

Multistandard-Videodekompression in Mehrkernsystemen

Henryk Richter, Erika Müller

Universität Rostock, Fakultät für Informatik und Elektrotechnik, Institut für Nachrichtentechnik, R.-Wagner-Str. 31, 18119 Rostock, (henryk.richter|erika.müller)@uni-rostock.de

1. Einleitung

Im Laufe der letzten Jahre hat sich ein gestiegenes Aufkommen digitaler Medien in verschiedensten Anwendungsbereichen, von Kleinstgeräten im Mobilfunksektor bis hin zu hochauflösendem Fernsehen (HDTV), ergeben. Zu diesem Zweck ist eine Vielzahl von Videokompressionsstandards mit anwendungsspezifischen Leistungsparametern entwickelt worden, die es in Hard- und Softwarelösungen zu unterstützen gilt.

Die bisherigen Standards basieren auf ähnlichen Paradigmen, unterscheiden sich jedoch in Datenstromsyntax, Prädiktionsalgorithmen, Signaltransformationen und Nachverarbeitungsschritten. Der bisher beschrittene Weg zur Unterstützung multipler Kompressionsverfahren basierte auf der getrennten Implementierung, Optimierung und anwendungsspezifischer Anpassung.

Im vorliegenden Beitrag wird ein Lösungsweg vorgeschlagen, welcher durch Trennung von syntaktischer Ebene und Bildsignalverarbeitung eine verallgemeinerte Darstellung von Vorschriften zur Rückgewinnung von Videodaten aus unterschiedlichen ursprünglichen Datenströmen ermöglicht. Die Basis dafür liefert ein Metadatenformat, welches die Charakteristiken aller untersuchten Standards vereint.

Der Vorteil dieses Ansatzes liegt darin, dass hardwarespezifische Optimierungen des Bildverarbeitungsteils allgemeingültig für alle unterstützten Standards durchgeführt werden können, so dass der Gesamtentwicklungsaufwand sinkt und potenziell die Laufzeitstabilität steigt. Aus der Trennung der Verarbeitungsstufen ergibt sich ein weiterer Vorteil in Hinblick auf die Ausrichtung aktueller Mikroprozessorsysteme mit multiplen Rechenkernen pro Chip. Der Entwurf von Videodecodern in Form einer mehrstufigen Struktur erlaubt für beliebige Eingangsdatenströme die Nutzung multipler Rechenkerne.

2. Multistandardverfahren – Stand der Technik

Alle bisher veröffentlichten Standards zur Kompression von Videodaten basieren auf den gleichen Grundprinzipien. Die zu komprimierenden Bilder werden in reguläre Blöcke von 16x16 Bildpunkten zerlegt (Abb. 1). Für jeden dieser Makroblöcke erfolgt eine Entscheidung über die Prädiktionsstrategie. Die Prädiktion kann innerhalb der Bilder aus Nachbarblöcken durchgeführt werden oder sich auf bereits übertragene Bilder beziehen (Bewegungskompensation). Das verbleibende Prädiktionsfehlersignal wird einer Signaltransformation unterzogen, wobei die DCT und deren Derivate für die Standards MPEG-1/-2/-4 und ITU-T H.263/H.264 relevant sind. Die quantisierten Symbole und Kontrolldaten werden mit Codes variabler Länge (VLC) als Bitstrom dargestellt.

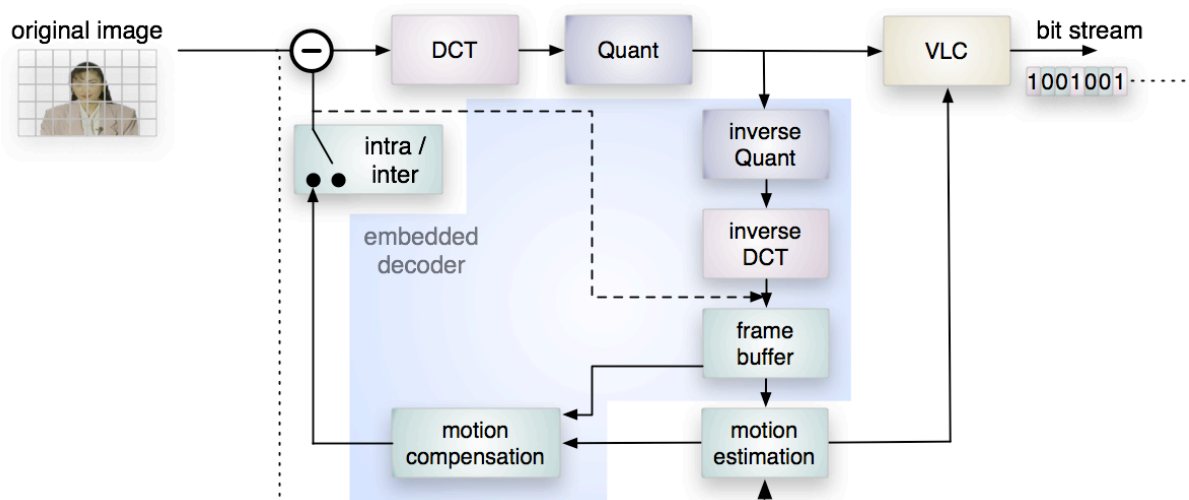


Abb. 1: Grundprinzip der Videodatenkompression: Zerlegung der Bilder, Prädiktion auf Basis zurückliegender Bilder, Transformation und Quantisierung des Fehlersignals

Im Zuge der Weiterentwicklung hinsichtlich gesteigerter Kompressionsleistung und Einsatzmöglichkeiten wurden die Prädiktions- und Codierungsalgorithmen sukzessiv verfeinert. Die erweiterte Semantik erforderte dafür entsprechende Syntaxelemente, so dass die resultierenden Bitströme einen rein standardspezifischen Aufbau besitzen. Diese Inkompatibilität der bisherigen Standards zueinander bewirkt bei der Implementierung für multistandardfähige Hard- und Software einige wesentliche Herausforderungen. Um robust und effizient verschiedene Standards zu unterstützen, bietet sich zunächst eine getrennte softwaregesteuerte Implementierung der Standards auf hinreichend generischer Hardware an. Zu dieser Kategorie zählt die direkte Umsetzung auf Prozessoren und DSPs sowie der Einsatz applikationsspezifischer Prozessoren mit zielgerichteten Hardware-Erweiterungen zur Beschleunigung von Bildverarbeitung und Bitstrom-Decodierung (Basoglu et al. 2002, van de Waerdt et al. 2005).

3. Metadatenformat für blockbasierte Videodatenkompression

In (Richter 2006) wurde ein alternatives Konzept zu Entwurf und Umsetzung der Dekompression von Videosequenzen vorgeschlagen. Anstelle des monolithi-

schen Entwurfs von Decodern erfolgt eine Trennung der sequenziellen Datenstromsyntax und der damit verbundenen rekursiven Prädiktion enthaltener Elemente von der blockweisen Bilddatenverarbeitung. Zu diesem Zweck werden die aus dem Datenstrom gewonnenen Elemente in einen Metadatenstrom überführt (Abb. 2). Das Metadatenformat besitzt die Eigenschaft, dass eine gleich lautende explizite Syntax für alle Ursprungsformate verwendet wird. Aspekte wie Timing, Laufzeitoptimierung und Speicherverwaltung sind nur einmalig der Stufe der generischen Decodierung zu realisieren.

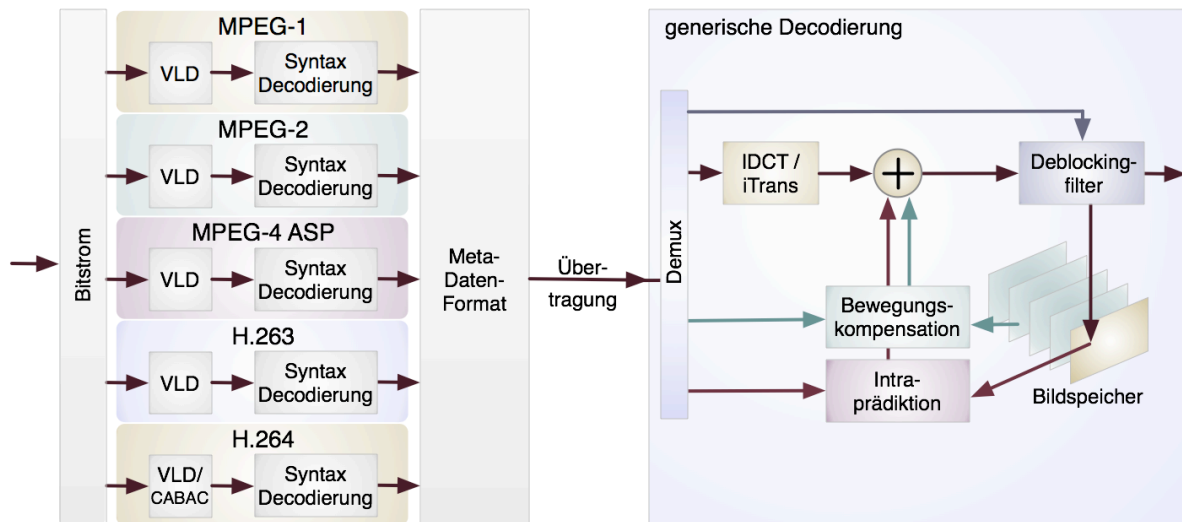


Abb. 2: Konzept des Metadatenformats: Generische Bildverarbeitung für eine Vielzahl unterstützter Kompressionsstandards

Innerhalb des Metadatenstromes wird auf implizite Verweise zu benachbarten Makroblöcken verzichtet. Auf diese Weise kann der Speicheraufwand in der generischen Decoderstufe auf die reinen Bildspeicher beschränkt werden, da die Steuerdaten fertig gestellter Makroblöcke verworfen werden können.

Aus dem in Abb. 2 gezeigten Konzept erwächst in Hinblick auf die aktuelle Entwicklung der Prozessoren ein weiterer Vorteil gegenüber monolithischen Decodern. Da die Verarbeitung der Bitstromsyntax zeitlich und örtlich unabhängig von der Bildverarbeitung durchgeführt werden kann, besteht an dieser Stelle bereits das Potenzial zur Nutzung von mehr als einem Rechenkern zur Lastverteilung. Im nachfolgenden Abschnitt wird ein Ansatz vorgestellt, auf Basis des multistandardfähigen Metadatenformats eine dynamisch verteilte Parallelverarbeitung zu erreichen.

3. Parallelisierung der Metadatenverarbeitung

Das Metadatenformat löst alle syntaktischen Abhängigkeiten zwischen benachbarten Makroblöcken auf, so dass diese unabhängig voneinander ausgewertet werden können. In Bezug auf die Standards bis einschließlich MPEG-4 Part 2 gilt diese Aussage auch für die Pixeldaten. Der Nachfolgestandard H.264 führte einige Neuerungen in Hinblick auf die Prädiktion benachbarter Bildpunkte ein.

Diese so genannte Intra-Prädiktion benötigt fertig gestellte Nachbarblöcke als Startpunkt für die Rekonstruktion aktueller Blöcke (Abb. 3a). Weiterhin ist bei H.264 ein Nachverarbeitungs-Filter zur Reduktion von Kompressions-Artefakten vorgeschrieben, welches ebenfalls rekursiv arbeitet (Abb. 3b). Das Deblocking-Filter verlangt neben der abgeschlossenen Rekonstruktion aller Blöcke in der Umgebung (gestrichelte Linie) auch die abgeschlossene Filterung der direkten Nachbarn (durchgezogene Linie). Da H.264 die höchste Rechenkomplexität der bisher erschienenen Standards aufweist (Ostermann et al. 2004), konzentrieren sich die folgenden Betrachtungen auf dessen Laufzeiteigenschaften.

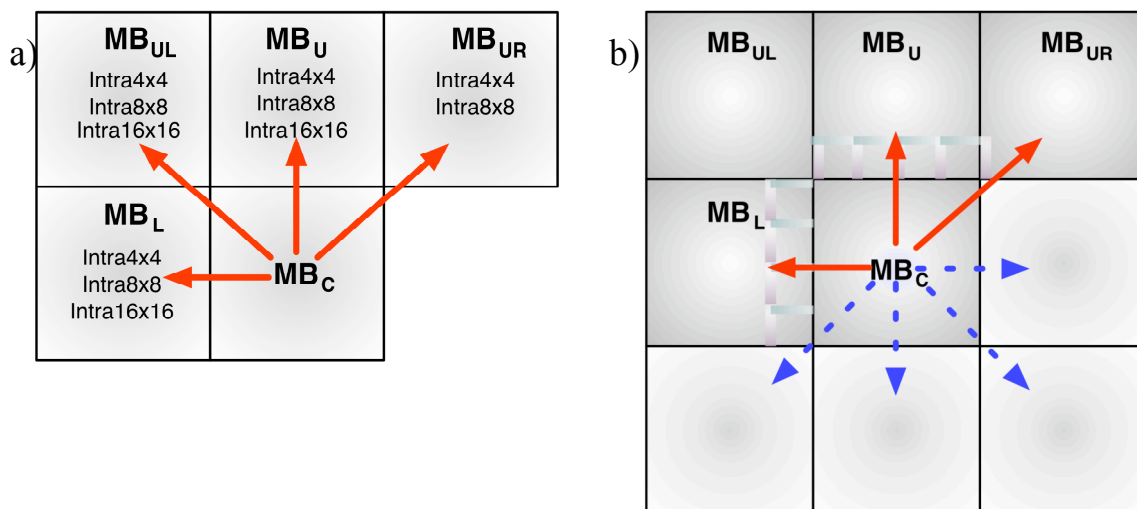


Abb. 3: a) Abhängigkeiten von Intra-Makroblöcken b) Abhängigkeiten im Deblocking-Filter

Die genannten Abhängigkeiten auf Pixelebene beschränken die Freiheitsgrade bei der parallelen Verarbeitung. Generell ist festzustellen, dass bei linearer Makroblockreihenfolge die Rekonstruktion des aktuellen Makroblockes MB_C gestartet werden kann, sobald der rechte obere Nachbar MB_{UR} fertig gestellt ist. Hinsichtlich optimaler Parallelisierung ergibt sich eine diagonale Hierarchie, bei der nach jedem zweiten Makroblock eine zusätzliche Instanz gestartet werden kann. In modernen Multitasking-Betriebssystemen stellen der Zeitbedarf von Taskwechsel-Vorgängen und die Interprozess-Kommunikation (IPC) weitere gewichtige Faktoren für begrenzte Gewinne bei der Parallelisierung dar (Härtig et al. 1997, Burghardt 2006).

Als Kompromiss zwischen Granularität und IPC wurde eine zeilenweise Zuordnung der Arbeitseinheiten auf die verfügbaren Instanzen vorgenommen, die insbesondere im Hinblick auf Cache-Effizienz weitgehend lineare Speicherzugriffe ermöglicht. Für Zustands-Abfragen und -Updates wird anstelle der Betriebssystemfunktionen für exklusive Zugriffe (mutex) auf kostengünstige Spinlocks zurückgegriffen, welche ohne Umweg über den Betriebssystemkern auskommen (Anderson 1990). Durch die Zugriffskontrolle auf die Verwaltungsebene der Arbeitseinheiten ist ein gemeinsames Management hinreichend. Die Instanz,

welche die Metadaten als Arbeitseinheiten aufbereitet, wird nach der Bereitstellung des letzten Makroblockes als zusätzlicher Worker eingesetzt. Damit sind zu jedem Zeitpunkt genau so viele Instanzen wie verfügbare Rechenkerne aktiv.

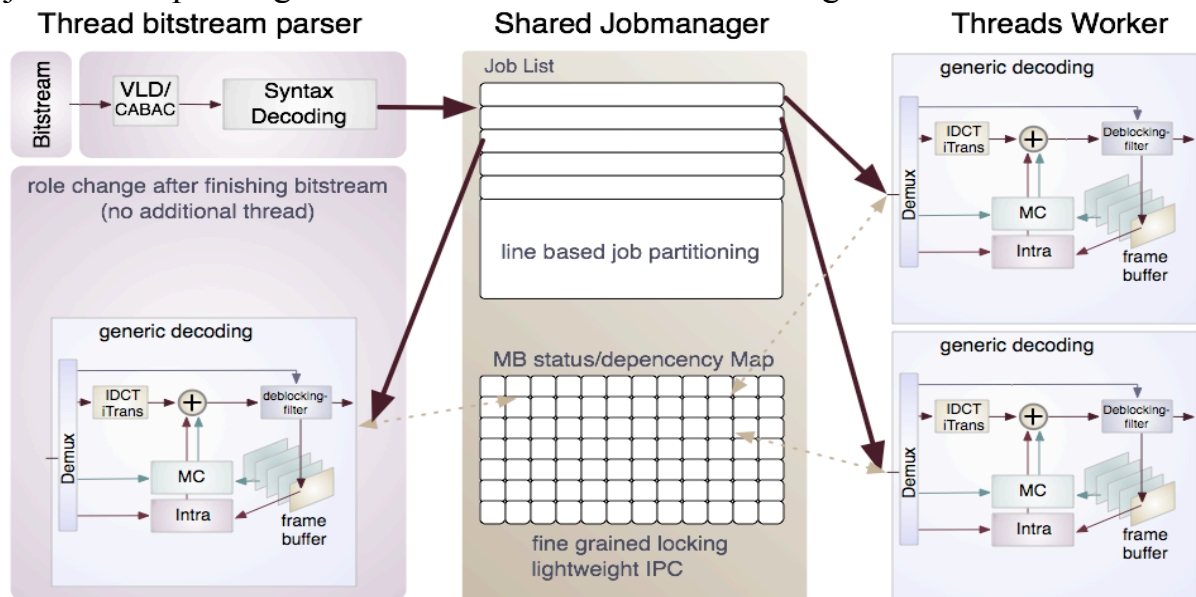


Abb. 4: Parallelisierung durch zeilenweise Verteilung der Arbeitseinheiten des Metadatenstromes auf die Arbeitsinstanzen

4. Ergebnisse / Zusammenfassung

Das vorgestellte Konzept dient der Beschleunigung der Dekompression von Videodaten im Echtzeitkontext. In Abb. 5 wird der Durchsatz in Bildern pro Sekunde für eine in H.264 codierte HDTV-Sequenz bei geringer Codierungskomplexität dargestellt (ohne B-Bilder, VLC, 1 Slice pro Bild).

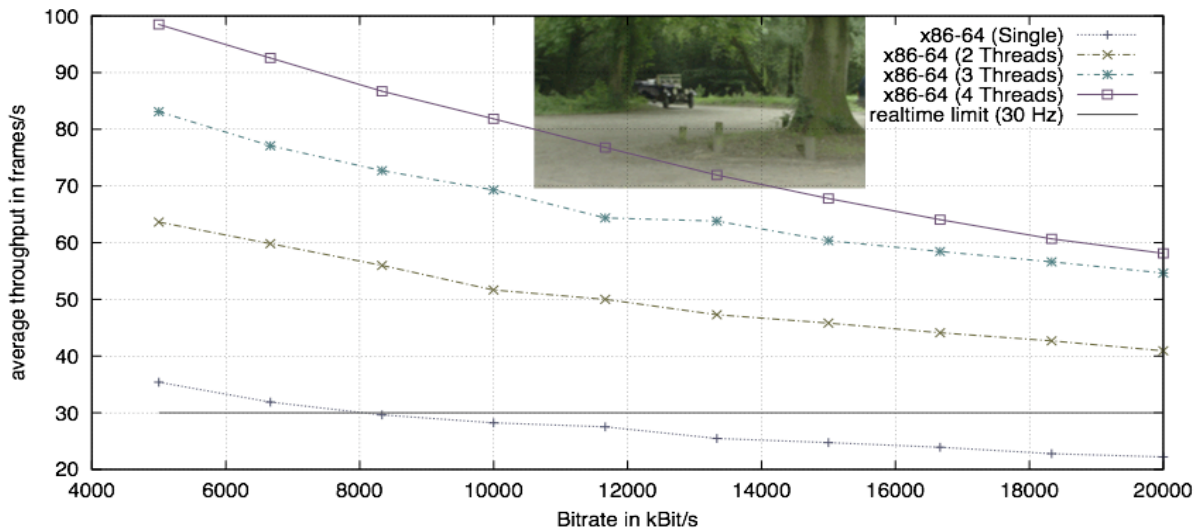


Abb. 5: Gewinn durch Parallelisierung in Bildern/s abhängig von der Eingangsdatenrate bei H.264 im Baseline-Modus. (Sequenz Vintagecar, 1080p, AMD Opteron 2 GHz)

Bei Aktivierung multipler Referenzbilder, arithmetischer Codierung und bidirektionaler Prädiktion erfordert H.264 einen deutlichen Mehraufwand. Weiter-

hin stellt die arithmetische Codierung bei hohen Datenraten den begrenzenden Faktor für die Skalierungsfähigkeit auf multiplen Kernen dar (Abb. 6).

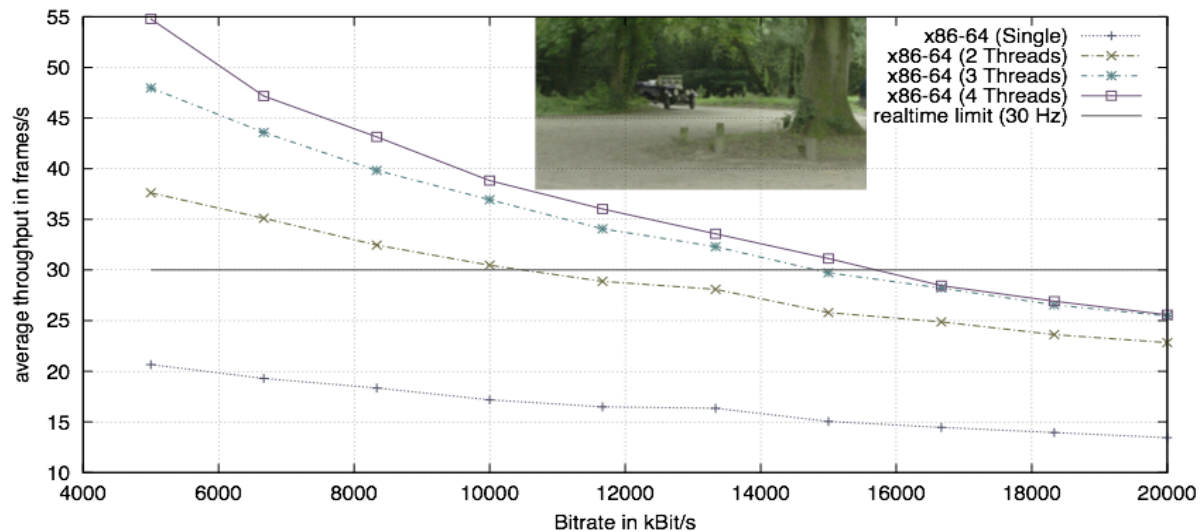


Abb. 6: Gewinn durch Parallelisierung in Bildern/s abhängig von der Eingangsdatenrate bei H.264 im High-Profil. (Sequenz Vintagecar, 1080p, AMD Opteron 2 GHz)

Das vorgestellte Konzept erlaubt trotz rekursiver Natur komprimierter Videodatenströme eine gewinnbringende Parallelverarbeitung auf multiplen Rechenkernen ohne spezielle Codierungsparameter vorauszusetzen. Dadurch kann auch bei extremen Datenraten und Coderparametern Echtzeitfähigkeit gegeben werden.

4. Literatur

- T. Anderson. (1990) The performance of spin lock alternatives for shared-memory multiprocessors. IEEE Trans. Parallel Distrib. Syst., vol. 01, no. 1, pp. 6–16.
- C. Basoglu, W. Lee, and J. O'Donnell. (2002) The Equator MAP-CA DSP: an end-to-end broadband signal processor VLIW. IEEE Trans. on Circuits and Systems for Video Technology, vol. 12, no. 8, pp. 646-659.
- C. Burghardt. (2006) Parallelisierung der Bildverarbeitung in blockbasierten Videodecodern. Diplomarbeit, Universität Rostock
- H. Härtig, M. Hohmuth, J. Liedtke, et al. (1997) The performance of μ -kernel-based systems. In SOSP '97: Proceedings of the sixteenth ACM symposium on OS principles. New York, NY, USA: ACM Press, pp. 66–77.
- J. Ostermann, J. Bormans, P. List, D. Marpe, et al. (2004) Video coding with H.264/AVC: tools, performance, and complexity, IEEE Circuits and Systems Magazine, vol. 4, no. 1, pp. 7–28.
- H. Richter. (2006) Standardübergreifende Konzepte für blockbasierte Videodecodierung. Dissertation, Universität Rostock
- J.W. van de Waerdt, S. Vassiliadis, S. Das, et al. (2005) The TM3270 Media-Processor. Proc. IEEE/ACM MICRO'05