

SOFTWARE-LIVE-DECODER FÜR DEN VIDEOSTANDARD DER ZUKUNFT: VERSATILE VIDEO CODING (VVC)

ADAM WIECKOWSKI, BENJAMIN BROSS, DETLEV MARPE

Der nachfolgende Artikel befasst sich mit dem aktuellen Entwicklungsstand des Nachfolgestandards H.266/VVC von H.265/HEVC und stellt die bisherigen Ergebnisse eines Projekts des Fraunhofer HHI zur Implementierung eines 4K-fähigen Software-Live-Decoders vor.

► The following article deals with the current state of development of the H.265/HEVC successor standard H.266/VVC and presents the results of a Fraunhofer HHI project on the implementation of a 4K-capable software live decoder.

Einführung

Komprimierte Videodaten bilden bereits heute den mit Abstand höchsten Anteil von Bits im Internet und im mobilen Datenverkehr – Tendenz steigend. Im April 2018 hat die Standardisierung des H.265/HEVC-Nachfolgestandards H.266/VVC begonnen (siehe Artikel in der FKT 04/2019, S. 43 ff.). Das Fraunhofer Heinrich-Hertz-Institut HHI hat seitdem, wie bei den Standards H.264/AVC und H.265/HEVC zuvor, zu vielen Kerntechnologien beigetragen und arbeitet

aktiv daran mit, den neuen Standard so leistungsfähig wie möglich und zugleich praktisch umsetzbar zu gestalten.

Um einen möglichst frühzeitigen kommerziellen Einsatz des neuen Standards zu ermöglichen und gleichzeitig mögliche Probleme bezüglich seiner Implementierbarkeit rechtzeitig zu erkennen, haben Mitarbeitende des Fraunhofer HHI parallel zur laufenden Standardisierungsarbeit von H.266/VVC damit begonnen, einen 4K-fähigen Software Live-Decoder zu implementieren. Den aktuellen Stand des Projekts präsentiert das Fraunhofer HHI auf der IBC 2019.

Versatile Video Coding

Das erste VVC-Testmodell VTM-1.0 wurde im April 2018 auf Grundlage des Konzepts der blockbasierten hybriden Videokodierung als sogenanntes „clean slate test model“ entworfen. Das heißt, dass es neben einer verbesserten Blockaufteilung, der sogenannten „quadtree plus multi-type tree“ (QT+MTT)-Partitionierung, keine über HEVC hinausgehenden, neuen Kodieralgorithmen enthielt. Als Basis für das VVC-Testmodell (VTM) wurde die Software des Fraunhofer HHI gewählt. Das Testmodell ist die gemeinsame Experimentierplattform, auf der alle vielversprechenden, neu vorgeschlagenen Kodieralgorithmen im Hinblick auf Effizienz- und Komplexitätsaspekte getestet werden. Das neue Testmodell

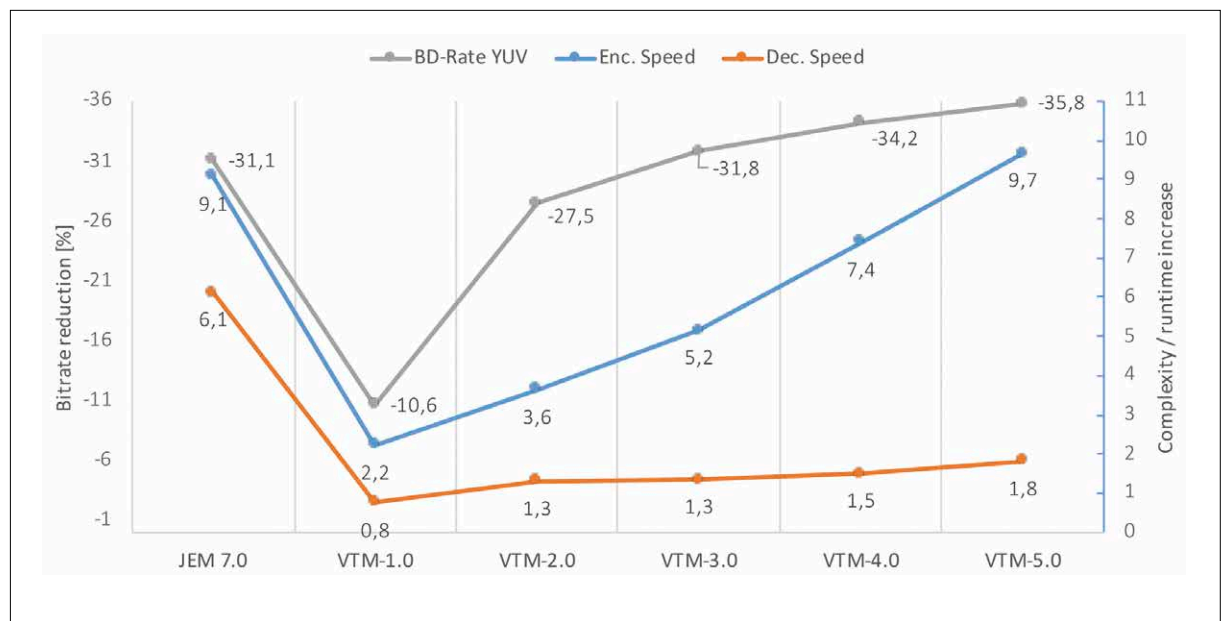


Bild 1. Entwicklung der Bitratensparnis (linke Ordinate, in Prozent BD-Ratenreduktion) sowie Encoder- und Decoderlaufzeiten (rechte Ordinate, in Faktor Laufzeit gegenüber HM) über die verschiedenen Versionen des Testmodells von VVC und das JEM gegenüber dem HEVC-Testmodell (HM). Laufzeiten und Bitratenreduktion sind über ein Testset von Videosequenzen unterschiedlicher Auflösung ermittelt.

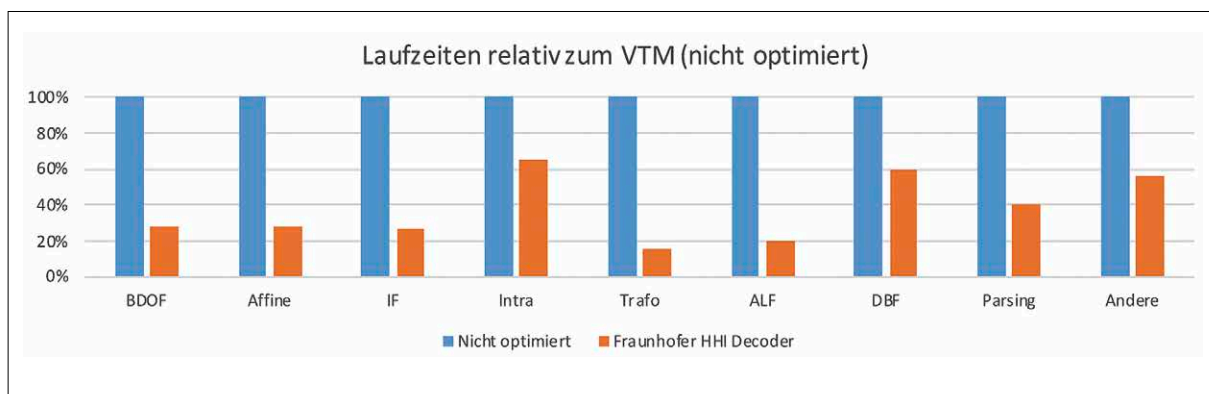


Bild: Fraunhofer HHI

Bild 2. Erreichte Beschleunigung der einzelnen Tools im Fraunhofer HHI-Decoder gegenüber dem Testmodell VTM ohne plattformspezifische Optimierung (IF = Interpolationsfilter, DBF = Deblocking Filter).

ersetzt das deutlich komplexere Joint Exploration Modell (JEM), das bis dahin als Basis genutzt wurde.

Bis zur Version VTM-5.0 wurden viele weitere Kodieralgorithmen in das Testmodell sowie den vorläufigen Entwurf der Standard-Spezifikation, dem sogenannten Draft, aufgenommen. Darunter sind auch viele der am Fraunhofer HHI entwickelten Algorithmen, wie z.B. eine Trellis-basierte Quantisierung inklusive einer effizienten Entropiekodierung der Transformationskoeffizienten oder eine mithilfe von Maschinellem Lernen entwickelte, matrix-basierte Intra-Prädiktion (MIP).

Viele der neuen Algorithmen bringen neben Bitratenreduktion auch zusätzliche Komplexität mit sich (siehe Entwicklung der Encoder- und Decoderlaufzeiten gegenüber der Bitratenreduktion in Bild 1). Wie bereits anhand der Grafik in Bild 1 ablesbar ist, wurde bei der bisherigen Entwicklung von VVC aber sorgfältig darauf geachtet, vor allem die Decoderkomplexität möglichst gering zu halten. Das steht in starkem Kontrast zur Vorstandardisierungsphase, die mit dem Joint Exploration-Modell JEM-7.0 abgeschlossen wurde und bei der eine sechsfache Decoderlaufzeit gegenüber dem HM-Testmodell zu beobachten war. Die bisher erreichte objektive, gemittelte PSNR-basierte Bitratenreduktion des aktuellen Testmodells VTM-5.0 von VVC beträgt fast 36 %. Dieser Gewinn an Kodiereffizienz kommt jedoch nicht umsonst. Vor allem neue Algorithmen zur Bewegungskompensation bringen Komplexität, d.h. zusätzlichen Rechenaufwand und Speicherbedarf mit sich. Um beispielsweise die Übertragung von Teilen der Bewegungsinformation im Bitstrom zu sparen, werden diese für die bewegungskompensierte Prädiktion decoderseitig in einer zusätzlichen Suche abgeleitet – eine Aufgabe, die in Vorgängerstandards ausschließlich dem Encoder überlassen wurde. So kann die Bewegungskompensation entweder mit einer Bewegungsvektorverfeinerung (DMVR) oder durch Lösung einer vereinfachten bidirektionalen optischen Flussgleichung (BDOF) verbessert werden. Andere Tools, wie etwa ein Modell für die Kompensation von nicht-translatorischen, affinen Bewegungsmustern (AFFINE), wie z.B. Rotationen, erhöhen die Granularität der Bewegungsparameter, wodurch ein zusätzlicher Aufwand für die Prädiktion entsteht.

Weiterhin wurden in VVC zwei zusätzliche Pipeline-Schritte eingeführt, die in H.265/HEVC so nicht existieren. Eine nicht separierbare Transformation (LFNST) wird nach der üblichen 2D-Transformation auf die niedrigen Frequenzanteile angewandt, um die Koeffizienten noch weiter zu kompaktieren. Dieser Schritt, der effektiv einer 2D-Re-Transformation und Umsortierung der niederfrequenten Trans-

formationskoeffizienten entspricht, erhöht die Latenz der Restfehlerrekonstruktion. Dazu kommt ein adaptives Loop-Filter (ALF), welches als zusätzliches Schleifenfilter dem Deblocking- und Sample-Adaptive-Offset-Filter aus H.265/HEVC nachgeschaltet ist und für jeden Abtastwert berechnet sowie angewandt werden muss.

Alle diese Algorithmen wurden im Laufe der bisherigen Standardisierung von VVC auf implementierungstechnische Aspekte hin sorgfältig untersucht und, soweit möglich, von ihrer ursprünglich angenommenen Version aus weitgehend vereinfacht. Nur so konnte mit VTM ein wesentlich besseres Kosten-Nutzen-Verhältnis im Sinne von Decoderkomplexität versus Codiereffizienz bzw. Bitratensparnis als mit JEM erzielt werden.

Software-Live-Decoder

Als Basis für den Live-Decoder wurde das Testmodell VTM gewählt, das, wie oben erwähnt, auf einer am Fraunhofer HHI entwickelten Softwareplattform basiert. Die gut konzipierten und bekannten Softwarestrukturen sowie die bereits vollständig implementierten Algorithmen erlaubten einen schnellen Start.

Optimierung der Implementierung

Im Testmodell werden Referenz-Decoder und -Encoder in einer einheitlichen Software-Basis gepflegt. Das hat zur Folge, dass die Software viele für den Decoder redundante Strukturen und Abläufe beinhaltet. Diese konnten in der ersten Entwicklungsphase entfernt werden.

In der zweiten Entwicklungsphase wurden die eigentlichen Sample-Operationen beschleunigt. Hier war der Spielraum insofern beschränkt, als dass die VTM-Referenzsoftware schon einige derartige Optimierungen enthält. Die Fortschritte bei der Beschleunigung der einzelnen Tools sind in Bild 2 dargestellt. Teile der Optimierung wurden aus der VVC-Referenzsoftware entnommen und ggf. weiter verbessert.

Bild 3 zeigt, inwieweit die einzelnen Komponenten des Decoders zur Gesamtlaufzeit beitragen. Dabei ist zunächst erkennbar, dass die prozentualen Anteile der Prädiktion, Transformation und der In-Loop-Filter gegenüber den jeweiligen Anteilen des nicht optimierten Testmodell-Decoders reduziert wurden. Auffällig ist auch, dass nahezu 30 % der Laufzeit nicht einer der Hauptkomponenten zuzuordnen ist. Dieser Anteil beinhaltet allgemeines Datenmanagement, Verwaltung der Strukturen und den Kontrollfluss auf höherer Ebene. Auch der Anteil von Parsing an der Gesamtlaufzeit ist leicht gestiegen, was zum einen auf die gute Implementie-

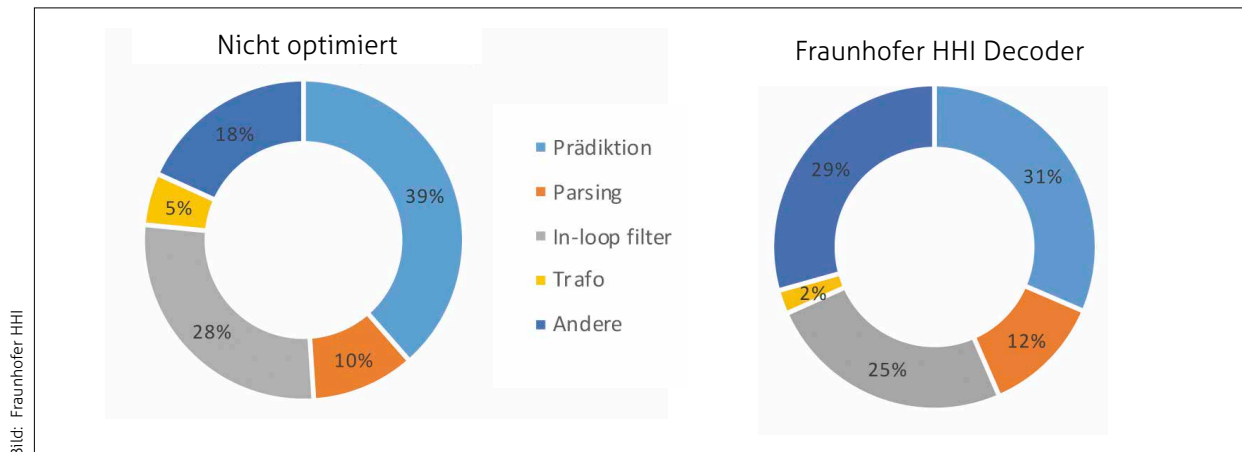


Bild: Fraunhofer HHI

Bild 3. Anteile der Laufzeiten der Hauptdecoderkomponenten im Testmodell (ohne plattformspezifische Optimierung) sowie in der aktuellen Version des Fraunhofer HHI-Live-Decoders.

zung der context-adaptiven Entropiekodierung im Testmodell zurückzuführen ist, andererseits ein potentielleres Indiz für möglichen weiteren Spielraum bei der Datenstrukturoptimierung darstellt. Für die Weiterentwicklung müssen vor allem in diesen beiden Bereichen noch weitere Optimierungen vorgenommen werden.

Durch die optimierte Implementierung konnte insgesamt eine weitreichende Beschleunigung um ungefähr den Faktor 2 gegenüber dem optimierten VTM-Decoder erreicht werden.

Parallelisierung

Die alleine durch die Optimierung der Decoder-Algorithmen erreichte zweifache Beschleunigung gegenüber dem Testmodell reicht nicht aus, um 4K/Ultra-HD oder, je nach Bitrate, sogar HD live zu decodieren. Nahezu alle modernen Software-Plattformen verfügen jedoch über mehrere CPU-Kerne und sind somit in der Lage, mehrere Aufgaben gleichzeitig zu bearbeiten. Auf den vom Fraunhofer HHI als Zielplattform gewählten mobilen x86-Geräten (z.B. Laptops) sind in der Regel zumindest vier solche Rechenkerne verfügbar.

Diese Leistung kann dann optimal ausgenutzt werden, wenn unabhängige Aufgaben definiert und mithilfe von mehreren Threads auf die vorhandenen Rechenkerne verteilt werden können. Die Fraunhofer HHI-Forschenden haben sich vorgenommen, von der üblichen bildweisen Parallelisierung abzusehen, weil diese eine höhere Latenz verursacht. Statt-

dessen nutzen die Forschenden ihren tiefen Einblick in die Strukturen und Abhängigkeiten des neuen Standards und definierten ein Parallelisierungsmodell, das sich auf Aufgaben in einem einzelnen Bild beschränkt. Verfügt ein Rechner über ausreichend Ressourcen, kann eine Parallelisierung auf höherer Ebene jedoch zusätzlich angewandt werden.

In Bild 4 wird die aktuelle Decodiergeschwindigkeit in Abhängigkeit der Anzahl der vom Decoder benutzten Threads als frames per second (fps, im Mittel über alle Bilder) gegenüber der Bitrate dargestellt. Dabei wird die Bitrate als grobes Indiz dafür verwendet, wie aufwändig ein bestimmtes Video zu decodieren ist.

Ausblick

Die Weiterentwicklung des VVC-Decoders wird sich auf zwei Hauptziele konzentrieren. Der Decoder schafft derzeit bei HD, wie in Bild 4 dargestellt, 60 fps für übliche Bitraten von bis zu 4 Mbps mit nur zwei Threads. Mit vier Threads kann eine Bildwiederholrate von 90 bis 120 fps erreicht werden. Um das Ziel von 4K-Decodierung bei 60 fps zu ermöglichen, muss noch weitere Entwicklungsarbeit geleistet werden. Weiterhin muss der Live-Decoder mit den neuesten Algorithmen, die in den VVC-Draft aufgenommen werden, aktualisiert werden.

Nach der initialen VVC-Entwicklungsphase, in der es vor allem um Bitratenreduktion ging, soll der Entwurf des Standards jetzt in der zweiten Phase bis zur Finalisierung im Juli

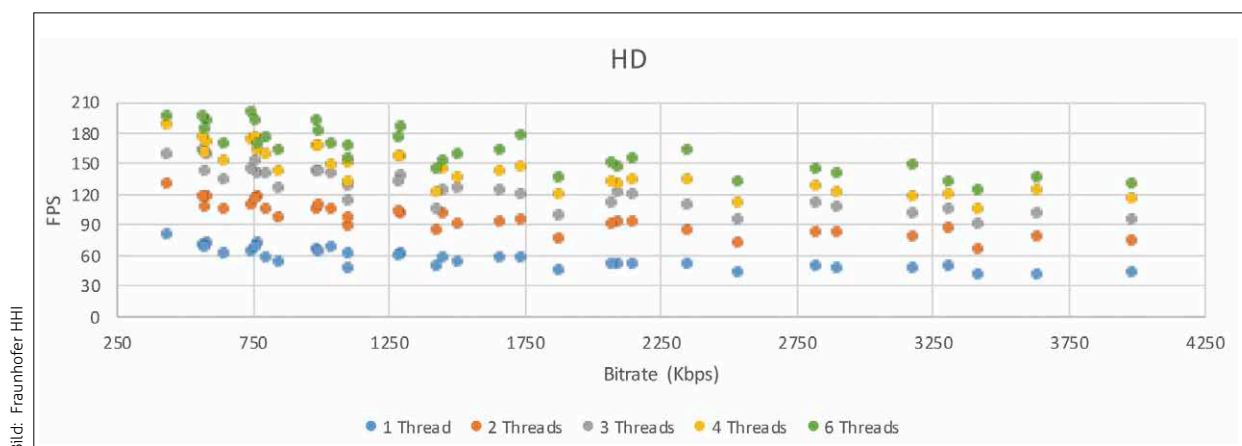


Bild: Fraunhofer HHI

Bild 4. Die Decodiergeschwindigkeit in Abhängigkeit von der Anzahl der parallel genutzten Threads sowie der Videobitrate.

Bild: Fraunhofer HHI



ADAM WIECKOWSKI

ist wissenschaftlicher Mitarbeiter und Projektleiter am Fraunhofer HHI. Er leitet die Entwicklungsarbeiten an der Software des VVC-Livedecoders.

🔗 www.hhi.fraunhofer.de

Bild: Fraunhofer HHI



BENJAMIN BROSS

ist Editor des VVC-Standards und Projektmanager am Fraunhofer HHI. Seine Veröffentlichungen zu H.265/HEVC wurden mehrfach ausgezeichnet

🔗 www.hhi.fraunhofer.de

Bild: Fraunhofer HHI



DR.-ING. DETLEV MARPE

ist Leiter der Abteilung „Videokodierung und Maschinelles Lernen“ am Fraunhofer HHI. 2015 wurde er zum IEEE Fellow ernannt.

🔗 www.hhi.fraunhofer.de

2020 konsolidiert werden. Dafür sollen einzelne Komponenten vereinfacht sowie vereinheitlicht werden. Somit werden beide Entwicklungsschwerpunkte für den Live-Decoder immer mehr zusammenfallen. Sobald der finale Standard verabschiedet worden ist, werden aufwändige Konformitätstests durchgeführt. Diese stellen sicher, dass der Decoder alle möglichen Bitströme, inklusive diverser Sonderfälle, korrekt decodiert.

Das Ziel der Standardisierung ist es, einen sehr leistungsfähigen, aber auch praktisch nutzbaren und breit einsetzbaren, sprich vielseitigen Standard „Versatile Video Coding“ zu spezifizieren. Das Fraunhofer HHI will sowohl mittels der aktiven Standardisierungsarbeit seiner Wissenschaftlerinnen und Wissenschaftler als auch mit dem hier vorgestellten und Mitte 2020 ausgereiften Software-Live-Decoder zu diesem Ziel beitragen. 🔗

no.33
**MEDIENTAGE
 MÜNCHEN**
 23-25 OKTOBER 2019

KONFERENZ UND EXPO TOPICS:
 TV, VoD & Streaming || 5G || Werbung & Marketing || Künstliche Intelligenz || Blockchain || Digital Publishing || Distribution || Journalism || VR, AR & 360° || Apps & Skills || Innovation || R U Ready Summit || Startup Area des Media Lab Bayern und vieles mehr...

NEXT DIGITAL LEVEL
 Let's build the Media we want!
 Die MEDIENTAGE diskutieren: Welche Macht haben Plattformen? Welche Rolle spielen Technologie, Ethik & Journalismus? Wie können die neuen Technologien gewinnbringend für die Gesellschaft genutzt werden?

Alle Infos unter www.medientage.de
 #MTM19

📍 Location: ICM
 Messe München

