

Adaptive Verfahren der Datenratenkontrolle für anwendungs- und netzübergreifende Videocodierung

Markus Zeller, Mike Heidrich

Fraunhofer Einrichtung für Systeme der Kommunikationstechnik (ESK), HansasträÙe 32, 80686 München

Email: markus.zeller@esk.fhg.de bzw. mike.heidrich@esk.fhg.de

Telefon: + 49 89 547088-352 bzw. + 49 89 541088-377

Kurzfassung

Neben der Steigerung der Codiereffizienz ist bei der Standardisierung aktueller Videocodierverfahren, wie z.B. H.263++ oder MPEG-4, berücksichtigt worden, dass die Codecs nicht mehr fest auf eine einzelne Zielapplikation ausgerichtet sind. Sie sind darauf ausgelegt, unterschiedlichste Anwendungen und Netze zu unterstützen. Das Spektrum reicht von niederbitratiger Videokommunikation über mobile Netze bis hin zu hochqualitativem Video über Breitbandnetze. Dadurch ergeben sich neue Anforderungen an die Steuerung der Videocodecs, um stets eine effektive Arbeitsweise sicherzustellen. Für die Datenratenkontrolle, als wichtiges qualitätsbestimmendes Element des Encoders, sind hierfür neue Ansätze notwendig, um unterschiedlichstes Bildmaterial mit jeweils passendem Ende-zu-Ende-Delay über zeitvariante mobile Netze mit mehreren Kilobits pro Sekunde oder aber über hochbitratige Festverbindungen effektiv übertragen zu können. In diesem Beitrag werden die sich daraus ergebenden Anforderungen aufgezeigt und darauf aufbauend wird ein einfaches adaptives Verfahren für die Datenratenkontrolle entworfen. Neben dem Vergleich anhand einer H.263-Implementierung mit den Verfahren aus der MPEG-4 und H.263+ Standardisierung wird auf Implementierungsfragen der verschiedenen Datenratenkontrollen eingegangen.

1 Einleitung

Der Bedarf an Mobilität und mehr Komfort für den Anwender sowie die Forderung nach ständiger Erreichbarkeit bzw. der Möglichkeit sich überall Informationen zu beschaffen, stellen die treibenden Kräfte für die steigende Marktrelevanz der audiovisuellen Kommunikation dar.

Bisher scheiterte eine vernünftige Übertragung von Videoinhalten oftmals an der zur Verfügung stehenden Bandbreite der Kommunikationsnetze. Durch die Einführung von höherbitratigen Netzzugängen, wie UMTS oder ADSL, und Verbesserungen der Videocodierverfahren selbst, sind die Grundlagen geschaffen, um qualitativ höherwertige Videokommunikation über unterschiedlichste Netze zu ermöglichen.

Im folgenden werden kurz wichtige Aspekte aktueller Verfahren der Videocodierung beschrieben und anhand von realistischen Anwendungsszenarios die Anforderungen an adaptive Datenratenkontrollen hergeleitet. Nach der Beschreibung der grundlegenden Arbeitsweise der Verfahren für die Datenratenkontrolle, wird auf wichtige Lösungsansätze näher eingegangen. Auf Basis der Anforderungen, wie sie in zukünftigen Kommunikationsszenarios mit zum Teil variierenden

Randbedingungen auftreten, wird ein adaptives Verfahren entworfen und anhand von Simulationen mit existierenden Verfahren verglichen. Erste Ergebnisse versprechen sowohl einen signifikanten Qualitätsgewinn über einen breiten Anwendungsbereich, unter Einhaltung der vorgegebenen Ende-zu-Ende-Verzögerung, als auch die rasche Adaption auf variierenden Randbedingungen, wie sie insbesondere in mobilen Netzen anzutreffen sind.

2 Aktuelle Verfahren der Videocodierung

Die Standardisierung eines Videocodecs soll gewährleisten, dass verschiedene Geräte miteinander kommunizieren können. Zusätzlich sollen die Standards offen sein, um z.B. Qualitätsverbesserungen durch technische Neuerungen zu erlauben. Aus diesem Grund liegt innerhalb der Standardisierung der Schwerpunkt auf der Definition des Decoders und des zu übertragenden Bitstroms, den der Decoder interpretiert und damit das codierte Videosignal rekonstruiert.

In der Vergangenheit wurden für jeden Einsatzbereich jeweils spezialisierte audiovisuelle Terminals standar-

disiert, so z.B. H.261 für Videokonferenzanwendungen und MPEG-1 für Video-CDs. Vor dem Hintergrund des raschen Zusammenwachsens der Telekommunikationsindustrie, der Computerindustrie und der TV-/Film-Wirtschaft, haben aktuelle und vor allem zukünftige Standardisierungsbestrebungen, wie MPEG-4 [3], H.263++ [4][7] oder H.26L [5], die Schaffung eines flexiblen Standards für unterschiedlichste, auch interaktive Multimediaanwendungen und -dienste zum Ziel. Darüber hinaus wird die Übertragung über unterschiedlichste Netze berücksichtigt.

Bei der Standardisierung von MPEG-4, aber auch in den Weiterentwicklungen von H.263, wie H.263+/, haben sehr umfangreiche Untersuchungen von unterschiedlichen Videocodiermethoden gezeigt, dass der bisher eingesetzte hybride DCT/DPCM-Core [8] auf Basis von H.263, wie er im Prinzip bereits bei früheren Standardisierungen, wie MPEG-1/2 oder H.261 verwendet wurde, einen sehr guten Kompromiss zwischen Aufwand und effektiven Nutzen darstellt. In aktuellen Codecs wird der Core durch neue Verfahren bzw. Werkzeuge/Tools an unterschiedliche Einsatzbereiche angepasst bzw. optimiert, so z.B. verbesserte Intraframecodierung oder zusätzliche Bewegungsmodelle. Ebenso spielen effektive Preprocessingverfahren, wie Global Motion Compensation oder die Unterstützung von Sprites eine wichtige Rolle.

3 Anwendungsszenarios und deren Anforderungen

Die Einsatzmöglichkeiten der flexiblen Codecs sollen beispielhaft an drei Anwendungsszenarios dargestellt werden:

- Mobile Endgeräte, wie z.B. ein PDA oder ein WebPad, sollen über unterschiedliche Netzanbindungen eine Videokommunikation realisieren (siehe Abbildung 1). Im mobilen Betrieb ist die Anbindung z.B. über UMTS oder im lokalen Netz über IEEE 802.11 vorgesehen bzw. im stationären Betrieb über LAN. Bedingt durch die unterschiedlichen Netzanbindungen soll und muss der Videocodec im Gerät je nach aktueller Situation sowohl mit unterschiedlichsten, sowie bei mobiler Anbindung auch mit stark variierenden, Netzdatenraten zurechtkommen.

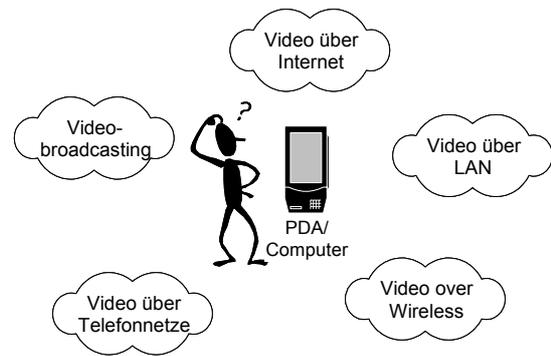


Abbildung 1: Kommunikationsanwendungen über unterschiedliche Netze

- Während es bisher üblich war, für jede Anwendung einen speziell optimierten und fest mit ihr verbundenen Videocodec zu entwickeln, wird die Codierung zunehmend als Basisdienst in Betriebssysteme integriert und kann damit von unterschiedlichen Anwendungen verwendet werden, so z.B. für eine Videokommunikation über ISDN und später als Codec für Videomails oder für die Aufzeichnung von TV-Programmen.
- Ein Videocodec soll für unterschiedliche Zielplattformen und -anwendungen realisiert werden. Hier bietet es sich an, die Implementierung soweit wie möglich anwendungs- und plattformunabhängig vorzunehmen, um sie dann erst in einem späteren Arbeitsschritt mit minimalem Aufwand an die unterschiedlichen Anwendungsbereiche anzupassen, ohne die Arbeitsweise für jede Applikation neu optimieren zu müssen.

Anhand dieser Beispiele werden der Bedarf und die neuen Möglichkeiten von „universellen“ Videostandards offensichtlich. Die drei wesentlichen Komponenten Anwendung, Videocodec und Netz eines modernen Kommunikationssystems stehen dabei in folgender Beziehung:

- Die Anwendung definiert zum einen die Anforderungen an das Videocodierungssystem, d.h. den Codec und das darunterliegende Transportnetz, wie z.B. Qualitätsanforderungen, akzeptables Ende-zu-Ende-Delay und zum anderen die Charakteristik des Eingangssignals, wie Anzahl und Geschwindigkeiten von bewegten Bildinhalten, Detailreichtum oder Häufigkeit von Szenenwechseln.
- Die Videocodierung stellt der Anwendung zahlreiche Funktionalitäten, wie z.B. Zugriffsmöglichkeit auf Objekte oder Streamingfähigkeit zur Verfügung. Zugleich hat die Videocodierung die

Aufgabe, das Eingangssignal effektiv zu codieren und ohne die Vorgaben des Gesamtsystems zu verletzen, so z.B. Delayanforderungen oder Vorgaben für die Datenrate. Für eine effektive Videoübertragung ist es notwendig, dass das Codierverfahren auch an die Eigenschaften des Übertragungs- oder Speichermediums angepasst wird. Eine "netzfreundliche" Quellencodierung unterstützt z.B. vereinfachtes Paketieren, erweiterte Schnittstellen für priorisierende bzw. zeitvariante Netze, Skalierbarkeit und effektiven Fehlerschutz.

- Das Netz ist für die eigentliche Übertragung der codierten Videosignale zuständig. Die Eigenschaften des Übertragungsnetzes sind von entscheidender Bedeutung für die Auswahl und die Optimierung geeigneter Videocodierverfahren bzw. von Fehlerschutzstrategien. In einigen Netzen, wie z.B. ATM oder in modernen IP-Netzen, ist es möglich, bereits vor der eigentlichen Übertragung Quality-of-Service-Garantien, wie Datenrate oder Fehlerrate auszuhandeln, d.h. die Übertragungscharakteristik des digitalen Netzes kann an die Anforderungen der Anwendung angepasst werden.

Für die Auslegung des Videoencoders, dessen Arbeitsweise die Effizienz der Quellencodierung und die tatsächliche Ende-zu-Ende-Verzögerung wesentlich festlegt, ergeben sich dadurch viele Einstellmöglichkeiten. Gleichzeitig entfällt die bisher praktizierte Möglichkeit, den Codec anhand fest vorgegebener Bedürfnisse im voraus auf feste Parameter zu optimieren. Vielmehr muss unterschiedliches Bildmaterial teilweise sogar im laufenden Betrieb mit einem einzigen Codec effektiv verarbeitet werden können. Gleichzeitig muss dabei die Übertragung über die unterschiedlichen, teilweise zeitvarianten Netze in möglichst hoher Qualität sichergestellt werden. Dadurch ergeben sich insbesondere bei der Codecsteuerung und der Datenratenkontrolle, auf die im folgenden näher eingegangen wird, neue Herausforderungen.

4 Lösungsansätze für die Datenratenkontrolle

4.1 Aufgaben und Arbeitsweise der Datenratenkontrolle

In Abbildung 2 ist das Grundkonzept von Videocodierverfahren dargestellt. Hierbei wird das Bildsignal nach der Dekorrelation quantisiert, d.h. je nach Quan-

tisierungsstufe wird hier vor der Entropiecodierung zur Erreichung eines höheren Komprimierungsfaktors eine Verzerrung festgelegt.

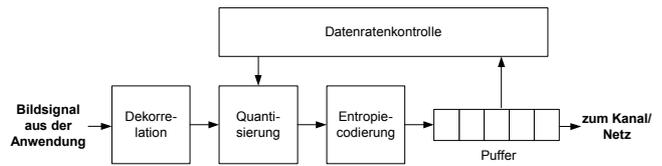


Abbildung 2: Grundkonzept der Videocodierung

Die durch die Anwendung einer Entropiecodierung auf das quantisierte Signal resultierende Bitrate, ist gemäss der Rate-Distortion-Funktion [8] von der Verteilungsdichtefunktion des Bildes (Frames) abhängig. Eine Entropiecodierung bei konstanter Verzerrung ergibt also grundsätzlich eine - vom Informationsgehalt des jeweiligen Bildes abhängige - variable Bitrate. Soll jedoch eine Übertragung mit zumindest abschnittsweise fester Bitrate erfolgen, so muss die Verzerrung in Abhängigkeit vom Detailgehalt des Bildes variiert werden. Das hierzu am häufigsten angewandte Verfahren benutzt einen Puffer im En- und Decoder, in dem der resultierende Bitstrom des Encoders vor der Übertragung zunächst zwischengespeichert wird. Droht der Puffer überzulaufen oder leerzulaufen muss gegengesteuert werden, indem die Verzerrung (Quantisierungsstufenhöhe) erhöht bzw. verringert wird. Ein großer (virtueller) Ausgangspuffer im Encoder ist für den Regelprozess weniger kritisch, allerdings hat die Größe des Ausgangspuffers im Encoder maßgeblichen Einfluss auf die Ende-zu-Ende-Verzögerung der Videokommunikation.

4.2 Existierende Verfahren

Das verwendete Verfahren zur Kontrolle der Datenrate bestimmt im starken Masse die Effizienz der Quellencodierung. Für die a priori Bestimmung des Quantisierungsfaktors vor der Entropiecodierung werden je nach Verfahren und dessen Zielanwendungsbereich die Rate-Distortion-Charakteristik des Encoders in Verbindung mit dem aktuell codierten Signals modelliert und anhand von einfachen Messgrößen, wie Signalvarianz oder mittlerer quadratischer Fehler, ermittelt. Meist wird hierfür ein zweistufiges Verfahren eingesetzt, das je nach Anforderung mehr oder weniger abgewandelt wird:

- Reservierung der Bits (Bitallocation) für die Codierung des gesamten Bildes, in Abhängigkeit von der Aktivität und des Pufferfüllstands für eine vorgegebene Datenrate, wobei die Komplexität der zukünftigen Bilder abgeschätzt und mit berücksichtigt wird.

- Festlegung des Referenzquantisierungsparameters auf Bild- bzw. Group of Blocks-Ebene, um die Zieldatenrate zu erreichen und die lokale Festlegung des Quantisierungsparameters auf Makroblock-Ebene in Abhängigkeit von der lokalen Aktivität und den vorhandenen Bits

In der Literatur wurden bereits unterschiedliche Ansätze für die Datenratenkontrolle vorgestellt (z.B. [2][11]). Für nicht echtzeitfähige Anwendungen, wie z.B. Codierung von Inhalten für Video-CDs, wurden zahlreiche Ansätze beschrieben, die jedes Bild mehrfach codieren und für die Auswahl der richtigen Parameter die Information der vorherigen Durchläufe nutzen. Bei echtzeitfähigen Codecs, wie sie z.B. bei Videokonferenzen eingesetzt werden, ist diese Vorgehensweise aufgrund der systembedingten Erhöhung der Ende-zu-Ende-Verzögerung bzw. der notwendigen Rechenleistung nicht möglich. Innerhalb der Standardisierung von H.263 und dessen Weiterentwicklungen bzw. bei MPEG-4 wurden verschiedene Algorithmen für diesen Anwendungsbereich miteinander verglichen und die vielversprechendsten Ansätze in die Referenzcodecs implementiert [1][3][9][10][11][14], wobei in der Literatur aktuell noch Erweiterungen der Verfahren, z.B. für mobile Netze, beschrieben wurden [13].

In der aktuellen Version von H.263++ wird ein Verfahren eingesetzt, das speziell für niedrige Datenraten und eine geringe Ende-zu-Ende-Verzögerung entworfen wurde. Der Ansatz basiert auf folgenden Annahmen:

- Abschätzung der Bits pro Makroblock (MB):

$$B_i = A \left(K \frac{\sigma_i^2}{Q_i^2} + C \right) \text{ mit}$$

- B_i : Anzahl der Bits für den i-ten MB
- A : Anzahl der Pixel pro MB (z.B. 16x16)
- K : Faktor für Anpassung des Modells an die Signalstatistik des Bildsignals
- σ_i : Standardabweichung des i-ten MBs
- Q_i : Quantisierungsfaktor für den i-ten MB
- C : Overhead pro Pixel bei der Codierung (z.B. Bits für Bewegungsvektoren, Steuerinformation, ...)

- Abschätzung der Verzerrung:

$$D = \frac{1}{N} \sum_{i=1}^N \alpha_i^2 \frac{Q_i^2}{12} \text{ mit}$$

- D : Mittlerer quadratischer Fehler für das aktuelle Bild
- N : Anzahl der MBs im Bild
- α_i : Gewichtungsfaktor für den i-ten MB
- Q_i : Quantisierungsfaktor für den i-ten MB

Die Berechnung der einzelnen Quantisierungsparameter basiert auf der Minimierung der Verzerrung des aktuellen Bildes unter Anwendung der Lagrange-Theorie, unter Einhaltung der vorgegebenen Bits. Die zur Verfügung stehende Bitmenge wird aus der aktuellen Bitrate und dem Pufferfüllstand ermittelt. Das Prinzip ist in Abbildung 3 für zwei Makroblöcke mit gegebener bzw. geschätzter Rate-Distortion-Funktion dargestellt und beruht darauf, dass die Gesamtverzerrung der beiden Blöcke durch unterschiedliche Bitzuweisung pro Makroblock und dem damit verbundenen Codierfehler minimiert werden kann.

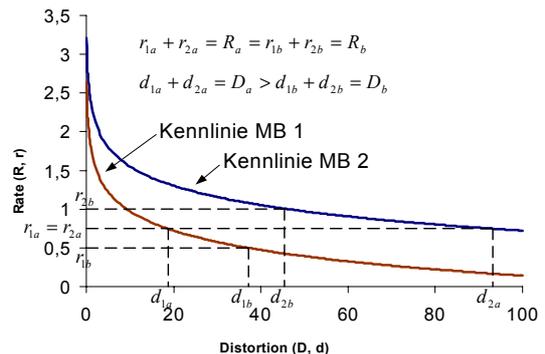


Abbildung 3: Grundprinzip zur Minimierung der Verzerrung bei vorgegebener Bitrate

Über den Gewichtungsfaktor für den i-ten Makroblock α_i kann der Algorithmus an spezielle Bedingungen, wie z.B. Berücksichtigung von subjektiv wichtigen Bildbereichen angepasst werden. Eine detaillierte Beschreibung des Algorithmus und dessen Implementierung befindet sich in [10].

Das bei MPEG-4 verwendete Modell basiert im Kern auf folgendem Ansatz:

$$B_i = \frac{X_1 * S}{Q_i} + \frac{X_2 * S}{Q_i^2} \text{ mit}$$

- B_i : Bits für das i-te Bild bzw. den i-ten MB
- X_1 : Modellparameter, der aus der Bildstatistik ermittelt wird
- S : Mittlere absolute Differenz des i-ten Bildes bzw. des i-ten MBs
- Q_i : Quantisierungsfaktor für das i-te Bild bzw. den i-ten MB
- X_2 : Modellparameter, der aus der Bildstatistik ermittelt wird

Je nach Anwendungsbereich wird dieses Modell eingesetzt, um für eine bestimmte Menge an Bits einen globalen Quantisierungsfaktor für das gesamte Bild bzw. Objekt zu berechnen oder aber die Faktoren pro Makroblock zu berechnen. Die Parameter X_1 und X_2 werden mittels Ausgleichsrechnung nach dem Prinzip des kleinsten quadratischen Fehlers aus den bereits codierten Bildern bzw. Makroblöcken berechnet. Die Beschreibung der Algorithmen bzw. deren Implementierung befindet sich in [3].

5 Anforderungen an flexible Datenratenkontrollen

Die bisherigen Anforderungen an die Verfahren für die Datenratenkontrolle, wie Sicherstellung von homogener, bestmöglicher Bildqualität und die Einhaltung von Anwendungs- und Netzvorgaben, wie maximale Ende-zu-Ende-Verzögerung und Einhaltung der mittleren Ausgangsdatenrate unter meist fest vorgegebenen bzw. beschränkten Einsatzbedingungen, werden durch die „universellen“ Anwendungsmöglichkeiten moderner Videostandards insbesondere um folgende Aspekte erweitert:

- Variables Bildmaterial, das sich je nach Anwendung sowohl im Format als auch in der Signalcharakteristik und damit der Signalstatistik, die auch im laufenden Betrieb stark variieren kann (z.B. während einer Videokonferenz wird eine Sportsequenz übertragen), ändert.
- Je nach Vorgaben der Anwendung variiert die maximale Ende-zu-Ende-Verzögerung und damit auch die maximale Ausgangspuffergröße.
- Einsatz von unterschiedlichen Bewertungskriterien für die Bildqualität. Während für viele Aufgaben minimale Verzerrung gewünscht wird, ist es insbesondere bei der Videokommunikation interessant, subjektiv wichtige Bildinformationen gezielt betonen zu können.

- Gewährleistung einer flexiblen Ausgangsdatenratensteuerung, um insbesondere bei mobilen Netzen die effektive Übertragung zu sichern.
- Unterstützung eines sehr breiten Spektrums von Ausgangsdatenraten (von einigen kBit/s bis hin zu mehreren Mbit/s).

Anhand der genannten Anforderungen, wie sie in zukünftigen Kommunikationsbedingungen auftreten, können für realistische Analysen der verschiedenen Verfahren der Datenratenkontrollen folgende Testbedingungen aufgestellt werden: Einsatz von Bildmaterial in verschiedenen Formaten mit stark unterschiedlicher Bildstatistik (z.B. Details, Anzahl und Geschwindigkeit von bewegten Objekten und globalen Szenenänderungen) bei Datenraten von 10 kBit/s bis über 1,5 MBit/s und Bildwiederholraten von 7,5 Bilder/s bis hin zu 30 Bilder/s. Die je nach Anwendung unterschiedliche maximale Ende-zu-Ende-Verzögerung wird dabei durch unterschiedliche Ausgangspuffergrößen von $0,1 \times$ bis $2 \times$ Ausgangsdatenrate \times Sekunde berücksichtigt (entspricht 100 ms - 2 s Pufferverzögerung).

Gerade die Unterstützung eines großen Bandbreitenspektrums, die effektive Codierung von Bildmaterial mit stark variierender Signalcharakteristik und wechselnde Netzdatenraten stellen hohe Anforderungen an die Datenratenkontrolle. Das Modell muss in der Lage sein, sich rasch an die unterschiedlichen Gegebenheiten anzupassen, um eine effektive Ausgangsdatenratenregelung sicherzustellen. In Abbildung 4 und Abbildung 5 sind beispielhafte Kennlinien für die Anzahl der Bits pro Pixel für die DCT-Codierung in Abhängigkeit von der Standardabweichung (Sigma) für unterschiedliche Quantisierungsfaktoren dargestellt, wie sie am Ausgang eines H.263-Codecs ermittelt wurden (Sequenz Foreman, konstanter Quantisierungsfaktor für gesamte Sequenz, Bild 72, P-codiert).

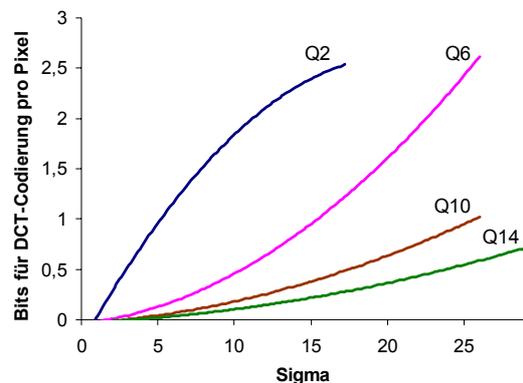


Abbildung 4: Kennlinie für die Anzahl der Bits für die DCT-Codierung pro Pixel in Abhängigkeit von Sigma für unterschiedliche Quantisierungsfaktoren

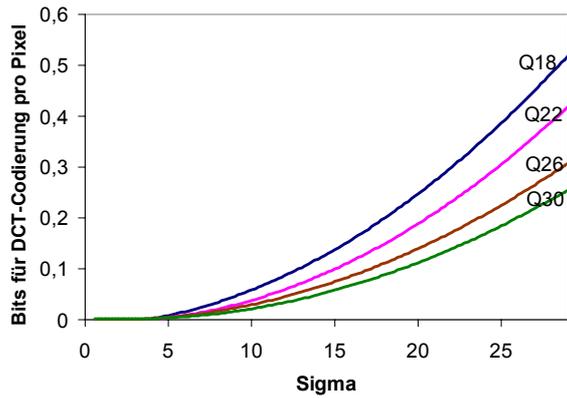


Abbildung 5: Kennlinie für die Anzahl der Bits für die DCT-Codierung pro Pixel in Abhängigkeit von Sigma für unterschiedliche Quantisierungsfaktoren

Neben der reinen Funktionsweise und der Stabilität der Verfahren sind im besonderen folgende Kriterien beim Entwurf von neuen Verfahren zu berücksichtigen:

- Möglichst geringe Komplexität und einfache Implementierbarkeit
- Sicherstellung einer effektiven Encodierung über den gesamten Anwendungsbereich.
- Rasche Adaptivität, um auf variierende Randbedingungen schnell reagieren zu können.
- Geringe systembedingte Verzögerung

Wesentlichen Einfluss auf die systembedingte Verzögerung des Videoencoders hat die Auswahl der Parameter, die als Eingangsparameter in die Datenratenkontrolle eingehen. Die Encodierung kann bei H.263 und MPEG-4, solange keine speziellen Optionen wie z.B. Global Motion Compensation eingesetzt werden, rein makroblockbasiert ablaufen. Ein Bild kann also hierzu zeilenbasiert von links nach rechts bearbeitet werden, ohne dass Informationen von zukünftigen Makroblöcken benötigt werden. Allerdings benötigen alle in Abschnitt 4.2 vorgestellten Verfahren framebasierte Kenngrößen, so dass hier eine Speicherung der Informationen innerhalb des Codierablaufes notwendig wird und eine zusätzliche Verzögerung entsteht (siehe Abbildung 6).

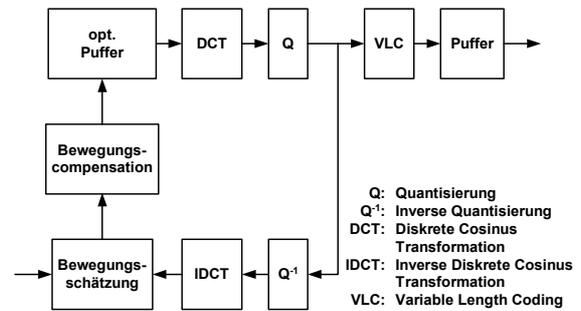


Abbildung 6: Funktionsblöcke eines hybriden DCT/DPCM-Videoencoders mit Zwischenspeicher für framebasierte Analyse

Für eine in Echtzeit ablaufende Codierung ist es notwendig, dass innerhalb der Zeit t_{COD} alle Funktionsblöcke des Encoders von allen MBs des zu codierenden Bildes (Anzahl der MBs = anz_{MB} , Anzahl der Funktionsblöcke ohne Rückkopplungsschleife und Puffer = anz_{BLOCK} , Anzahl der bei der Codierung gegenüber der Referenzframerate t_{REF} ausgelassenen Frames = m) durchlaufen werden. Mit der Annahme einer makroblockbasierten Pipelineverarbeitung ergibt sich damit für alle MBs die Durchlaufzeit t_{COD} für die Codierung ohne Zwischenpufferung und für Echtzeitbedingung die Durchlaufzeit eines MBs für einen Funktionsblock im Encoder $t_{BLOCKMB}$ wie folgt:

$$t_{COD} = [(anz_{MB} - 1) + anz_{BLOCK}] t_{BLOCKMB} = (m + 1) t_{REF}$$

$$t_{BLOCKMB} = \frac{(m + 1) t_{REF}}{(anz_{MB} - 1) + anz_{BLOCK}}$$

Für die Codierung mit Zwischenpufferung ergibt sich mit der gleichen Zeit $t_{BLOCKMB}$ die Durchlaufzeit t_{CODF} und damit die systembedingte Erhöhung der Durchlaufzeit t_V im Encoder zu:

$$t_{CODF} = [2(anz_{MB} - 1) + anz_{BLOCK}] t_{BLOCKMB}$$

$$t_V = (anz_{MB} - 1) t_{BLOCKMB}$$

Die zusätzliche Verzögerung bei der Encodierung hat auf die Ende-zu-Ende-Verzögerung der Videokommunikation die gleiche Auswirkung wie ein größerer Ausgangspuffer und wirkt sich insbesondere bei der niederbitratigen Kommunikation stark störend aus, da sich z.B. für QCIF mit 7.5 Frames/s (fps) und einem

Encoderaufbau nach Abbildung 6 ein zusätzliches Delay von über 100 ms ergibt.

6 Lösungsansatz für eine adaptive Datenratenkontrolle

Als Basis für experimentelle Untersuchungen der Anforderungen im breiten Anwendungsbereich von „universellen“ Datenratenkontrollen wurde ein Lösungsansatz basierend auf das Modell von H.263++ [10] hergeleitet und implementiert. Im wesentlichen wurden dabei folgende Punkte modifiziert:

- Bei der Reservierung der Bits für das aktuelle Bild wurde ein Modell hergeleitet, das die augenblickliche Ausgangsdatenrate, die Bildkomplexität und den aktuellen Pufferstand beinhaltet.
- Als Aktivitätsmass wird anstelle der Standardabweichung die mittlere Energie der DCT-Koeffizienten herangezogen.
- Anhand datenraten- und bildstatistikspezifischer Berechnung des Gewichtungsfaktors α wurde eine gezielte Optimierung des Basismodells auf unterschiedliche Ausgangsdatenraten und Bildeigenschaften durchgeführt.
- Die (optimale) Berechnung der Quantisierungsfaktoren bezieht sich im Gegensatz zu der Implementierung in [10] stets auf das aktuelle Frame und nicht nur auf die noch zu codierenden Makroblöcken im aktuellen Frame.
- Organisation der Codierung auf Basis einer Makroblock-Zeilensstruktur, um die Berechnung der sequenz- und datenratenabhängigen Modellparameter nicht nur intra- sondern auch interframebasiert durchzuführen.
- Der Füllstand des Ausgangspuffers wird für eine effektive Kontrolle von Modellfehlern ebenso wie die augenblickliche Ausgangsdatenrate pro Makroblockzeile kontrolliert.

7 Simulationsergebnisse

Für die Untersuchungen ihrer Leistungsfähigkeit auf Basis der in Abschnitt 5 definierten Anforderungen, sind die einzelnen Verfahren aus der Standardisierung von MPEG-4 (MPEG-4-framebasiert (M4), MPEG-4-makroblockbasiert (M4MB)) und H.263++ (TMN8)

ebenso wie der neue Ansatz (FlexR) in einen H.263-Core implementiert worden.

Hierfür wurden 11 Sequenzen aus der MPEG4- bzw. H.263/H.26L-Standardisierung mit unterschiedlichsten Eigenschaften ausgewählt und 30 verschiedene Codierszenarios zusammengestellt, die jeweils mit 6 verschiedenen Ausgangspuffergrößen (2/1/0,75/0,5/0,25/0,1 * Zieldatenrate * Sekunde) näher untersucht wurden. Eine Auswahl von repräsentativen Ergebnissen ist in Tabelle 1 zu finden.

Container, QCIF, 10 kBit/s, 7,5 fps				
Ausgangspuffergröße	M4	M4MB	TMN8	FlexR
0,1	32,32	31,72	32,58	32,64
0,25	32,70	32,61	32,54	33,30
0,5	32,65	32,60	32,54	33,20
0,75	32,69	32,52	32,54	32,96
1	32,68	32,59	32,47	33,06
2	32,75	32,59	32,25	33,02
Weather, CIF, 128 kBit/s, 15 fps				
Ausgangspuffergröße	M4	M4MB	TMN8	FlexR
0,1	32,20	31,77	31,65	32,20
0,25	32,24	31,83	31,62	32,82
0,5	31,33	31,83	31,57	32,53
0,75	31,18	31,84	31,53	32,53
1	31,30	31,79	31,51	32,67
2	32,09	31,82	31,42	32,61
Mobile&Calender, 1024 kBit/s, 30 fps				
Ausgangspuffergröße	M4	M4MB	TMN8	FlexR
0,1	26,60	26,51	26,70	26,74
0,25	26,74	26,51	26,67	26,73
0,5	26,74	26,51	26,63	26,71
0,75	26,75	26,51	26,59	26,73
1	26,75	26,51	26,54	26,76
2	26,77	26,51	26,31	26,78

Tabelle 1: Auswahl von repräsentativen Ergebnissen anhand von mittleren PSNR-Werten der Sequenzen

Das Verfahren von TMN8 erzielt dabei insbesondere bei sehr kleinen Puffergrößen und bei niederbitratigen Testläufen gute Ergebnisse, wohingegen das framebasierte MPEG-4-Verfahren bei hohen Bitraten und grossen Ausgangspuffergrößen Performancegewinne erzielen kann. Das sehr rechenintensive makroblockbasierte MPEG-4-Verfahren erzielt dagegen im gesamten Anwendungsbereich stets eine durchschnittliche Codiereffizienz. Aufgrund der hohen Adaptivität, die mit geringerer Rechenleistung als der makroblockbasierte MPEG-4-Ansatz erreicht wird, erfüllt der neue Ansatz alle in Abschnitt 5 definierten Anforderungen und erzielt dabei über alle 180 Testläufe im

Durchschnitt bis zu 0,25dB (PSNR) Performancegewinn, wobei bei einzelnen Sequenzen bis über 2,5 dB (PSNR) erreicht werden. Bei weiteren Untersuchungen zeigte sich, dass das Verfahren auch bei variablen Datenraten wie sie in mobilen Netzen auftreten, effektiv eingesetzt werden kann.

8 Zusammenfassung und Ausblick

Aktuelle Codiervorgänge, wie z.B. H.263++ oder MPEG-4, sind nicht mehr fest auf eine Zielapplikation ausgerichtet, sondern unterstützen unterschiedlichste Anwendungen und Übertragungsnetze. Dabei reicht das Spektrum von niederbitratiger Videokommunikation über mobile Netze bis hin zu hochqualitativem Video über Breitbandnetze. Für die Datenratenkontrolle, als wichtiges qualitätsbestimmendes Element des Videoencoders, sind adaptive Verfahren notwendig, um unterschiedlichstes Bildmaterial mit jeweils passendem Ende-zu-Ende-Delay über zeitvariable mobile Netze mit mehreren Kilobits pro Sekunde oder aber über hochbitratige Festverbindungen effektiv übertragen zu können. Dieser Beitrag zeigt die sich daraus ergebenden Anforderungen auf und stellt ein neues einfaches adaptives Verfahren für die Datenratenkontrolle vor. Der Ansatz erzielt im Vergleich mit den Verfahren aus der MPEG-4 und H.263++ Standardisierung einen mittleren Performancegewinn von bis zu 0,25dB (PSNR), wobei bei einzelnen Sequenzen bis über 2,5 dB (PSNR) erreicht werden.

Aktuell wird ein erweiterter Ansatz untersucht, der durch ein verbessertes Rate-Distortion-Modell verstärkt die variierenden statistischen Eigenschaften innerhalb eines Bildes, aber auch die hohe zeitliche Abhängigkeit der Einzelbilder ausnutzt. Zielsetzung ist dabei über den gesamten Anwendungsbereich eine noch effektivere Codierung unter Einhaltung der Randbedingungen, wie z.B. mittlere Ausgangsdatenrate und maximal vorgegebene Ende-zu-Ende-Verzögerung, sicherzustellen. Darüber hinaus soll der Einfluss auf die Bildqualität untersucht werden, wenn subjektiv wichtige Bildbereiche gezielt bei Auswahl der Quantisierungsfaktoren berücksichtigt werden.

9 Literatur

- [1] Anthony Vetro, Huifang Sun, Yao Wang, "MPEG-4 Rate Control for Multiple Video Objects", IEEE Trans. Circuits Syst. Video Technol., vol. 9, pp 186-, Feb. 1996
- [2] H.-M. Hang and J.-J. Chen, "Source Model for Transform Video Coder and Its Application--- Part I: Fundamental Theory", IEEE Trans. Circuits Syst. Video Technol., vol 7, pp. 287-298, Apr. 1997
- [3] ISO/IEC JTC1/SC29/WG11 Dokument M5477, "MPEG-4 Video Verification Model version 14.2", Maui, Dec. 1999
- [4] ITU Telecom. Standardization Sector of ITU, STUDY GROUP 16, "DRAFT FOR "H.263++" ANNEXES U, V, AND W TO RECOMMENDATION H.263", Nov. 2000
- [5] ITU Telecom. Standardization Sector of ITU, STUDY GROUP 16, "H.26L Test Model Long Term Number 6 (TML-6) draft0", Eibsee, Germany, Jan. 2001
- [6] ITU Telecom. Standardization Sector of ITU, STUDY GROUP 16, "Test model 11", Monterey, Finland, Feb. 1999
- [7] ITU Telecom. Standardization Sector of ITU, "Video coding for lowbitrate communication," Draft ITU-T Recommendation H.263 Version 2, Sept. 1997.
- [8] Jens-Rainer Ohm; „Digitale Bildcodierung“; ISBN 3-540-58579-6, Springer Verlag, Berlin Heidelberg 1995
- [9] Jordi Ribas-Corbera, Shaw-Min Lei, "A Frame-Layer Bit Allocation for H.263+", IEEE Trans. Circuits Syst. Video Technol., vol. 10, pp 1154-1158, Oct. 2000
- [10] Jordi Ribas-Corbera, Shawmin Lei, "Rate Control in DCT Video Coding for Low-Delay Communications", IEEE Trans. Circuits Syst. Video Technol., vol 9, pp. 172-185, Apr. 1999
- [11] L.-J. Lin, A. Ortega, and C.-C. J. Kuo, "Rate Control using Spline-Interpolated R-D Characteristics", in Proc. VCIP, Orlando, FL, Mar. 1996, pp. 111-122
- [12] Minhua Zhou, „Optimization of MPEG-2 Video Encoding“, ISBN 3-8265-2817-4, Shaker Verlag, Aachen, 1997
- [13] Supavadee Aramvith, I-Ming Pao, Ming-Ting Sun, "A Rate-Control Scheme for Video Transport over Wireless Channels", <http://pierce.ee.washington.edu/~supava/csvtrev7.pdf> über <http://citeseer.nj.nec.com/cs>
- [14] W. Ding and B. Liu, "Rate Control of MPEG Video Coding and Recording by Rate-quantization Modelling", IEEE Trans. Circuits Syst. Video Technol., vol. 6, pp 12-19, Feb. 1999