

IP Video CCTV und H.264

H.264, der neueste Standard zur Videokompression, stellt eine wesentliche Weiterentwicklung der IP Videosysteme dar. Barry Keepence, CTO von IndigoVision, erklärt den Standard und dessen Vorteile.

H.264 bezeichnet den neuesten offiziellen Videokompressionsstandard, der auf die äußerst erfolgreichen MPEG-2 und MPEG-4 Video-Standards folgt und Verbesserungen der Videoqualität wie auch der Kompression bietet. Der entscheidende Vorteil für IP Videosysteme ist die Möglichkeit, die gleiche digitale Video-Übertragung mit hoher Qualität bei gleichzeitig niedriger Latenz zu realisieren, dabei aber die Bandbreitenbelastung und Speicherkapazitäten um 25% bis 50% zu senken. Anders ausgedrückt, die Videoqualität ist bei gleicher Bandbreite erheblich besser als bei MPEG-4.



Was ist H.264?

H.264 ist ein Video-Codec (**Compressor** und **Decompressor**) Standard. Ein Video-Codec dient der Kompression und Dekompression der digitalen Videosignale und damit zur Reduzierung der benötigten Bandbreite bei der Übertragung und der Kapazität bei der Speicherung. Dies ist notwendig, da die Ausgangsdatenrate des unkomprimierten Videosignals nach ITU-601 (720x576 Pixel, 4:2:2 Abtastung, 8 Bit Digitalisierung und bei 25 B/s) über 165,9 Mbit/s beträgt – dies ist eine ca. 330-fache Überschreitung der Kapazität eines 512 kbit/s ADSL-Anschlusses und ergibt eine Aufzeichnungszeit von knapp über einer Stunde auf einer 80GB Festplatte.

Als Beispiel; eine Reduzierung des Videosignals auf SIF-Auflösung (352x288 Pixel, 4:2:0, 8 Bit, 25 B/s) und dem „Packen“ mittels einem standardmäßigen PC Programmen wie WinZip* oder gzip könnte eine Kompression von 10:1 erreicht werden. Um aber eine Übertragung von Live-Videoaufnahmen über eine ADSL-Verbindung und damit eine Aufzeichnungszeit von ca. 300 Stunden auf einer 80GB Festplatte zu erreichen, wird jedoch eine Kompression von mindestens 300:1 benötigt. Derartige Kompressionswerte sind mit MPEG-4 und H.264 erreichbar.

* WinZip® der WinZip International LCC

Den Standard implementieren

Bevor man sich eingehender mit H.264 befasst, ist es wichtig, den Unterschied eines Vergleiches verschiedener Standards und dem Vergleich zwischen verschiedenen Implementierungen eines Standards zu verstehen. Diese beiden Punkte unterscheiden sich wesentlich. Die Aussage: „*H.264 bietet eine bessere Videoqualität als MPEG-2*“ ist also etwas irreführend bzw. falsch.

H.264 ist ein Standard zur Videokompression. Er definiert die Syntax eines konformen Bitstromes den ein entsprechender Decoder genau „verstehen“ muss und der alle notwendigen, vom Standard definierten Funktions-Elemente enthält, damit er den Bitstrom entsprechend dekodieren kann.

Umgekehrt ist es auch möglich, dass ein H.264 Encoder eine Teil- oder Mehrmenge der vom Standard definierten Syntax enthält und dennoch einen konformen Bitstrom produziert. Encoder für CCTV-Anwendungen können

SEITENKASTEN

SIF oder CIF? Welches ist richtig?

Beide sind richtig, doch das eine ist gültig für MPEG-4, das andere für die alten H.261/H.263-basierenden Kompressionen. Was ist was?

CIF und SIF sind Formate für die Videoauflösung. Grundsätzlich beträgt die CIF-Auflösung 352x288 Pixel, unabhängig davon, ob der Videostandard NTSC oder PAL ist. Die SIF Auflösung beträgt dagegen 352x288 Pixel für PAL Kameras, aber 352x240 für NTSC Kameras. Dies wird manchmal in anderen Dokumenten auch als 320x240 Pixel gekennzeichnet – auch dies ist eine gültige SIF Auflösung. Die Zahlen sind jedoch irreführend, denn CIF ist SIF nicht überlegen.

CIF ist mit H.261/H.263 verbunden und SIF mit den MPEG-Versionen. Auf Grund der phonetischen Ähnlichkeit und der Tatsache, dass die Werte für PAL-Quellen identisch sind, werden leider die beiden Begriffe manchmal miteinander vertauscht, obgleich sie streng genommen unterschiedlich sind. Es kommt also vor, dass man über CIF spricht, aber eigentlich SIF meint.

Wenn Sie also ein CIF (oder 2CIF oder 4CIF) System haben, dann benutzen Sie einen auf dem H.261/263-Standard basierenden Codec, und diese sind mit MPEG-4 oder H.264 nicht konform! Alle MPEG-4 und H.264 Systeme basieren auf der Definition SIF. Wenn Sie also das richtige MPEG-4 wollen, ob Teil 2 oder Teil 10, dann brauchen Sie das „S!“

daher auch zusätzliche Funktionen (Implementierungen und Algorithmen) enthalten, die so im Standard nicht definiert, aber auch nicht ausgeschlossen sind und damit individuellen Strategien für verschiedene Anwendungen folgen. H.264 Encoder von verschiedenen Anbietern erzeugen also bei gleicher Bitrate Videobilder in unterschiedlicher Qualität oder bei gleicher Qualität mit unterschiedlichen Bitraten.

Die richtigere Aussage lautet also: *“H.264 bietet eine erheblich reichere Syntax und ein umfangreicheres Angebot an Möglichkeiten als z. B. MPEG-2 und ermöglicht daher die Implementierung eines weit leistungsstärkeren Video Encoder, der bei gleicher Bitrate qualitativ höherwertiges Video und qualitativ gleichwertiges Video bei einer viel niedrigeren Bitrate erzeugen kann“*, als alle anderen Kompressions-Verfahren.

Dies kann anhand des Vergleichssoftware-Encoder (JM11), der von der Internationalen Organisation zur Normung (ISO) frei erhältlich ist, veranschaulicht werden. Mit Hilfe dieses Vergleichs-Encoder kann ein Betreiber die Elemente zur Enkodierung einer bestimmten Videosequenz auswählen. Die nachstehende Tabelle zeigt die Ergebnisse mehrerer Enkodierungen derselben Videosequenz durch den H.264 Vergleichs-Encoder, wenn jeweils verschiedene Elemente eingesetzt werden. Die Ausgangs-Bitströme der jeweiligen Tests sind konforme H.264 Bitströme und weisen jeweils vergleichbare Videoqualität auf.

Elemente	Bitrate (kbit/s)	Gesamtausführungszeit der benötigten Kodierungs-elemente (relativer Bezug zueinander)
ausschließlich I-Bild Enkodierung	2279	1
I und P-Bilder, jedoch ohne Bewegungsvorhersage (kein Suchbereich)	1055	1,5
I und P-Bilder mit einer +/-16 Blocksuche bei Verwendung eines vereinfachten Such-Algorithmus	453	14
I und P-Bilder bei Verwendung des kompletten Suchalgorithmus und mit variabler Blockgröße für die Bewegungskompensation	421	56

Aus dieser Tabelle geht klar hervor, dass bei gleicher Videoqualität der Kompressionswert zunimmt, wenn mehr Elemente und Algorithmen eingesetzt werden. Es wird jedoch auch deutlich, dass der Einsatz von zusätzlichen Elementen eine Zunahme an Komplexität mit sich bringt – in diesem Fall ist dieses anhand des markanten Anstiegs der relativen Durchführungszeit des Enkodierungsprozesses zu erkennen. Diese Zunahme an Komplexität ist auch der Grund dafür, dass beim Entwurf eines H.264 Encoders oft einige Elemente oder Algorithmen weggelassen werden.

Bezug zu MPEG-4 Teil 2

MPEG-4 (ISO/IEC 14496) ist eine Übersicht der einzelnen Standards, die die Kodierung von audio-visuellen Objekten definiert. Die Liste enthält eine Anzahl von *Teilen*, die die Video- und Audiokompressions-Standards definieren, sowie Teile, die die Systemebene z. B. die Merkmale des MPEG-4 Dateiformat beschreiben. Die in vielen heutigen Produkten eingesetzte Videokompression basiert auf dem traditionell mit DCT (Direct Cosinus Transformation) arbeitendem MPEG-4 Teil 2 (ISO/IEC 14496-2) Standard.

Der H.264 Videokompressionsstandard wurde in MPEG-4 als MPEG-4 Teil 10 (ISO/IEC 14496-10) aufgenommen. Dies bedeutet, dass MPEG-4 nun zwei Videokompressionsstandards anbietet. Diese beiden Videokompressionsstandards sind jedoch nicht vollständig kompatibel, da jeder Standard unterschiedliche Methoden zur Kompression und Wiedergabe der Daten einsetzt, d. h. ein MPEG-4 Teil 10 (H.264) Dekodierer kann keinen MPEG-4 Teil 2 Bitstrom dekodieren und umgekehrt.

IP Video und H.264

Am deutlichsten erkennt man die Vorteile von H.264 an Hand einer vorliegenden Implementierung des Standards, wie z. B. in der neuen Serie 9000 von IndigoVision.

In einem Transmitter 9000 werden die analogen Videobilder (z. B. einer Kamera) digitalisiert und an einen internen H.264 Encoder zur Kompression geleitet. Für die Kompression stehen für jedes Videobild zwei Möglichkeiten zur Verfügung, das I-Bild oder das P-Bild.

Ein I-Bild ist ein Videobild, das ohne Bezug zu einem anderen Videobild enkodiert wurde. Ein Videostrom oder eine Aufzeichnung beginnt immer mit einem I-Bild. In einem kontinuierlichen Datenstrom werden in regelmäßigen Abständen immer wieder neue I-Bilder übertragen. Diese I-Bilder, die auch *intra frames, key*

frames (Schlüsselbilder) oder access points (Zugriffspunkte) genannt werden, sind für die uneingeschränkte Bildwiedergabe in aufgezeichneten H.264 Dateien unabdingbar, z. B. beim Vor- und Rücklauf oder bei Suchvorgängen. Der regelmäßige Abstand dieser I-Bilder wird als I-Bild Intervall bezeichnet. Der Nachteil der I-Bilder besteht jedoch darin, dass sie im Allgemeinen sehr viel größer als P-Bilder sind und damit den wesentlichen Anteil des Datenstromes ausmachen.

P-Bilder sind Rahmen, die mit *Bewegungskompensation (Motion Compensation)* arbeiten, d. h. der Encoder nutzt den Unterschied zwischen dem aktuell zu verarbeitendem Bild und einem vorherigen Videobild und sorgt dafür, dass Informationen die sich nicht ändern (z. B. ein unveränderter Hintergrund) nicht mehrmals übertragen werden. Im Gegensatz zu rein auf Bilddifferenzen basierten Verfahren wie z. B. M-JPEG sucht H.264 nicht nur nach Unterschieden, sondern auch nach Bewegungen die innerhalb der Videobilder vorhanden sind. Dies bedeutet bei vorliegenden Bildänderungen, dass bewegungskompensierende Verfahren den einfachen, auf reine Bildunterschiede basierenden Verfahren überlegen sind. Den Vorgang der Bewegungssuche nennt man *Bewegungsvorhersage oder Motion Estimation (ME)*.

Innerhalb des H.264 Encoder ist die Funktion Bewegungsvorhersage das rechenaufwendigste und für die Leistung das am meisten entscheidende Element. Die Bewegungsvorhersage ist insgesamt ein sehr komplexer Vorgang. Viele Encoder auf Basis Echtzeit-Software oder DSP reduzieren daher häufig die Suchbereiche der Bewegungserkennung oder sie benutzen nur einen eingeschränkten Suchalgorithmus, damit die Echtzeitleistung erhalten bleibt. Dies führt jedoch häufig zu schlechter Videoqualität und zu einer deutlich verringerten Kompression.

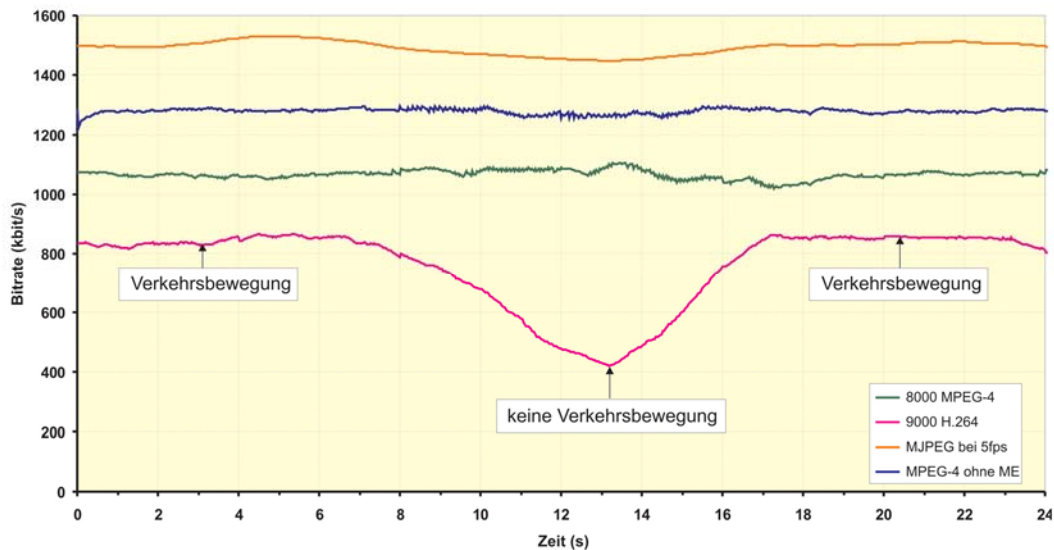


Die für die Bewegungsvorhersage verwendete CCTV-Szene

Die folgende Grafik verdeutlicht die möglichen Datenraten bei einer Videobildsequenz (Bild rechts). In diesem Beispiel enkodieren vier verschiedene Encoder dieselbe Videosequenz:

- IndigoVision 8000 MPEG-4 Teil 2
- der neue IndigoVision 9000 MPEG-4 Teil 10, H.264
- ein MPEG-4 Encoder ohne Bewegungsvorhersage
- ein M-JPEG Encoder.

Alle arbeiten mit 25B/s (Ausnahme der M-JPEG mit 5B/s), alle zeigen eine subjektiv ähnliche Videoqualität.



MPEG-4 und H.264 Bitraten bei subjektiv gleicher Videoqualität.

Die Grafik zeigt, dass im Vergleich zu MPEG-4 der Einsatz von H.264 eine Reduzierung der Bandbreite um etwa 20 bis 25% bringt und bis zu 50% wenn keine/wenig Veränderungen im Bild vorliegen – z. B. hier bei geringem Verkehrsaufkommen. Es wird jedoch nicht nur die Gesamtbeanspruchung des Netzwerkes durch das IP Videosystem verringert, sondern noch ausschlaggebender ist u. U. die signifikante Verringerung der

zur Aufzeichnung des Videos benötigten Speicherkapazität, die oft eine der teuersten Komponenten des Systems ist.

Überlegungen zur Hardware

Wenn man die Möglichkeiten zur Implementierung von H.264 betrachtet, wird deutlich, dass die Ansprüche an die Verarbeitungsleistung des Codecs entscheidend sind. Dies gilt besonders wenn das alle Verarbeitungselemente des H.264 Standards verwendet werden und die gesamten Vorteile der Technologie umgesetzt werden sollen. H.264 ist ein Videokompressionsstandard für universelle Anwendungen und ist nicht speziell für digitale CCTV-Anwendungen entworfen worden. Durch den Einsatz eines sehr speziellen FPGA-basierenden Konzeptes, wie z. B. in der Serie 9000 von IndigoVision, wird die hier notwendige hohe Verarbeitungsleistung bereitgestellt und es kann zusätzlich optimal auf CCTV-Applikationen zugeschnitten werden. Bei geringer Aktivität im Video – einer typischen Situation in vielen Überwachungsapplikationen - ist z. B eine noch höhere Kompression möglich. Der maßgeschneiderte FPGA-Ansatz weist eine Reihe weiterer Vorteile auf:

- Eine qualitativ hochwertige Videoaufzeichnung kann, unabhängig von der Bitrate und der Bildinhalte, auch bei vielen, schnellen Bildänderungen ohne Auslassen von Bildern aufrecht erhalten werden. Dies ist z. B. für Anwendungen bei der Überwachung von Spieltischen in Casinos entscheidend.
- Eine kostengünstige, hohe Leistung bei der Enkodierung von 4SIF@25B/s Videoaufzeichnungen, die mit dem H.264 Standard vollständig konform ist.
- Ein Vorort-Upgrade ist mit der Weiterentwicklung des Kompressionsstandards möglich; ein Vorteil, denn H.264 steht erst am Anfang seiner Möglichkeiten.
- Echtzeit-Analytics Algorithmen können statt in Software in spezieller, dafür entwickelter Hardware ausgeführt werden. Durch deren flexiblen Einsatz, d. h. auch im Livemodus auf der Kameraseite, erhält man eine umfassend und sehr flexible Lösung.

Zusammenfassung

H.264 bietet für Anwender und Planer eine Reihe neue Möglichkeiten und Vorteile. Der zusätzliche Aufwand und die komplexere Implementierung führen jedoch auch zu höheren Produktkosten. MPEG-4 wird deshalb nicht über Nacht durch H.264 ersetzt werden. Beide werden auf längere Zeit parallel zum Einsatz kommen, was dem Endanwender ein breiteres Spektrum an Lösungen zur Verfügung stellt und zwar technische wie kostenmäßig. IndigoVision wird daher noch für viele Jahre die neue H.264 Serie 9000 und die derzeitige 8000 MPEG-4 Serie weiterführen und auch weiterentwickeln. Obwohl die MPEG-4 und H. 264 Kompressionsstandards nicht direkt kompatibel sind, ermöglichen die IP Netzwerke und das IndigoVision Systemkonzept eine Koexistenz und den uneingeschränkten Mischbetrieb der beiden Systeme. Da sowohl die Videomanagement- und die Aufzeichnungssoftware von IndigoVision beide Bitstrom-Typen direkt „verstehen“, ist eine schrittweise Auf- oder Umrüstung der MPEG-4 Systeme mit H.264 basierten Produkten möglich.

Über den Autor

Barry Keepence trat 1999 als Chief Technical Officer (CTO) bei IndigoVision ein. Vorher arbeitete er über 10 Jahre lang in der Weltraumindustrie und bei 3Com als Systemarchitekt.

*Barry Keepence, CTO,
IndigoVision Group plc,
The Edinburgh Technopole
Bush Loan, Edinburgh EH26 0PJ
Tel: + 44 131 475 7200
Fax: +44 131 475 7201
www.indigovision.com*