

- 42 AKTUELL
- 43 Versatile Video Coding: Mehr Effizienz und Vielseitigkeit für die Videokomprimierung
BENJAMIN BROSS UND DETLEV MARPE



VERSATILE VIDEO CODING: MEHR EFFIZIENZ UND VIELSEITIGKEIT FÜR DIE VIDEOKOMPRIMIERUNG

BENJAMIN BROSS UND DETLEV MARPE

Komprimierte Videodaten nehmen schneller zu als jemals zuvor. Bereits heute bilden sie den mit Abstand höchsten Anteil von Bits im Internet und im mobilen Datenverkehr. Dies veranschaulicht den Bedarf nach noch effizienterer Komprimierung über den aktuellen HEVC-Standard hinaus. Im April 2018 hat die Standardisierung des HEVC-Nachfolgestandards VVC begonnen. Der Beitrag geht auf den aktuellen Stand der Entwicklung ein.

► Compressed video data is growing faster than ever before. Already today, they make up by far the highest proportion of bits on the Internet and in mobile data traffic. This illustrates the need for even more efficient compression beyond the current HEVC standard. In April 2018, standardization of the HEVC successor standard VVC began. The article deals with the current state of development.

Einleitung

Die globalen Wachstumsraten aller Arten von Internet-basierten Videodaten steigen beständig. Jevons' Paradoxon besagt, dass jede effizientere Nutzung eines Rohstoffs mit sich bringt, dass der Konsum dieses Rohstoffs letztendlich erhöht statt gesenkt wird. Dieser sogenannte Rebound-Effekt lässt sich ebenfalls bei Internet-basierten Videodaten beobachten. Trotz immer effizienterer Codierung der Videodaten wird bis zum Jahr 2021 mit einer Verdopplung des derzeitigen Datenaufkommens für Internet-Videos gerechnet. Dabei soll der Anteil von IP-basiertem Video am globalen Datenaufkommen auf über 80 Prozent steigen. Außerdem ändert sich auch die häusliche Landschaft und es wird prognostiziert, dass der Anteil an Fernsehgeräten mit 4K-Auflösung in privaten Haushalten auf nahezu zwei Drittel anwächst. All diese Entwicklungen verlangen – trotz oder gerade wegen des Rebound-Effekts – nach neuen Forschungsbeiträgen im Hinblick auf die effiziente Codierung, den Transport, die Verarbeitung und die Analyse von zunehmend höher aufgelösten Videosignalen. Die Abteilung „Videokodierung und Maschinelles Lernen“ am Fraunhofer-Institut für Nachrichtentechnik, Heinrich-Hertz-Institut, HHI, befasst sich mit allen relevanten Aspekten der damit verbundenen Forschungsbereiche. Sie hat bereits in der Vergangenheit mit ihren maßgeblichen Beiträgen zu den internationalen Videocodierungsstandards H.264/MPEG-AVC und H.265/MPEG-HEVC dazu beigetragen, die Schlüsseltechnologien für die zunehmende Verbreitung von digitalem Video in unserem alltäglichen Leben bereitzustellen. Diese Beiträge wurden bereits viermal mit einem Emmy Award gewürdigt.

Um die neuen Herausforderungen zu bewältigen, ha-

ben sich die Experten der beiden führenden Standardisierungsorganisationen, die „ITU-T Video Coding Expert Group (VCEG)“ und die „ISO/IEC Moving Pictures Expert Group (MPEG)“, in einem „Joint Video Experts Teams (JVET)“ zusammengeschlossen und mit der Arbeit an einem gemeinsamen Nachfolgestandard von H.265/MPEG-HEVC begonnen.

Explorations- und Einreichungsphase

Die Vorarbeiten zu dem neuen Standard hatten bereits 2015 begonnen. Aus diesen Arbeiten entstand das „Joint Video Exploration Model (JEM)“. Modell heißt in diesem Fall: Eine Software mitsamt Encoder und Decoder zur Quantifizierung der Bitratenreduktion darin enthaltener Kompressionsalgorithmen. Das JEM bot der Forschungsgemeinde damit gewissermaßen eine „Spielwiese“ zum Testen neuer Techniken, die geeignet waren, eine Verbesserung der Codiereffizienz relativ zum aktuellen HEVC-Standard zu erzielen. Das vollentwickelte JEM-Modell konnte am Ende auch eine durchschnittlich 34-prozentige Bitratenreduktion bei gleicher objektiver Qualität gegenüber HEVC realisieren. Aufgrund dieser gewonnenen Erkenntnisse und der damit nachgewiesenen Evidenz für eine deutliche Steigerung der Codiereffizienz wurde der Entschluss gefasst, die Entwicklung eines neuen Standards zu beginnen. Dazu haben die ITU-T VCEG und die ISO/IEC MPEG im Oktober 2017 einen gemeinsamen

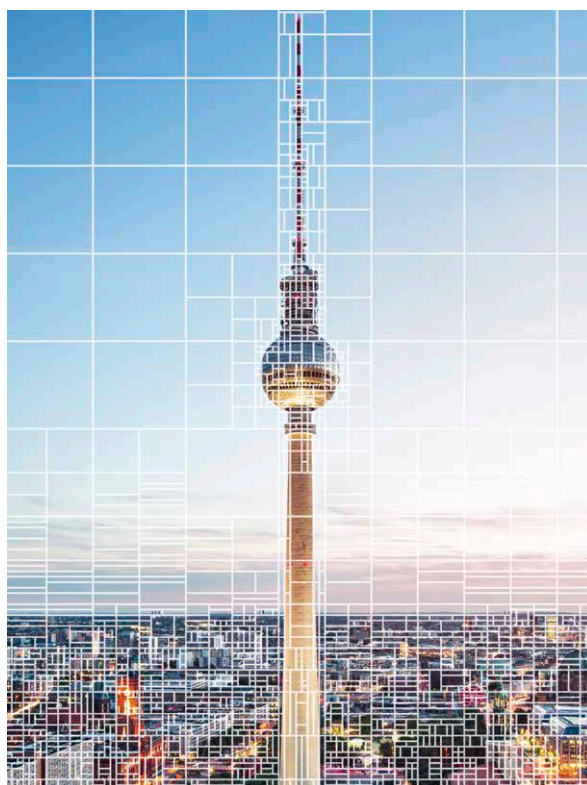


Bild 1. Beispiel für die neue, extrem flexible Blockaufteilung von VVC alle Bilder:

Fraunhofer HHI

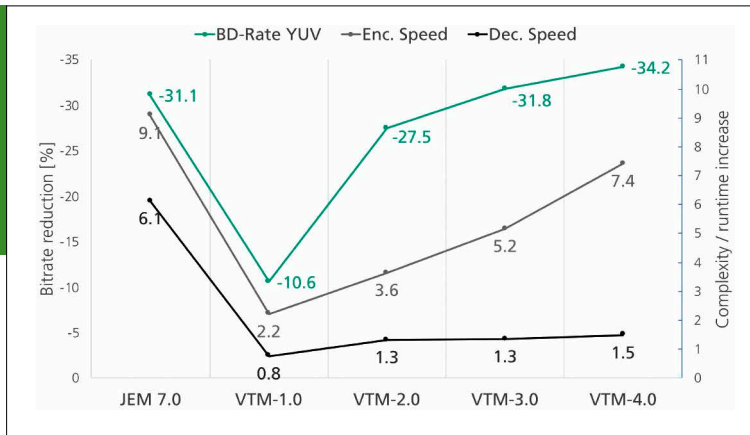


Bild 2. Diagramm mit Verlauf der Bitratenreduktion sowie Encoder-/Decoder-Laufzeiten für die einzelnen Test Modell-Versionen und JEM.

Aufruf zur Einreichung von Vorschlägen, den sogenannten „Joint Call for Proposals“ für neue Videocodier-Technologien, veröffentlicht.

Auf diesen Aufruf hin reichten das Fraunhofer HHI und andere führende Technologieunternehmen vielversprechende Vorschläge ein. Die eingereichten Vorschläge und Konzepte wurden entsprechend dreier Kategorien von unterschiedlichen Testvideodaten bewertet. Die erste Kategorie bestand aus Standard-Dynamic-Range (SDR) Video mit fünf Sequenzen in High Definition (HD) Auflösung, die bei Bitraten von 400 Kbit/s bis 3,8 Mbit/s codiert werden sollten, und fünf Sequenzen in Ultra-HD (UHD) Auflösung mit Bitraten von 950 Kbit/s bis 10 Mbit/s. Die zweite Gattung von Testvideos beinhaltete High-Dynamic-Range (HDR) Video mit vier Perceptual Quantizer (PQ) Sequenzen in HD und Bitraten von 350 Kbit/s bis 3 Mbit/s sowie drei Hybrid-Log-Gamma (HLG) Sequenzen in UHD mit Bitraten von 640 Kbit/s bis 10 Mbit/s. In der dritten Kategorie wurden schließlich 360-Grad Videos untersucht, wobei diese eine Sequenz mit 6K x 3K Auflösung und Bitraten von 2 Mbit/s bis 10 Mbit/s sowie vier Sequenzen mit 8K x 4K Auflösung und Bitraten von 400 kbit/s bis 7 Mbit/s umfassten.

Insgesamt reichten 33 Unternehmen Vorschläge ein, von denen die leistungsstärksten durchschnittlich 40 oder mehr Prozent an Bitrateneinsparungen bei gleichem objektiven Bildgütemaß (PSNR gegenüber HEVC) realisieren konnten. In Bezug auf die subjektive, d.h. durch Testpersonen wahrgenommene Qualität waren die eingereichten Vorschläge in vielen Fällen sogar besser als HEVC bei der doppelten Bitrate und in einer aussagekräftigen Anzahl auch besser als JEM bei derselben Bitrate. Der Beitrag des Fraunhofer HHI gehörte zu den leistungsstärksten Einreichungen in allen drei Kategorien und enthielt neben der konsequenten Weiterentwicklung bereits bekannter Technologien ebenfalls eine Reihe neuer Kompressionstechniken, von denen einige beispielsweise mithilfe Maschinellen Lernens entwickelt wurden.

Standardisierungsphase

Nach den positiven Erkenntnissen aus dem Call for Proposals wurde im April 2018 die kollaborative Forschungs- und Testphase für den neuen Versatile Video Coding (VVC) Standard eingeleitet. Das erste VVC-Testmodell VTM-1.0 wurde auf Grundlage des blockbasierten hybriden Codierparadigmas als sogenanntes „clean slate test model“ konzipiert. Das heißt, dass es neben einer verbesserten Blockaufteilung, der sogenannten „quadtree plus multi-type tree“ (QT+MTT) Partitionierung keine über HEVC hinausgehenden, neuen Codier-

werkzeuge enthielt. Diese Entscheidung basierte auf der Tatsache, dass diese Partitionierungstechnik das am häufigsten verwendete Aufteilungsschema unter allen Einreichungen war und zudem grundlegende Wechselwirkungen auf alle neu hinzukommenden Codierwerkzeuge ausübt. Neben QT+MTT wurde die Software des Fraunhofer HHI als Basis für das VVC-Testmodell verwendet. Dabei ist das Testmodell die gemeinsame Experimentierplattform, auf der alle vielversprechenden, zusätzlichen Codierwerkzeuge im Hinblick auf Effizienz- und Komplexitätsaspekte getestet werden.

Die neue Blockaufteilung im VVC-Entwurf zeichnet sich dadurch aus, dass das Bildsignal zunächst, wie in HEVC, in quadratische Coding Tree Units (CTU) unterteilt wird, die aber im Unterschied zu HEVC eine Luma-Blockgröße von bis zu 128x128 Bildpunkten (engl.: Samples) umfassen. Jede CTU kann mittels rekursiver Quadtree-Partitionierungsoptionen (QT) und verschachtelten rekursiven Multi-Type-Tree Partitionierungsoptionen (MTT) mit Binär- oder Ternärsplit in Codiereinheiten mit variabler Größe (CU) unterteilt werden. Ein Beispiel für das Ergebnis dieses flexiblen Verfahrens der Blockaufteilung zeigt Bild 1. Es ist gut erkennbar, wie die neuen, schmalen Blöcke des MTT dazu führen, dass sich die Blockstruktur sehr gut an den Bildinhalt anpassen kann.

Im weiteren Verlauf der bisherigen Standardisierungsphase wurden weitere Codiertechniken der besten Einreichungen genauer untersucht und, abhängig von ihrer Leistungsbewertung, in den Entwurf des VVC-Standards aufgenommen. Zu den Tools, die bis dato in den derzeit aktuellen VVC-Entwurf aufgenommen wurden, gehören neben verbesserten Methoden zur Prädiktion folgende Verbesserungen bzw. Neuerungen gegenüber HEVC: Multiple Auswahl von Transformationen, Trellis-codierte Quantisierung, verbessertes Deblocking-Filter, neuartiges adaptives Wiener-Filter in der Schleife nach dem Deblocking sowie CABAC-Entropiecodierung mit verbesserter Wahrscheinlichkeitsschätzung und verbesserter Transformationskoeffizientencodierung.

Neue Methoden zur Intra- und Inter-Prädiktion

Die Methoden der Prädiktion in der Videocodierung werden grundsätzlich in zwei Kategorien unterteilt und zwar abhängig davon, ob die prädizierten Bildinhalte mit Hilfe von bereits rekonstruierten Samples desselben Bildes oder anderer Bilder generiert werden; im ersten Fall spricht man von Intra-Prädiktion während der letztere Fall Inter-Prädiktion genannt wird. Die folgende Aufstellung gibt einen Überblick über die entsprechenden neuen Codiertools im aktuellen VVC-Testmodell VTM-4.0 (Stand Februar 2019). Dabei werden im Folgenden zunächst einige der neuen Methoden der Intra-Prädiktion im aktuellen VVC-Entwurf kurz vorgestellt:

- 65 Prädiktionsrichtungen: Die Anzahl der Prädiktionsrichtungen wird von 33 (wie in HEVC) auf 65 erhöht, um ein größeres Spektrum an Kantenrichtungen repräsentieren zu können. Zusätzlich werden für nichtquadratische Blöcke sogenannte „wide angle modes“ eingeführt.
- Interpolationsfilter mit 4 Taps: Durch Verwendung eines 4-Tap-Interpolationsfilters für die gerichteten Intra-Prädiktoren wird eine erhöhte Genauigkeit bei der Berechnung des Prädiktionssignals gegenüber HEVC erzielt.
- Prädiktionskombination: Für eine bestimmte Teilmenge der Intra-Prädiktions-Modi wird das Prädiktionsignal aus einer Linearkombination von gefilterten und ungefilterten Referenzsamples gebildet.
- Multi-Referenzzeilen: Zusätzlich zu der direkt benachbarten Zeile bzw. Spalte von rekonstruierten Referenzsamples können optional 2 weitere Zeilen bzw. Spalten

von Referenzsamples aus der Nachbarschaft des aktuellen Blocks benutzt werden.

- **Komponentenübergreifendes lineares Modell:** Diese Methode beinhaltet die Prädiktion von Chroma-Samples unter Verwendung von rekonstruierten Luma-Samples auf Basis eines linearen Modells.
- **Intra-Sub-Partitionen:** Die zusätzliche Unterteilung eines Blocks in 2 oder 4 Teilblöcke erlaubt es, die Teilblöcke unabhängig voneinander mit demselben Intra-Prädiktions-Mode zu präzisieren und zu rekonstruieren.

Folgende neue Codiertools für die Inter-Prädiktion (auch bewegungskompensierte Prädiktion genannt) sind Bestandteil des aktuellen VVC-Entwurfs:

- **Affine Bewegungskompensation:** Aus zwei oder drei Bewegungsvektoren am Rand eines Blocks wird für jeden 4x4-Subblock ein eigener Bewegungsvektor interpoliert.
- **Adaptive Bewegungsvektor-Auflösung:** Die Bewegungsvektoren können pro Block wahlweise in einer Auflösung von $\frac{1}{4}$ Sample, 1 Sample oder 4 Samples angegeben werden.
- **Merge-Mode mit zusätzlicher Bewegungsvektor-Differenz:** Die Merge-Bewegungsvektoren, die aus benachbarten Bewegungsvektoren hergeleitet statt codiert werden, können hiermit in horizontaler oder vertikaler Richtung verschoben werden.
- **Dreieckige Partitionen:** Ein Block kann entlang der Diagonalen in zwei Partitionen mit unterschiedlichen Bewegungsvektoren unterteilt werden.
- **Kombinierte Intra/Inter-Prädiktion:** Inter- und Intra-Prädiktion können in einem Block kombiniert verwendet werden.

- **Bi-direktionaler optischer Fluss:** Für das Bi-Prädiktions-signal wird eine Verfeinerung bestimmt, die sich aus der optischen Flussgleichung ergibt.
- **Decoderseitige Bewegungsvektor-Verfeinerung:** Mittels einer einfachen decoderseitigen Bewegungssuche können bei Bi-Prädiktion die Bewegungsvektoren verfeinert werden.

Ergebnisse und Ausblick

Vergleicht man die aktuelle Version des VVC-Testmodells VTM-4.0 mit der letzten JEM-Version JEM 7.0 bezüglich Codiereffizienz und Komplexität von Encoder und Decoder, zeigt sich, dass die Bitratenreduktion gegenüber HEVC von 31,1 % auf 34,2 % erhöht und gleichzeitig die Encoder-Laufzeit von 9.1x auf 7.4x und die Decoder-Laufzeit von 6.1x auf 1.5x (jeweils relativ zu HEVC) reduziert werden konnte. Die Grafik in Bild 2 veranschaulicht auch die entsprechenden Kennzahlen der drei VVC-Testmodell-Vorgängerversionen von VTM-4.0 im Vergleich mit JEM 7.0.

Der Zeitplan für das VVC-Standardisierungsprojekt sieht vor, dass der finale Standard im Oktober 2020 offiziell verabschiedet werden soll. Dabei ist es das erklärte Ziel für VVC, eine durchschnittliche Bitratenreduktion von 50 % gegenüber HEVC für typische Videoinhalte zu realisieren. Auf der NAB Show 2019 wird das Fraunhofer HHI die aktuelle Version des VVC-Testmodells VTM-4.0 der Öffentlichkeit präsentieren. Diese Version zeigt bereits signifikante Bitratengewinne im Vergleich zu HEVC für ein breites Spektrum an Videoinhalten von SDR HD bis hin zu HDR UHD und das bei einem relativ moderaten Zuwachs an Komplexität. ◀

FILME NACHHALTIG DIGITALISIEREN

Die Empfehlungen zur Digitalisierung von kinematografischem Film, DIN SPEC 15587, ist erschienen, das teilte der DIN-Normenausschuss Veranstaltungstechnik, Bild und Film (NVBF) jetzt mit.

Die Veröffentlichung legt die Anforderungen an einen nachhaltigen Digitalisierungsworkflow von kinematografischem Film und magnetischen Filmtongträgern fest. Hierzu werden Anforderungen für die Vorarbeiten zur Digitalisierung, die Digitalisierung selbst und die nachfolgenden Bearbeitungsstufen dargestellt. So gilt Nachhaltigkeit dann als gegeben, wenn nach Kenntnisstand zum Zeitpunkt der Arbeiten eine Auswahl der qualitativ besten Ausgangsmaterialien stattgefunden hat und diese Materialien optimal zur Digitalisierung aufbereitet wurden. Darüber hinaus sind auch qualitative Eigenschaften des analogen Ausgangsmaterials im Digitalisierungsprozess korrekt zu berücksichtigen und für das Digitalisat geeignete Formate beziehungsweise Codierungen zu wählen. Bei den Nacharbeiten am Digitali-

sat müssen zudem restaurierungsethische Grundsätze eingehalten werden und die Dokumentation der Arbeiten archivari-schen Prinzipien entsprechen.

Die Empfehlung beschäftigt sich nicht mit der Erfassung von in digitaler Form auf Film aufbelichteten Tonsignalen oder aufbelichteten Steuersignalen zur Synchronisierung externer digitaler Tonquellen.

Sie soll Orientierung für Digitalisierungsprojekte, bei denen weniger die Nachhaltigkeit der Digitalisierung, sondern vorwiegend die Publikation beziehungsweise Präsentation im Vordergrund steht, geben (zum Beispiel im Rahmen von Digitalisierungen für das Fernsehen).

Die DIN SPEC (Fachbericht) wurde vom DIN-Normenausschuss Veranstaltungstechnik, Bild und Film (NVBF), Arbeitsausschuss NA 149-00-03 AA „Aufnahme, Bearbeitung und Wiedergabe von Bewegtbild und Ton“ erstellt.

Das Dokument ist bei der Beuth Verlag GmbH, Berlin <http://www.beuth.de> erhältlich.



RICHTLINIEN ZUM DOWNLOAD VERFÜGBAR

Das Institut für Rundfunktechnik (IRT) ist Herausgeber der Technischen Richtlinien und Arbeitsgruppenergebnisse der öffentlich-rechtlichen Rundfunkanstalten in Deutschland (ARD, ZDF, Deutschlandradio), in Österreich (ORF) und in der Schweiz (SRG/SSR).

Diese Dokumente stehen ab sofort auf der Homepage des IRT kostenfrei und ohne Anmeldung zum Download zur Verfügung. Des Weiteren ist auf der Homepage ein Archiv mit älteren Versionen der Richtlinien verfügbar. Ausgewählte Richtlinien sind auch auf Englisch erhältlich.

◀ www.irt.de/richtlinien