SIF 352x240 30fps MPEG 2 Half D-1 352x480 30fps	MPEG 1 QSIF 176x120 30fps MPEG 1 SIF 352x240 30fps MPEG 2 Half D-1 352x480 30fps MPEG 2 Full D-1 720x480 30fps
NTSC Film	MPEG 1 SIF 352x240 30fps MPEG 2 Half D-1 352x480 30fps	MPEG 1 SIF 352x240 30fps MPEG 2 Half D-1 352x480 30fps MPEG 2 Full D-1 720x480 30fps
PAL Video	MPEG 1 QSIF 176x120 25fps MPEG 1 SIF 352x240 25fps MPEG 2 Half D-1 352x480 25fps	MPEG 1 QSIF 176x120 25fps MPEG 1 SIF 352x240 25fps MPEG 2 Half D-1 352x480 25fps MPEG 2 Full D-1 720x480 25fps
Bit Rates	MPEG 1 QSIF 0,3 –10 Mbps ³⁴ MPEG 1 SIF 1-5 Mbps MPEG 2 Half D-1 1-5 Mbps	MPEG 1 QSIF 0,3 –10 Mbps MPEG 1 SIF 1-5 Mbps MPEG 2 Half D-1 1-5 Mbps MPEG 2 Full D-1 1-15 Mbps
File Formates	Elementary streams, System, Video-CD, Program, Transport streams	
Audio Formats	Mono, Dual Mono, Stereo, Joint Stereo	

Wie aus Tabelle 5-1 zu entnehmen, folgen beide Karten den ISO Standards für MPEG1 und MPEG2. Weiterhin unterscheiden sich beide Karten nicht in der Vielfalt der Anschlussmöglichkeiten und unterstützter File- und Audioformaten. Die einzigen Unterschiede finden sich in der Leistungsfähigkeit der Karten. Die MPEG FUSION

³³ fps - frames pro Sekunde

³⁴ Mbps - Megabit pro Sekunde

unterstützt im Gegensatz zur MPEG FORGE MPEG 2 Full D1 (Auflösungen bis zu 720 x 480 Pixel mit bis zu 30 fps)

5.3 Der Time Base Corrector (TBC)

Wenn aus mehreren Video Quellen ein Film geschnitten wird, oder z.B. ein Live-Bild aus mehreren Kameras zusammengeschnitten wird, ist das Problem des richtigen Timings der Video Quellen zu beachten.

Damit es während der Aufnahmen nicht zu Synchronisationsproblemen, zu „time base errors“ kommt, muss ein Gerät im gesamten Aufnahmeszenario den Synchronisationstakt, den „burst“, vorgeben. Normalerweise wird diese Aufgabe von einem zentralen Punkt, einem Mischpult oder einem Übertragungswagen, übernommen.

In unserem Fall sind die 2 Videokarten, die MPEG-Forge und die VideoPlex, und eine Audiokarte miteinander zu synchronisieren. Diese Aufgabe übernimmt der TBC.

Der TBC schickt allen angeschlossenen Karten einen Burst-Impuls. Auf diesen Impuls loggen sich die Karten ein und laufen somit synchron.

Der Videorecorder (VCR) wird indirekt synchronisiert. Wie in Abbildung 5-3 zu sehen, leitet dieser sein Videosignal (rot) über den I/O-Adapter des TBC und wird somit auch synchronisiert.

Der Time Base Corrector ist eine PC-Steckkarte und findet im ISA-Bus platz.

5.4 Die peripheren Geräte

In diesem Kapitel werden die peripheren Geräte des Schnittplatzes, Videorecorder und Monitor, beschrieben.

5.4.1 Der Videorecorder (VCR) [JVC1]

Als Videorecorder wurde das Model BR-S522E der Firma JVC³⁵ benutzt.

Einige Features der VCR's seien hier stichpunktartig genannt:

- Hochwertige SVHS Bildqualität durch einen digitalen Dropout-Kompensator
- Einfache Systemerweiterung durch Steckkartensystem
- Fernsteuerungseingang im RS 422A Protokoll
- Fernsteuerungseingang im RS 232C Protokoll (für Computersteuerung)
- Hochwertiger Schnitt-Feeder
- 32 -fache Suchgeschwindigkeit
- HiFi Stereo Audio (mit einem dynamischen Bereich >87 dB, zuschaltbares Dolby B Rauschunterdrückung³⁶)

5.4.2 Der Monitor [JVC2]

Der Monitor, welcher zum Darstellen der Videos benutzt wurde, ist ein TM-1500PS Farbmonitor der Firma JVC.

³⁵ JVC -ist ein eingetragenes Warenzeichen der Firma Victor Company of Japan, Limited

³⁶ Dolby B -Rauschunterdrückung unter Lizenz der Dolby Laboratories Licensing Corporation. DOLBY und das Doppel-D Symbol sind Warenzeichen der Dolby Laboratories Licensing Corporation.

Dieser 15“ Monitor zeichnet sich durch seine reichhaltigen Ausstattung (FBAS³⁷- und Y/C-Eingang, Underscan-, Impulskreuzschalter) aus. Er ist für professionelle Anwendungen geeignet.

³⁷FBAS Signal -Farb Bild Austast Synchron Signal (auch Composit), gemultiplextes Analogsignal, welches aus einem Synchron-, einem Burst- und einem Bildinhaltanteil (Chrominanz und Luminanz) besteht.

angeschlossen. Zur Ausgabe der Audiodaten wurden handelsübliche PC Speaker, an die AV 6000 angeschlossen, genutzt.

5.5.1 Die Aufnahme von Audio- Videostreams

Wie in Abbildung 5-3 zu sehen, werden bei der Aufnahme der Audio- Video- Streams die Videodaten unkomprimiert von VCR, durch den TBC synchronisiert, zur MPEG 6000 mittels eines 75 Ohm BNC Kabels (FBAS Signal) übertragen. Die Audiodaten werden ebenfalls unkomprimiert zur AV6000 per XLR-Kabel (Balanced Audio) übertragen.

Diese Daten werden jeweils durch die Karten komprimiert und auf die Harddisk geschrieben.

Gleichzeitig können durch das Einsetzen der VideoPlex-Karte die Daten von der Harddisk gelesen und ohne merkliche Verzögerung am Monitor angezeigt werden. Diese Übertragung geschieht per 75 Ohm Y/C Kabel .

Da gleichzeitig der VCR direkt per 75 Ohm BNC Kabels (FBAS Signal) mit dem Monitor verbunden ist und der Monitor 2 Eingänge besitzt, kann während der Aufnahme zwischen dem Signal des VCR's und dem der VideoPlex – Karte umgeschaltet werden. So kann die Qualität der Aufnahme schon während des Aufnehmens kontrolliert werden.

5.5.2 Die Wiedergabe der Audio- Videostreams

Beim Wiedergeben der komprimierten Audio-Video-Ströme werden die Daten durch die VideoPlex von der Harddisk geholt, dekomprimiert und über den Monitor angezeigt.

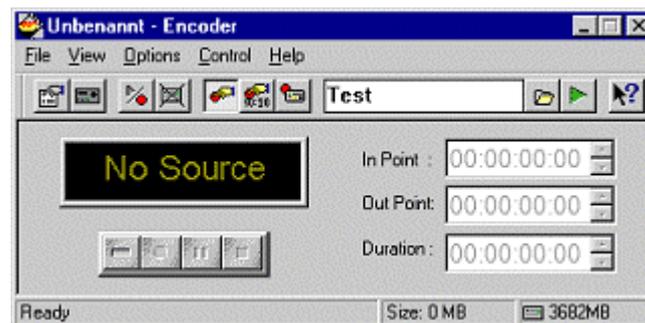
5.6 Die Software

Als Betriebssystem kam Windows³⁸ NT4.0 mit Servicepack 3 zum Einsatz.

Die Firma Optibase liefert ein Softwarepaket , bestehend aus Encoder, Player und Schnittsoftware, zu der MPEG Forge. Weiterhin steht eine Fernsteuersoftware auf dem Schnittplatz zur Verfügung.

5.6.1 Der Encoder

Abbildung 5-4 Bedienungsansicht der Encoder Software

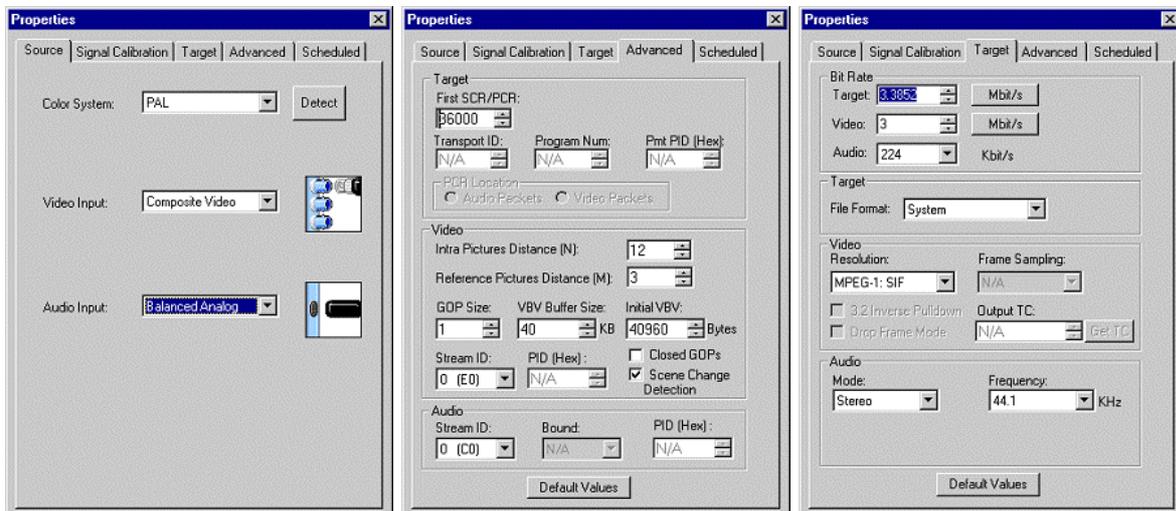


Wie in Abbildung 5-4 zu sehen, ist die Encode Software zur MPEG FORGE wenig komfortabel ausgestattet. Dennoch bietet die Software alles was man zum Aufnehmen der Videos braucht. Im rechten Bereich der GUI³⁹ findet sich die Steuerung der VCR Fernbedienung, wird in 5.6.3 näher erläutert, im oberen Bereich sind die Bedienelemente zur Einstellung des Encoders und im Bereich unten links die klassischen VCR- Bedienelemente zu finden. Mittels diesen klassischen VCR- Bedienelementen lassen sich die Aufnahme starten, anhalten und beenden.

³⁸ Windows -die Marke Windows ist eingetragenes Warenzeichen der Firma Microsoft Inc.

³⁹ GUI -Graphical User Interface (grafische Bedienoberfläche)

Abbildung 5-5 Einstellungen des Encoders, Registerkarte eins, drei und vier



Wie in Abbildung 5-5 zu sehen ist, lassen sich die Eingangs- sowie auch die Ausgangsparameter des Audio-Videostreams innerhalb des Leistungsumfangs der MPEG Forge variieren.

Auf der ersten Registerkarte können die Eingangsparameter eingestellt werden, wobei das Farbsystem PAL oder NTSC bei anliegendem Videostream am Karteneingang vom Encoder selbst erkannt werden kann.

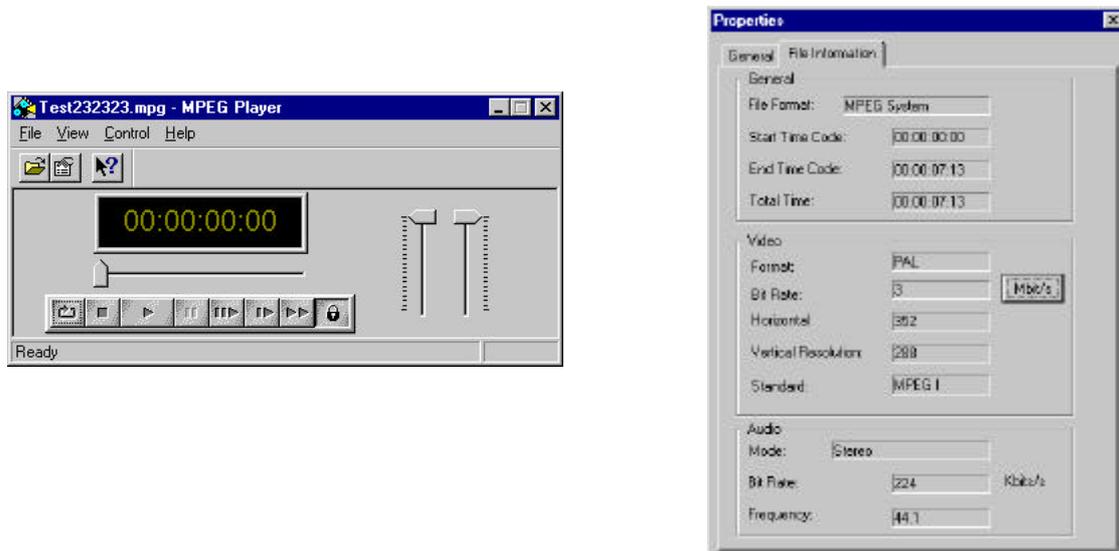
Die nächste Registerkarte ermöglicht das Einstellen der Kompressionsparameter wie Kompressionsrate, Art der Kompression (MPEG 1-SIF, MPEG 2-HalfSIF) und Art und Auflösung des Audiosignals. Die Möglichkeiten der Einstellung der genannten Parameter ergibt sich aus Tabelle 5-1.

In der vierten Registerkarte können die Einstellungen zur Videokompression noch genauer justiert werden. Es werden u.a. die Art und Beschaffenheit der GOP (I und P Pictures Distance) und das „Scene change detection“ eingestellt.

Die Registerkarten 2 und 5 sind für Betrachtungen innerhalb dieser Arbeit nicht von Bedeutung.

5.6.2 Der Player

Abbildung 5-6 MPEG Player mit File Information Registerkarte



Wie in Abbildung 5-6 zu sehen, wird der MPEG Player wie ein Videorecorder bedient. Durch das genaue Zählwerk des VCR's wird es ermöglicht, den Videostream Bild für Bild anzuschauen (für spätere Qualitätsanalysen).

Die File Information Registerkarte gibt statistische Auskünfte zum aktuellen Videostream.

5.6.3 Die Fernsteuersoftware

Abbildung 5-7 zeigt das GUI der Fernsteuerung mit deren Hilfe das VCR Bild genau positioniert werden kann. Es können Anfangspunkt (Mark In) und Endpunkt (Mark Out) zum Steuern einer Aufnahme festgelegt werden. Dies veranlasst VCR und Encoder Software, welche die Einstellungen der Fernsteuerung übernehmen kann, genau diejenigen Frames aufzuzeichnen, welche sich zwischen den Zeitpunkten Mark In und Mark Out befinden. Diese speziellen Auswertemöglichkeiten durch die

Encoder Software machen überhaupt erst das vergleichende Aufnehmen der Videostreams möglich.

Abbildung 5-7 Die Fernsteuersoftware



5.6.4 Der Mainactor Video Editor

Mit Hilfe des Mainactor Video Editor [MAI2.05] konnten die aufgenommenen Audio - Video - Streams genauer analysiert werden.

Es wurden die Anzahl der Bilder im Videostream, die genaue Größe des Videostreams und die durchschnittliche Größe der Bilder festgestellt.

Weiter wurden mit Hilfe des Mainactors einzelne Bilder des Videostreams extrahiert, um die Bilder miteinander vergleichen zu können.

5.6.5 Das Optibase RTE SDK⁴⁰

Das von der Firma Optibase mit freundlicher Genehmigung zur Verfügung gestellte RTE SDK V2.2 Beta Build 5 ist ein einfach zu handhabendes MPEG Encoder Tool zur Entwicklung von native 32 bit Tools für Windows 98 und Windows NT4.0. Es arbeitet mit einer Reihe von Optibase's Hardware Encodern zusammen, darunter auch MPEG Forge und MPEG Fusion.

Das Optibase RTE SDK und Microsoft's C++ ermöglichten es, ein Tool (Time Measure Tool) zur Messung der Dauer des Codierens zu entwickeln. Das Time Measure Tool wird in Kapitel 5.6.8 genauer erklärt.

5.6.6 Der Borland C++ Builder

Ein weiteres Tool, welches im Rahmen dieser Arbeit zum Einsatz gekommen ist, ist das PSNR⁴¹ Tool [FIE99] und wird in Kapitel 5.6.9 näher beschrieben. Dieses Tool wurde mit Hilfe des Borland C++ Builder V3.0 entwickelt .

5.6.7 Das Microsoft Visual Studio C++

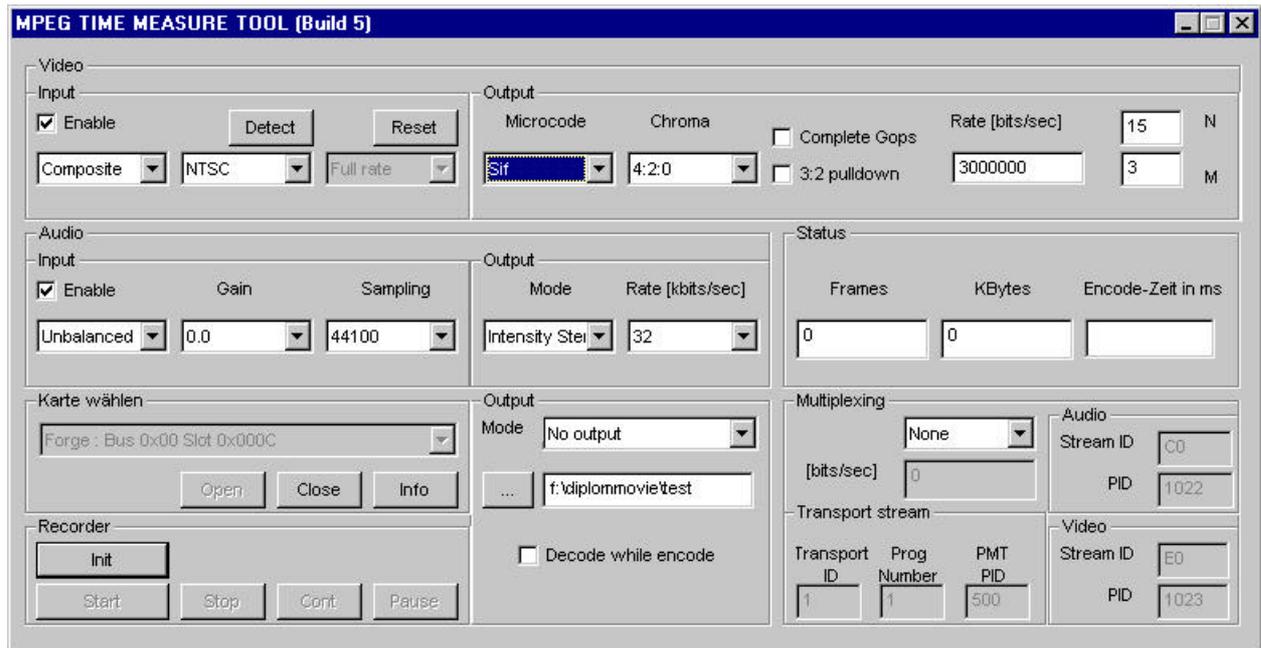
Zum Entwickeln des Time Measure Tool wurde Visual Studio C++ [MSD97] benutzt.

⁴⁰ RTE SDK -Real Time Encoder Software Development Kit

⁴¹ PSNR -Peak Signal to Noise Ratio, Rauschabstand zwischen geometrisch gleichgelegenen Pixel zweier zu vergleichender Bilder.

5.6.8 Das Time Measure Tool (TMT)

Abbildung 5-8 GUI des TMT Time Measure Tools



Das TMT (Abbildung 5-8) wurde mit Hilfe des Microsoft C++ Visual Studios und dem RTE SDK im Rahmen der vorliegenden Arbeit entwickelt. Es kann die Encodezeit und die Anzahl der codierten Bilder zählen, was eine Berechnung der durchschnittlichen Encode -Zeit pro Bild ermöglicht.

Im „Input“ Bereich (links) der GUI des TMT werden die Eigenschaften des Eingangs-Streams eingestellt. Im „Output“ Bereich (rechts), werden die Ausgangsparameter eingestellt.

Abbildung 5-9 Einteilung der GUI in funktionelle Bereiche

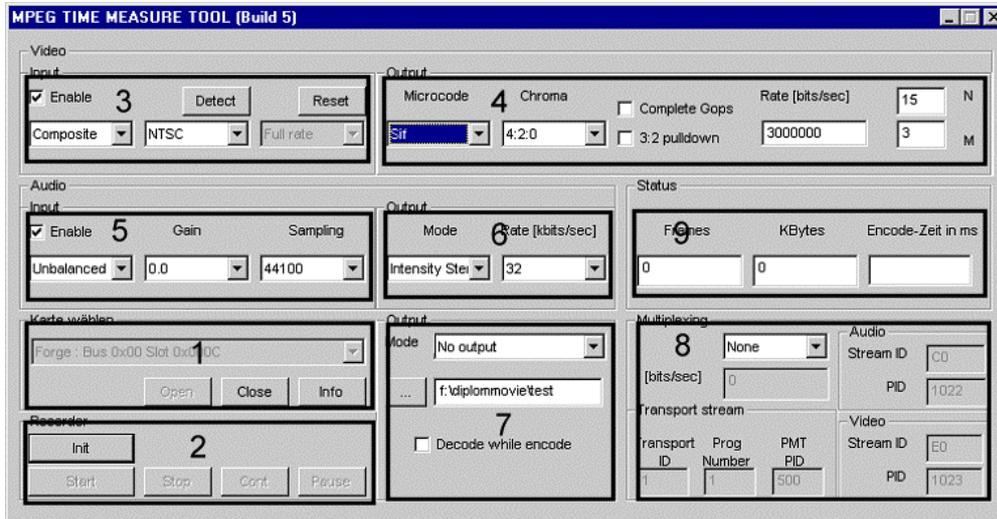
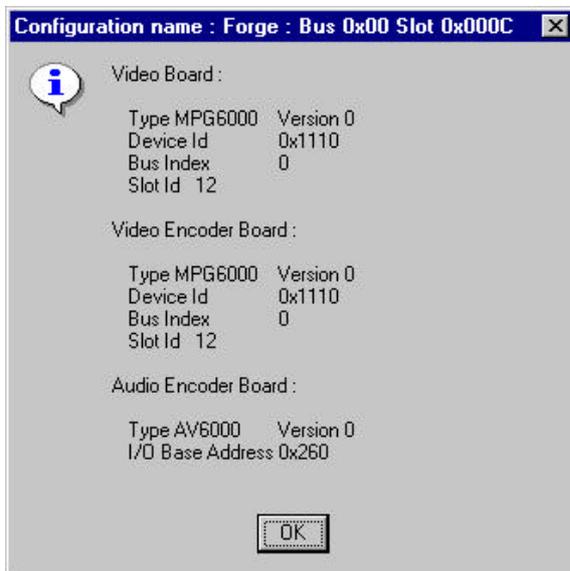


Abbildung 5-10 Infowindow des TMT



Die nachfolgende Tabelle 5-2 stellt die einzelnen Bereiche welche in Abbildung 5-9 gekennzeichnet wurden vor und erklärt die Funktion der Elemente in den Bereichen.

Tabelle 5-2 Funktionsbeschreibung der Bereiche im GUI des TMT

Bereich	Element	Beschreibung
1 "Encoder Board wählen"	"open"	wählt eine von eventuell mehreren eingebauten MPEG Encoder Karten
	"close"	wählt die mit dem "open" Button gewählte Karte wieder ab
	"info"	zeigt ein Infofenster mit detaillierten Informationen zur gewählten Karte (Abbildung 5-10)
2 "Videorecorder Funktionen"	"init"	initialisieren des Recorders (Übernahme der Einstellungen in 3-8)
	"start"	starten der Aufnahme
	"stopp"	stoppen der Aufnahme
	"continue"	weiter nach "pause"
	"pause"	Aufnahme kurzfristig unterbrechen
3 "Video Input"	"enable"	sollen Video-Input Einstellungen automatisch oder per Hand gemacht werden
	Eingang	Wählt die Eingangsquelle aus: Composite ⁴² , SVideo ⁴³ , Betacam ⁴⁴ oder Digital Video ⁴⁵
	"detect"	versucht das Farbsystem (PAL/SECAM) automatisch einzustellen, kann auch manuell eingestellt werden
	"reset"	setzt alle unter 3 gemachten Einstellungen zurück
4 "Video Output"	"Microcode"	stellt die Auflösung des Videos ein: SIF, HalfSIF
	"Complete GOPs"	Wenn Complete GOPs gewählt wurde, so beziehen sich B und P Frames jeder GOP nur auf I und P Frames der selben GOP.
	"Rate"	stellt die Bitrate ein (bits/s)
	"M"	stellt den Abstand der I Frames ein
	"N"	stellt den Abstand der P und B Frames ein
5 "Audio Input"	Eingang	stellt die Eingangsquelle ein: Balanced ⁴⁶ , Unbalanced ⁴⁷ , Digital Audio ⁴⁸

⁴² Composite -Videosignal, welches alle Komponenten des Signals in einem Stream kombiniert

⁴³ SVideo -von Videorecordern unterstütztes Signal mit separaten Luminance und Crominance Ein- und Ausgängen

⁴⁴ Betacam -(Component Analog) analoges Videosignal mit separierten Y, R-Y und B-Y Farbsignalen

⁴⁵ Digital Video -(Digital Component) digitales Videosignal mit separierten Y, R-Y und B-Y Farbsignalen

⁴⁶ Balanced -professionelles Analogsignal, Signale werden über 3 seperate Wege (Ground, Plus, Minus) übertragen, um das Grundrauschen zu vermindern

⁴⁷ Unbalanced -Analogsignal, welches die Übertragung mittels einem einfachen Spannungsunterschieds zur Masse durchführt

⁴⁸ Digital Audio -ein digitales, von der ANSI (American National Standard Group) genormtes Audiosignal

	“gain“	zum Kalibrieren des Eingangssignals um Lautstärkeunterschiede auszugleichen (-12 bis +10,5)
	“Sampling“	stellt die Sampling- Frequenz ein (32000, 44100,48000)
6 “Audio Output“	“Mode“	zum Einstellen der verschiedener Audio- Streams: Stereo, Mono, Dual Mono ⁴⁹ , Intensity Stereo ⁵⁰
	“Rate“	stellt die Audiobitrate ein, Tabelle 5-3 bietet einen Überblick über alle Möglichen Kombinationen
7 “File Output“	“Mode“	Stellt ein, ob ein File auf Platte geschrieben werden soll oder nicht. Wenn ja, dann kann im darunter liegenden Button den Pfad wählen.
	“Decode while Encode“	stellt ein, ob die Videoplexkarte während der Aufnahme als Kontrolle benutzt werden soll
8 “Multiplexing“	Format	stellt das Ausgabeformat ein: MPEG System ⁵¹ , Program ⁵² , Transport ⁵³ , Video-only ⁵⁴ , Audio-only ⁵⁵ , Video+Audio ⁵⁶ , Video CD ⁵⁷ , CD-i ⁵⁸
	“Video Stream ID“	eine ID dem Video- Stream zuweisen, ein MPEG File kann bis zu 16 Video- Streams beinhalten
	“Video PID“	stellt die ID für die Video Pakete innerhalb des Tranport- Streams ein
	“Audio Stream ID“	eine ID dem Audio- Stream zuweisen, ein MPEG File kann bis zu 32 Audio- Streams beinhalten
	“Audio PID“	stellt die ID für die Audiopakete innerhalb des Tranport- Streams ein
	“Transport ID“	ein 16 Bit Wert welcher den eigentlichen Transport-Stream identifiziert

⁴⁹ Dual Mono -zwei Mono Streams werden in einem Mono Stream encodet (Mehrsprachig)

⁵⁰ Intensity Stereo -eine Methode bei welcher hohe Subbänder (links+ rechts) miteinander kombiniert werden (Ausnutzug der Redundanzen des rechten und linken Kanals)

⁵¹ MPEG System -MPEG1 kompatibles gemultiplextes File, *.mpg Extension

⁵² Program -MPEG2 kompatibles File, welches von vielen Programmen unterstützt wird. *.mpg Extension

⁵³ Transport -MPEG2 kompatibles File, welches für Übertragungen entwickelt wurde. (Error correction, encryption)

⁵⁴ Video-only -File welches nur Video Daten enthält. *.mpv Extension

⁵⁵ Audio-only -File welches nur Audio Daten enthält. *.mpa Extension

⁵⁶ Video+Audio -Es werden 2 Files auf Platte geschrieben, ein Video-only und ein Audio-only File

⁵⁷ Video CD -MPEG1 kompatibles gemultiplextes Video CD Whithe Book File, *.mpg Extension

⁵⁸ CD-i -MPEG1 kompatibles gemultiplextes CD-i File, *.mpg Extension, CD-i ist ein eingetragenes Warenzeichen von Royal Philips Electronics, Niederlande

	“Program ID”	stellt die ID für jedes Programm innerhalb des Transport- Streams ein
	“PMT PID”	stellt die ID für die PMT ⁵⁹ Pakete ein
9 “Kontrolle“	“Frames“	zeigt die Anzahl der erzeugten Frames
	“KBytes“	zeigt die Menge der auf die Platte geschriebenen Daten
	“Encode Zeit“	zeigt die seit dem Start der Aufnahme vergangene Zeit im ms an

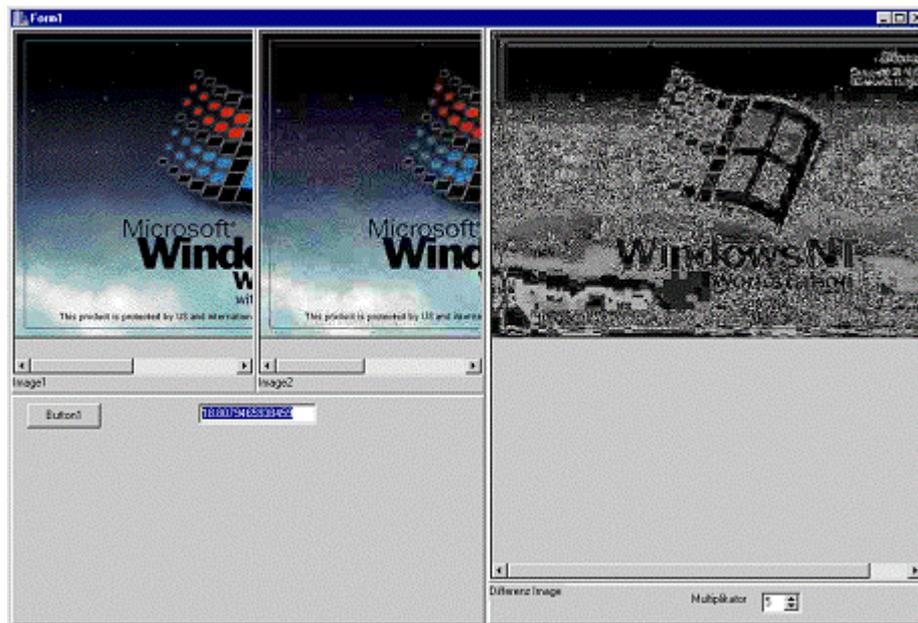
Tabelle 5-3 mögliche Audiobitraten

Bit Rate (Kbit/s)	Mögliche Audio Modes
32	Mono
48	Mono
56	Mono
64	alle Modes
80	Mono
96	alle Modes
112	alle Modes
128	alle Modes
160	alle Modes
192	alle Modes
224	Stereo, Intensity Stereo, Dual Mono
256	Stereo, Intensity Stereo, Dual Mono
320	Stereo, Intensity Stereo, Dual Mono
384	Stereo, Intensity Stereo, Dual Mono

⁵⁹ PMT -Program Map Tables tragen wichtige Systeminformationen welche für zufälligen Zugriff auf den Transport- Stream und für das Recover bei verloren gegangenen Daten sind.

5.6.9 Das PSNR-Tool

Abbildung 5-11 Das GUI des PSNR-Tools



Zur quantitativen Bewertung der Bildqualität wird die PSNR (Peak Signal to Noise Ratio), siehe Gleichung 5-1, welche den Rauschabstand zwischen zwei geometrisch gleich gelegenen Pixel zweier zu vergleichender Bilder darstellt, zugrunde gelegt. Das PSNR Tool erzeugt diese PSNR aus den zwei zu vergleichenden Bildern.

Es werden diese zwei zu vergleichenden Bilder im linken Bereich der GUI und das Differenzbild, im rechten Bereich, angezeigt. Das Differenzbild zeigt die Unterschiede zwischen dem ersten und dem zweiten Bild in Form von Graustufen, wobei je heller ein Punkt dargestellt wird, desto stärker weicht der entsprechende Punkt im zweiten Bild von dem im ersten ab.

Gleichung 5-1 Peak Signal to Noise Ratio, PSNR

$$\text{PSNR} := \frac{\sum_{i=1}^M \sum_{j=1}^N (F_{i,j}^{\text{orig}})^2}{\sum_{i=1}^M \sum_{j=1}^N (F_{i,j}^{\text{orig}} - F_{i,j}^{\text{dec}})^2}$$

$i, j, M, N \in \mathbb{I}$

In Gleichung 5-1 sind F_{ij}^{orig} die Farbwerte der Pixel des Original Bildes und F_{ij}^{dec} die Farbwerte der Pixel des decodierten Bildes. Wie aus der Gleichung hervor geht, steigt der Rauschabstand mit der pixelweisen Übereinstimmung der zwei Bilder.

Dieser Rauschabstand wird als objektives Maß der Übereinstimmung zweier Bilder in die Auswertungen einfließen.

Das PSNR Tool wurde in [FIE99] programmiert und für diese Arbeit zur Verfügung gestellt.

6 Qualitäts- und Kompressionszeitmessungen

6.1 Die Messung der Kompressionszeit

6.1.1 Die Vorstellung der Testmethode

Zur Einschätzung der Leistung der Kompressionskarte wurde die Kompressionszeit gemessen.

Zu diesem Zwecke wurde das in Abschnitt 5.6.8 vorgestellte TMT Tool entwickelt. Es misst die Zeitdauer vom Start der Aufnahme bis zum Bereitstellen der Bilder durch die Videokarte.

Es wurden 18 Videofilme mit unterschiedlicher Kompressionsrate, von 300.000 bis 10.000.000 bit/s und unterschiedlichen GOP's aufgenommen (siehe Tabelle 6-1, Bandbreite in bits/s, I-Picture Distance, P-Picture Distance). Es wurde die Dauer der Aufnahme und die Anzahl der aufgenommenen Frames gemessen (Tabelle 6-1, Frames , Encode Zeit in ms).

Für den Parameter GOP wurden die gebräuchlichsten GOP- Strukturen wie 12-1, 12-3, 30-1 und 30-3 (jeweils I- Picture - P- Picture - Abstand) verwand. Alle Messungen wurden mehrfach wiederholt und führten immer zu gleichen Ergebnissen.

Die Anzahl der Frames pro Sekunde ist der Quotient aus der Anzahl der Frames (Tabelle 6-1, Frames/s) und der Dauer der Aufnahme, Gleichung 6-1.

Gleichung 6-1 Berechnung der Anzahl der erzeugten Frames pro Zeiteinheit

$$\frac{\text{Frames}}{\text{sek}} = \frac{\text{Anzahl der erzeugten Frames}}{\text{Dauer der Aufnahme (in sek)}}$$

6.1.2 Die Ergebnisse

Tabelle 6-1 Messungen der Kompressionszeit

Filename	Bandbreite in bit/s	I-Picture Distance	P-Picture Distance	Frames	Kbytes	Encode Zeit in ms	Frames/s
Test7	300000	12	1	252	4371	10080	25
Test13	400000	12	1	261	1011	10440	25
Test12	500000	12	1	253	1015	10120	25
Test14	400000	12	3	253	920	10120	25
Test4	500000	12	3	258	1149	10320	25
Test3	1000000	12	3	252	1679	10080	25
Test2	2000000	12	3	252	3030	10080	25
Test1	3000000	12	3	252	4389	10080	25
Test5	5000000	12	3	258	7271	10320	25
Test6	10000000	12	3	312	17101	12480	25
Test16	400000	30	1	254	934	10160	25
Test11	500000	30	1	252	985	10080	25
Test8	3000000	30	1	252	4371	10080	25
Test17	10000000	30	1	312	17101	12480	25
Test15	400000	30	3	252	855	10080	25
Test10	500000	30	3	252	989	10080	25
Test9	3000000	30	3	252	4391	10080	25
Test18	10000000	30	3	312	16956	12480	25

Alle Videos, welche in der Tabelle 6-1 aufgeführt sind, wurden mit den in Tabelle 6-2 genannten Parametern aufgenommen. Die Reihenfolge der Filme in Tabelle 6-1 ergibt sich aus der Sortierung nach I-Picture und P-Picture Distance.

Tabelle 6-2 Video Parameter

Audioparameter:	
Samplerate	44,1 KHz
Audio Mode	Stereo
Ausgangs Bitrate	224 kbit/sek
Videoparameter:	
Eingangssignal	Composite PAL
Kompressionsverfahren	MPEG 1-SIF
Complete GOP's	Off

6.1.3 Diskussion der Ergebnisse

Wie aus Tabelle 6-1 zu entnehmen ist, verhält sich die Karte immer entsprechend dem PAL Standard, die Karte erzeugt alle 40 ms ein Bild. Das entspricht exakt der PAL-Bildwechselfrequenz (25 Bildern pro Sekunde). Es konnte mittels des Optibase SDK kein Tool zur Veränderung des Bilderzeugungstakt von 25 Bildern pro Sekunde programmiert werden, weil dieses in den Spezifikationen dieses SDK's nicht erlaubt und möglich war.

Es konnte auch bei Ausreizung der gesamten Leistungsfähigkeit (bis zu 10 Mbit/s) der Kompressionskarte kein abweichendes Verhalten beobachtet werden.

Als Fazit dieses Versuches bleibt festzuhalten, dass sich die MPEG FORGE innerhalb ihres gesamten Leitungsbereiches an die PAL Spezifikation von 25 Bildern pro Sekunde hält und sich keine Möglichkeit bot, bei abweichenden Parametervorgaben andere Taktzeiten zu erhalten.

6.2 Die Qualitätsmessungen

6.2.1 Die Vorstellung der Testmethode

Um verschiedene Netzwerkbandbreiten zu simulieren, wurden Videos mit unterschiedlichen Bitraten erzeugt. Es wurden Videos mit den Bitraten 760, 2000 und 3000 kbit/s aufgenommen, dies deckt gut das Leistungsspektrum der verwendeten Videohardware ab. Von diesen Filmen wurden die für die Qualitätstests relevanten Bilder extrahiert.

6.2.2 Die Testfilme im Überblick

Um die Qualität der aufgenommenen Bilder miteinander vergleichen zu können, wurden aus den Videofilmen über 2400 Bilder extrahiert. Eine tabellarische Übersicht dieser Bilder findet sich im Anhang unter 9.3 (Tabelle 9-1 bis Tabelle 9-9).

Die neun Tabellen erhalten alle erhobenen statistischen Auswertungen und errechneten Daten pro aufgenommenem Film. Die erhobenen Daten sind Videobitrate, I- und P- Picture Distance, Dateigröße, durchschnittliche Bildgröße, Anzahl der Bilder und die PSNR. Die errechneten Daten sind die Größe des Videoteils (Bilderanzahl x Bildergröße) und der Overhead (wird in 6.3.2 näher beschrieben). Es wurden pro Tabelle immer die Filme zusammengefasst, welche mit gleicher Kompressionsart (MPEG 1 oder MPEG 2) und gleicher Kompressionsrate aufgenommen wurden. Die Reihenfolge der Filme in den Tabellen ergibt sich aus der Sortierung nach I-Picture und P-Picture Distance.

Die Vergleichsbilder sind die Startbilder der Filme (Tabelle 9-8 und Tabelle 9-9) in bestmöglicher Bitrate jeweils mit MPEG 1 und MPEG 2 aufgenommen. Diese Bilder liefern die Vergleichswerte zur subjektiven und objektiven Qualitätsbestimmung der Bilder in den aufgenommenen Filmen im Abschnitt 6.3.4 und 6.3.5.

6.2.3 Vergleich der Qualität der Bilder in den Videofilmen

Die Tabellen Tabelle 6-3 bis Tabelle 6-7 zeigen die aufgenommenen Videofilme im qualitativen Vergleich der Startbilder. Es wurde ein subjektiver Qualitätseindruck, durch betrachten der Bilder im direktem Vergleich ermittelt und ein objektiver Qualitätswert, mittels des in Kapitel 5.6.9 vorgestelltem PSNR – Tools gewonnen, pro Videofilm ermittelt.

Die erste Spalte jeweils zeigt die Nummer des Filmes, die Zweite Spalte zeigt den PSNR Wert und die Dritte Spalte eine subjektive Qualitätsnote des ersten Bildes im Film im Vergleich zu den Filmen aus den Tabellen Tabelle 9-8 und Tabelle 9-9. Wobei die Qualitätsnote Werte zwischen 1, beste Qualität und 15, schlechteste Qualität, annehmen kann. Die Tabelle 6-8 zeigt eine genaue Übersicht der Einteilung in Qualitätsnoten.

Tabelle 6-3 MPEG 1-SIF mit 760,4 kbit/s /Qualitätsvergleich

Titel	PSNR	Qualitäts-Note
16	18,21551543	9
17	19,12427193	9
18	19,22345538	9
19	19,75125843	9
20	19,17163349	10
21	19,1800811	10
22	19,32235801	10
23	19,25583429	10
24	19,2906975	12
25	19,32858905	10
26	19,30517488	10
27	19,4562912	10
28	19,13397535	10
29	19,27687932	10
30	19,0935589	10

Tabelle 6-4 MPEG 1-SIF mit 2000 kbit/s /Qualitätsvergleich

Titel	PSNR	Qualitäts- Note
31	21,40010368	2
32	21,3258593	5
33	21,33322355	4
34	21,35128648	4
36	21,22091957	3
37	21,21543913	3
38	21,25526805	2
39	20,98248679	2
40	21,04786037	4
41	21,33611947	6
42	21,17046245	4
43	21,06128932	5
44	21,39306578	6
45	21,17447354	4
61	21,29884395	8
62	21,37368535	7
63	21,16295431	7
64	21,39358136	7
65	21,37972394	7

Tabelle 6-5 MPEG 1-SIF mit 3000 kbit/s /Qualitätsvergleich

Titel	PSNR	Qualitäts- Note
71	27,56482677	2
72	27,48253819	2
73	27,58250984	2
74	27,75448739	4
75	27,58672638	2
76	27,58126511	2
77	27,38792541	2
78	27,30254789	2
79	27,42387544	1
80	26,35153581	1
81	26,22673409	1
82	27,36357326	1
83	27,81152903	1
84	27,28100637	2
85	27,3447044	6
86	27,35481688	2
102	27,90447228	2
103	26,43916775	1

Tabelle 6-6 MPEG 2 Half D-1 mit 2000 kbit/s /Qualitätsvergleich

Titel	PSNR	Qualitäts- Note
46	21,48223145	9
47	21,5438618	9
48	21,55497363	9
49	21,55497363	9
50	21,56219563	8
51	21,52540963	8
52	21,5572013	8
53	21,56954342	8
54	21,67672524	9
55	23,95643047	12
56	23,94750481	12
57	23,93753382	9
58	23,97684056	9
59	23,96225684	9
60	24,25988412	9
66	23,92185493	9
67	23,94515267	11
68	23,91857293	11
69	23,88936657	12
70	23,93153266	12

Tabelle 6-7 MPEG 2 Half D-1 mit 3000 kbit/s /Qualitätsvergleich

Titel	PSNR	Qualitäts- Note
87	19,63455161	3
88	19,59801902	3
89	19,60338749	2
90	19,5716366	2
91	19,59375603	2
92	19,54697955	2
93	19,56667019	2
94	19,53954507	2
95	20,78273151	10 ^a
96	19,48334256	2
97	20,72445987	10 ^a
98	19,51941348	2
99	19,53752066	2
100	19,49894123	5
101	19,4903165	6
104	19,0399814	2
105	18,35043523	2
106	18,36080029	1
107	18,32668349	1
108	18,12843785	1

^a Softwarefehler (viele kurze Querstreifen im Bild) beim extrahieren der Bilder aus dem Film

Tabelle 6-8 Festlegen von Qualitätsstufen

Noten	Qualitätsstufe
1	keine Unterschiede zum Original
2-3	geringe Unterschiede (nur im direkten Vergleich sichtbar)
4-9	akzeptable Unterschiede zum Original erkennbar (vereinzelt Artefakte)
10-14	nicht akzeptable Unterschiede zum Original
15	Original nicht mehr erkennbar

Die Abbildungen Abbildung 6-1 bis Abbildung 6-4 verdeutlichen das in der vorangegangenen Tabelle vorgestellte Notensystem. Es wird als erstes das Originalbild vorgestellt und danach werden die einzelnen Qualitätsstufen gezeigt.

Abbildung 6-1 Originalbild, 10.000 kbit/s I/P=12/3 (keine Artefakte)



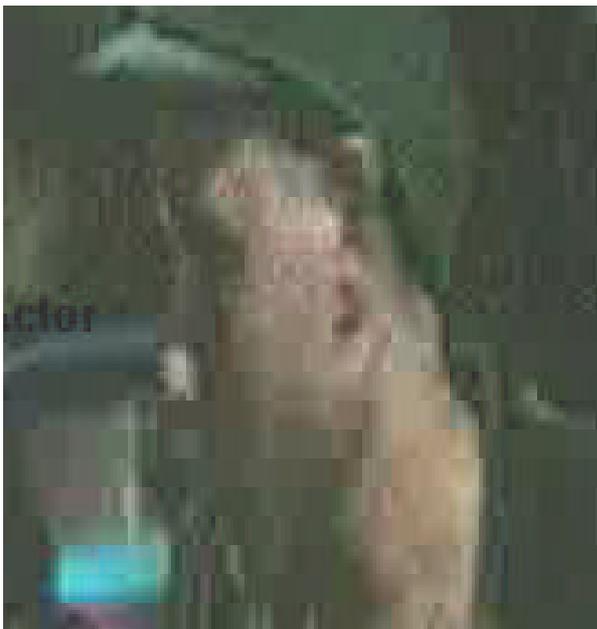
Abbildung 6-2 Bild der Note 2, 3.000 kbit/s I/P=90/1 (leichte Artefakte am oberen Ohransatz)



Abbildung 6-3 Bild der Note 6, 3.000 kbit/s I/P=1/1 (starke Artefakte am Ohr)



Abbildung 6-4 Bild der Note 10, 2.000 kbit/s I/P=3/3 (starke Artefakte im ganzen Bild)



6.3 Auswertung der erzielten Ergebnisse

6.3.1 Die durchschnittliche Framegröße

Aus Abbildung 1 und 2 ist zu entnehmen, dass die durchschnittliche Framegröße um weniger als 100 Bytes im Verlauf einer Messreihe schwankt. Dies war in über 90% aller Versuchsreihen zu beobachten. Einige wenige Messreihen haben Ausreißer mit Schwankungen von über 300 Bytes, was mit Messungenauigkeiten und Statistischen Schwankungen zu erklären ist. Dies erlaubt die Schlussfolgerung, dass der von der Firma Optibase implementierte MPEG 1 und MPEG 2 Codec über den gesamten Leistungsbereich in Bezug auf die Kompressionsrate bei Veränderung der GOP konstant arbeitet.

Abbildung 6-5 Schwankung der durchschnittliche Framegröße bei Veränderung der GOP (MPEG2 2000 kbit/s)

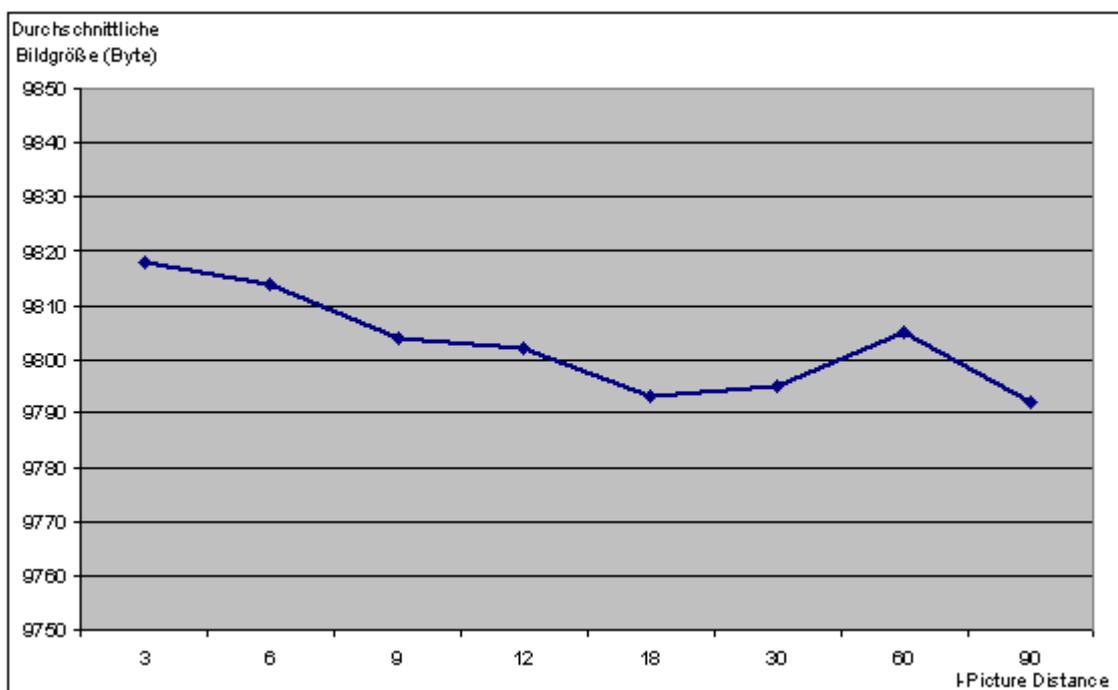
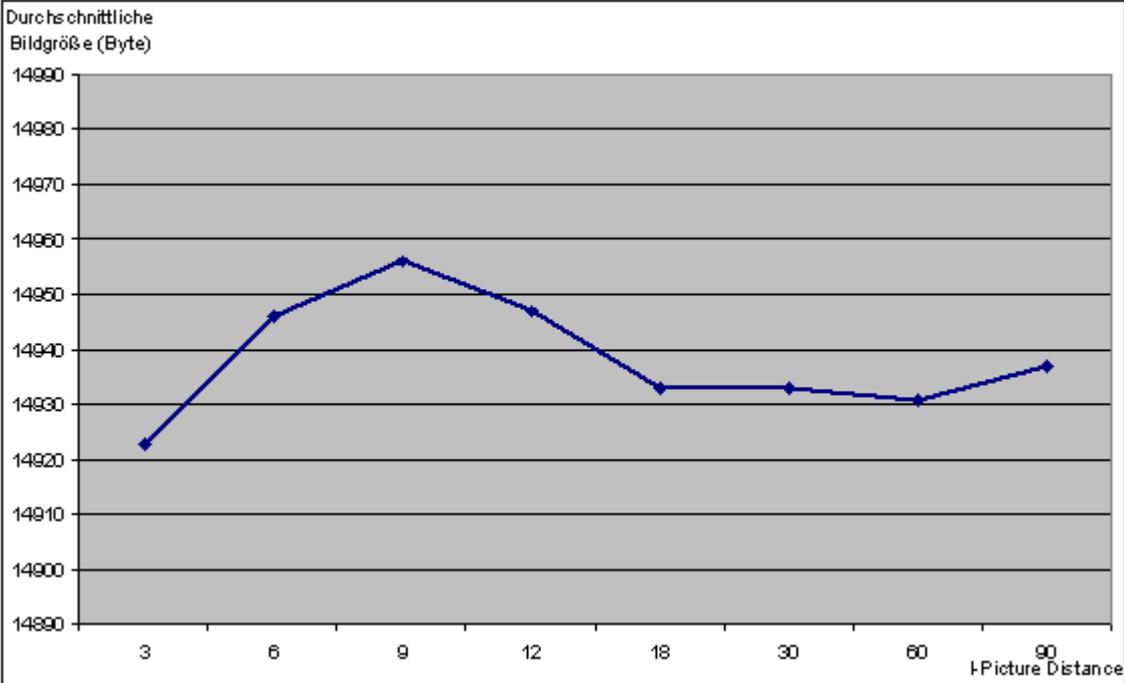


Abbildung 6-6 Schwankung der durchschnittliche Bildgröße bei Veränderung der GOP (MPEG1 3000 kbit/s)



6.3.2 Overhead bei der Codierung

Um einen Einblick in die Effizienz der verwendeten Codecs zu bekommen, wurde der erzeugte Overhead berechnet. Der Overhead errechnet sich aus der tatsächlichen Dateigröße, der Anzahl der erzeugten Bilder und der durchschnittlichen Bildergröße im Videofilm, Gleichung 6-2.

Bei dem berechneten Overhead müsste natürlich noch eine konstante Größe für den im Audio– Video Stream enthaltenem Audio Anteil abgezogen werden. Es war aber leider nicht möglich die Größe des codierten Audioteils zu bestimmen, da keinerlei Werkzeug zur Messung dieser von Optibase zu Verfügung gestellt wird. Da aber der Audioanteil (224 kbit/s bei 20s pro Film ca. 560Kbyte unkomprimiert) immer gleich ist, bleibt diese Konstante unberücksichtigt. Die Aussage wird dadurch nicht verfälscht.

Gleichung 6-2 Errechnung des Overheads der erzeugten Audio-Video Streams

Bildanteil [byte] = Bilderanzahl * durchschnittlichen Bildergröße [byte]

Overhead [byte] = Dateigröße [byte] – (Bildanteil [byte]) – konstantem Audioanteil [byte]

Da die Größe der erzeugten Files nur sehr gering schwankt, die Unterschiede liegen im Bereich von 10000 – 50000 Bytes, und die Anzahl der Bilder immer 501⁶⁰ beträgt, ist der erzeugte Overhead weitestgehend gleichbleibend. Das spricht für den implementierten Codec. Der Codec erzeugt, egal wie die Beschaffenheit der GOP ist immer einen gleichbleibend niedrigen Overhead.

⁶⁰ 501 Bilder - bei manchen Filmen wurden 502 Bilder erzeugt, dies ist durch die Beschaffenheit der GOP bedingt, wenn die GOP den Aufbau 12 - 2 oder 6 – 2 (jeweils I- und P-Bilderabstand) hat. Hat keinen Einfluss auf den Overhead.

6.3.3 Entwicklung der Qualitätswerte innerhalb der GOP

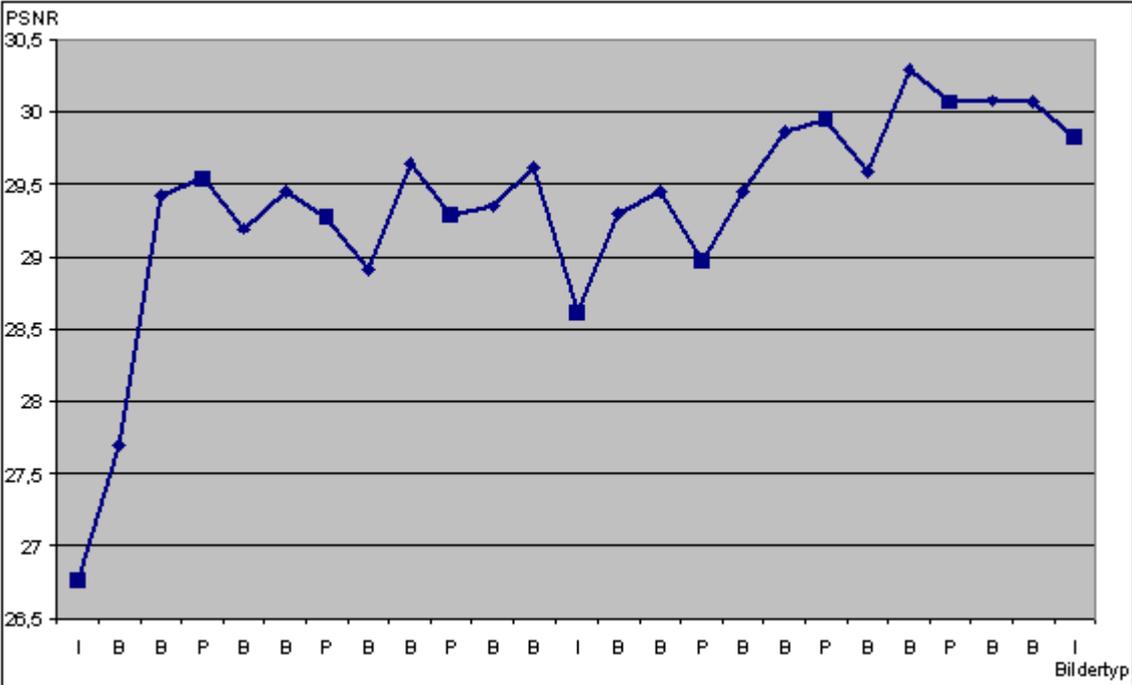
Um die Qualität der Bilder miteinander Vergleichen zu können, wurden von jedem Frame die PSNR- Werte gemessen.

Wie in Abbildung 3 zu sehen, entwickelt sich die Qualität der einzelnen Bilder eines Filmes im Laufe mehrerer GOPs zu besseren Werten hin.

Das erste I-Bild hat einen PSNR- Wert von unter 27. Innerhalb der ersten GOP sind Werte bis zu einer PSNR von 29,6 möglich. Die zweite GOP startet wieder mit einem recht niedrigem PSNR- Wert eines I-Bildes und erzielt dann schon deutlich bessere Werte als die erste GOP. Die letzten Bilder dieser Sequenz erreichen Werte von über 30, dies sind im Vergleich zu den Werten um die 27 am Anfang dieser Sequenz sehr viel bessere Werte. Solche PSNR Unterschiede sind auch mit dem bloßem Auge wahrnehmbar.

Diese PSNR- Wert Schwankungen innerhalb einer Sequenz aufeinander folgender Bilder sind mit der verstärkten Ausnutzung von Redundanzen in den vorangegangenen und nachfolgenden Bildern zu erklären (Absatz 3.4.2). Was deutlich auffällt sind die im Vergleich zu B- und P- Bilder schlechten PSNR Werte der I- Bilder. Der Grund dafür ist, dass eventuelle Redundanzen zwischen aufeinanderfolgenden Bildern nicht genutzt werden. Daraus ergibt sich nur eine begrenzte Kompressionsmöglichkeit der I-Bilder und das decodierte Bild bekommt geringere PSNR- Werte, hat also eine schlechtere Qualität.

Abbildung 6-7 PSNR Entwicklung I=12 P=3 (MPEG1 2000kbit/s)



6.3.4 Der Qualitätsvergleich mittels PSNR

6.3.4.1 Messergebnisse im Überblick

Die Abbildungen in diesem Kapitel sind wie folgt aufgebaut. Auf der Y- Achse sind die PSNR Werte angetragen. Auf der X-Achse wurden bei den Abbildungen Abbildung 6-8, Abbildung 6-11, Abbildung 6-12, Abbildung 6-15, Abbildung 6-16, Abbildung 6-18, Abbildung 6-20, Abbildung 6-21, Abbildung 6-22 und Abbildung 6-24 der P- Picture- Abstand und bei den Abbildungen Abbildung 6-9, Abbildung 6-10, Abbildung 6-13, Abbildung 6-14, Abbildung 6-17, Abbildung 6-19 und Abbildung 6-23 der I- Picture- Abstand dargestellt. Der Graph kennzeichnet den PSNR- Wert zur jeweiligen GOP.

Abbildung 6-8 PSNR- Werte für MPEG 1, 760,4 kbit/s, (I = 12, P = 1-3)

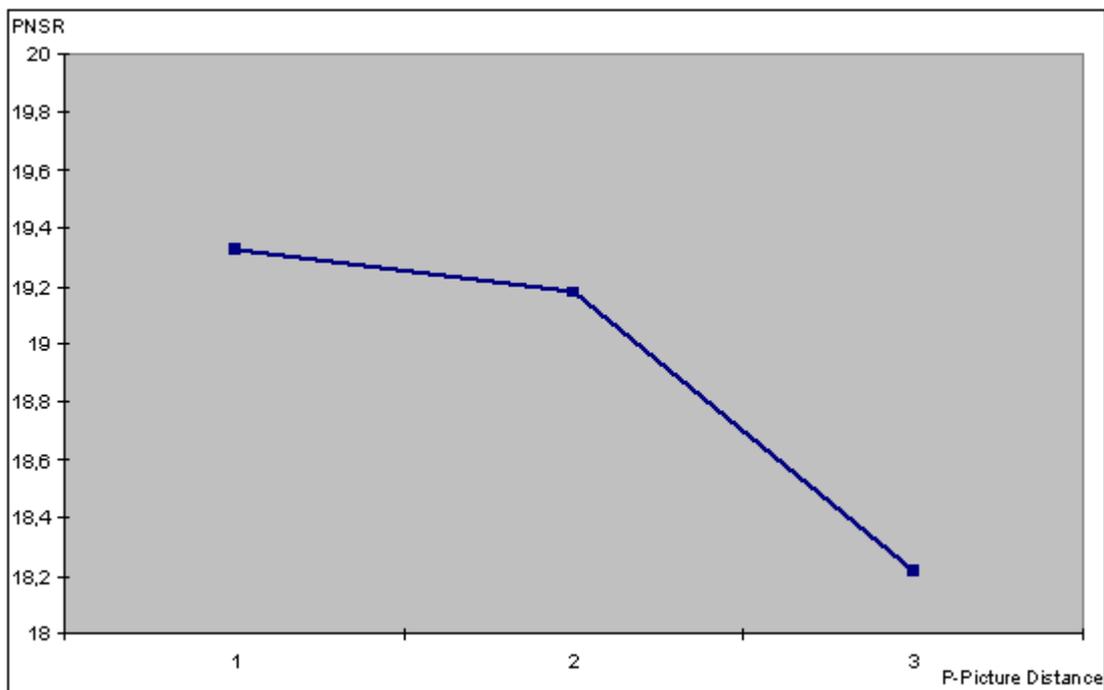


Abbildung 6-9 PSNR- Werte für MPEG 1, 760,4 kbit/s, (I = 3-90, P = 3)

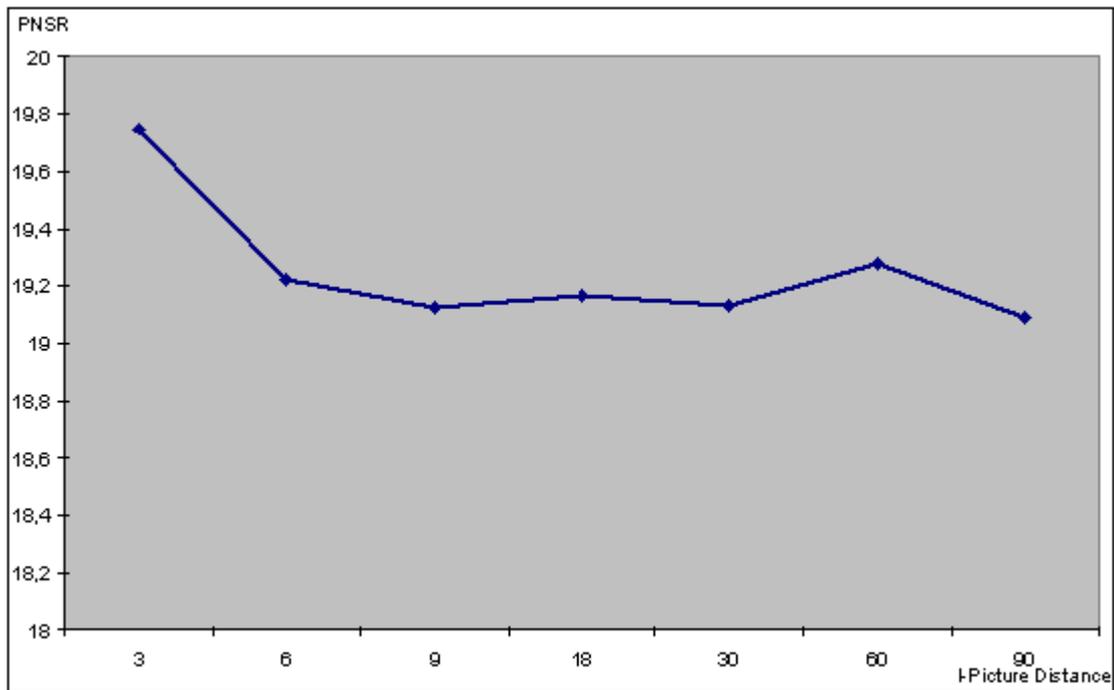


Abbildung 6-10 PSNR- Werte für MPEG 1, 760,4 kbit/s, (I = 12-90, P = 1)

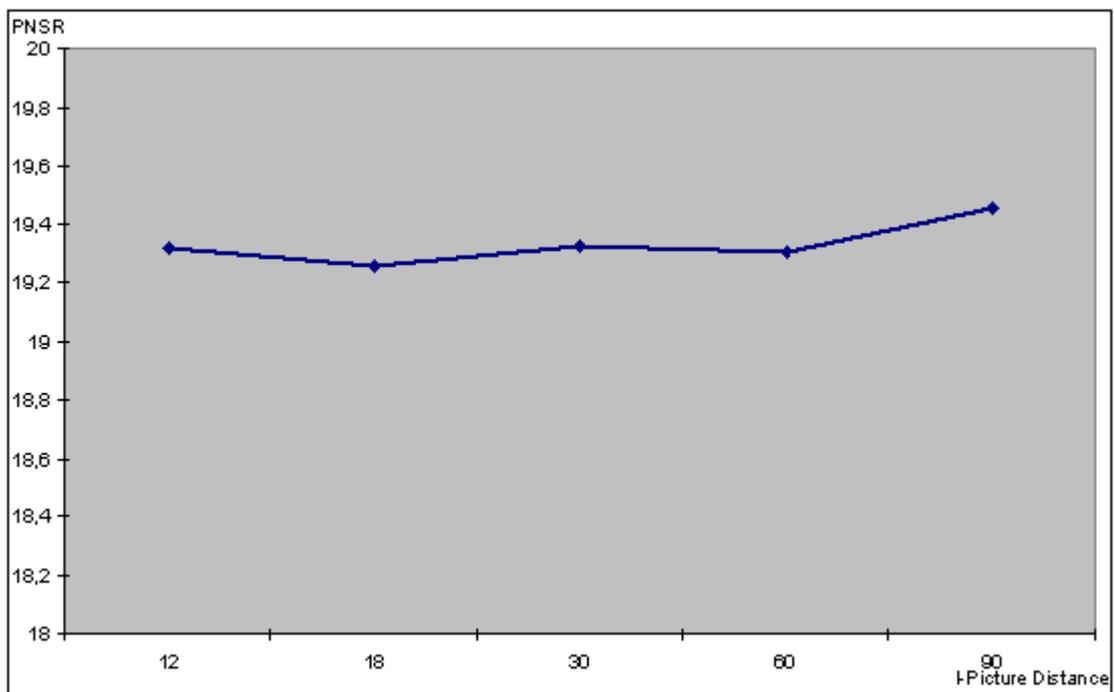


Abbildung 6-11 PSNR- Werte für MPEG 1, 2000 kbit/s, (I = 1-30, P = 1)

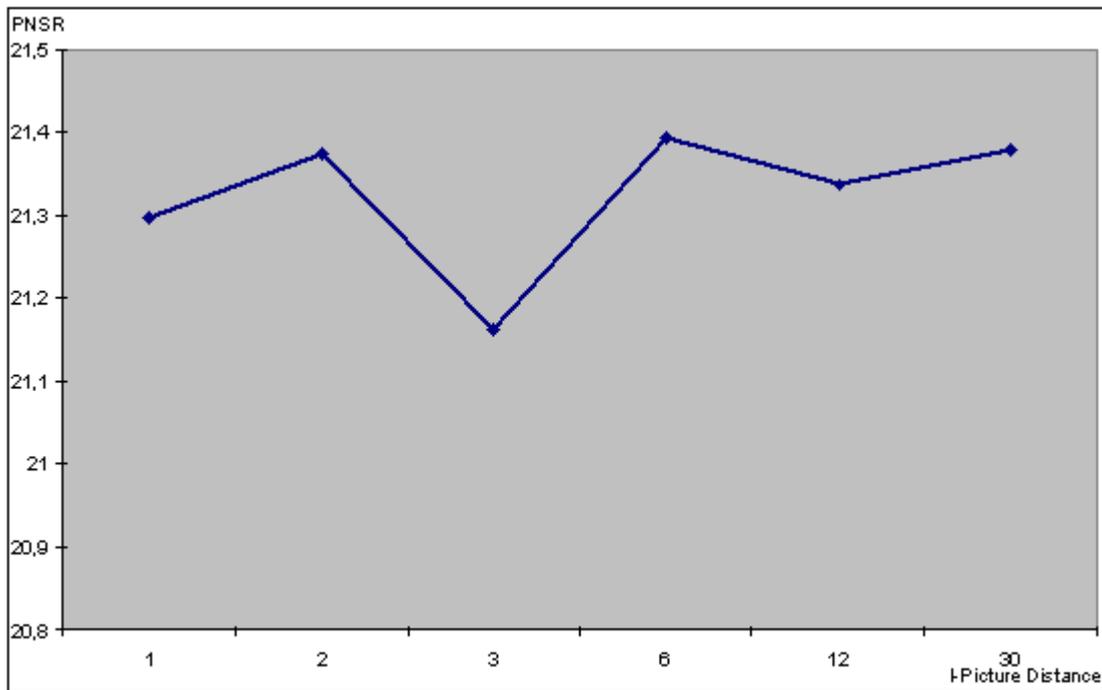


Abbildung 6-12 PSNR- Werte für MPEG 1, 2000 kbit/s, (I = 3-90, P = 3)

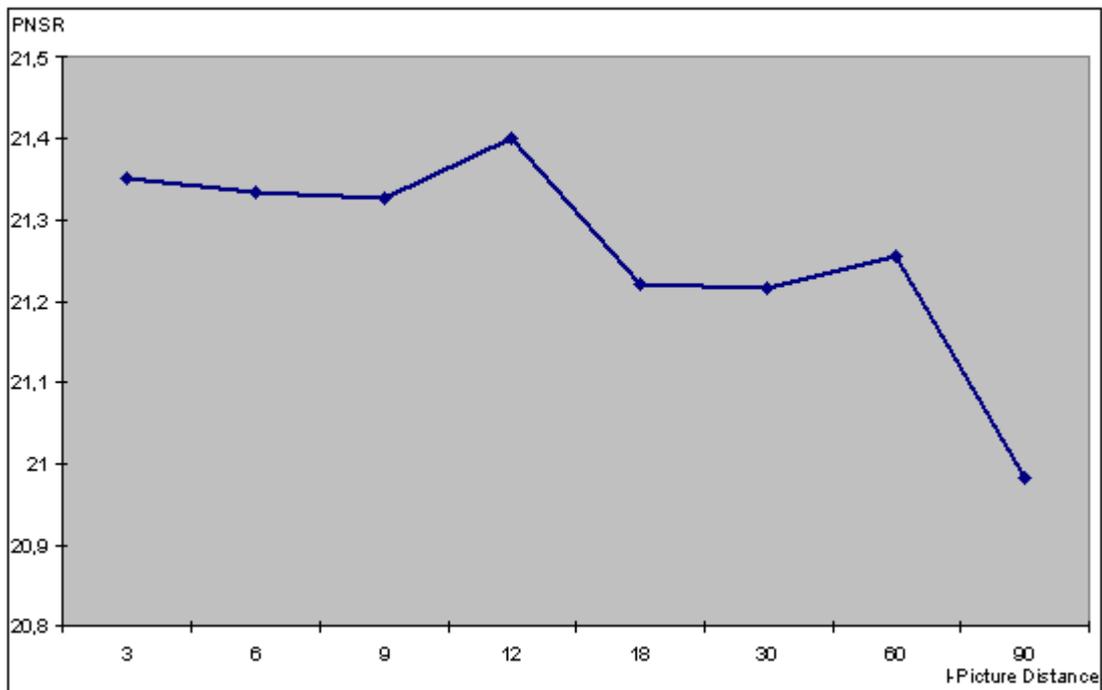


Abbildung 6-13 PSNR- Werte für MPEG 1, 2000 kbit/s, (I = 12, P = 1-3)

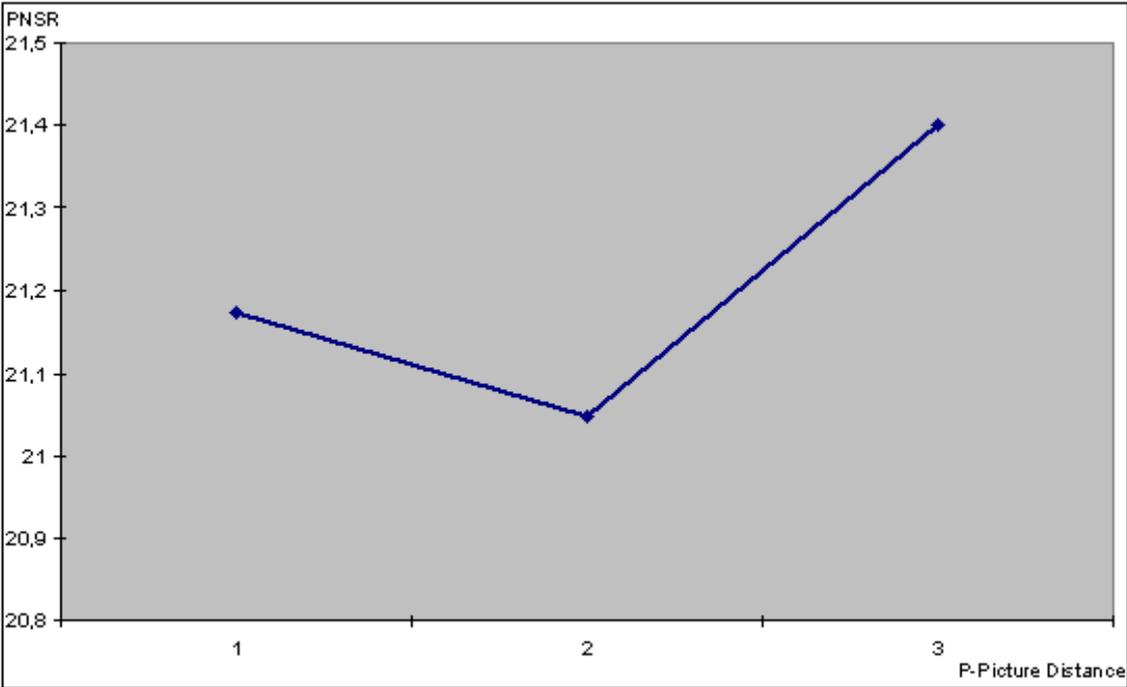


Abbildung 6-14 PSNR- Werte für MPEG 1, 3000 kbit/s, (I = 12, P = 1-3)

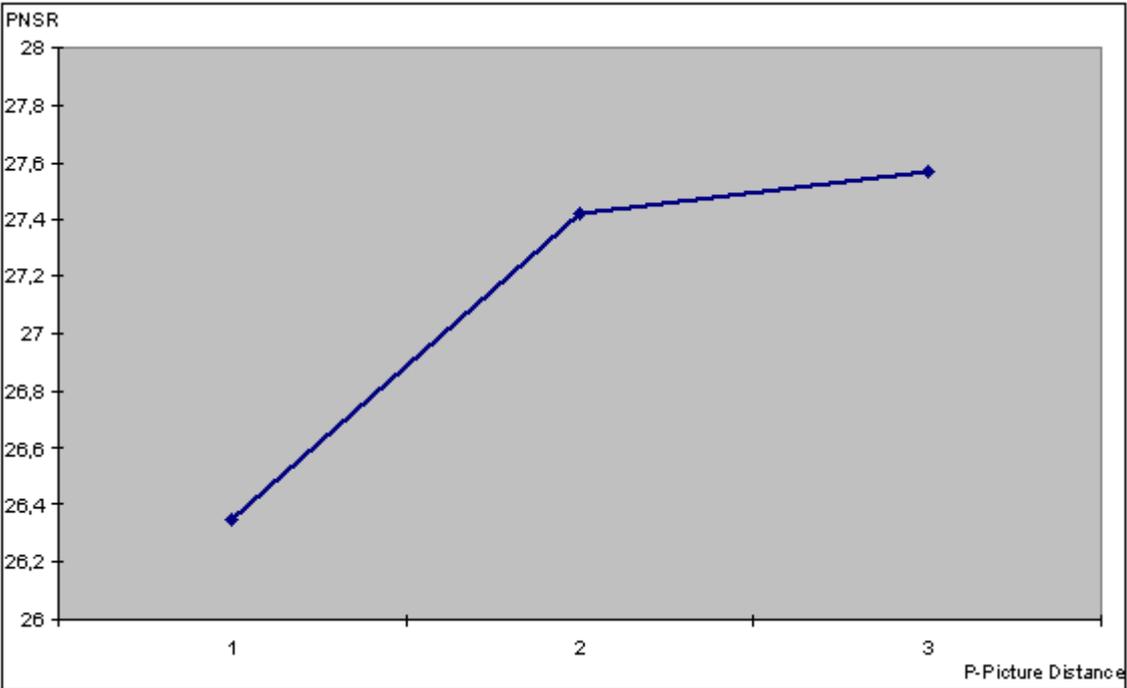


Abbildung 6-15 PSNR- Werte für MPEG 1, 3000 kbit/s, (I = 3-90, P = 3)

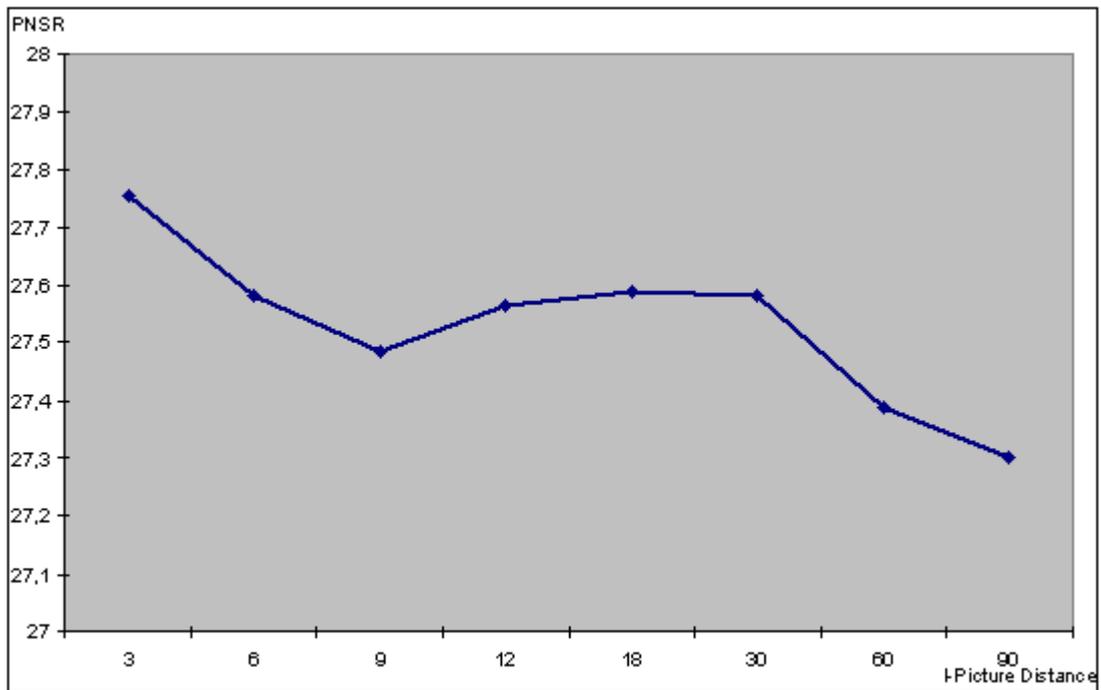


Abbildung 6-16 PSNR- Werte für MPEG 1, 3000 kbit/s, (I = 1-30, P = 1)

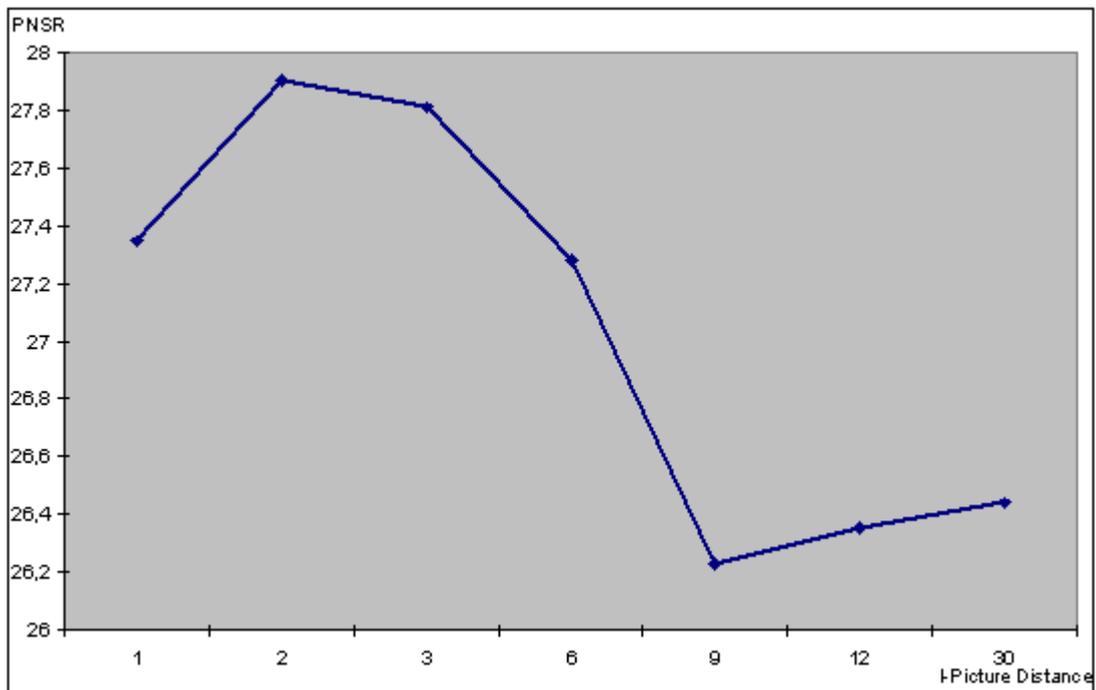


Abbildung 6-17 PSNR- Werte für MPEG 1, 3000 kbit/s, (I = 6, P = 1-3)

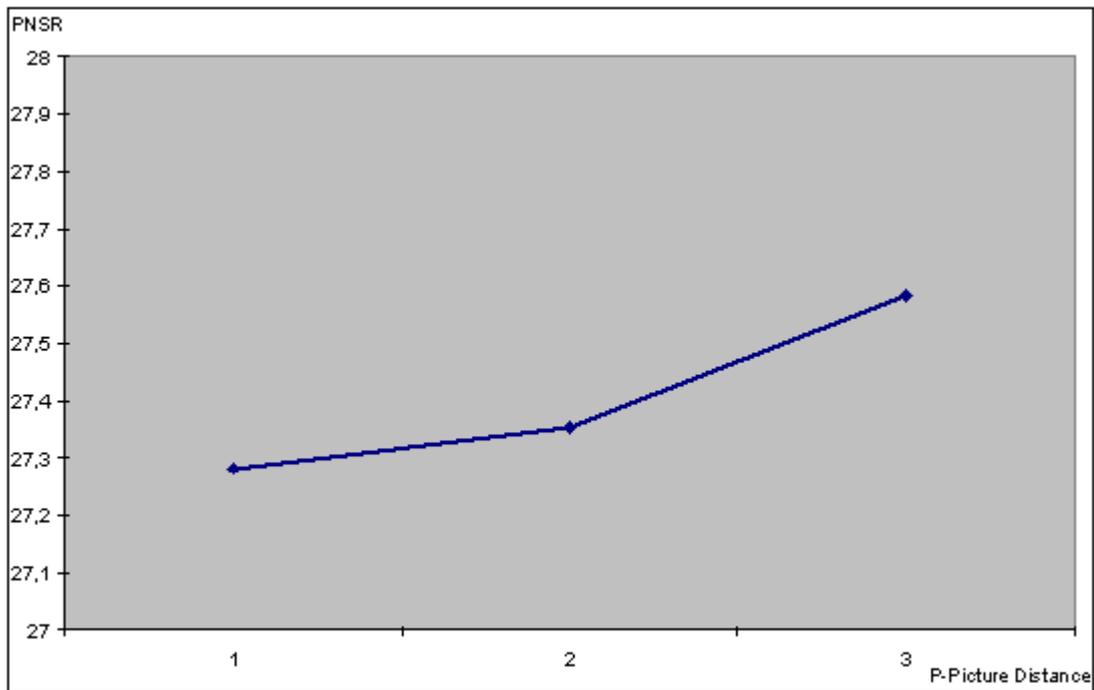


Abbildung 6-18 PSNR- Werte für MPEG 1, 3000 kbit/s, (I = 6-12, P = 2)

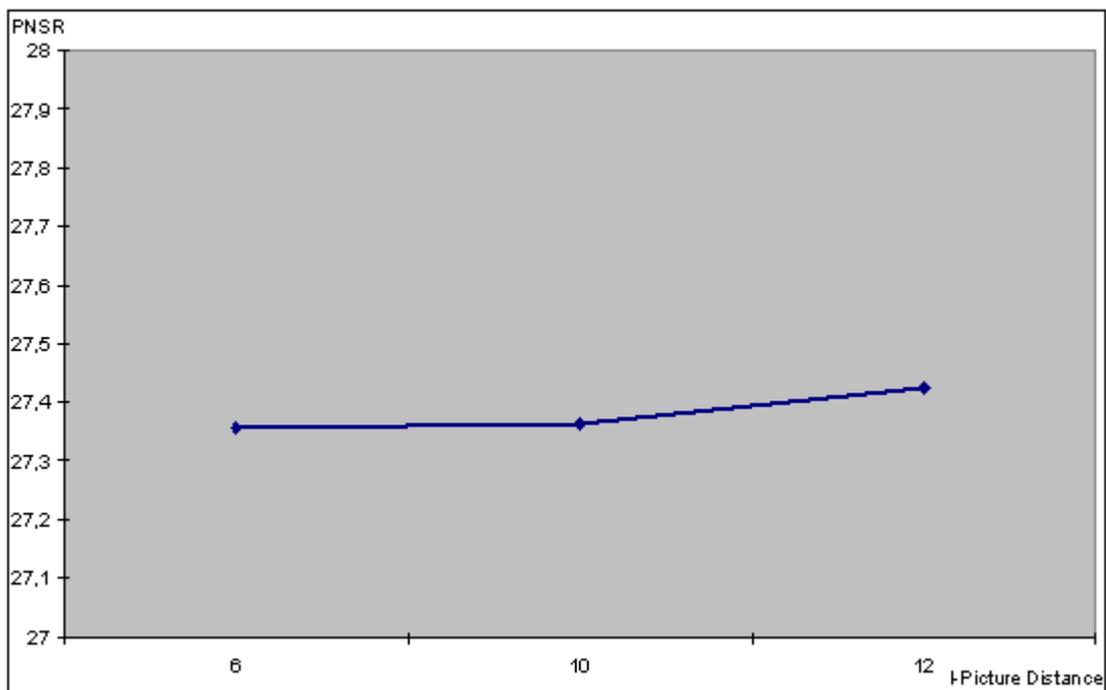


Abbildung 6-19 PSNR- Werte für MPEG 2, 2000 kbit/s, (I = 12, P = 1-3)

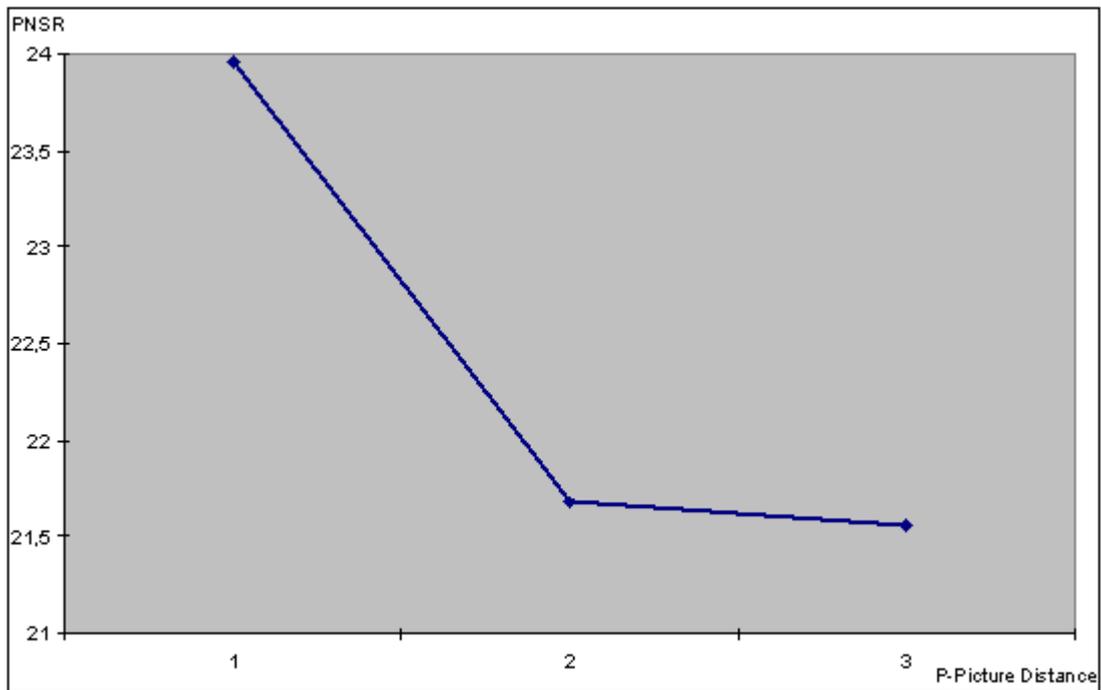


Abbildung 6-20 PSNR- Werte für MPEG 2, 2000 kbit/s, (I = 3-90, P = 3)

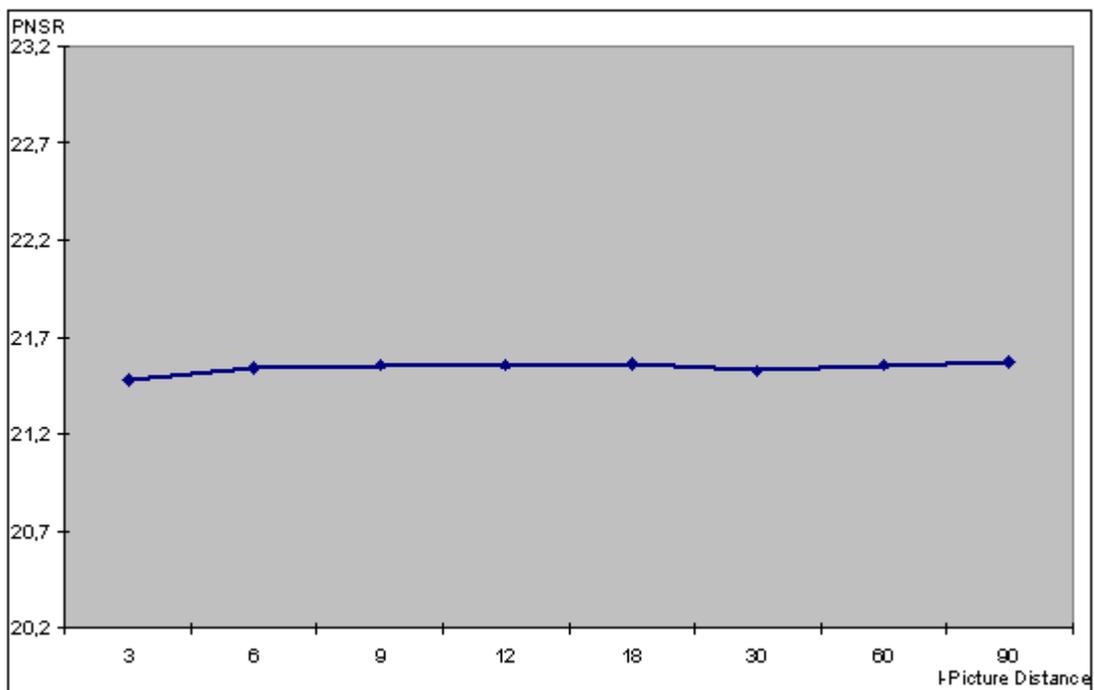


Abbildung 6-21 PSNR- Werte für MPEG 2, 2000 kbit/s, (I = 1-100, P = 1)

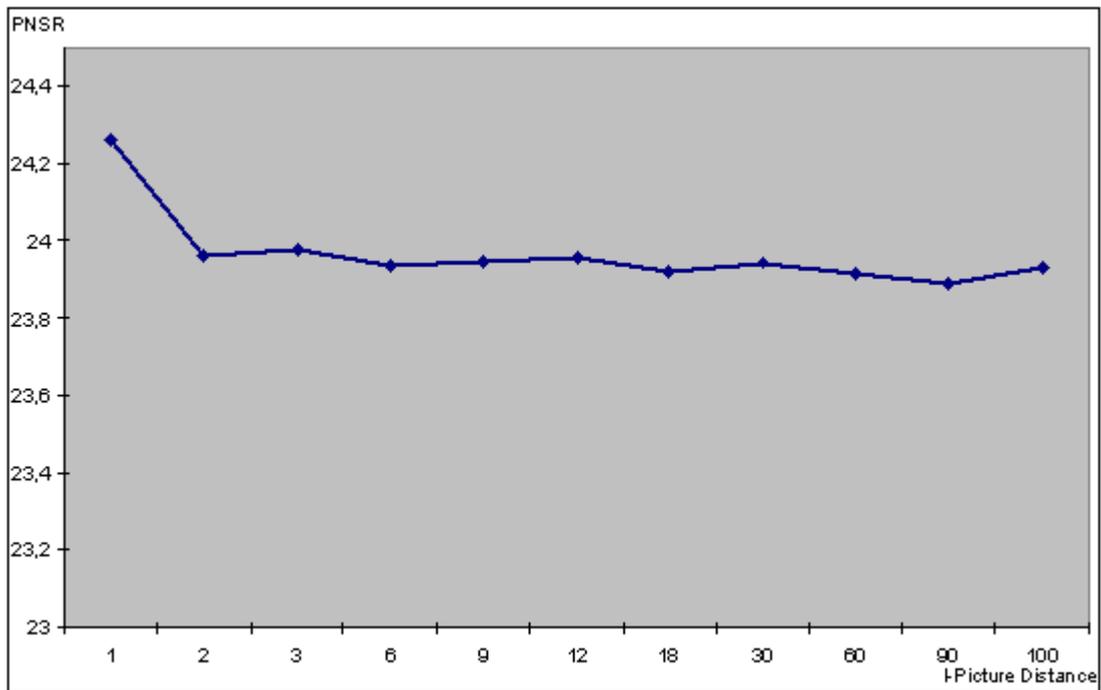


Abbildung 6-22 PSNR- Werte für MPEG 2, 3000 kbit/s, (I = 3-100, P = 1)

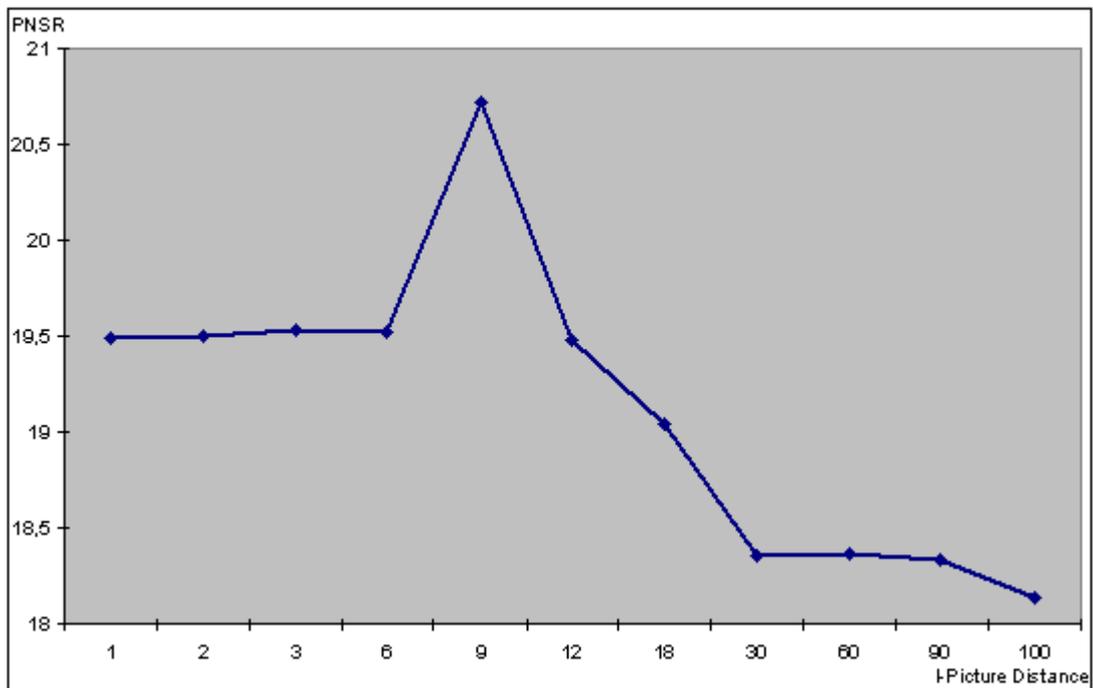


Abbildung 6-23 PSNR- Werte für MPEG 2, 3000 kbit/s, (I = 12, P = 1-3)

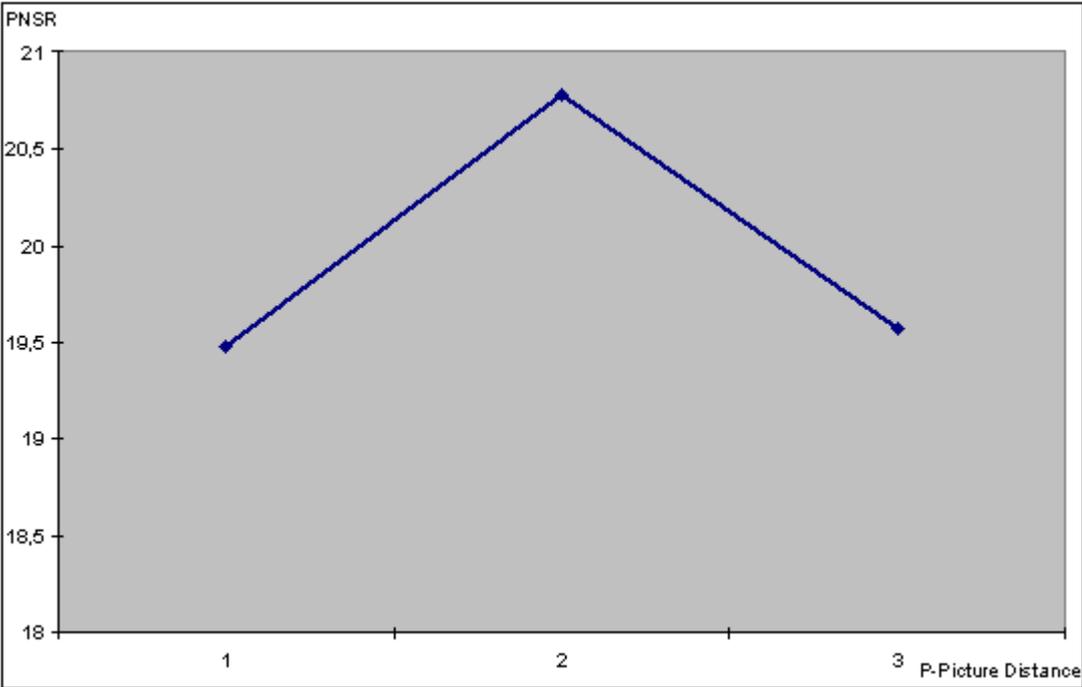
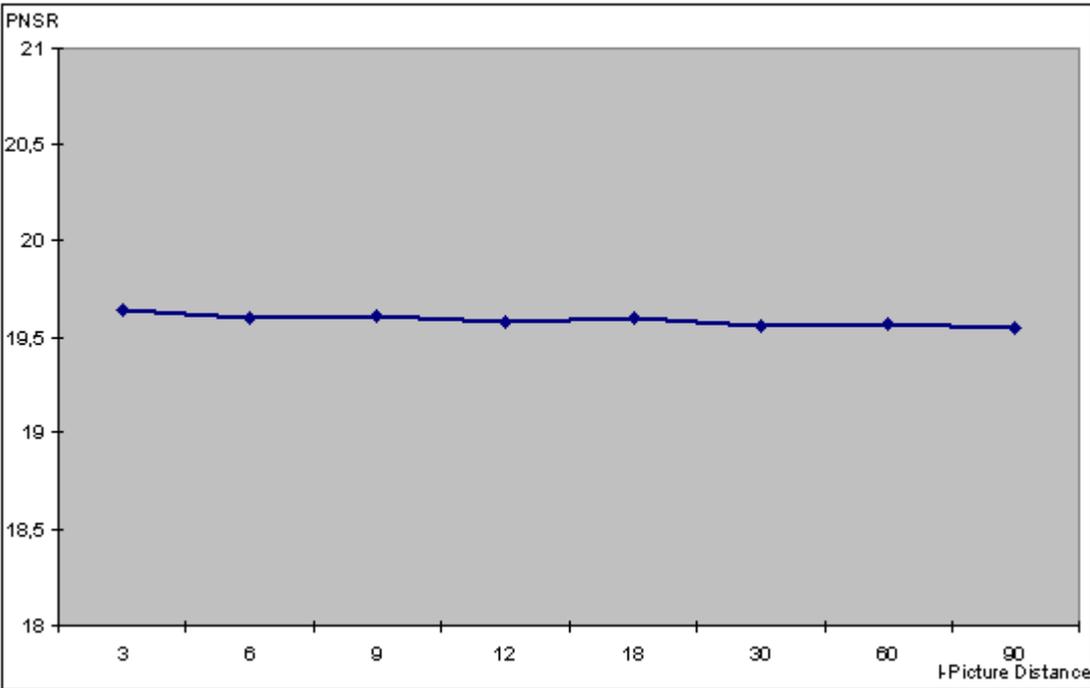


Abbildung 6-24 PSNR- Werte für MPEG 2, 3000 kbit/s, (I = 3-90, P = 3)



6.3.4.2 Diskussion der Ergebnisse

Die Abbildungen Abbildung 6-7 bis Abbildung 6-24 sind in zwei Gruppen zu ordnen. Die erste Gruppe (I-Gruppe) ist die mit den Diagrammen, bei welchen die I- Picture Distance verändert und die P- Picture Distance fest gehalten wurde. Zu dieser Gruppe gehören die Abbildungen Abbildung 6-8, Abbildung 6-11, Abbildung 6-12, Abbildung 6-15, Abbildung 6-16, Abbildung 6-18, Abbildung 6-20, Abbildung 6-21, Abbildung 6-22 und Abbildung 6-24. Zu der zweiten Gruppe (P-Gruppe) gehören die Diagramme bei denen die P- Picture Distance verändert und die I- Picture Distance fest gehalten wurde. Das sind die Abbildungen Abbildung 6-9, Abbildung 6-10, Abbildung 6-13, Abbildung 6-14, Abbildung 6-17, Abbildung 6-19 und Abbildung 6-23.

Tabelle 6-9 Einteilung der Diagramme anhand der Neigung der Graphen⁶¹

Abbildung	Kompr. Verfahren	I/P- Picture Distance nicht fest (Gruppe)	Neigung des Graphen
Abbildung 6-8	MPEG 1	P	fallend
Abbildung 6-9	MPEG 1	I	fallend
Abbildung 6-10	MPEG 1	I	gleichbleibend
Abbildung 6-11	MPEG 1	I	gleichbleibend
Abbildung 6-12	MPEG 1	I	fallend
Abbildung 6-13	MPEG 1	P	steigend
Abbildung 6-14	MPEG 1	P	steigend
Abbildung 6-15	MPEG 1	I	fallend
Abbildung 6-16	MPEG 1	I	fallend
Abbildung 6-17	MPEG 1	P	leicht steigend
Abbildung 6-18	MPEG 1	I	gleichbleibend
Abbildung 6-19	MPEG 2	P	fallend
Abbildung 6-20	MPEG 2	P	fallend
Abbildung 6-21	MPEG 2	I	gleichbleibend
Abbildung 6-22	MPEG 2	I	leicht fallend
Abbildung 6-23	MPEG 2	I	fallend
Abbildung 6-24	MPEG 2	P	gleichbleibend

Tabelle 6-9 zeigt eine Bewertung der Diagramme. Es zeigt sich, dass bei allen Diagrammen der I-Gruppe entweder fallende oder gleichbleibende Tendenz festzustellen ist. Das bedeutet, dass bei Vergrößerung der GOP, des I-Bilder

Abstandes, die PSNR- Werte der Bilder schlechter werden oder zumindest gleich bleiben. Bei der P-Gruppe sind die Werte entweder gleichbleibend oder steigend. Dies bedeutet, dass mit zunehmendem Abstand der P-Bilder sich die Qualität der Bilder verbessert oder die Qualität gleich bleibt.

Diese Ergebnisse zeigen, je größer der Abstand der I-Bilder wird, desto schlechter wird die Qualität der Bilder im Film. Dies ist mit dem zunehmenden zeitlichem Abstand der I-Bilder zu erklären, dadurch können nachfolgende I- und B-Bilder keine Redundanzen mehr in den I-Bildern finden, es verringert sich die Kompressionsrate und die Bilder werden, weil die Ausgangs Bitrate ja gleich bleibt, insgesamt mit einer schlechteren Qualität abgespeichert.

Andererseits, wird der I-Bilderabstand (hier $I = 12$ oder $I = 6$) fest gehalten und der Abstand der P-Bilder erhöht, erhöht sich die Qualität der Bilder im Film.

Dieses Verhalten erklärt sich dadurch, dass wenn der I-Bilder Abstand klein gewählt wurde, in einem Film durch die Ausnutzung der Redundanzen erheblich bessere Kompressionsraten erzielt werden können. Denn durch die Verkleinerung der Anzahl der P-Bilder in einer GOP kommt es zur Erhöhung der Anzahl der B-Bilder in der GOP. Wodurch eine Erhöhung der Kompressionsrate erzielt wird, da B-Bilder auch Redundanzen der nachfolgenden P- und I-Bilder ausnutzen, wie in Absatz 6.3.3. (Entwicklung der Qualitätswerte innerhalb der GOP) gezeigt.

Dies zeigt, dass Filme mit einem GOP Aufbau I-Bilder Abstand von 12 – 18 und P-Bilder Abstand = 3, die besten Ergebnisse erzielen. Weiter bleibt festzustellen, dass

⁶¹ Neigung der Graphen-ist gleich bedeutend mit der Entwicklung der PSNR- Werte in Richtung größer werdender I oder P - Picture Distance

6.3.5 Der Qualitätsvergleich durch Betrachten der Bilder

Wie man aus den Tabellen im Abschnitt (6.2.3) entnehmen kann, deckt sich die Qualitätsnote durch Betrachten nicht mit den PSNR- Werten. Dies ist durch Pixelfehler beim Grabben der Bilder, welche den PSNR- Wert stören, und durch Artefakte, welche die Qualitätsnote beeinflussen zu erklären.

Die Pixelfehler welche beim Grabben der Bilder entstehen, durch schlechtes Videoband oder Dreck auf dem Videokopf, verändern stärker den PSNR- Wert als Veränderungen durch Artefakte. Da die Pixel der Artefakte ja ähnlicher dem Originalpixel als ein Pixelfehler sind. Andererseits werden vom Betrachter die Pixelfehler fast gar nicht wahrgenommen, Artefakte jedoch als störend empfunden.

Betrachtet man die Qualitätsnoten losgelöst von den PSNR- Werten, fällt auf, dass wenn überhaupt Unterschiede wahrnehmbar waren für Filme mit großem (über 12) I-Bilder Abstand schlechtere Noten verteilt wurden als für Filme mit kleinerem Abstand (kleiner, einschließlich 12). Beispiele für dieses Verhalten sind die Filme 55, 56, 67, 68, 69 und 70, Tabelle 9-13, und die Filme 20 – 30, Tabelle 9-10, mit großem I-Bilder Abstand. Diese Filme haben schlechtere Noten als alle anderen Filme der gleichen Versuchsreihe. Weiterhin zeigt sich, dass Filme, welche mit P-Bilder Abstand = 1 aufgenommen wurden ebenfalls schlechtere Noten haben als solche mit größerem Abstand. Beispiele hierfür sind die Filme 41 und 44 aus der Tabelle 6-4 und die Filme 55, 56, 67, 68, 69 und 70, Tabelle 6-6.

Dieses Ergebnis deckt sich mit den Ergebnissen aus dem Vergleich mittels PSNR- Werten aus dem vorhergehenden Abschnitt.

7 Ausblick auf künftige Standards

7.1 JPEG 2000

Seit dem Jahr 1997 wird beim JPEG 2000-Komitee (offiziell: ISO/IEC JTC1/SC29/WG1) über einen neuen Standard zur Kompression von Bilddaten diskutiert.

Dieser neue Standard soll JPEG 2000 sein. JPEG 2000 soll die Nachteile des alten JPEG Standards ausgleichen und gleichzeitig für neue Entwicklungen gerüstet sein.

Als JPEG Ende der achtziger Jahre entwickelt wurde, genügte es einen Standard zu entwickeln, welcher

- unabhängig von der Art der Bildinhalte ist,
- geringe Ansprüche an die Hardware stellt,
- und dem damaligen Stand der Technik entsprach.

Die explosionsartige Entwicklung im Sektor der Datenverarbeitung und Datenkommunikation brachte enorm viele Anwendungsmöglichkeiten für Bildkompression mit sich, welche sehr schnell die Grenzen des JPEG Standards aufzeigten.

Mit diesem Hintergrund wurde 1997 mit der Schaffung eines neuen Kompressionsstandards, des JPEG 2000, begonnen.

Die Hauptkritikpunkte am alten Standard, wie die geringe Bildqualität bei mittleren bis hohen Kompressionsraten, die schlechte progressive Darstellung und die mangelnde Fehlertoleranz für Online- und Mobilfunkanwendungen, sollen durch den neuen Standard abgeschafft werden. Tabelle 7-1 listet die durch das JPEG-2000-Komitee aufgestellten Anforderungen an JPEG 2000 auf.

Tabelle 7-1 Anforderungen an JPEG 2000 nach [c't2699]

bessere Bildqualität als JPEG bei mittleren und hohen Kompressionsraten
Erfassung von Echtfarb-, Graustufen- und Schwarzweißbildern
wahlweise verlustbehaftete oder verlustfreie Kompression
progressive Übertragung
Arbeit mit begrenzten Ressourcen (Übertragungs- oder Speicherkapazität sowie Arbeitsspeicher) und in Echtzeitumgebungen
Freier Zugriff auf Teilbereiche eines kodierten Bildes
Robustheit gegenüber Übertragungsfehlern
Möglichkeit der Definition besonders wichtiger Bildregionen (ROI's, Regions of Interest), die mit höherer Genauigkeit kodiert werden
Mitübertragung von Randinformationen wie etwa stufenlose Transparenz
offene Architektur, die Optimierungen auf spezielle Bildklassen erlaubt
Rückwärtskompatibilität zu JPEG und Verträglichkeit mit MPEG-4 sowie bestehenden Fax-Standards
Unterstützung von Metadaten, etwa zur Beschreibung der Bildinhalte, sowie von Mechanismen zum Schutz der Eigentums- und Verwertungsrechte

Weil bereits zu Beginn des Standardisierungsprozesses deutlich wurde, dass die Wavelet - Kompression am besten die Anforderungen an JPEG 2000 erfüllt, beruht der Kern des JPEG 2000 Kompressionsalgorithmus auf der Wavelet - Transformation. Die Kompressionsraten von JPEG 2000 liegen bei verlustbehafteter Wavelet - Kompression bei 200:1, bei verlustfreier Wavelet - Kompression bei immerhin noch 2:1. Durch die Technik der Wavelet - Kompression erscheinen die Bilder bei zu hohen Kompressionsraten lediglich schwammig oder verwischt, es kommt nicht zu den unerwünschten Artefakten (Abbildung 7-1).

Abbildung 7-1 Überkomprimiertes JPEG (links) bei 70:1 und überkomprimiertes Wavelet Kodiertes Bild (rechts) bei 150:1 [MEN00]



Anders als bei JPEG bietet das neue JPEG 2000 insgesamt 256 Informationskanäle zur Beschreibung des gespeicherten Bildes an. So lassen sich nun viele Metadaten abspeichern. Es können zum Beispiel hohe Auflösungen eines Bildes mittels Passwort geschützt werden.

Damit ist JPEG 2000 vor allem für Internetanwendungen und Webseiten besonders geeignet.

Der Zeitplan des JPEG 2000-Komitees sah die offizielle Verabschiedung des Standards für die erste Hälfte des Jahres 2001 vor. Ein offizieller Standard für Motion JPEG sollte im November 2001 vorliegen. Bis zum heutigem Tag wurde JPEG 2000 noch nicht komplett zum verbindlichen Standard erhoben (Tabelle 7-2).

Tabelle 7-2 Teile und Status des JPEG 2000 Standards [JPEG]

Teil	Beschreibung	Status
1	Verabschiedung des lizenzfreien (nicht patentfrei) und verbindlichen Kodierungsschemas	Publiziert als internationaler Standard (Dezember 2000)
2	Ergänzungen des Standards. Diese können lizenzbelastet sein und sind nicht verbindlich. Erscheinungstermin des Drafts: Juli 2001	komplett oder fast komplett
3	Motion JPEG 2000.	
4	Konformität (Tests)	
5	Referenzsoftware , Java und C	
6	Draft für vermischte Bildformat Inhalte (Grafik und Text), wie z.B. Faxe	
7	wurde aufgegeben	
8	JPSEC Sicherheitsfeatures	in Entwicklung
9	JPIP interaktive Protokolle und eine API	
10	JP3D 3D Features, volumetric imaging	
11	JPWL wireless applications	

8 Zusammenfassung

Die Treiber der MPEG Forge scheinen optimal auf das Betriebssystem Windows NT 4.0 abgestimmt zu sein, es kam im gesamten Versuchszeitraum zu keinerlei Abstürzen. Des weiteren war auch das Betriebssystem nach Installation des Servicepacs 3 völlig stabil.

Für die Praxis bedeutet das, dass mit der in dieser Arbeit vorgestellten Hard und Software ein optimales arbeiten als Schnittplatz möglich ist. Es ist nicht mit Instabilitäten im Betrieb der MPEG Forge zu rechnen. Der Rechner, auch mit seiner eher geringen Leistung kann im Zusammenspiel mit der MPEG Forge semiprofessionellen Anforderungen gerecht werden. Es können Videofilme in unterschiedlichen Auflösungen (MPEG 1 und MPEG 2) und mit verschiedenen Bitraten gecaptured werden. Dabei sollten GOPs mit folgenden Parametern benutzt werden um optimale Leistung aus der Videohardware herauszuholen. Der I-Bilder Abstand ist möglichst klein zu wählen, optimal sind Abstände von 9 bis 12. Der P-Bilder Abstand ist mit 3 optimal eingestellt. Abweichende Werte liefern schlechtere Qualität. Wie man sieht ist die GOP 12 – 3 (I-Bilder Abstand – P-Bilder Abstand), ein von der Literatur bevorzugter Wert, durch diese Arbeit bestätigt wurden.

In zukünftigen Arbeiten könnte das modernere Motion JPEG 2000 hinsichtlich dieser Parameter untersucht werden. Dazu ist aber eine Erweiterung des Videoschnittplatzes um neue Hardware erforderlich, da die vorhandenen Videokarten dieses Format nicht unterstützen. Sehr interessant ist wäre zu sehen, wie sich der Kompressionsalgorithmus auf der Basis von Wavelet – Transformation im Gegensatz zum JPEG –Algorithmus des MPEG 1 oder MPEG 2 verhält.

9 Anhang

9.1 Literaturverzeichnis

[CCIR601] CCIR 601: Encoding parameters of digital Television for Studios, Recommendation, 1982

[c't1300] Peter Röbbke-Doerr, Audio im Sixpack, c't magazin für computer und technik 13/2000:,Verlag Heinz Heise GmbH &Co KG, Hannover, 2000

[c't2699] Berghorn W., Boskamp T., Jung K.: Schlanke Bilder, Der zukünftige Bildkompressionsstandard JPEG 2000, c't magazin für computer und technik 26/1999:, Verlag Heinz Heise GmbH &Co KG, Hannover, 1999

[FIE99] Fiebig M., Untersuchung der Online-Videoübertragung unter Windows NT über Dual-Video-Systeme, Diplomarbeit an der Fakultät für Mathematik und Informatik der Universität Leipzig, Leipzig, 1999

[FORA] FA – 30P Time Base Corrector Operation Manual, FOR-A Company Limited, Japan

[ISO10918] ISO/IEC 10918: Digital Compression and Coding of Continuous Still Images (JPEG), Genf, 1992

[ISO11172] ISO/IEC 11172: Information Technology – Coding of moving pictures and associated audio for digital storage media up to 1,5 mbit/s (MPEG-1), Genf, 1993

[ISO13818] ISO/IEC 13818: Information Technology – Generic coding of moving pictures and associated audio (MPEG-2), Genf, 1994

[JPEG] www.jpeg.org, Onlineauftritt der Joint Photographic Experts Group

[JVC1] Bedienanleitung des VCR BR-S522E der Firma JVC, JVC Japan, (Jahr)

[JVC2] Bedienanleitung des Monitors TM-1500PS der Firma JVC, JVC Japan, (Jahr)

[KYA96] Kyas, O.: ATM - Netzwerke Aufbau – Funktion - Performance,
DATACOM Buchverlag, Bergheim 1996 (4)

[MAI2.05] Mainactor Video Editor V2.05 der Firma Main Concept GbR Moemig/Zabel

[MEI94] Meissner, H.: Digitale Multimediasysteme, Verlag Technik 1994

[MEN00] Menn, Ralph, JPEG 2000: Das neue Format für Bilder
(www.tecchannel.de/multimedia/67/6.html), Onlineauftritt der Zeitung tecchannel vom
IDG Communications Verlag AG, 28.11.2000 (Update am 09.08.2001)

[MIL95] Milde, T.: Videokompressionsverfahren im Vergleich. JPEG, MPEG, H.261,
Waveletts, Fraktale, dpunkt - Verlag, Heidelberg, 1995

[MIT96] Mitchel, J. L., Pennebaker, W. B., Fogg, C. E., LeGall, D.J.: MPEG Video
Compression Standart, Chapman & Hall, New York 1996

[MPEG] www.cselt.it/mpeg/index.html, Internetauftritt der Moving Picture Experts
Group, 2000/2001

[MSD97] Microsoft Developer Studio 97, Microsoft Inc 1997

[Network] www.networkworld.de, Internetauftritt der Zeitung networkworld mit einem
Begriffs-Erklärregister der Informatik

[Ohm95] Ohm, Jens-Rainer: Digitale Bildcodierung: Repräsentation, Kompression
und Übertragung von Bildsignalen, Springer 1995

[OPT] www.optibase.de, Internetauftritt der Firma Optibase Inc., 2000

[ORZ98] Orzessek, M., Sommer, P.: ATM & MPEG2: integrating digital video into broadband networks, Hewlett-Packard Professional Books, New Jersey 1998

[PRY96] de Prycker, M.: Asynchronous Transfer Modus - Die Lösung für Breitband - ISDN, München, 1996

[SCH97] Schill, A., Hess, R., Kümmel, S., Hege, D., Lieb, H.: ATM - Netze in der Praxis, Einsatzszenarien, Produktkategorien, aktuelle Standards, ADDISON-WESLEY-LONGMAN, Bonn, 1997

9.2 Verzeichnis der Abbildungen, Tabellen und Gleichungen

9.2.1 Abbildungen

Abbildung 1-1 Beispielszenario einer Dialoganwendung	1-6
Abbildung 2-1 Würfeldarstellung des RGB-Farbsystem.....	2-10
Abbildung 3-1 Zuordnung der Farb- und Helligkeitskomponenten vor und nach der Farbraumkonvertierung	3-21
Abbildung 3-2 Kantenglättung bei der Filterung	3-22
Abbildung 3-3 Abtastung der RGB-Komponenten bei der Digitalisierung	3-23
Abbildung 3-4 Abtastung der Helligkeits- und Farbwerte bei der Digitalisierung....	3-23
Abbildung 3-5 Veranschaulichung der Frequenzzunahme über der Hauptdiagonalen (übernommen aus [MIL95] Seite 20)	3-27
Abbildung 3-6 Darstellung eines 8x8 Bildes mit 64 Basisbildern $\{B(k,n)\}$ (nach [MIL95]).....	3-28
Abbildung 3-7 Anschauliche Darstellung eines 8x8 Bildes mit 64 Basisbildern (nach [c't2699]).....	3-29
Abbildung 3-8 Bedeutung der DCT-Koeffizienten, Unterteilung in wichtige (blau markiert) und unwichtige Werte (nach [MEI94]).....	3-30
Abbildung 3-9 Bearbeitungsreihenfolge der DCT-Koeffizienten.....	3-31
Abbildung 3-10 Darstellung der Blockbildung bei geringer Komprimierung	3-39
Abbildung 3-11 Darstellung der Blockbildung bei großer Komprimierung, es sind deutlich Artefakte zu sehen.....	3-39
Abbildung 3-12 Schritte bei der MPEG 1 Komprimierung	3-41
Abbildung 3-13 Ausnutzen räumlicher und zeitlicher Redundanzen bei P-Bildern	3-42
Abbildung 3-14 Ausnutzen zeitlicher und räumlicher Redundanzen der vorhergehenden und nachfolgenden I- oder P-Bilder bei B-Bildern	3-43
Abbildung 3-15 GOP einer MPEG komprimierten Bilderfolge	3-43
Abbildung 3-16 MPEG kodiertes Video als Abfolge von GOP's dargestellt	3-44
Abbildung 3-17 Zeitliche-Scalability	3-50
Abbildung 3-18 Qualitäts-Scalibility.....	3-51
Abbildung 4-1 Generelles Blockschaltbild eines MPEG 1 Audiokodierers	4-56
Abbildung 5-1 MPEG FORGE von Optibase Inc.(entnommen aus [OPT])	5-61
Abbildung 5-2 MPEG Fusion von Optibase Inc. (entnommen aus [OPT])	5-61

Abbildung 5-3 Die Komponenten im Zusammenspiel	5-68
Abbildung 5-4 Bedienungsansicht der Encoder Software	5-70
Abbildung 5-5 Einstellungen des Encoders, Registerkarte eins, drei und vier	5-71
Abbildung 5-6 MPEG Player mit File Information Registerkarte	5-72
Abbildung 5-7 Die Fernsteuersoftware.....	5-73
Abbildung 5-8 GUI des TMT Time Measure Tools.....	5-75
Abbildung 5-9 Einteilung der GUI in funktionelle Bereiche.....	5-76
Abbildung 5-10 Infofenster des TMT.....	5-76
Abbildung 5-11 Das GUI des PSNR-Tools	5-80
Abbildung 6-1 Originalbild, 10.000 kbit/s I/P=12/3 (keine Artefakte)	6-90
Abbildung 6-2 Bild der Note 2, 3.000 kbit/s I/P=90/1 (leichte Artefakte am oberen Ohransatz).....	6-90
Abbildung 6-3 Bild der Note 6, 3.000 kbit/s I/P=1/1 (starke Artefakte am Ohr)	6-91
Abbildung 6-4 Bild der Note 10, 2.000 kbit/s I/P=3/3 (starke Artefakte im ganzen Bild)	6-91
Abbildung 6-5 Schwankung der durchschnittliche Framegröße bei Veränderung der GOP (MPEG2 2000 kbit/s).....	6-92
Abbildung 6-6 Schwankung der durchschnittliche Bildgröße bei Veränderung der GOP (MPEG1 3000 kbit/s).....	6-93
Abbildung 6-7 PSNR Entwicklung I=12 P=3 (MPEG1 2000kbit/s)	6-96
Abbildung 6-8 PSNR- Werte für MPEG 1, 760,4 kbit/s, (I = 12, P = 1-3)	6-97
Abbildung 6-9 PSNR- Werte für MPEG 1, 760,4 kbit/s, (I = 3-90, P = 3)	6-98
Abbildung 6-10 PSNR- Werte für MPEG 1, 760,4 kbit/s, (I = 12-90, P = 1)	6-98
Abbildung 6-11 PSNR- Werte für MPEG 1, 2000 kbit/s, (I = 1-30, P = 1)	6-99
Abbildung 6-12 PSNR- Werte für MPEG 1, 2000 kbit/s, (I = 3-90, P = 3)	6-99
Abbildung 6-13 PSNR- Werte für MPEG 1, 2000 kbit/s, (I = 12, P = 1-3)	6-100
Abbildung 6-14 PSNR- Werte für MPEG 1, 3000 kbit/s, (I = 12, P = 1-3)	6-100
Abbildung 6-15 PSNR- Werte für MPEG 1, 3000 kbit/s, (I = 3-90, P = 3)	6-101
Abbildung 6-16 PSNR- Werte für MPEG 1, 3000 kbit/s, (I = 1-30, P = 1)	6-101
Abbildung 6-17 PSNR- Werte für MPEG 1, 3000 kbit/s, (I = 6, P = 1-3)	6-102
Abbildung 6-18 PSNR- Werte für MPEG 1, 3000 kbit/s, (I = 6-12, P = 2)	6-102
Abbildung 6-19 PSNR- Werte für MPEG 2, 2000 kbit/s, (I = 12, P = 1-3)	6-103
Abbildung 6-20 PSNR- Werte für MPEG 2, 2000 kbit/s, (I = 3-90, P = 3)	6-103
Abbildung 6-21 PSNR- Werte für MPEG 2, 2000 kbit/s, (I = 1-100, P = 1)	6-104

Abbildung 6-22 PSNR- Werte für MPEG 2, 3000 kbit/s, (I = 3-100, P = 1)	6-104
Abbildung 6-23 PSNR- Werte für MPEG 2, 3000 kbit/s, (I = 12, P = 1-3)	6-105
Abbildung 6-24 PSNR- Werte für MPEG 2, 3000 kbit/s, (I = 3-90, P = 3)	6-105
Abbildung 7-1 Überkomprimiertes JPEG (links) bei 70:1 und überkomprimiertes Wavelet Kodiertes Bild (rechts) bei 150:1 [MEN00]	7-111
Abbildung 9-1 Die Struktur der CDs	9-127

9.2.2 Tabellen

Tabelle 2-1 Eigenschaften digitaler Bildformate (auszugsweise aus [Ohm95])	2-9
Tabelle 2-2 Vergleich des Datenaufkommens eines Schwarz-Weiß-Standbildes, eines RBG Standbildes und eines 1s Videostreams in gleicher Auflösung (EQTV).....	2-12
Tabelle 2-3 Dauer der Übertragung von 2,8Mbyte über verschiedene Netzwerkbandbreiten (abgewandelt übernommen aus [KYA96] Seite 31)	2-13
Tabelle 3-1 Gliederung der Kompressionsverfahren, aus [MIL95].....	3-18
Tabelle 3-2 Kategorisierung von Kodierungstechniken, aus [MIL95]	3-19
Tabelle 3-3 Blockbildung für die diskrete Kosinus-Transformation	3-26
Tabelle 3-4 Ergebnismatrix nach Anwendung der DCT (Integerwerte).....	3-27
Tabelle 3-5 Quantisierungsmatrix mit dem Qualitätsfaktor 2	3-32
Tabelle 3-6 DCT-Koeffizienten nach der Quantisierung.....	3-33
Tabelle 3-7 Profile und Level Organisation in MPEG 2 Video nach [MIL95].....	3-49
Tabelle 4-1 Tonformate auf DVD Medien [c't1300].....	4-59
Tabelle 5-1 Technical Specifications des MPEG FORGE Audio-Video Systems der Firma Optibase im Vergleich mit dem MPEG FUSION System vom selben Hersteller [OPT].....	5-63
Tabelle 5-2 Funktionsbeschreibung der Bereiche im GUI des TMT.....	5-77
Tabelle 5-3 mögliche Audiobitraten.....	5-79
Tabelle 6-1 Messungen der Kompressionszeit.....	6-83
Tabelle 6-2 Video Parameter	6-83
Tabelle 6-3 MPEG 1-SIF mit 760,4 kbit/s /Qualitätsvergleich	6-86
Tabelle 6-4 MPEG 1-SIF mit 2000 kbit/s /Qualitätsvergleich	6-87
Tabelle 6-5 MPEG 1-SIF mit 3000 kbit/s /Qualitätsvergleich	6-87
Tabelle 6-6 MPEG 2 Half D-1 mit 2000 kbit/s /Qualitätsvergleich	6-88
Tabelle 6-7 MPEG 2 Half D-1 mit 3000 kbit/s /Qualitätsvergleich	6-88
Tabelle 6-8 Festlegen von Qualitätsstufen.....	6-89
Tabelle 6-9 Einteilung der Diagramme anhand der Neigung der Graphen	6-106
Tabelle 7-1 Anforderungen an JPEG 2000 nach [c't2699].....	7-110
Tabelle 7-2 Teile und Status des JPEG 2000 Standards [JPEG].....	7-112
Tabelle 9-1 MPEG 1-SIF mit 3385,2 kbit/s.....	9-123
Tabelle 9-2 MPEG 1-SIF mit 1000 kbit/s.....	9-123

Tabelle 9-3 MPEG 1-SIF mit 760,4 kbit/s.....	9-124
Tabelle 9-4 MPEG 1-SIF mit 2000 kbit/s.....	9-124
Tabelle 9-5 MPEG 1-SIF mit 3000 kbit/s.....	9-125
Tabelle 9-6 MPEG 2 Half D-1 mit 2000 kbit/s	9-125
Tabelle 9-7 MPEG 2 Half D-1 mit 3000 kbit/s	9-126
Tabelle 9-8 Vergleichsbilder MPEG 2 Half D-1 mit 10000 kbit/s.....	9-126
Tabelle 9-9 Vergleichsbilder MPEG 1-SIF mit 5000 kbit/s	9-126
Tabelle 9-10 Erklärung der auf den CDs gespeicherten Daten.....	9-127

9.2.3 Gleichungen

Gleichung 2-1 Umwandlung von RGB-Farbwerten in das innerhalb Europas übliche PAL und SECAM YUV-Komponentenmodell	2-11
Gleichung 2-2 Berechnung der Datenmenge eines Bildes, wenn mittels eines Luminanzvektors und zwei Farbvektoren digital gespeichert wird	2-11
Gleichung 2-3 Datenrate einer Bilderfolge in Mbit/s.....	2-11
Gleichung 3-1 Umwandlung der RGB Werte ins YUV- Farbmodell	3-20
Gleichung 3-2 Die zweidimensionale DCT	3-26
Gleichung 3-3 Bearbeitungsreihenfolge der DCT-Koeffizienten.....	3-30
Gleichung 3-4 Die inverse zweidimensionale DCT	3-33
Gleichung 5-1 Peak Signal to Noise Ratio, PSNR	5-81
Gleichung 6-1 Berechnung der Anzahl der erzeugten Frames pro Zeiteinheit	6-82
Gleichung 6-2 Errechnung des Overheads der erzeugten Audio-Video Streams ..	6-94

9.3 Messergebnisse im tabellarischem Überblick

Tabelle 9-1 MPEG 1-SIF mit 3385,2 kbit/s

Titel	Video Bitrate in kbit/s	I-Pics Dist	P-Pics Dist	Datei-Größe in bytes	Durchschn. Bildgröße (Byte)	Bilder	Bilder * Bildergröße	Größe - (Bilder * Bildergröße)
4	3000	3	3	8708960	16875	501	8454375	254585
3	3000	6	3	8708960	16870	501	8451870	257090
2	3000	9	3	8708960	16870	501	8451870	257090
7	3000	12	1	8686080	16853	501	8443353	242727
6	3000	12	2	8686080	16831	503	8465993	220087
1	3000	12	3	8708960	16869	501	8451369	257591
5	3000	18	3	8708960	16866	501	8449866	259094

Tabelle 9-2 MPEG 1-SIF mit 1000 kbit/s

Titel	Video Bitrate in kbit/s	I-Pics Dist	P-Pics Dist	Datei-Größe in bytes	Durchschn. Bildgröße (Byte)	Bilder	Bilder * Bildergröße	Größe / (Bilder * Bildergröße)
11	728,4	3	3	2575040	4959	501	2484459	90581
10	728,4	6	3	2575040	4966	501	2487966	87074
9	728,4	9	3	2575040	4966	501	2487966	87074
14	728,4	12	1	2560480	4962	501	2485962	74518
13	728,4	12	2	2568800	4947	502	2483394	85406
8	728,4	12	3	2575040	4964	501	2486964	88076
15	728,4	18	1	2585440	4954	501	2481954	103486
12	728,4	18	3	2575040	4960	501	2484960	90080

Tabelle 9-3 MPEG 1-SIF mit 760,4 kbit/s

Titel	BitRate in kbit/s	I-Pics Dist	P-Pics Dist	Datei-Größe in bytes	Durchschn. Bildgröße (Byte)	Bilder	Bilder * Bildergröße	Größe - (Bilder * Bildergröße)	PSNR
19	760,4	3	3	1959360	3790	501	1898790	60570	19,751
18	760,4	6	3	1959360	3811	501	1909311	50049	19,223
17	760,4	9	3	1959360	3789	501	1898289	61071	19,124
22	760,4	12	1	1946880	3758	501	1882758	64122	19,322
21	760,4	12	2	1955200	3773	502	1894046	61154	19,180
16	760,4	12	3	1959360	3788	501	1897788	61572	18,215
23	760,4	18	1	1946880	3756	501	1881756	65124	19,255
20	760,4	18	3	1959360	3756	501	1881756	77604	19,171
25	760,4	30	1	1955200	3754	501	1880754	74446	19,328
26	760,4	60	1	1946880	3757	501	1882257	64623	19,305
27	760,4	90	1	1955200	3754	501	1880754	74446	19,456
28	760,4	30	3	1959360	3773	501	1890273	69087	19,133
24	760,4	30	3	1959360	3768	501	1887768	71592	19,290
29	760,4	60	3	1959360	3764	501	1885764	73596	19,276
30	760,4	90	3	1959360	3778	501	1892778	66582	19,093

Tabelle 9-4 MPEG 1-SIF mit 2000 kbit/s

Titel	Video Bitrate in kbit/s	I-Pics Dist	P-Pics Dist	Datei-Größe in bytes	Durchschn. Bildgröße (Byte)	Bilder	Bilder * Bildergröße	Größe - (Bilder * Bildergröße)	PSNR
61	1680,8	1	1	5133440	10053	501	5036553	96887	21,298
62	1680,8	2	1	5114720	10032	501	5026032	88688	21,373
63	1680,8	3	1	5114720	10052	501	5036052	78668	21,162
34	1680,8	3	3	5145920	10053	501	5036553	109367	21,351
64	1680,8	6	1	5114720	10053	501	5036553	78167	21,393
33	1680,8	6	3	5114720	10055	501	5037555	108365	21,333
44	1680,8	9	1	5114720	10062	502	5051124	63596	21,393
42	1680,8	9	1	5114720	10064	501	5042064	72656	21,170
32	1680,8	9	3	5145920	10069	501	5044569	101351	21,325
43	1680,8	10	2	5145920	10063	502	5051626	94294	21,061
45	1680,8	12	1	5114720	10050	502	5045100	69620	21,174
41	1680,8	12	1	5133440	10047	501	5033547	99893	21,336
40	1680,8	12	2	5133440	10048	502	5044096	89344	21,047
31	1680,8	12	3	5145920	10054	501	5037054	108866	21,400
36	1680,8	18	3	5145920	10050	501	5035050	110870	21,220
65	1680,8	30	1	5114720	10048	501	5034048	80672	21,379
37	1680,8	30	3	5145920	10050	501	5035050	110870	21,215
38	1680,8	60	3	5145920	10051	501	5035551	110369	21,255
39	1680,8	90	3	5145920	10053	501	5036553	109367	20,582

Tabelle 9-5 MPEG 1-SIF mit 3000 kbit/s

Titel	Video Bitrate in kbit/s	I-Pics Dist	P-Pics Dist	Datei-Größe in bytes	Durchschn. Bildgröße (Byte)	Bilder	Bilder * Bildgröße	Größe - (Bilder * Bildgröße)	PSNR
85	26332	1	1	7668960	14926	501	7477926	191034	27,344
102	26332	2	1	7698080	15133	501	7581633	116447	27,904
83	26332	3	1	7668960	14935	501	7482435	186525	27,811
74	26332	3	3	7716800	14923	501	7476423	240377	27,754
84	26332	6	1	7668960	14944	501	7486944	182016	27,281
86	26332	6	2	7716800	15142	502	7601284	115516	27,354
73	26332	6	3	7716800	14946	501	7487946	228854	27,582
81	26332	9	1	7668960	14949	501	7489449	179511	26,226
72	26332	9	3	7716800	14956	501	7492956	223844	27,482
82	26332	10	2	7716800	15170	501	7600170	116630	27,363
80	26332	12	1	7668960	14930	501	7479930	189030	26,351
79	26332	12	2	7716800	15135	502	7597770	119030	27,423
71	26332	12	3	7716800	14947	501	7488447	228353	27,564
75	26332	18	3	7716800	14933	501	7481433	235367	27,586
103	26332	30	1	7668960	14937	501	7483437	185523	26,439
76	26332	30	3	7716800	14933	501	7481433	235367	27,581
77	26332	60	3	7716800	14931	501	7480431	236369	27,387
78	26332	90	3	7716800	14937	501	7483437	233363	27,302

Tabelle 9-6 MPEG 2 Half D-1 mit 2000 kbit/s

Titel	Video Bitrate in kbit/s	I-Pics Dist	P-Pics Dist	Datei-Größe in bytes	Durchschn. Bildgröße (Byte)	Bilder	Bilder * Bildgröße	Größe - (Bilder * Bildgröße)	PSNR
60	1680,8	1	1	5114720	10101	501	5060601	54119	24,259
59	1680,8	2	1	5133440	9844	501	4931844	201596	23,962
58	1680,8	3	1	5133440	9826	501	4922826	210614	23,976
46	1680,8	3	3	5145920	9818	502	4928636	217284	21,482
57	1680,8	6	1	5114720	9807	501	4913307	201413	23,937
47	1680,8	6	3	5145920	9814	502	4926628	219292	21,543
56	1680,8	9	1	5133440	9800	501	4909800	223640	23,947
48	1680,8	9	3	5145920	9804	502	4921608	224312	21,554
55	1680,8	12	1	5133440	9803	501	4911303	222137	23,956
54	1680,8	12	2	5145920	9794	502	4916588	229332	21,676
49	1680,8	12	3	5145920	9802	502	4920604	225316	21,554
66	1680,8	18	1	5133440	9793	501	4906293	227147	23,921
50	1680,8	18	3	5145920	9793	502	4916086	229834	21,562
67	1680,8	30	1	5114720	9807	501	4913307	201413	23,945
51	1680,8	30	3	5145920	9795	502	4917090	228830	21,525
68	1680,8	60	1	5114720	9812	501	4915812	198908	23,918
52	1680,8	60	3	5145920	9805	502	4922110	223810	21,557
69	1680,8	90	1	5133440	9790	501	4904790	228650	23,889
53	1680,8	90	3	5145920	9792	502	4915584	230336	21,569
70	1680,8	100	1	5114720	9790	501	4904790	209930	23,931

Tabelle 9-7 MPEG 2 Half D-1 mit 3000 kbit/s

Titel	Video Bitrate in kbit/s	I-Pics Dist	P-Pics Dist	Datei-Größe in bytes	Durchschn. Bildgröße (Byte)	Bilder	Bilder * Bildgröße	Größe - (Bilder * Bildgröße)	PSNR
101	26332	1	1	7668960	14901	501	7465401	203559	19,490
100	26332	2	1	7668960	14878	501	7453878	215082	19,498
99	26332	3	1	7668960	14892	501	7460892	208068	19,537
87	26332	3	3	7716800	14894	501	7461894	254906	19,634
98	26332	6	1	7698080	14898	501	7463898	234182	19,519
88	26332	6	3	7716800	14894	501	7461894	254906	19,598
97	26332	9	1	7668960	14898	501	7463898	205062	20,724
89	26332	9	3	7716800	14896	501	7462896	253904	19,603
96	26332	12	1	7668960	14886	501	7457886	211074	19,483
95	26332	12	2	7716800	14879	502	7469258	247542	20,782
90	26332	12	3	7716800	14889	501	7459389	257411	19,571
104	26332	18	1	7668960	14889	501	7459389	209571	19,039
91	26332	18	3	7716800	14803	501	7416303	300497	19,593
105	26332	30	1	7668960	14889	501	7459389	209571	18,350
92	26332	30	3	7716800	14886	501	7457886	258914	19,546
106	26332	60	1	7689080	14896	501	7462896	226184	18,360
93	26332	60	3	7716800	14900	501	7464900	251900	19,566
107	26332	90	1	7668960	14890	501	7459890	209070	18,326
94	26332	90	3	7716800	14885	501	7457385	259415	19,539
108	26332	100	1	7689080	14892	501	7460892	228188	18,128

Tabelle 9-8 Vergleichsbilder MPEG 2 Half D-1 mit 10000 kbit/s

Titel	Video Bitrate in kbit/s	I-Pics Dist	P-Pics Dist	Datei-Größe in bytes	Durchschn. Bildgröße (Byte)	Bilder	Bilder * Bildgröße	Größe - (Bilder * Bildgröße)
109	9300	12	3	25712960	49818	501	24958818	754142

Tabelle 9-9 Vergleichsbilder MPEG 1-SIF mit 5000 kbit/s

Titel	Video Bitrate in kbit/s	I-Pics Dist	P-Pics Dist	Datei-Größe in bytes	Durchschn. Bildgröße (Byte)	Bilder	Bilder * Bildgröße	Größe - (Bilder * Bildgröße)
110	4538	12	3	12858560	24895	501	12472395	386165

9.4 Anlagen

Als Anlagen wurden 4 CDs beigefügt. Auf den CDs sind die aufgenommenen Filme, CD3 und CD4, und die daraus extrahierten Bilder, CD1 und CD2, gespeichert.

Abbildung 9-1 Die Struktur der CDs

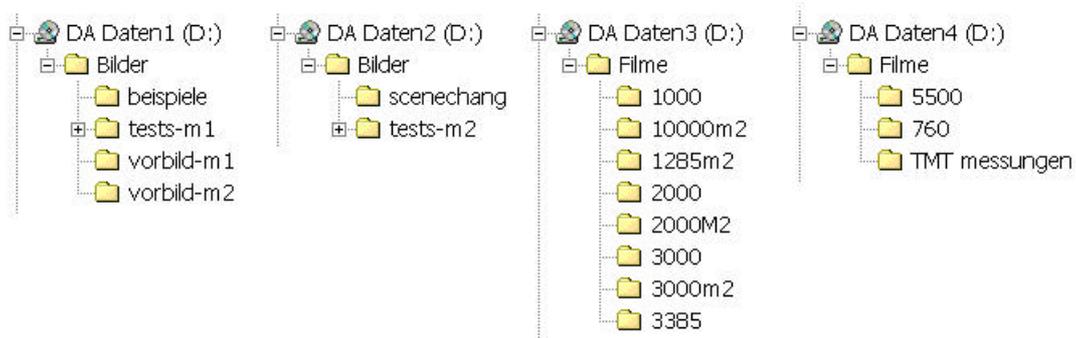


Tabelle 9-10 Erklärung der auf den CDs gespeicherten Daten

CD	Verzeichnis	Erklärung
1	Bilder/beispiele	Beispielbilder für die Notenvergabe im Kapitel Fehler! Verweisquelle konnte nicht gefunden werden.
	Bilder/tests-m1	extrahierte Bilder der mit MPEG 1 aufgenommenen Filme
	Bilder/vorbild-m1	alle Bilder des MPEG 1 Vorbildfilmes
	Bilder/vorbild-m2	alle Bilder des MPEG 2 Vorbildfilmes
2	Bilder/scenechange	extrahierte Bilder des Filmes für den Scenechange- Test
	Bilder/tests-m1	extrahierte Bilder der mit MPEG 2 aufgenommenen Filme
3	Filme/1000	Filme MPEG 1, 1000 kbit/s
	Filme/1000m2	Filme MPEG 2, 1000 kbit/s
	Filme/1285m2	Filme MPEG 2, 1285 kbit/s
	Filme/2000	Filme MPEG 1, 2000 kbit/s
	Filme/2000m2	Filme MPEG 2, 2000 kbit/s
	Filme/3000	Filme MPEG 1, 3000 kbit/s
	Filme/3000m2	Filme MPEG 2, 3000 kbit/s
	Filme/3385	Filme MPEG 1, 3385 kbit/s
4	Filme/5500	Filme MPEG 1, 5500 kbit/s
	Filme/760	Filme MPEG 1, 760 kbit/s
	Filme/TMT Messungen	Filme für die Messungen im Kapitel 6.1

10 Danksagung

An Optibase

An Prof. Dr. Imscher

An Dr. Hänßgen

...

11 Erklärung

Ich versichere, dass ich die vorliegende Arbeit selbständig und nur unter Verwendung der angegebenen Quellen und Hilfsmittel angefertigt habe.