

H.264 Video-Komprimierungsstandard. Neue Möglichkeiten bei der Videoüberwachung.

Inhalt

1. Einleitung	3
2. Entwicklung von H.264	3
3. Funktionsweise von Videokomprimierung	4
4. H.264-Profiles und -Level	5
5. Frames	5
6. Grundlegende Methoden der Datenreduzierung	6
7. Komprimierungseffizienz von H.264	8
8. Zusammenfassung	9

1. Einleitung

Der neueste Videokomprimierungsstandard, H.264 (auch als MPEG-4 Part 10/AVC für "Advanced Video Coding" bezeichnet) wird sich nach Meinung von Experten in den nächsten Jahren als Videostandard der Wahl durchsetzen.

H.264 ist ein offener, lizenzierter Standard, der die meisten heute üblichen, wirksamen Videokomprimierungsverfahren unterstützt. Ohne Einbußen bei der Bildqualität erreicht der H.264-Encoder bei digitalen Videodateien im Vergleich zu Motion JPEG eine 80 % höhere Komprimierung. Gegenüber dem Standard MPEG-4 Part 2 wird eine um 50 % höhere Komprimierung erzielt. Das bedeutet, dass Videodateien wesentlich weniger Netzbandbreite und Speicherplatz beanspruchen. Umgekehrt kann bei gleicher Bitrate eine deutlich höhere Videoqualität erzielt werden.

Aufgrund der gemeinsamen Norm der internationalen Normierungsgremien für Telekommunikation und IT wird erwartet, dass H.264 eine weitere Verbreitung als die früheren Normen und Standards findet.

H.264 hat sich bereits in nutzerorientierten Geräten wie Handys und digitalen Videoplayern bewährt und sich bei den Endverbrauchern in kurzer Zeit durchgesetzt. Dienstanbieter wie Online-Videospeicher- und Telekommunikationsunternehmen haben ebenfalls mit der Implementierung von H.264 begonnen.

In der Videoüberwachungsbranche wird sich H.264 vermutlich zuerst in Anwendungen durchsetzen, die hohe Ansprüche an Bildfrequenzen und Auflösung stellen, so bei der Überwachung von Autobahnen, Flughäfen und Spielhallen, bei der Bildfrequenzen von 30 bzw. 25 fps (NTSC bzw. PAL) die Norm sind. In diesen Bereichen ist das Potenzial von Kosteneinsparungen durch geringere Anforderungen an Bandbreite und Speicher am größten.

Ebenso wird erwartet, dass H.264 die Akzeptanz von Megapixel-Kameras in der Überwachungsbranche beschleunigt, da das hocheffiziente Komprimierungsverfahren deutlich reduzierte Dateigrößen und Bitraten ohne Einbußen bei der Bildqualität ermöglicht. Das Verfahren birgt jedoch auch Nachteile. H.264 sorgt zwar für niedrigere Kosten bei Netzwerkbandbreite und Speicher, erfordert jedoch Netzwerkkameras und Überwachungsstationen mit höherer Rechenleistung.

2. Entwicklung von H.264

H.264 ist aus einem gemeinsamen Projekt der ITU-T-Videocodierungsgruppe (VCEG) und der ISO/IEC MPEG-Gruppe (Moving Picture Experts Group) hervorgegangen. Im Sektor ITU-T werden die Telekommunikationsnormen der Internationalen Fernmeldeunion (ITU) koordiniert. ISO ist das Kürzel der "International Organization for Standardization" (Internationale Organisation für Normung). Die IEC (Internationale elektrotechnische Kommission) ist für das Normierungswesen für Elektrotechnik und Elektronik zuständig. Die Bezeichnung "H.264" wird von der ITU-T genutzt, während ISO/IEC den Standard als "MPEG-4 Part 10/AVC" bezeichnen, da sie ihn als neuen Bestandteil der MPEG-4-Normenreihe führt. Die Normenreihe MPEG-4 enthält zum Beispiel MPEG-4 Part 2, eine Norm, die bisher für IP-basierte Video-Encoder und Netzwerk-Kameras üblich war.

H.264 sollte verschiedene Schwachstellen bei den bisherigen Videokomprimierungsstandards beheben. Dies ist in mehrerlei Hinsicht gelungen:

- > Implementierungen mit einer durchschnittlichen Senkung der Bitrate um 50 % gegenüber allen anderen Standards bei gleicher Bildqualität
- > Hohe Fehlertoleranz, dadurch Ausgleich von Übertragungsfehlern über mehrere Netzwerke hinweg
- > Niedrige Latenz und bessere Qualität bei höherer Latenz
- > Einfache Syntax-Spezifikation, dadurch erleichterte Implementierung
- > Präzises Decodierungsverfahren; definiert das Verfahren für numerische Berechnungen im Encoder und Decoder zur Vermeidung von Fehlerakkumulationen

H.264 zeichnet sich zudem durch die flexible Unterstützung einer hohen Bandbreite von Anwendungen mit sehr unterschiedlichen Bitratenanforderungen aus. Bei Videoanwendungen aus dem Unterhaltungsbereich zum Beispiel (wie terrestrische Übertragung, Satelliten-, Kabel- und DVD-Anwendungen), kann H.264 1 bis 10 Mbit/s bei hoher Latenz bieten, während bei Telekommunikationsdiensten Bitraten unter 1 Mbit/s bei niedriger Latenz möglich sind.

3. Funktionsweise von Videokomprimierung

Die Videokomprimierung sorgt durch die Reduzierung und Eliminierung redundanter Videodaten für eine effiziente Übertragung und Speicherung digitaler Videodateien.

Hierbei wird das Quellvideo über einen Algorithmus in eine komprimierte Datei umgewandelt, die dann übertragen oder gespeichert werden kann. Zur Wiedergabe der komprimierten Datei wird ein inverser Algorithmus angewendet. Dadurch wird Video erzeugt, das im Wesentlichen denselben Inhalt bietet wie das ursprüngliche Quellvideo. Die Zeit, die zur Komprimierung, Übertragung, Dekomprimierung und Wiedergabe einer Datei erforderlich ist, wird als Latenz bezeichnet. Je weiter der Komprimierungsalgorithmus entwickelt ist, desto höher ist bei gegebener Verarbeitungsleistung auch die Latenz.

Ein zusammengehöriges Algorithmenpaar wird als Video-Codec (Encoder/Decoder) bezeichnet. Video-Codex, die unterschiedliche Standards umsetzen, sind normalerweise nicht miteinander kompatibel, d. h. dass Videoinhalte, die mit einem Standard komprimiert worden sind, nicht mit einem anderen Standard dekomprimiert werden können. Zum Beispiel passt ein MPEG-4 Part 2 Decoder nicht mit einem H.264-Encoder zusammen, da ein bestimmter Algorithmus nicht das Ergebnis eines anderen Algorithmus entschlüsseln kann. Man kann allerdings mehrere verschiedene Algorithmen in derselben Software oder Hardware implementieren, so dass verschiedene Formate komprimiert werden können.

Die verschiedenen Videokomprimierungsstandards nutzen unterschiedliche Methoden zur Datenreduzierung. Daher unterscheiden sie sich hinsichtlich Bitrate, Qualität und Latenz.

Ebenso können Encoder mit demselben Komprimierungsstandard unterschiedliche Ergebnisse erzeugen, da der Encoder-Programmierer unterschiedliche Kombinationen von Tools implementieren kann, die für den Standard definiert sind. Solange die Ausgabe eines Encoders mit dem Standardformat und dem Decoder übereinstimmt, können unterschiedliche Implementierungen realisiert werden. Dies ist von Vorteil, da sich die Ziele und die Kostenanforderungen der verschiedenen Implementierungen voneinander unterscheiden. Professionelle Software-Encoder zum Mastering optischer Medien ohne Echtzeiterfordernisse müssen Video besser codieren können als Echtzeit-Hardware-Encoder für Videokonferenzen in Handheld-Geräten. Ein bestimmter Standard kann daher keine bestimmte Bitrate oder Qualität garantieren. Zudem lässt sich die Leistung eines Standards nur dann objektiv mit anderen Standards oder auch mit anderen Implementierungen desselben Standards vergleichen, wenn zuerst definiert wird, wie er implementiert wird.

Anders als beim Encoder müssen im Decoder alle erforderlichen Bestandteile eines Standards implementiert sein, damit eine Decodierung normgerechter Bitströme gewährleistet ist. Dies ist erforderlich, weil der Standard genau festlegt, auf welche Weise ein Dekomprimierungsalgorithmus die einzelnen Bits von dekomprimiertem Video wiederherstellt.

Das nachstehende Diagramm vergleicht die Bitraten der folgenden Videostandards bei gegebener Bildqualität: Motion JPEG, MPEG-4 Part 2 (ohne Bewegungsausgleich), MPEG-4 Part 2 (mit Bewegungsausgleich) und H.264 (Grundprofil).

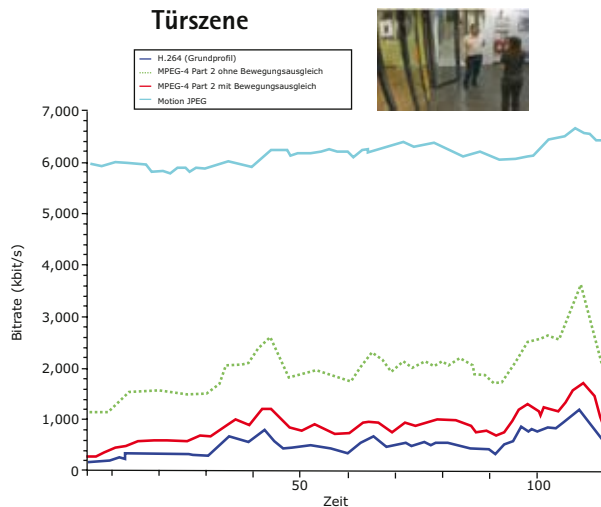


Abbildung 1. Der H.264-Encoder erzeugt aus einer Muster-Videosequenz bis zu 50 % weniger Bit pro Sekunde als der MPEG-4-Encoder mit Bewegungsausgleich. Der H.264-Encoder erwies sich als mindestens dreimal so effizient wie ein MPEG-4-Encoder ohne Bewegungsausgleich und mindestens sechsmal so effizient wie Motion JPEG.

4. H.264-Profil und -Level

Die gemeinsame Arbeitsgruppe zur Definition von H.264 legte großen Wert auf eine einfache und saubere Lösung und reduzierte daher die Optionen und Funktionen auf ein Mindestmaß. Ein wichtiger Aspekt des Standards ist, ebenso wie bei anderen Videostandards, die Definition von Profilen (Sätze von algorithmischen Funktionen) und Levels (Leistungsklassen), die gängige Produktionen und gemeinsame Formate unterstützen.

H.264 bietet sieben Profile, die jeweils auf eine bestimmte Anwendungsklasse ausgerichtet sind. Jedes Profil definiert die Funktionen, die der Encoder nutzen kann und begrenzt die Komplexität der Decoder-Implementierung.

Für Netzwerk-Kameras und Video-Encoder dürfte sich in erster Linie das Grundprofil eignen, das vor Allem für Anwendungen mit begrenzten Rechenressourcen konzipiert wurde. Das Grundprofil ist am besten auf die Leistungsfähigkeit von Echtzeit-Encodern in Netzwerk-Videoprodukten abgestimmt. Es ermöglicht eine niedrige Latenz, die eine wichtige Anforderung für Überwachungsvideo darstellt und für die PTZ-Steuerung von Netzwerk-Kameras in Echtzeit besonders wichtig ist.

H.264 stellt 11 Levels oder Leistungsstufen bereit, die die Anforderungen an Leistung, Bandbreite und Speicher begrenzen. Jeder Level definiert die Bitrate und Codierungsrate in Makroblöcken pro Sekunde für Auflösungen von QCIF bis HDTV und höhere Anforderungen. Je höher die Auflösung, desto höher ist der erforderliche Level.

5. Frames

Je nach H.264-Profil können in einem Encoder verschiedene Frame-Arten wie I-Frames, P-Frames und B-Frames genutzt werden.

Ein I-Frame oder Intra-Frame ist ein unabhängiger Frame, der separat und ohne Verweis auf andere Bilder decodiert werden kann. Das erste Bild in einer Videosequenz ist stets ein I-Frame. I-Frames werden als Ausgangspunkte für neue Betrachter oder als Synchronisierungspunkte benötigt, wenn der übertragene Bit-Strom beschädigt ist. I-Frames eignen sich zur Implementierung von Vorlauf, Rücklauf und anderen Funktionen mit Direktzugriff. Encoder fügen I-Frames automatisch in regelmäßigen Abständen ein, aber auch auf Anforderung, wenn erwartet wird, dass sich neue Clients in die Wiedergabe des Datenstroms einklinken. Nachteil der I-Frames ist der weitaus höhere Bitverbrauch. Allerdings sind sie relativ unanfällig für Artefakte.

P-Frames ("Predictive Inter Frame") verweisen bei ihrer Codierung auf Bestandteile vorangegangener I- oder P-Frames. P-Frames erfordern weniger Bits als I-Frames, sind jedoch aufgrund der komplexen Abhängigkeit von früheren P- und I-Frames relativ anfällig für Übertragungsfehler.

B-Frames ("Bi-Predictive Inter Frames") verweisen zugleich auf einen vorausgehenden und auf einen späteren Frame.

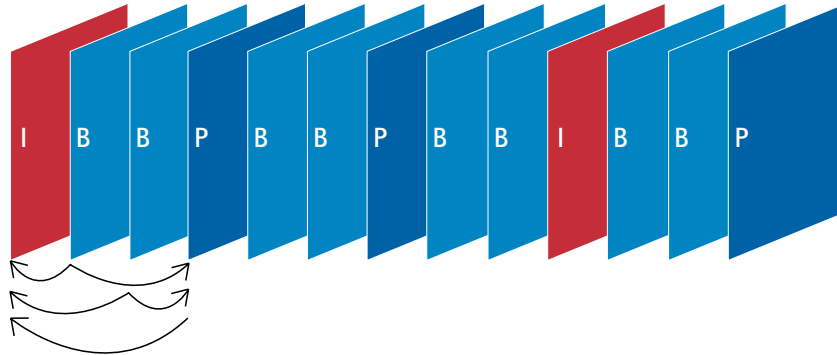


Abbildung 2. Typische Sequenz mit I-, B- und P-Frames. P-Frames verweisen nur auf vorausgehene I- und P-Frames, B-Frames ggf. auch auf nachfolgende I- oder P-Frames.

Wenn ein Video-Decoder ein Video wiederherstellt, indem er den Bitstrom Frame für Frame decodiert, muss er stets an einem I-Frame beginnen. Vorhandene P-Frames und B-Frames müssen zusammen mit ihren Referenz-Frames decodiert werden.

Im H.264-Grundprofil werden nur I- und P-Frames verwendet. Dieses Profil eignet sich ideal für Netzwerk-Kameras und Video-Encoder, da durch den Verzicht auf B-Frames niedrige Latenzen erzielt werden.

6. Grundlegende Methoden der Datenreduzierung

Zur Reduzierung von Videodaten stehen verschiedene Methoden innerhalb einzelner Bild-Frames und zwischen den Frames einer Bildfolge zur Verfügung.

Im Bildframe lassen sich Daten durch einfache Eliminierung unnötiger Daten reduzieren. Dies wirkt sich auf die Bildauflösung aus.

In einer Reihe von Frames können Videodaten durch Methoden wie der Differenzcodierung reduziert werden, die von den meisten Videokomprimierungsstandards genutzt werden, darunter auch H.264. Bei der Differenzcodierung wird der Frame mit einem Referenz-Frame, also einem vorausgegangenen I- oder P-Frame verglichen. Codiert werden nur die Pixel, die sich nach dem Referenz-Frame verändert haben. Auf diese Weise wird die Anzahl der Pixelwerte reduziert, die codiert und übertragen werden müssen.



Abbildung 3. Im Motion JPEG-Format werden die drei Bilder in der obigen Folge codiert und als separate, eindeutige Bilder (I-Frames) übertragen, die voneinander unabhängig sind.

— Übertragen - - - Nicht übertragen



Abbildung 4. Bei der Differenzcodierung (in den meisten Videokomprimierungsstandards einschließlich H.264 gebräuchlich) wird nur das erste Bild (I-Frame) vollständig codiert. In den beiden folgenden Bildern (P-Frames) wird bei den statischen Bestandteilen, in diesem Fall das Haus, nur auf das erste Bild verwiesen. Nur die bewegten Bestandteile, hier der Läufer, werden mit den Bewegungsvektoren codiert. Dadurch wird die gesendete und gespeicherte Datenmenge verringert.

Die Codierungsmenge kann weiter reduziert werden, wenn die Erkennung und Codierung von Differenzen auf Pixelblocks (Makroblocks) und nicht auf einzelnen Pixeln basiert. Hierbei werden größere Bereiche verglichen, und nur Blocks mit wesentlichen Differenzen werden codiert. Der Datenaufwand zum Verweis auf die geänderten Bildbereiche wird ebenfalls reduziert.

Die Differenzcodierung kann die Datenmenge aber nur unwesentlich reduzieren, wenn der Bewegungsanteil im Video hoch ist. In diesem Fall sind Verfahren wie der blockbasierte Bewegungsausgleich von Nutzen. Beim blockbasierten Bewegungsausgleich wird berücksichtigt, welche Anteile des neuen Frames einem früheren Frame in der Videosequenz entnommen werden können, ggf. aus einer anderen Stelle. Bei diesem Verfahren wird der Frame in eine Reihe von Makroblöcken unterteilt. Ein neuer Frame, z. B. ein P-Frame, kann dann Block für Block zusammengesetzt oder durch die Suche nach einem übereinstimmenden Block in einem Referenz-Frame vorhergesagt werden. Wird eine Übereinstimmung gefunden, codiert der Encoder nur die Position des passenden Blocks im Referenz-Frame. Die Codierung dieses Bewegungsvektors erfordert eine geringere Bitmenge als der eigentliche Inhalt des codierten Blocks.

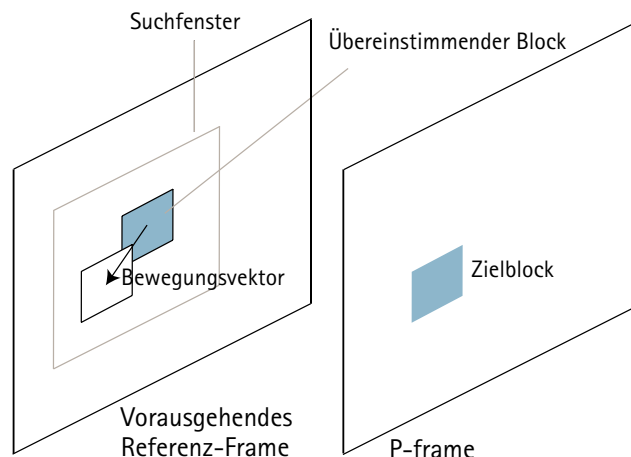


Abbildung 5. Blockbasierter Bewegungsausgleich.

7. Komprimierungseffizienz von H.264

Mit H.264 erreicht die Videokomprimierung ein neues Niveau. H.264 nutzt bei der Codierung von I-Frames ein neues, modernes Verfahren zur Vorhersage von Differenzdaten. Hierdurch lässt sich die Bitgröße bei hoher Bildqualität erheblich reduzieren. Hierzu werden in den einzelnen Makroblocks des Frames kleinere Pixelblocks sukzessive vorhergesagt. Hierzu wird nach Übereinstimmungen in den zuvor bereits codierten Pixeln gesucht, die an einen neuen 4x4 Pixel-Block angrenzen, der intra-codiert werden soll. Durch die Wiederverwendung von bereits codierten Pixelwerten lässt sich die Bitgröße drastisch reduzieren. Die neue Intra-Vorhersage ist ein Schlüsselement des H.264-Verfahrens, das sich als äußerst effizient erwiesen hat. Selbst wenn in einem H.264 ausschließlich I-Frames verwendet würden, wäre die Datei erheblich kleiner als bei einem Motion JPEG-Strom, der ausschließlich I-Frames nutzt.

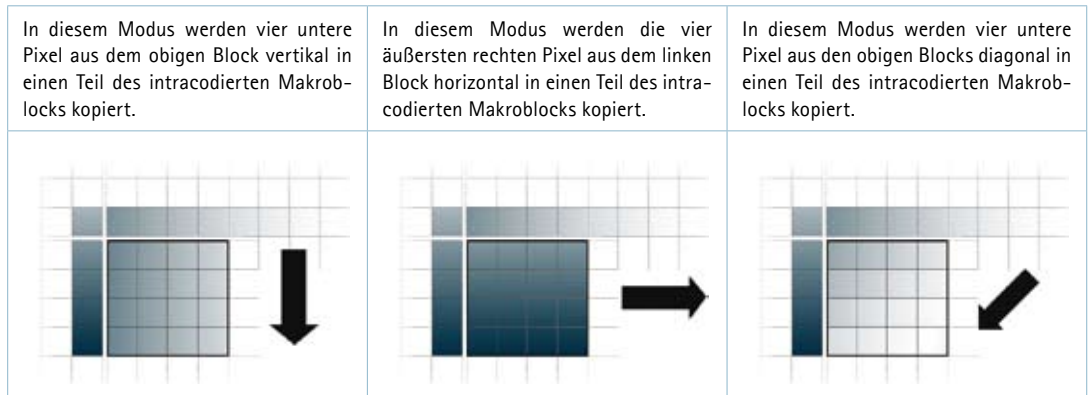


Abbildung 6. Einige der Modi für die Intra-Vorhersage zur Codierung von 4x4 Pixeln in einem der 6 Blocks, aus denen sich ein Makroblock zusammensetzen kann. Jeder der 16 Blocks in einem Makroblock kann in einem anderen Modus codiert werden.

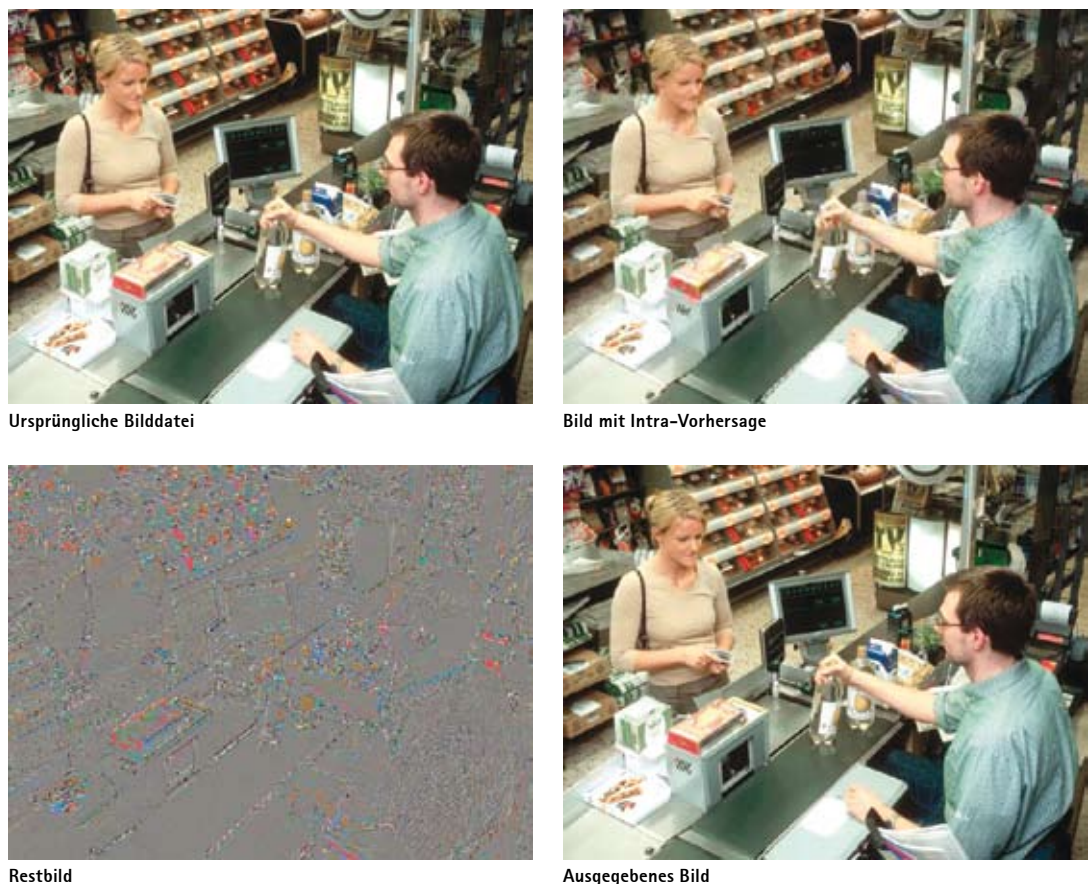


Abbildung 7. Die obigen Bilder veranschaulichen die Effizienz des Intra-Vorhersageschemas von H.264, bei dem das vorhergesagte Bild aufwandlos übermittelt wird. Für die Bildausgabe müssen nur der Restinhalt und die Zwischenbildvorhersage codiert werden.

Der blockbasierte Bewegungsausgleich – verwendet beim Codieren von P- und B-Frames – wurde ebenfalls bei H.264 verbessert. Der H.264-Encoder kann – bis zu einer Genauigkeit von Teilpixeln – in einigen oder zahlreichen Bereichen von einem oder mehreren Referenz-Frames nach übereinstimmenden Blocks suchen. Blockgröße und -form können zugunsten einer genaueren Übereinstimmung angepasst werden. In Bereichen ohne passende Blocks in Referenz-Frames werden intracodierte Makroblocks genutzt. Die hohe Flexibilität des blockbasierten Bewegungsausgleichs von H.264 zahlt sich besonders in dicht gedrängten Überwachungsszenen aus, bei denen die Qualität für anspruchsvolle Anwendungen aufrechterhalten werden kann. Der Bewegungsausgleich ist die größte Herausforderung für den Video-Encoder. Die verschiedenen Methoden und Stufen, mit der er vom H.264-Encoder implementiert werden kann, beeinflussen die Komprimierungsstärke des Videostroms.

Unter H.264 können die typischen blockartigen Artefakte, die bei hoch komprimiertem Video im JPEG-Standard und den bisherigen MPEG-Standards auffallen, mit einem internen Deblocking-Filter reduziert werden. Der Filter glättet Blockränder mit regelbarer Stärke und sorgt für eine sehr hohe Qualität der dekomprimierten Bilddaten.



Abbildung 8. Blockartige Artefakte im hoch komprimierten Bild links werden durch einen Deblocking-Filter reduziert, wie im rechten Bild gezeigt.

8. Zusammenfassung

H.264 bedeutet in der Videokomprimierung einen großen Schritt nach vorn. Durch verbesserte Vorhersageverfahren und die geringere Fehleranfälligkeit kann der neue Standard Verfahren mit höherer Komprimierungswirkung bereitstellen. Er eröffnet neue Möglichkeiten zur Programmierung besserer Video-Encoder, die Videoströme höherer Qualität, höhere Bildfrequenzen und höhere Auflösungen bei (im Vergleich zu älteren Standards) gleicher Bitrate oder umgekehrt, dieselbe Bildqualität bei reduzierter Bitrate bieten.

H.264 ist der erste gemeinsame, internationale Standard der Normierungsorganisationen ITU, ISO und IEC für die Videokomprimierung. Aufgrund seiner Flexibilität kommt H.264 in unterschiedlichen Bereichen wie High Definition-DVD (z. B. Blu-ray), digitales Fernsehen einschließlich HDTV, Online-Videospeichern (z. B. YouTube), in der Mobiltelefonie der dritten Generation und in Software wie QuickTime, Flash und dem Apple-Betriebssystem MacOS X zum Einsatz, aber auch in Videospielekonsolen wie der PlayStation 3. Mit der Unterstützung vieler Branchen und Anwendungen für den Kunden- und Profibedarf wird H.264 nach Meinung von Experten an die Stelle von heute üblichen Komprimierungsstandards und -verfahren treten.

Durch die wachsende Verfügbarkeit von H.264 in Netzwerk-Kameras, Video-Encodern und Videoverwaltungssoftware müssen Systemdesigner und -integratoren bei der Auswahl von Produkten und Anbietern vermehrt auf eine Unterstützung dieses neuen offenen Standards achten. Bis dahin sind Netzwerk-Video-Produkte, die sowohl H.264 als auch Motion JPEG unterstützen, die ideale Wahl für maximale Flexibilität und Integrationsmöglichkeiten.

Hintergrund Axis Communications

Axis ist ein IT-Unternehmen, das Netzwerk-Videolösungen für professionelle Installationen anbietet. Das Unternehmen ist der weltweite Marktführer im Bereich Netzwerk-Video und treibt den Wechsel von analoger zu digitaler Videoüberwachungs-Technologie an. Die Produkte und Lösungen von Axis konzentrieren sich auf Anwendungen wie Sicherheits- und Fernüberwachung und basieren auf einer innovativen und offenen Technologie-Plattform.

Axis ist an der OMX Nordic Exchange, Large Cap, Information Technology notiert. Das Unternehmen ist weltweit mit Niederlassungen in 18 Ländern tätig und arbeitet mit Vertriebspartnern und Systemintegratoren in 70 Ländern zusammen. Weitere Informationen über Axis finden Sie unter www.axis.com