

Anforderungen an den idealen CCTV Video Codec

1. Einführung

Dem Thema Videokompression kommt in digitalen CCTV Systemen eine zentrale Bedeutung zu. Kernkomponente der Videoverarbeitung ist dabei der aus Encoder- (Kompressor) und Decoder (Dekompressor) bestehende Codec (Bild 1).

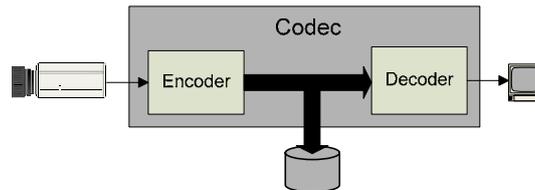


Bild 1: Vereinfachtes Video-Codec System

Die in einem digitalen CCTV System verwendeten Codecs bestimmen maßgeblich dessen Leistungsfähigkeit, Flexibilität in Bezug auf die Vielfalt von Sicherheitsszenarien und natürlich auch die Kosten für Speicherung und Übertragung der digitalen Videodatenströme. Es zeigt sich, dass die Anforderungen an CCTV Video Codecs weit über das Thema hohe Kompression bei hoher Qualität hinausgehen. Neben der eigentlichen Kompression haben CCTV Codecs noch Aufgaben wie:

- Video- und Audio-Sensorik (z. B. Bewegungskennung)
- Video- und Audio-Processing (z. B. Wasserzeichen)

Heute gibt es eine große Vielfalt von Video Codecs - teilweise standardisiert und teilweise proprietär. Die Entwicklung der den Codec-Implementierungen zugrunde liegenden Standards wurde jedoch nicht von den Erfordernissen des CCTV Marktes als kleiner Nische für Bildanwendungen sondern von den Anforderungen des Multimedia-Massenmarktes und speziell des Fernseh- bzw. DVD-Marktes getrieben. Dies hat zur Folge, dass die standardisierten Videokompressionsverfahren die Bedürfnisse von CCTV nur ungenügend widerspiegeln. Selbst wenn der jeweilige Standard Freiheitsgrade definiert, die für CCTV Anwendungen sinnvoll sind, werden diese bei den aufwändigen Hard- und Software-Implementierungen der Codecs nicht berücksichtigt. Die Verwendung derartiger Multimedia-Codecs für CCTV Zwecke führt oft zu großen Kompromissen bei der Lösung von CCTV Praxisproblemen bzw. zu starken Einschränkungen des Einsatzbereiches von darauf basierenden CCTV Produkten.

Der Artikel soll zeigen, wie unterschiedlich und teilweise konträr die Anforderungen an Video Codecs im Multimedia und CCTV Bereich meist sind. Davon ausgehend wird gezeigt, welche Anforderungen im Design von CCTV Codecs berücksichtigt werden müssen, um ein möglichst breit einsetzbares, kostengünstiges Verfahren zu erhalten. Dies bildet die Basis für die Beurteilung von Kompressionswerkzeugen standardisierter Videokompressions-Verfahren wie MPEG-1/2/4 oder H.263/264 auf Erfüllbarkeit dieser Anforderungen. Grundsätzlich zeigt sich, dass die Kompressionsstandards zwar CCTV Anforderungen Rechnung tragen können, die verfügbaren Codec-Implementierungen diese Möglichkeiten aber wiederum stark beschränken. Dies zwingt die Hersteller professioneller CCTV-Produkte zu Eigenentwicklungen von Codecs.

Wegen der großen Vielfalt an Kompressionsstandards und der riesigen Freiheitsgrade, die diese Standards den darauf basierenden Codec-Implementierungen lassen, herrscht breite Verunsicherung welches Verfahren sich für CCTV am besten eignet. Ist es nun MPEG-2, MPEG-4/SP, MPEG-4/ASP, MPEG-4/AVC, H.263/264 oder doch besser ein proprietäres MJPEG oder MWavelet-Verfahren? Die Freiheitsgrade der Standards drücken sich in einer Vielzahl von genannter Profile und Level-Kombinationen aus, die sich je nach verwendeten Kompressionswerkzeugen und Wertebereichen für z. B. Bildraten und Auflösungen

unterscheiden. Viele Bestandteile der Standards sind optional und müssen von den Encodern nicht implementiert werden, um standardkonform zu bleiben. Die Familie der MPEG-Standards und speziell MPEG-4 erscheint als Sammelsurium von Werkzeugen zur Kompression, aus welchem sich die Codec-Entwickler beliebig bedienen können. Dies erschwert extrem die Vergleichbarkeit von Codecs und schafft darüber hinaus erhebliche Kompatibilitätsprobleme. Damit ist z. B. die alleinige Aussage, dass ein CCTV System den MPEG-4 Standard nutzt, relativ wertlos, um seine Leistungsfähigkeit zu beurteilen. Man hat damit weder eine Aussage zum Grad der erreichbaren Kompression noch zur Performance beim Multikanalbetrieb, zur Bildqualität oder zur Flexibilität für die Realisierbarkeit von CCTV typischen Anforderungen.

Im Beitrag sollen die Anforderungen von CCTV an Video-Codecs analysiert und gezeigt werden, wie sich durch Ausnutzung von Prozesskenntnissen die Eigenschaften der Codecs grundlegend verbessern lassen ohne den Spezifikationsrahmen der Standards aufzugeben. Um zu einer optimalen Anpassbarkeit an wechselnde Forderungen zu kommen, sollten CCTV Systeme sowohl die Nutzung von Multimedia-Chip basierten Encodern als auch über Firmware flexibel anpassbare DSP- (Digital Signal Processor) Implementationen unterstützen. Betrachtet man die Anwendungsbreite für CCTV Probleme sind DSP basierte Systeme trotz niedrigerer Rechenleistung der DSPs heute schon Multimedia-Chip basierten mit starren Algorithmen weit überlegen. Nichtsdestotrotz gibt es auch bei CCTV Anwendungsbereiche, bei denen der Einsatz von Multimedia-Encodern zu besseren Ergebnissen führt. Diese Tatsache lässt CCTV-Produkte als optimal erscheinen, welche beide Codec Implementations-Varianten unterstützen können.

2. CCTV versus Multimedia --- Anforderungen an Video Codecs

Wichtigstes Ziel der Anwendung eines Video-Kompressionsverfahrens für Multimedia Anwendungen ist die Verringerung der Kosten für die Speicherung und Übertragung bei guter Bildqualität. Bei CCTV kommt hinzu, dass das Kompressionsverfahren die eigentliche Zielstellung von CCTV - die Lösung eines Sicherheitsproblems - optimal unterstützen muss. Im Gegensatz zu Multimedia, wo die Mediendaten selbst das Produkt darstellen, ist das Produkt von CCTV bei richtig verstandener Zielstellung ein optimal gelöstes Sicherheitsproblem. Die Mediendaten stehen also oft nicht im Zentrum sondern sind eigentlich nur Mittel zum Zweck. Dies rechtfertigt bestimmte Kompromisse in Bezug auf Qualität und Bildraten wie sie bei Multimedia nicht akzeptiert werden können. Andererseits sind Multimedia-typische Kompressions-Kompromisse oft nicht für CCTV zulässig.

So führt die bei Multimedia oft verwendete CBR (Constant Bitrate) Kompression zu einer Verschlechterung der Bildqualität gerade dann, wenn sich Bewegung in der Szene ereignet und diese damit aus Sicherheitssicht interessant wird. Ein anderer Multimedia üblicher Kompromiss ist die Akzeptanz hoher Übertragungsverzögerungen (Latenzen) zugunsten einer verbesserten Kompression, welche bei vielen CCTV Anwendungen ebenfalls nicht tragbar sind.

Ein Beispiel für einen typischen CCTV-Kompromiss sind die Bilder einer Parkhauszufahrt. Wird ein Fahrzeug detektiert, sollen Bilder hoher Qualität erzeugt werden um z. B. Nummernschilder erkennen bzw. auch automatisch analysieren zu können. Im Normalfall reicht aber eine niedrige Qualität für Überblickszwecke zu Vorgängen in der Szene aus. Man stelle sich einen Fernsehfilm mit stellenweise guter und schlechter Bildqualität vor. Ein anderes, extremes Beispiel sind CCTV Systeme, die aufgrund intelligenter Bildsensorik autonome Entscheidungen treffen können. Der Mensch als Betrachter und subjektiver Qualitätsmaßstab entfällt hier vollständig oder teilweise. Die Qualität des Bildmaterials muss optimal in Bezug auf die Anforderungen des Sensors aber nicht in Bezug auf die Anforderungen eines menschlichen Betrachters sein.

Sämtliche Anforderungen an den CCTV-Codec werden vom Sicherheitsszenario bestimmt. Ein allgemeines Sicherheitsszenario und damit einen universell sinnvoll einsetzbaren Codec gibt es aber nicht. So vielfältig wie die Szenarien sind so vielfältig sind die Strategien das entstehende Bilddatenmaterial optimal darauf zugeschnitten aufzubereiten. Im Vergleich zu CCTV sind die Anforderungen bei Multimedia vergleichsweise einfach generalisierbar.

Bild 2 zeigt einige wichtige Kompromissgrößen eines Video-Codex. Die Gewichtungen dieser Kennwerte sind für CCTV und Multimedia in der Masse der Anwendungen stark verschoben.

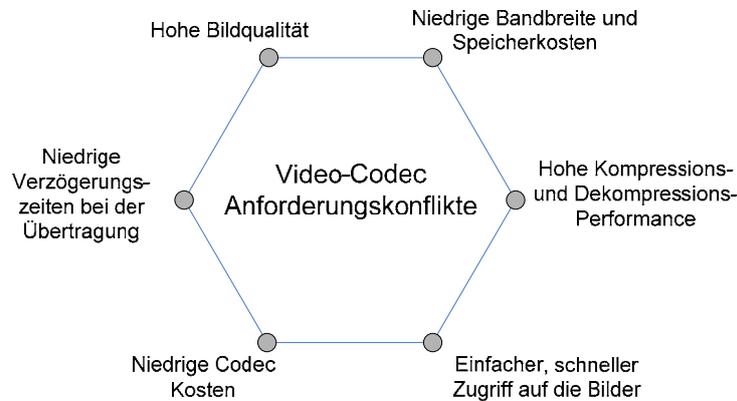


Bild 2: Anforderungskonflikte von Video-Codex

Die Übertragung und Speicherung unkomprimierten Materials ist sowohl physikalisch als auch kostenseitig nicht beherrschbar. Die Situation wird mit der gegenwärtig stattfindenden Einführung hochauflösender Bildformate (HTDV, Megapixel IP-Kameras) noch dramatischer. CCTV treibt die Bandbreiteproblematik durch die Vielzahl der Kamerakanäle weiter auf die Spitze. Mit welchen Größenordnungen man es hier zu tun hat, kann z. B. [1] entnommen werden. Im Ziel hoher Kompression stimmen die Anforderungen von CCTV und Multimedia/TV überein. Unterschiedlich sind die Mittel und Wege, wie diese Ziele erreicht werden sollen. Diese werden sinnvollerweise von den meist völlig unterschiedlichen Anwendungsszenarien und Bildinhalten bestimmt.

2.1 Asymmetrie der Encoder/Decoder Infrastruktur

Bild 1 zeigt eine vereinfachte digitale Videoübertragungsinfrastruktur bestehend aus Video-Encodern, Übertragungstechnik und Video-Decodern. Während bei Multimedia/TV die Anzahl der Decoder (z. B. Fernsehgeräte) die Anzahl der Encoder (Fernsehkamäle, professionelles DVD-Mastering) bei weitem überwiegt, ist diese Asymmetrie für CCTV Anwendungen im allgemeinen genau umgekehrt - ist es doch gerade eines des wichtigsten Ziele von CCTV die Informationsflut einer Vielzahl von Kanälen zu verringern und die Präsentation auf das für die Sicherheitsaufgabe notwendige Maß zu beschränken.

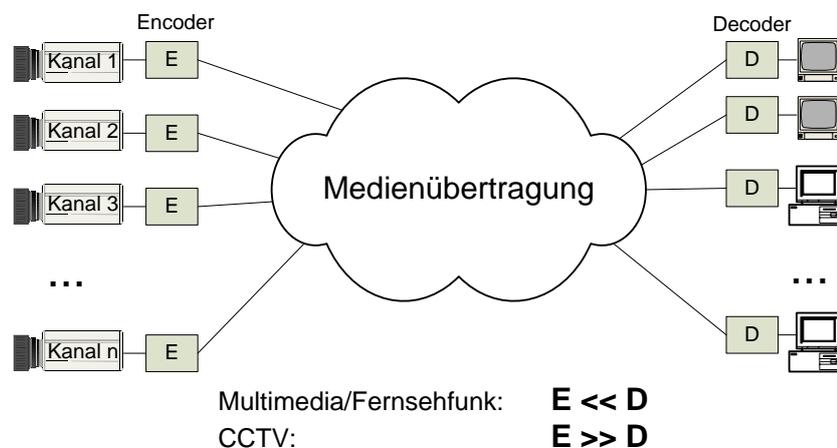


Bild 1: Asymmetrie der Codec-Infrastruktur von Multimedia/TV und CCTV Anwendungen

Dies hat eine Reihe von Konsequenzen für das Codec Design zur Folge:

- Während bei CCTV Anwendungen durch die Vielzahl der Kanäle auf der Encoder-Seite möglichst kostengünstige Lösungen gefunden werden müssen, spielt die Kostenfrage bei professionellen TV-Studio-Encodern eine geringere Rolle.
- Umgekehrt ist die Situation auf der Decoder-Seite. Während bei den Wiedergabegeräten im TV-Bereich ein extremer Kostendruck herrscht, tritt dieser bei CCTV zugunsten der Funktionalität oft zurück. Beispiel sind die hohen Kosten für professionelle Großbildsysteme in Kontrollräumen.

Das heißt aber wiederum nicht, dass CCTV Encoder wegen des Kostendrucks schlechter arbeiten als Multimedia-Encoder. Sie arbeiten nur anders und problemangepasster basierend auf optimalen Kompromissen zwischen so gegensätzlichen Anforderungen wie sie z. B. Bild 2 zeigt. Multimedia Encoder nutzen ein Bündel an Werkzeugen (Abschnitt 3) zur Erreichung einer hohen Kompression. Diese beanspruchen oft eine gewaltige Rechenleistung. DSP basierte CCTV Encoder müssen derartige Werkzeuge nicht mit den gleichen Algorithmen wie Multimedia Encoder implementieren. Wie in Abschnitt 3 gezeigt wird, können oft mit vergleichsweise einfachen Mitteln nur durch Kenntnis der Kameraszenen alternative Encoder-Werkzeuge implementiert werden, die die Effizienz von üblichen Kompressionsalgorithmen für den Einzelfall bei weitem übersteigen können. Damit erreicht man trotz Kostendrucks oftmals erheblich bessere Resultate als sie mit dem Einsatz von auf den allgemeinen Einsatz zugeschnittenen Multimedia-Encodern zu erreichen wären.

2.2 Echtzeitanforderungen

Die Forderungen bezüglich des Echtzeitverhaltens sind ein weiterer wesentlicher Unterschied zwischen CCTV und TV-Anwendungen. Nahezu alle CCTV Anwendungen erfordern vom Video-Encoder eine Kompression in Echtzeit, d. h. der Encoder muss in der Lage sein alle eintreffenden Bilder mit voller Auflösung (25 pro Sekunde mit 704x576 bei PAL Kameras) zu komprimieren. Dies betrifft sowohl die Speicherung als auch die Live-Übertragung. Bei vielen TV-Anwendungen ist eine solche **Echtzeitfähigkeit** von niedrigerer Bedeutung. Legt man z. B. Wert darauf einen Film in DVD Qualität mit hohem Kompressionsfaktor zu komprimieren, spielt es eine untergeordnete Rolle, wie lange der Encoder dafür braucht. Bei qualitativ hochwertigen Filmen wird der Encodierungsprozess sogar manuell gesteuert, um einen optimalen Kompromiss von Qualität und Kompression zu erzielen. So kann es im professionellem Mastering Prozess Stunden oder auch Tage kosten, eine Stunde hochwertiges DVD Material zu erzeugen.

Selbst wenn TV-Material zum Zwecke der Live-Übertragung in Echtzeit encodiert werden muss, sind CCTV- und TV-Anforderungen stark unterschiedlich. „Echtes“ Live im Sinne einer CCTV Überwachungsanlage gibt es bei TV-Übertragungen nicht. Durch spezielle Verfahren im Encodierungs-Prozess von TV-Material kommt es zu **Übertragungsverzögerungen**, die bei Fernseh Anwendungen toleriert werden, die aber für viele CCTV Anwendungen nicht brauchbar sind. Man kann die bei TV-Live-Übertragungen auftretenden Verzögerungen sehr gut beobachten, wenn z. B. ein Radiobericht und die digital übertragene Fernsehsendung zu einem Fußballspiel simultan laufen. Die Übertragungslatenzen digitaler CCTV Systeme müssen sich an den Verzögerungen analoger Kreuzschiensysteme messen lassen. Latenzen größer als 200 ms sind in vielen CCTV-Anwendungen nicht zulässig. Dies gilt verschärft, wenn Domekameras bedient werden sollen. Bei großen Übertragungslatenzen ist eine manuelle Bedienung nur noch sehr schlecht oder gar nicht mehr möglich.

Ein anderes Problem der für TV-Übertragungen verwendeten Encoder sind die **Aufschaltverzögerungen** beim Wechsel oder auch erstmaligen Aufschalten von Kanälen. Im TV-Bereich hat man sich daran gewöhnt, dass das Zappen durch die Kanäle bei digitaler Übertragung sehr langsam ist. Grund ist die Struktur der MPEG-Daten von Fernsehübertragungen. Um den Dekodierungsprozess starten zu können, müssen die Decoder auf das Auftreten so genannter Referenzbilder im Datenstrom warten. Da der Konsument keinerlei Einfluss auf den Sender hat, muss im schlimmsten Fall eine so genannte GOP (Group

of Pictures) gewartet werden, bis ein Bild erscheint. Ein Kennzeichen für gute CCTV Encoder ist, dass diese Verzögerungen sehr viel kleiner sind. Dies wird zum einen durch spezielle Eigenschaften der komprimierten Datenströme erreicht, zum anderen durch die direkte Steuerung des Encoders durch den Decoder. So kann der Encoder z. B. beim Umschalten die Anweisung erhalten, den normalen Bildfluss abzubrechen und unmittelbar ein neues Referenzbild zu schicken.

2.3 Präsentationsszenarien

Während bei Multimedia Videowiedergaben von Ausnahmen abgesehen (z. B. Bild in Bild) immer nur ein Kanal angezeigt und evtl. gesteuert (Vorlauf/Rücklauf...) wird, sind die Präsentationsanforderungen von CCTV erheblich vielfältiger. Eine Wiedergabe ist fast immer eine szenariorientierte **Multikanalwiedergabe**. So werden mehrere Kanäle, die ein Sicherheitsszenario repräsentieren, in einer PC-Oberfläche simultan dargestellt. Es kann zwischen verschiedenen dieser Szenarien entweder manuell oder automatisch hin- und hergeschaltet werden. Die Kombination von Live- und Speicherbildern (z. B. Alarmbilder) ist gewünscht. Gespeicherte Bilder mehrerer Kanäle sollen synchronisiert wiedergegeben werden. Diese Basisforderungen von CCTV gehen weit über TV übliche Präsentationen hinaus. Dies ist auch der Grund warum übliche Medienplayer für CCTV Zwecke nur sehr beschränkt einsetzbar sind. Obwohl z. B. der MPEG Standard eine Mehrkanalspeicherung innerhalb einer MPEG-Datei - einem so genannten Program Stream - zulässt, gibt es praktisch keine Wiedergabesoftware oder auch Geräte, die derartige Funktionen in einem CCTV tauglichen Rahmen liefern.

Die CCTV üblichen Präsentations-Szenarien stellen höchste Anforderungen an die Decoder-Software. Im Gegensatz zu Multimedia müssen die Decoder eine Performance bieten, die es erlaubt viele Kanäle gleichzeitig synchronisiert zur Darstellung zu bringen. Während z. B. die Wiedergabe einer DVD auf einem PC heute mit ihren 25 Bildern pro Sekunde recht wenig Performance beansprucht, müssen CCTV Decoder im Multikanal-Betrieb Lasten von einigen hundert Bildern pro Sekunde verarbeiten können. So erfordert eine 8 Kanal-Wiedergabe mit voller Video-Bildrate die Dekompression von 200 Bildern pro Sekunde. Dies erzwingt besondere Optimierungen, die sich auch auf das Design der Encoder und auf die Eigenschaften der Bilddatenströme auswirken. Ein Beispiel dafür ist eine automatisierte Auflösungssteuerung. Die Bilddaten sollten vom Encoder auch nur mit der jeweils benötigten Auflösung des PC-Wiedergabefensters erzeugt werden. Haben die Wiedergabefenster eines Systems QCIF-Auflösung, ist die Übertragung von 4CIF großen Bildern nicht sinnvoll. Eine derartige Auflösungssteuerung muss mit hoher Dynamik automatisch und für den Anwender transparent erfolgen. Wird z. B. ein Wiedergabefenster vergrößert, so soll der Encoder automatisch auf die dafür optimale Auflösung umschalten. Multimedia Codecs stellen derartige Möglichkeiten praktisch nicht oder nur mit sehr großen Umschaltverzögerungen bereit.

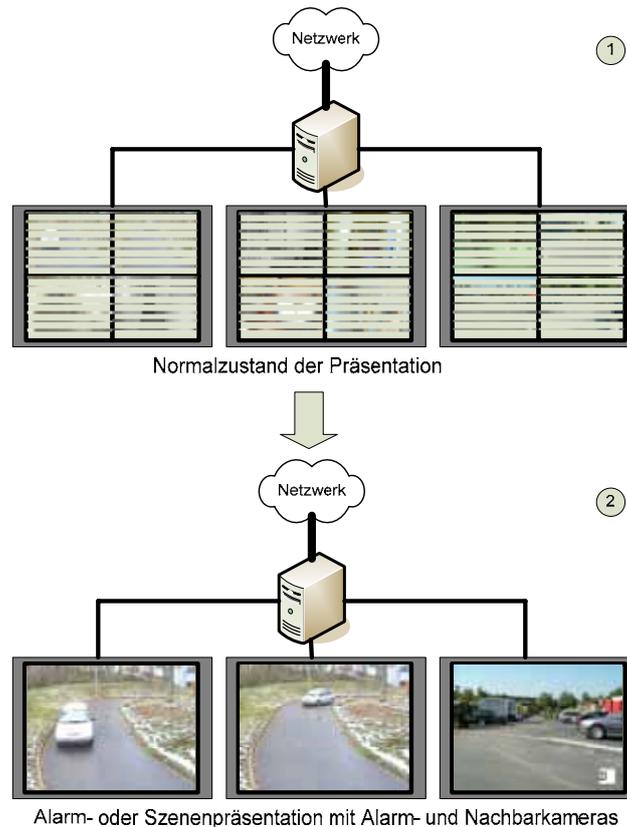


Bild 3: Auflösungssteuerung in Präsentationsszenarien

Bild 3 zeigt die Bedeutung einer automatisierten Auflösungssteuerung des Encoders unter Performancegesichtspunkten. Die Präsentation 1 zeigt im Grundzustand eine Überblicksszene mit 12 Kamerakanälen. Jedes Wiedergabefenster im PC-Monitor hat eine Auflösung von CIF. Bei voller Bildrate von 25 Bildern pro Sekunde wird über das Netz eine Last von 300 Bildern in CIF Auflösung übertragen, welche auch der Decoder der Wiedergabestation, die die Monitore betreibt in Echtzeit dekodieren muss. Präsentation 2 zeigt eine automatische durch einen Sensor ausgelöste Alarmaufschaltung. Die 3 Kameras werden in hoher 4CIF Auflösung präsentiert. Dazu erhält der Encoder entsprechende Kommandos. Die Netzwerklast ist in Fall 1 und 2 gleich. Ebenso die Last für den Decoder. Erfolgt diese dynamische Steuerung durch die Präsentation in einem CCTV System nicht, muss von Worst Case Anforderungen ausgegangen werden. In diesem Fall müsste der Encoder für alle 12 Kanäle 4CIF Auflösungen erzeugen um sicherzustellen, dass eine Anforderung nach hoher 4CIF Qualität auf Decoder Seite erfüllt werden kann. Der Performanceverlust bei dieser Vorgehensweise ist erheblich. Bei der dynamischen Auflösungssteuerung kann es zu Konflikten kommen, wenn mehrere Wiedergabefenster den gleichen Kanal mit unterschiedlicher Auflösung anfordern. Hier müssen intelligente Konfliktlösungsstrategien im CCTV System implementiert sein.

Voraussetzung für eine sinnvolle Realisierung derartiger Steuerungen ist eine direkte Kommunikation zwischen Decoder und Encoder und niedrige Latenzen bei den damit verbundenen Umschaltvorgängen.

Auch die Anforderungen an den **Wiedergabekomfort gespeicherter Bilder** sind bei CCTV höher als bei üblichen TV Anwendungen. So hat die Masse der DVD-Player und DVD-Software-Player Produkte nur einen geringen Komfort was Wiedergabemöglichkeiten wie Zeitlupe oder Zeitraffer betrifft. Ein schneller Vor- und Rücklauf erfolgt nur in großen Sprüngen. CCTV Player müssen einen Einzelbildbasierten Vor- und Rücklauf anbieten. Eine synchrone

Zeitlupen/Zeitraffer-Wiedergabe mehrerer Kanäle muss möglich sein. Von Decodern aus dem Multimedia Bereich werden solche Möglichkeiten nicht bereitgestellt. Für den Einzelbild-Zugriff benötigt man hier spezialisierte Videoschnittprogramme. Diese Beschränkungen von Multimedia-Decodern sind keine Einschränkungen der Kompressions-Standards. Sie sind vielmehr Beschränkungen der jeweiligen Standard-Implementation.

Die hohen Anforderungen von CCTV machen auch den Export von CCTV Daten aus dem System problematisch. Obwohl MPEG Program Streams es zulassen, mehrere synchronisierbare Kanäle in einer Datei zu erfassen, gibt es für derartige MPEG Dateien keine Wiedergabe-Software. Hier relativieren sich die Vorteile einer Forderung nach Konformität bezüglich eines Kompressionsstandards sehr schnell. Man möge versuchen eine synchrone Mehrkanalwiedergabe eines Sicherheitsvorfalles inklusive eingebetteter Alarminformation mittels des Windows Media Players zu realisieren. Das führt zur paradoxen Situation, dass ein CCTV System zwar standardisierte Kompressionsverfahren anbietet für die Wiedergabe nichtsdestotrotz bei entsprechendem Anspruch an die Leistungsfähigkeit proprietäre Software eingesetzt werden muss.

Bezüglich der Steuerung der Wiedergabe (**Interaktivität**) ist das Multimedia-Schlagwort „Video on Demand“ bei CCTV weitgehend (von sehr einfachen Produkten abgesehen) Praxis. So können mehrere CCTV Arbeitsplätze völlig unabhängig voneinander auf gespeichertes Video sogar parallel zur laufenden Aufzeichnung zugreifen. Während auf einem Arbeitsplatz eine synchronisierte Multikanalwiedergabe läuft, zeigt ein anderer Arbeitsplatz die gleichen Kanäle im Zeitraffer oder arbeitet eine Recherche ab. Die entsprechenden Streaming Schnittstellen von CCTV Systemen müssen solche Möglichkeiten unterstützen.

2.4 Eingebettete Textinformationen

Auch in dieser Anforderung gibt es erhebliche Unterschiede zwischen den Möglichkeiten von Multimedia- und CCTV Codecs. Im Rahmen von Multimedia-Anwendungen gibt es textuelle Begleitinformationen in Form von Untertiteln. Diese können in die Datenströme eingebettet werden. Diese Untertitel werden z. B. bei MPEG-2 in Form von Grafiken erzeugt und beim Dekodieren den Bilddaten überlagert.

CCTV Begleitinformationen müssen jederzeit aus den Datenströmen separiert werden können. Liegen die Texte in Form von Grafiken wie bei MPEG-2 vor, so können die originalen Daten nur mittels aufwändiger - so genannter OCR Software - wieder rekonstruiert werden. In dieser Form sind diese Begleitdaten z. B. für Recherchezwecke unbrauchbar. Ein CCTV Codec muss die Einbettung von Begleitdaten mit hoher Dynamik erlauben. Diese können sich gegebenenfalls sogar von Bild zu Bild (z. B. Zeitstempel im Millisekundenbereich) ändern. Weiterhin müssen diese Daten für Recherchezwecke separierbar bleiben. Die Präsentation dieser Daten im Decoder soll flexibel anpassbar sein. Neben Zu- und Abschalten der Anzeige, sollte Positionierung, Fonttyp, Zeichengröße und gegebenenfalls Farbe einstellbar sein.

Obwohl die Möglichkeiten von Untertiteln gerade im MPEG-4 Standard stark erweitert wurden - u. a. gibt es im Gegensatz zu den grafikbasierten Untertiteln auch textbasierte - erscheinen sie für CCTV Zwecke immer noch zu unflexibel bzw. sind Software-Werkzeuge für deren Nutzung kaum verbreitet.

2.5 Audio

Obwohl hier die Anforderungen an digitale CCTV Systeme in den letzten Jahren gewachsen sind, ist eingebettetes Audio noch immer die Ausnahme. Meist wird die Audio-Information in CCTV Systemen in separaten Kanälen erfasst und erst bei der Präsentation mit den entsprechend ausgewählten Videokanälen synchronisiert wiedergegeben. Existiert eine Audiofunktion, so ist auch bei CCTV eine lippensynchrone Wiedergabe wie bei Multimedia gewünscht. Da das hierfür notwendige Buffering allerdings zu zusätzlichen Latenzen führt, sollte der Decoder eine Wahlfreiheit zwischen lippensynchroner Wiedergabe und niedriger Latenz ermöglichen. Dies ist

ebenfalls eine sehr CCTV spezifische Anforderung die von Multimedia Codecs nicht befriedigt wird.

Interessant sind neue mit Audio verbundene Möglichkeiten der Sensorik. Diese können die übliche Videosensorik unterstützen. So kann z. B. ein Audio-Pegelmesser zum Einschalten einer Beleuchtung bei Geräuschen genutzt werden. Die notwendigen Analysen werden auch hier sinnvollerweise im Encoder vor der Kompression durchgeführt.

Mehrkanal-Audio ist bei CCTV Produkten nur selten im Einsatz. Insgesamt sind die Anforderungen an den Codec bezogen auf die Audiofunktion für CCTV heute noch geringer als bei Multimedia-Anwendungen.

2.6 Speichieranforderungen

Geht es bei der Speicherung von Multimedia-Inhalten meist nur darum Filme vergleichsweise kurzer Dauer (2-3 Stunden) entweder auf DVD oder Festplatte als Dateien zu speichern und diese als Einzelkanäle sequentiell wiederzugeben, liegen die Anforderungen an CCTV Speicher weit höher. Multikanalspeicherungen mit Archivierungszeiträumen über Monate sind Stand der Technik. Wahlfreiheit des Zugriffs auf jedes einzelne Bild trotz Mediendatenbanken weit im Terabyte Bereich ist gewünscht.

Die Anforderungen an CCTV Codecs sind auch in diesem Bereich weit differenzierter als bei Multimedia. Auch die Werkzeuge, die für eine effektive Kompression eingesetzt werden können, unterscheiden sich stark. Im Gegensatz zu Multimedia, wo natürlich alle Bilder eines Filmes gespeichert und auch wiedergegeben werden müssen, beschränken sich CCTV Systeme sinnvollerweise auf die Speicherung notwendiger Inhalte. Notwendige Inhalte sind die Bilder/Begleitdaten die der Erfüllung des definierten Sicherheitszieles der Anlage dienen. In der Masse der CCTV-Anwendungen ist der weitaus größte Teil der Bilddaten reiner Datenmüll. Bei einfacher permanenter Speicherung mit voller Bildrate und Auflösung würden oft mehr als 99% der Datenbanken aus unnützer Information bestehen - mit den entsprechenden Kosten abgesehen von den enormen Zeiten für die Auswertung diese Materials. Wie auch in analogen Zeiten speichern deshalb entsprechend ausgelegte CCTV Systeme bei Bedarf Daten im Timelapse Modus mit variabler Bildrate wobei trotzdem jeder Kanal auf Anforderung die volle Bildrate bieten sollte. Andere Möglichkeiten zu einer problemangepassten Speicherung zu kommen, sind ereignisbasierte:

- Qualitätssteuerung
- Auflösungssteuerung

Mittels dieser CCTV-typischen Mechanismen erzielt man im allgemeinen weitaus bessere Kompressionsergebnisse als mit den klassischen Kompressionsmethoden von Multimedia-Encodern. Erschwerend für die Encoder-Implementation kommt hinzu, dass die Steuerung dieser Stream-Eigenschaften mit geringer Latenz erfolgen muss. Was nutzt die Umschaltung auf eine bessere Qualität im Alarm-Fall, wenn die zugehörigen Bilder erst mit großer Verzögerung erzeugt werden und die wichtigen ersten Momente eines Ereignisses noch mit schlechter Qualität komprimiert werden? Auch wenn es eine Reihe von CCTV Szenarien gibt, die eine Daueraufzeichnung mit hoher Geschwindigkeit erfordern (z. B. Spieltischaufzeichnung in Kasinos) ist es für die große Menge der Anwendungen wichtig derartige Möglichkeiten bereitzustellen. Warum sollte man auch kostenintensiv Daten aufzeichnen, die für die Erfüllung der Zielstellung keine Rolle spielen ja sogar kontraproduktiv sind, da die eigentlichen Nutzinformationen verdeckt werden?

Gerade in den Möglichkeiten der Streamsteuerung zeigen sich fundamentale Unterschiede zwischen Multimedia-Codecs und CCTV Codecs. In Teilen unterstützen die Videokompressions-Standards zwar derartige Forderungen es gibt aber praktisch keine Massenmarkt-Implementationen für die Unterstützung solcher Werkzeuge. Damit sind CCTV Hersteller, welche adäquate CCTV Systeme entwerfen wollen, auf die Eigenentwicklung leistungsfähiger Codecs angewiesen.

2.7 Substreaming - eine Quelle / mehrere Kanäle

Ein idealer CCTV Encoder sollte die Möglichkeit bereitstellen mehrere Streams unterschiedlicher Bandbreite aus den gleichen Quelldaten zu erzeugen. Dies ermöglicht ein bandbreitenangepasstes Streaming je nach verfügbarer Übertragungsbandbreite. Man spricht von skalierbaren Datenströmen. Es gibt MPEG Profile und Levels die derartige Skalierbarkeitseigenschaften unterstützen. Praktisch verfügbare Encoder-Implementationen mit für CCTV tragbarem Preis/Leistungsverhältnis gibt es nicht.

Als Minimalforderung sollte ein CCTV Encoder eine Möglichkeit bereitstellen die Qualität der Aufzeichnung und der Live-Übertragung (Kreuzschienenfunktion) separat zu steuern. So könnte man z. B. mit niedriger Qualität Live Bilder übertragen, wenn das Netzwerk nur über eine niedrige Bandbreite verfügt, aber gleichzeitig mit hoher Qualität aufzeichnen. Technische Lösungen des Problems können wie in Bild 4 gezeigte realisiert werden.

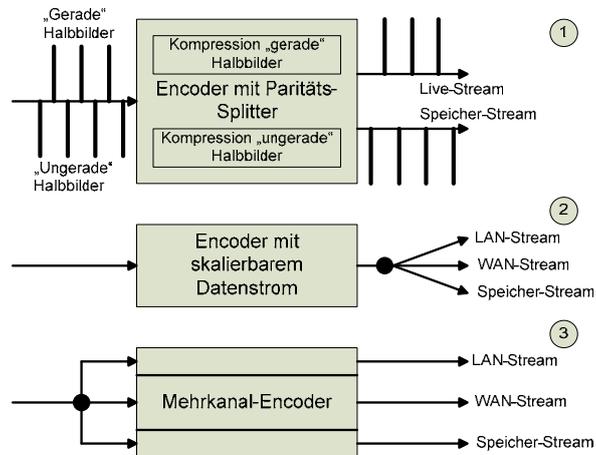


Bild 4: Technische Möglichkeiten der Realisierung von Substreaming

Variante 1 trennt die eingehenden Videodaten in 2 Datenströme auf. Diese enthalten jeweils nur eine Parität des Zeilensprung-basierten Eingangs-Videos. Beide Paritäten können mit unterschiedlichen Bitraten komprimiert und für verschiedene Anwendungsfälle genutzt werden. Vorteil ist die niedrige benötigte Kompressions-Performance und damit die niedrigen Kanalkosten. Nachteilig ist, dass die Ausgangsdatenströme nur Auflösungen bis maximal 2CIF (704x288) haben. Paradoxerweise verkehrt sich dieser Nachteil aufgrund der Eigenschaften von Videomaterial oft in einen Vorteil wegen des so genannten Lattenzauneffektes von Zeilensprungvideo [1] und der damit verbundenen oft schlechten Qualität von 4CIF Videomaterial. Das Verfahren lässt sich standardkonform im Rahmen der Möglichkeiten der MPEG-Spezifikation betreiben.

Variante 2 nutzt einen Kompressionsalgorithmus mit der so genannten Skalierbarkeitseigenschaft. Dies ist eine Eigenschaft von Wavelet basierten Encodern. Aus den bereits komprimierten Daten lassen sich Datenströme verschiedener Bitraten für verschiedene Anwendungsfälle ableiten. Normalerweise werden hierfür sehr teure Transcoder eingesetzt. Wavelet Verfahren zur Videokompression sind nicht standardisiert. Damit sind alle darauf basierenden Codecs proprietär.

Variante 3 komprimiert den eingehenden Video-Datenstrom mehrfach mit verschiedenen Bitraten. Dies ist zwar die flexibelste aber auch teuerste Variante um ein Substreaming zu realisieren.

2.8 Echtzeit-Videosensorik

Eine Problemstellung, die im Bereich von Multimedia Encodern gar keine Rolle spielt, ist der gesamte Bereich der Videosensorik. CCTV Encoder stellen eine ganze Palette von Sensoren zur Verfügung (z. B. Bewegungs- und Aktivitäts-Detektion), welche eine Echtzeitanalyse des Bildinhaltes ermöglichen. Die Ergebnisse dieser Sensoren können gezielt zur Alarmierung, Präsentation und Qualitäts- bzw. Aufzeichnungssteuerung genutzt werden. Echtzeitsensoren verlangen den Zugriff auf die unkomprimierten Rohbilder, sind also wie Bild 5 zeigt im Übertragungspfad am sinnvollsten direkt nach der Digitalisierung angesiedelt. Bei Multimedia-Encoder Chips findet man derartige Möglichkeiten gar nicht. CCTV Systeme, die auf derartigen Encodern aufsetzen, verfügen damit nur über sehr beschränkte Sensormöglichkeiten, die allenfalls zur Analyse langsamer Vorgänge geeignet sind. Ganz besonders kritisch wird die Situation hier, wenn eine Vielzahl von Kanälen simultan überwacht werden soll. In diesem Fall muss der Auswerte-Rechner neben der eigentlichen Sensorik noch die Dekompressionslast einer Vielzahl von Kanälen verarbeiten. Obwohl dies ein für langsame Vorgänge sinnvoller Ansatz ist, schränkt der Einsatz von Multimedia-Encodern den Anwendungsbereich eines CCTV Systems gerade in Hinsicht Sensorik sehr stark ein.

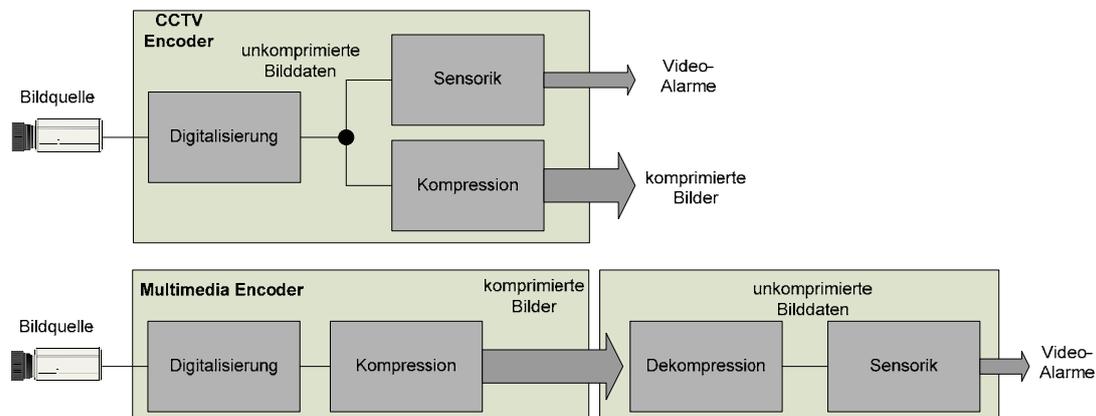


Bild 5: Sensorik im Zusammenspiel mit CCTV- und Multimedia-Encodern

2.9 Bildinhalte

Multimedia Codecs sind auf die optimale Kompression möglichst universeller Bildinhalte hin optimiert. Durch diese generalisierte Zielstellung werden die individuellen, kanalabhängigen Eigenschaften von CCTV Überwachungssystemen nur unzureichend ausgenutzt, um die Kompression zu verbessern. Während Multimedia Inhalte sich im Allgemeinen durch permanente Bewegung und andauernde Szenenwechsel auszeichnen, entspricht diese Annahme der Situation bei CCTV oft nicht. Die weitaus meisten installierten Kameras zeigen Szenen mit relativ seltenen oder sehr langsamen Änderungen. Beispiele sind Kameras entlang von Zäunen oder Beobachtungskameras in Korridoren oder Tresorräumen. Andere Szenen zeigen tageszeitabhängig verschieden hohe Anteile an Bewegung. Natürlich gibt es auch Extrembeispiele für Kamerasituationen mit viel Bewegung (bewegte Domekameras, die Spieltischbeobachtung in Casinos oder belebte Fußgängerzonen). Im Mittel gesehen sind Änderungen in CCTV Bildern jedoch eher die Ausnahme - dies sind die eigentlichen Nutzdaten einer CCTV Überwachung - während sie bei Multimedia Inhalten die Regel sind (siehe z. B. Filme wie „Matrix“).

Nutzt man die a priori Kenntnis über die zu erwartenden Bildinhalte einer Kamerasicht aus, um einen CCTV Encoder entsprechend anzupassen, können Kompressionsgewinne in Größenordnungen erzielt werden, wie sie bei Universal-Encodern für Multimedia-Zwecke nicht

einmal mit extremsten Rechenaufwand realisierbar wären. Als eines unter vielen denkbaren Beispielen zeigt Bild 7 ein dem Inhalt einer Szene angepasstes Profil der Kompressionsqualität. Mittels einer so genannten Region of Non Interest kann hier z. B. ein CCTV Encoder so eingestellt werden, dass bestimmte Bereiche eines Bildes, in denen sich ständig irrelevante Bewegung ereignet, mit schlechterer Qualität komprimiert werden. Diese Bildanteile bleiben dann zwar erkennbar, sie tragen aber nicht mehr so stark zur Datenlast bei wie sie es im Falle eines nicht adaptierbaren Multimedia Encoders tun würden. Im Extremfall könnte man derartige irrelevante Bereiche auch komplett ausblenden oder diese Bereiche nur in größeren Zeitabständen auffrischen (partielles Timelapse). Die MPEG Standards bieten grundsätzlich auch derartige Möglichkeiten, Multimedia Encoder stellen sie aber nicht bereit. Die Anzahl denkbarer Ansätze, die Kenntnis über eine Szene für die Optimierung der Kompression auszunutzen ist mannigfaltig. Beispiele sind:

- Die Kenntnis von zu erwartenden Geschwindigkeiten relevanter Vorgänge und die Anpassung der entsprechenden Bildraten.
- Die Festlegung interessanter Bereiche im Bild und die Definition einer Region of Interest mit besserer Kompressionsqualität.
- Die Festlegung von Zeiten besonderer Aktivität und die Steuerung von Qualität und Bildraten entsprechend derartiger Zeitpläne.
- Die automatische Adaption von Qualitäten und Bildraten entsprechend verschiedener Kritikalitätsstufen, die auf Basis von im Encoder arbeitenden Echtzeitsensoren ermittelt werden.

Die Herausforderung für den Codec-Entwickler bei der Ausnutzung derartiger Möglichkeiten besteht darin, dies in einer standardkonformen Weise zu tun. Dafür müssen die von Multimedia-Encodern normalerweise nicht implementierten Freiheitsgrade der Standards viel intensiver genutzt werden (Abschnitt 3).

2.10 Schutz der Inhalte

Das in Multimedia Codecs in zunehmendem Maße implementierte DRM (Digital Rights Management) hat mit den Anforderungen von CCTV praktisch keine Gemeinsamkeiten. CCTV hat zwei grundsätzliche Anforderungen bezogen auf den Content-Schutz:

- Erfüllung datenschutzrechtlicher Bestimmungen bzw. projektspezifischer Festlegungen zur Sichtbarkeit bestimmter Bildinhalte
- Sicherung der Manipulationsfreiheit der Inhalte. Da CCTV Daten in zunehmendem Maße für juristische Zwecke eingesetzt werden, hat diese Anforderung eine wachsende Bedeutung.

CCTV Codecs sollten Werkzeuge bereitstellen, um derartige Forderungen erfüllen zu können. Da die Algorithmen von Multimedia-Codecs im Allgemeinen starr und vom Anwender nicht veränderbar sind, bieten Systeme, die auf dieser Basis arbeiten, nur sehr eingeschränkte Möglichkeiten obige Forderungen zu erfüllen.

Für den Zugriffsschutz bieten sich Verschlüsselungen an. Die verschlüsselte Information kann nur mit entsprechenden Passwörtern wieder sichtbar gemacht werden. Dabei kann entweder das ganze Bild oder nur Teilbereiche (Privatbereiche) durch Verschlüsselung geschützt werden (Bild 6).



Bild 6: Verschlüsselte Privatbereiche im Bereich von Arbeitsplätzen

Mittels Privatbereichen können Szenenanteile definiert werden, die nur von autorisierten Personen eingesehen werden können. Derartige Verschlüsselungen (speziell wenn nur Teile eines Bildes betroffen sind) werden sinnvollerweise schon beim Encodieren in Echtzeit realisiert, d. h. die Werkzeuge zur Bildbearbeitung sollten wie die Sensorik in Bild 5 im digitalen Übertragungspfad direkt auf die Digitalisierung folgen.

Ähnlich ist die Situation beim Schutz der Inhalte vor Manipulationen. Ein CCTV Encoder sollte über eine Echtzeitfunktion zum Einbetten von Wasserzeichen in die Bilddaten verfügen. Da es keinen Standard für Wasserzeichen gibt, sollten verschiedene projektangepasste Algorithmen zum Einsatz kommen können. CCTV Wasserzeichen gehören zu den so genannten fragilen Wasserzeichen. Diese haben die Eigenschaft, dass sie bei Manipulationen der Bilder (z. B. Dekomprimieren und erneutes Komprimieren) zerstört werden, wodurch der Nachweis einer Manipulation erbracht ist. Ein CCTV-Decoder muss den Bruch von Wasserzeichen detektieren können und dazu entsprechende Informationen anzeigen. Dies betrifft insbesondere exportiertes Videomaterial über welches das CCTV System keine Kontrolle mehr hat. Eine derartige Funktion im Decoder erfordert natürlich proprietäre Software, da Multimedia-Decoder eingebettete Wasserzeichen nicht auswerten können, weil diese nicht Bestandteil der Kompressionsstandards sind.

Demgegenüber verwendet das DRM von Multimedia so genannte robuste Wasserzeichen. Diese sollen selbst bei Manipulationen (z. B. Rippen/Transkodieren einer DVD) erhalten bleiben, damit die Quelle nachgewiesen werden kann.

3. Kompressionswerkzeuge unter dem Blickwinkel von CCTV

Im Folgenden sollen überblicksweise die wichtigsten Werkzeuge vorgestellt werden, die für die Kompression in Video-Encodern Verwendung finden. Eine erschöpfende Behandlung würde Bände füllen. Für detailliertere Beschreibungen sei auf [2] verwiesen. Im Zuge des Standardisierungsprozesses wurden diese Werkzeuge immer vielfältiger und verfeinert, um zu Verbesserungen der Kompression zu kommen. Es zeigt sich jedoch, dass viele der von Multimedia Encoder-Implementationen genutzten Kompressionswerkzeuge für CCTV entweder aus Kostengründen oder aus funktionalen Gründen keine praktische Bedeutung haben bzw. einer anderen Herangehensweise als üblich bedürfen um den Anforderungen von CCTV besser zu entsprechen. Viele der Werkzeuge höherer MPEG-Standards z. B. aus dem MPEG-4 Umfeld sind für CCTV-Echtzeitkompression nicht oder nur bedingt geeignet. Die meisten auch für CCTV breit in Anwendung befindlichen Kompressionswerkzeuge wurden schon mit den MPEG-1 bzw. H.263 Standards definiert. Es zeigt sich, dass durch Modifikationen dieser Werkzeuge unter stärkerer Ausnutzung der a priori Information über die Bildeigenschaften eines Kanals erheblich

bessere Kompressionsergebnisse erzielt werden können als sie die teilweise extrem rechenaufwändigen Werkzeuge von Multimedia Encodern erbringen können. Neben einer Verbesserung der Kompressionseffizienz liefern diese modifizierten Werkzeuge niedrigere Latenzzeiten. Ideal sind CCTV spezifische Kompressionswerkzeuge, die sich durch Ausnutzung der verfügbaren Freiheitsgrade in die Standards einbetten lassen. Daneben gibt es eine Vielzahl an Werkzeugen, die nur in proprietären Codecs Verwendung finden können, da die Standards die notwendigen Freiheiten nicht zulassen.

3.1. DCT und Quantisierung - Qualitätssteuerung

Von wenigen Ausnahmen abgesehen, ist dieses Kompressionswerkzeug in allen heutigen Encodern in mehr oder weniger abgewandelter Form im Einsatz (eine Alternative sind Wavelet basierte proprietäre Encoder). D. h. sowohl Motion-JPEG, MPEG-1, MPEG-2 als auch MPEG-4 oder die H.26x Standards verwenden dieses Werkzeug in ihrem Kern. Es gibt Unterschiede in den Verfahrenseigenschaften im Grunde arbeiten aber alle Encoder hier gleich.

Die DCT dient dazu die originalen Bilddaten in ein Format zu transformieren, welches die Entfernung von Informationsanteilen vereinfacht, die vom menschlichen Auge entweder gar nicht wahr genommen werden bzw. deren Entfernung mit einer stufenweisen Verschlechterung der Qualität einhergeht. Man kann es als eine Vergrößerung des Detailreichtums der Vorlage betrachten. Durch die Entfernung von feinen Details der Vorlage wird der Informationsgehalt des Bildes geringer, wodurch es durch spezielle Algorithmen sehr gut komprimiert werden kann. Die reine DCT basierte Einzelbildkompression liefert im Mittel einen Kompressionsfaktor von 15-20.

Bei Motion JPEG oder bei so genanntem I-Frame-Only-MPEG wird jedes einzelne Bild unabhängig von den anderen Bildern einer Sequenz mit diesem Werkzeug behandelt. Obwohl der Rechenaufwand relativ groß ist, stellt diese Stufe heute auch für die DSP basierte Kompression mit voller Video-Bildrate kein Problem mehr dar. [1] enthält eine detaillierte Beschreibung dieses Kompressionswerkzeuges in seiner Anwendung im MJPEG Verfahren.

Sowohl Multimedia als auch CCTV-Codecs nutzen dieses Kompressionswerkzeug in ihrem Kern. Das Verhältnis von Kompressionsrate und Qualität wird durch die so genannten Quantisierungsfaktoren bestimmt. Diese Kompressionsstufen sind bei Multimedia Codecs meist nur in engen Grenzen als globale Qualitätseinstellungen des jeweiligen Kanals zugänglich. Die Umschaltung der Qualität eines Kanals in Abhängigkeit eines Ereignisses erfolgt wegen der in Abschnitt 3.2 beschriebenen Verzögerungseffekte einer Differenzkompression mit großer Latenz. Die Anforderungen an einen CCTV Encoder bezüglich der Einstellung von Qualitätsstufen sind erheblich vielfältiger und dynamischer. Folgende Werkzeuge sind denkbar:

1. Dynamisch wechselnde globale Qualitätsstufen eines Kanals in Abhängigkeit von externen oder auch internen Ereignissen (Bewegungssensorik oder externe Alarme). Diese Wechsel sollten mit geringer Verzögerung möglich sein. So kann ein CCTV Encoder z. B. mehrere dynamisch umschaltbare Kompressions-Stufen bereitstellen, die von einem im Encoder betriebenen Echtzeit-Videobewegungssensor gesteuert werden. In Ruhephasen wird mit niedriger Qualität komprimiert, in Aktivitätsphasen mit einer mittleren Qualität und in Alarmphase mit hoher Qualität.
2. Bildregionen innerhalb eines Kanals mit verschiedenen lokalen Qualitätsstufen je nach Wichtigkeit der in den Regionen erwarteten Information. Diese Regionen sollten dem Encoder als Parametersatz für die Steuerung der Kompression in diesen Bereichen übergeben werden. Eine dynamische Umschaltbarkeit dieser Parametersätze wie in Punkt eins für globale Qualitätsstufen ist denkbar. Bild 7 zeigt ein Beispiel für die Parametrierung eines Qualitätsprofils für einen Videokanal.

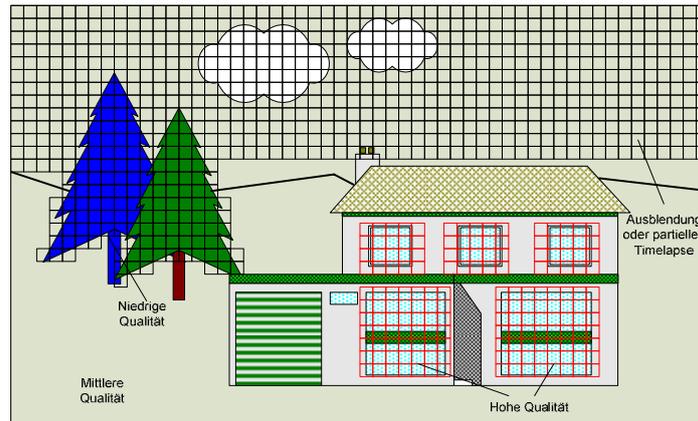


Bild 7: Problemangepasstes Qualitätsprofil einer CCTV Kameraszene

Die hier angedeuteten Möglichkeiten einer dynamischen Qualitätssteuerung lassen sich in den MPEG bzw. H.26x Standards normkonform implementieren.

3.2 Differenzkompression

Dieses Werkzeug wird auch als DPCM bezeichnet. Hier werden Ähnlichkeiten aufeinander folgender Bilder genutzt, um die Kompression zu verbessern. Das Werkzeug ist typisch für die MPEG basierten Encoder. Vereinfacht wird die Differenz zwischen aktuell zu komprimierendem Bild und zugehörigen Referenzbildern berechnet. Diese Differenz wird wieder mittels des DCT Werkzeuges aus 3.1 komprimiert. Bei sich ähnelnden Bildern sind die Differenzen der einzelnen Bildelemente natürlich klein. Dies führt dazu, dass die komprimierten Differenzbilder ebenfalls kleiner werden als würde man das originale Bild komprimieren.

Je nach Art der Implementation des Verfahrens gewinnt man gegenüber einer reinen Einzelbildkompression sogar Performance-Vorteile, da die Anzahl der vom Encoder und Decoder zu behandelnden Bildblöcke im Allgemeinen stark reduziert wird. Damit wird eine DSP basierte Differenz- bzw. MPEG Kompression ohne Berücksichtigung evt. weiterer Kompressionswerkzeuge performanter als eine reine MJPEG Einzelbildkompression.

Das Verfahren der Differenzkompression hat eine große Anzahl an Implementationsdetails, die von den Encodern implementiert werden können oder auch nicht. So gibt es so genannte P- und B-Bilder, welche durch unterschiedliche Vorgehensweisen bei der Differenzbildung gewonnen werden. Nicht jeder MPEG-Encoder unterstützt alle zulässigen Differenzbildungsmethoden und damit die Erzeugung von P- bzw. B-Bildern. Encoder, welche nach MPEG-4/SP Vorgaben implementiert werden, liefern z. B. keine B-Bilder. Arbeitet der Encoder nach MPEG-4/ASP kann er - muss es aber nicht - auch B-Bilder zur Steigerung der Kompressionsrate nutzen.

Für CCTV Anwendungen hat die Art der Differenzbildung im Encoder wichtige Konsequenzen bezüglich des Latenzverhaltens der Datenströme. Dies betrifft die Live-Übertragungsverzögerung als auch die Umschaltverzögerungen beim z. B. Kanalwechsel oder bei Qualitätsänderungen der Kanalparameter. Entstehende Probleme sind prinzipieller Natur und können weder durch schnellere Übertragungsnetze oder schnellere Computer beseitigt werden. Einzig ein sinnvoll an die Anforderung angepasstes Encoder-Design kann die Probleme auflösen. Verursacher von Übertragungs-Latenzen ist hauptsächlich die Art und Weise der Implementation der B-Bilder. Bild 8 soll das Problem verdeutlichen:

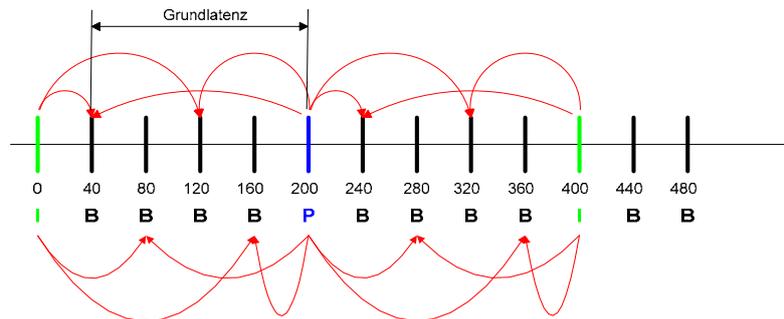


Bild 8: Kompressionslatenz verursacht durch B-Bilder

B-Bilder werden aus jeweils 2 Referenzbildern berechnet. Eines der Referenzbilder liegt in der Vergangenheit und eines in der Zukunft. Die Pfeile in Bild 8 zeigen dies. Problematisch für die Latenz der Kompression sind die in der Zukunft liegenden Referenzbilder. Betrachtet man z. B. das erste B-Bild in Bild 8, so liegt dessen zukünftige Referenz 160 ms entfernt. Der Encoder muss diese Zeit abwarten bevor er das B-Bild komprimieren kann. Erst danach kann es z. B. im Netzwerk versendet werden. Durch diese Art der Differenzierung hat der Encoder also eine prinzipbedingte Übertragungslatenz von 160 ms, die zu allen anderen Verzögerungseffekten hinzu addiert werden muss. Diese tritt bei dieser Art der Vorgehensweise immer auf und kann ohne Verfahrensumstellung nicht verkleinert werden. Bei Einzelbild-Kompression oder nur der Kompression von I- und P-Bildern tritt dieses Problem nicht auf. Je größer die Anzahl direkt aufeinander folgender B-Bilder ist desto größer ist diese Grundlatenz der Kompression. Da B-Bilder im Allgemeinen aber kleiner als P-Bilder sind und damit zu einem besseren Kompressionsfaktor führen, muss für die Anwendung eine Kompromissentscheidung zwischen Kompressionseffizienz und Latenz getroffen werden.

Das Problem der Aufschalt- bzw. Kanalwechsel-Latenz und der Latenz bei Änderung der Qualitätsparameter der Kompression aus Abschnitt 2.2 wird ebenfalls hauptsächlich durch die Differenzkompression hervorgerufen. Bild 9 verdeutlicht die Ursachen. Besonders ärgerlich sind solche Latenzen z. B. für das häufig in CCTV Anlagen verwendete zyklische Durchschalten mehrerer Kanäle in einem Wiedergabefenster bzw. -monitor oder die automatische Aufschaltung von Kanälen durch Alarme, was eigentlich möglichst verzögerungsfrei erfolgen soll.

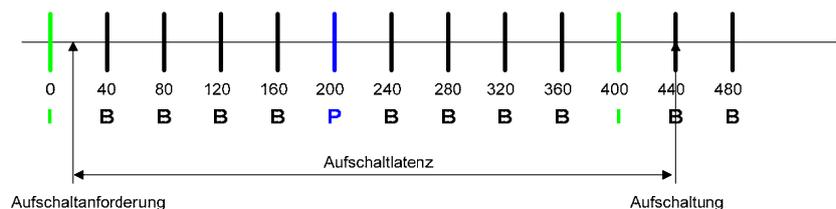


Bild 9: Aufschaltlatenz durch Warten auf I-Referenzbild

Der Video-Decoder hat wie Bild 9 zeigt im Moment der Aufschaltanforderung kein I-Referenzbild. Damit können die folgenden B- und P-Bilder nicht korrekt dekomprimiert werden. Die Wiedergabe kann erst beginnen, wenn der Decoder die für die Dekompression eines Differenzbildes jeweils notwendigen Referenzbilder empfangen hat. Im schlimmsten Fall wird dies zu einer Verzögerung der Länge einer GOP führen. Bei Video mit 25 Bildern pro Sekunde und MPEG-4 üblichen GOV Längen von mehr als 20 Bildern kommt es also leicht zu Verzögerungen im Sekundenbereich, wenn der verwendete CCTV Codec nicht flexibel und schnell durch Erzeugung eines neuen Referenzbildes reagieren kann.

Ein optimaler CCTV Encoder lässt die Einstellung eines individuellen Kompromisses für jeden einzelnen Kanal zu. Um die beiden Parameter Kompressionsfaktor und Latenz durch Einstellung des Differenzbildungsverfahrens zu beeinflussen, gibt es z. B. folgende Möglichkeiten:

1. Verwendung von B-Bildern ausschalten. Der Encoder arbeitet entweder nur mit I-Bildern oder mit Sequenzen aus I und P-Bildern und hat damit geringe Latenz aber auch höhere Bandbreite.
2. Einstellbarkeit der maximalen Anzahl direkt aufeinander folgender B-Bilder für kanalangepasstes optimales Latenz/Kompressionsfaktor Verhältnis
3. Unidirektionale B-Bilder: Deaktivierung der bidirektionalen Referenzierung von B-Bildern. B-Bilder werden wie P-Bilder nur auf ein Referenzbild aus der Vergangenheit bezogen. Die Wartezeit und damit die Latenz entfällt wie bei P-Bildern. Der Kompressionsgewinn der B-Bilder entfällt damit allerdings auch. Der Unterschied zur Verwendung von P-Bildern besteht aber in einer einfacheren logischen Verarbeitung. Für die Dekompression eines P-Bildes braucht neben dem zugehörigen I-Bild alle zeitlich vorangegangenen P-Bilder einer GOP. Für die Dekompression eines nur einseitig vorwärts prädierten B-Bildes braucht man nur das zugehörige I- oder P-Bild.
4. Dynamisches Ändern von GOP Längen, Strukturen und Einfügen von I-Referenzbildern. Dafür benötigt das CCTV System eine Rückkopplung von Decoder auf den Encoder, die den Encoder z. B. über Aufschaltvorgänge informiert und zum Erzeugen von I-Bildern veranlasst. Diese Parameter sollten auch durch Ereignisse beeinflussbar sein, welche z. B. im Alarmfall den Encoder auf eine hohe Qualitätsstufe umschalten.

Im Mittel erreicht man durch Differenzkompression ohne Betrachtung anderer Werkzeuge bzw. kanalabhängiger Besonderheiten eine Halbierung der Bandbreite gegenüber der reinen Einzelbildkompression nach dem JPEG Verfahren. Die beschriebenen Maßnahmen zur Optimierung der Differenzkompression sind normkonform zu den MPEG-Standards von einem CCTV Encoder implementierbar.

3.3 Bewegungskompensation

Die Differenzkompression aufeinander folgender Bilder arbeitet dann gut, wenn sich kaum Änderungen ergeben. Durch Bewegung von Objekten oder auch der ganzen Kamera (Dome) gibt es jedoch mehr oder weniger starke Änderungen zwischen den Bildern, welche die Differenzkompression nachteilig beeinflussen.

Die Verfahren der Bewegungskompensation versuchen die Bewegungen aufeinander folgender Bilder dadurch zu kompensieren, dass übereinstimmende Bildblöcke von aktuellem Bild und Referenzbild gesucht werden. Die Differenzbildung erfolgt dann auf Basis dieser gefundenen Blöcke. Das Differenz-komprimierte P- oder B-Bild enthält Informationen an welcher Stelle in den Referenzbildern der jeweils beste passende Block gefunden wurde. Aus dieser Verschiebungsinformation und der Differenz errechnet der Decoder die rekonstruierten Bilddaten.

Die MPEG-Standards definieren nicht, welcher Algorithmen sich die Bewegungskompensation bedienen muss. Sie legen nur fest, wie die P- und B-Bilder dem Decoder die durch Bewegungen verursachten Blockverschiebungen mitteilen, damit dieser die richtigen Blöcke der Referenzbilder für die Rekonstruktion der originalen Bilddaten verwendet.

Der rechentechnische Aufwand der normalerweise verwendeten Verfahren der Bewegungskompensation ist gewaltig und übersteigt bei weitem den Aufwand aller bisher beschriebenen Kompressionswerkzeuge. Der Aufwand steigt je größer der Suchbereich nach sich ähnelnden Blöcken ist. Größere Suchbereiche können gegenwärtig nur in hochpreisigen DSPs genutzt werden. Für DSPs im CCTV Bereich kommen, wenn überhaupt nur einfachste Verfahren einer Bewegungskompensation in Frage. Die Alternative der Verwendung von starren Multimedia Encodern nimmt der Anwendung aber die Flexibilität des DSP Ansatzes. Wie auch schon bei den Verfahren der Differenzierung muss man zu einer Kompromissentscheidung zwischen Kompressionseffizienz und Flexibilität kommen. Für Szenen mit hohen Bewegungsanteilen, die mit voller Bildrate komprimiert werden sollen, sind Multimedia Encoder DSPs gegenwärtig noch überlegen. Abschnitt 2.9 beschreibt aber die besonderen Eigenschaften

von CCTV Bildmaterial. Bewegung ist hier eher die Ausnahme und damit ist auch der Kompressionsgewinn einer Bewegungskompensation in den meisten Fällen eher gering verglichen mit CCTV typischen Werkzeugen, die Vorteil aus der Kenntnis der Bildinhalte ziehen können. Die MPEG Standards definieren zwar die Möglichkeit einer Bewegungskompensation als Werkzeug zur Verbesserung der Kompressionsrate. Die Anwendung dieser Möglichkeit ist aber optional. So sind auch Encoder ohne dieses Werkzeug oder mit sehr einfachen Implementationen einer Bewegungskompensation durchaus standardkonform.

Unter bestimmten Umständen können CCTV Encoder auch bezüglich der Anwendung des Werkzeuges Bewegungskompensation Vorteile aus a priori Kenntnissen zu den erwarteten Bildinhalten ziehen. Anders als ein Universal-Encoder, der die gesamte Umgebung eines Bildblockes gleichberechtigt nach Ähnlichkeiten durchsuchen muss, können gegebenenfalls Kenntnisse über erwartete Bewegungsrichtungen von Objekten und deren Geschwindigkeiten dazu genutzt werden, den Suchbereich der Bewegungskompensation einzugrenzen und damit den Rechenaufwand auf ein beherrschbares Maß zu reduzieren.

3.4 Bitratenregelung - CBR versus VBR

Multimedia Encoder komprimieren meist mit konstanter Bitrate (CBR). Dazu werden aufwändige Regelalgorithmen im Encoder implementiert. Dies benötigt man um zu deterministischen Annahmen über benötigte Kanalbandbreiten bei der Übertragung zu kommen. Eine CBR basierte Kompression regelt durch Bewegung entstehende variable Bitraten durch Verschlechterung der Bildqualität aus. Dies läuft den Zielstellungen vieler CCTV Probleme direkt zuwider. In den Momenten in denen die eigentliche Nutzinformation eines CCTV Systems produziert wird, wird die Qualität schlechter. Im umgekehrten Fall der Ruhe wird eine höhere Bitrate erzeugt, als eigentlich nötig.

CCTV-Encoder sollten deshalb auch eine VBR (Variable Bitrate) Kompression zulassen. Während Zeiten stärkerer Bewegung im Bild wird hier die Bitrate ansteigen, verglichen mit Zeiten ohne Bewegung. Der Nachteil von VBR ist eine schwankende Bandbreite des Videokanals und damit auch des Speicherbedarfs bei Aufzeichnung. Dies erschwert die Bestimmung von Speicherkapazitäten und erfordert Worst Case bzw. Mittelwertanalysen, welche nur durch intime Kenntnis der jeweiligen Bildsituation zu einigermaßen korrekten Ergebnissen führen.

3.5 Bildratensteuerung - VFR versus CFR

Variable Framerate ist eines der am häufigsten von CCTV genutzten Mittel um Bandbreiten und Speicherbedarf zu senken. Schon klassische Timelapse Rekorder nutzten diese Möglichkeit zur problemangepassten Langzeitarchivierung. Die Möglichkeiten digitaler DVRs in dieser Richtung sind extrem erweitert und flexibilisiert worden. So kann die Bildrate mit hoher Geschwindigkeit umgeschaltet werden. Wie Bild 10 zeigt, wird in bewegungsfreien Zeiten mit niedriger Rate aufgezeichnet bzw. übertragen. Detektiert ein Sensor wird auf hohe Bildrate umgeschaltet.

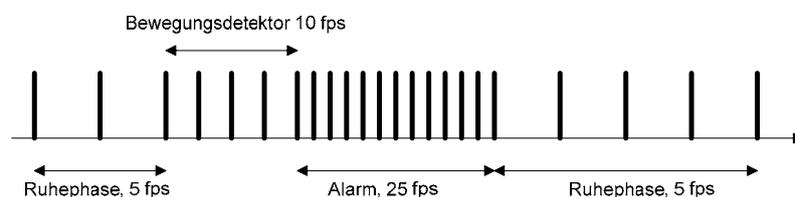


Bild 10: Variable Bildrate durch Videosensorik

Multimedia Encoder unterstützen VFR nur sehr eingeschränkt. Dynamische Umschaltungen der Rate sind praktisch gar nicht möglich. Bildratenwechsel beziehen sich allenfalls auf den ganzen Film oder längere Szenen. MPEG-2 geht obwohl es kleinere Freiheitsgrade gibt fest von

videoüblichen Bildraten von 25 Bildern pro Sekunde aus. Encodiert ein CCTV Encoder nach MPEG-2 Norm aber mit niedriger Bildrate so kommt es bei Wiedergabe gespeicherter Sequenzen auf einem Standard-Decoder zu Problemen. Ohne besondere Maßnahmen werden die Bilder im Zeitraffer wiedergeben. Der originale Zeitfluss - insbesondere bei ständig wechselnder Bildrate bei der Aufzeichnung - kann dann normalerweise nicht rekonstruiert werden. Dies bereitet besondere Probleme, wenn die Bilddaten mehrerer Kanäle evt. sogar noch mit Audio synchronisiert wiedergegeben werden sollen.

MPEG-4 erlaubt auch andere Bildraten als 25 fps. Die meisten Codecs gehen aber wiederum von einer festen Rate innerhalb eines aktuell verarbeiteten Streams aus. Diese kann zwar niedriger als 25 fps sein aber nicht variabel innerhalb eines MPEG-4 Streams schwanken.

Diese Probleme können CCTV Encoder aber bei intelligenter Implementation auflösen, so dass es gelingt auch MPEG-2 bzw. MPEG-4 kompatible Datenströme mit dynamisch wechselnder Bildrate zu erzeugen und diese auch mit Standard-Decodern im originalen Zeitfluss wiederzugeben.

3.6 GOP-Längen- und Struktursteuerung

Die MPEG-Standards haben erhebliche Freiheitsgrade bezüglich:

- der in den Datenströmen verwendeten Bildtypen (I-, P-, B-Bilder)
- GOP-Strukturen (z. B. Anzahl aufeinander folgender B-Bilder)
- GOP-Längen

So kann ein CCTV Encoder die GOP Länge dynamisch steuern. In bewegungsarmen Zeiten sind die GOPs länger und es werden weniger bandbreitenhungrige Referenzbilder erzeugt. Ereignen sich Bewegungen, so gibt es mehr Referenzbilder. Bei Auslösen eines Sensors setzt der Encoder sofort mit einem neuen I-Referenzbild auf und gestattet es damit unmittelbar mit einem neuen Qualitätsprofil zu arbeiten.

Durch Ausnutzung dieser Freiheitsgrade kann es ein CCTV Encoder die negativen Seiteneffekte einer Differenzkompression z. B. bezüglich von Latenzen oder der Zugriffsflexibilität auf Einzelbilder weitgehend beseitigen.

Speziell die MPEG-Kompressionsstandards treffen keine Festlegungen zu GOP-Längen also dem zeitlichen Abstand von I-Referenzbildern. Je länger eine GOP ist, desto mehr Differenzbilder enthält sie. Sind die Änderungen der Bilder auch über lange Zeiten von Alarmsituationen abgesehen nur klein, so kann die Kompressionsrate durch Verlängerung der GOPs weiter verbessert werden. Auch dieses Werkzeug muss mit Bedacht situationsabhängig eingesetzt werden. Lange GOPs haben Nachteile wie:

- Nicht alle Decoder können damit umgehen, so dass man an bestimmte Decoder, die diese Möglichkeit unterstützen, gebunden ist.
- Ein normkonformer MPEG-Export für DVDs schreibt eine maximale GOP-Länge von 15 vor. Nutzt man längere GOPs können gegebenenfalls einige Exportmöglichkeiten der Bilder aus dem CCTV System nicht mehr oder nur noch mit aufwändigem Transcodieren genutzt werden.
- Ist die Länge der GOPs statisch vorgegeben, arbeitet das Werkzeug dann ineffizient, wenn sich während der Dauer der GOP starke Änderungen ereignen, welche die Größe der Differenzbilder so lange in die Höhe treiben bis das nächste I-Referenzbild planmäßig erzeugt wird. Diesen Effekt kann man durch Erhöhung der Eigenintelligenz des Encoders durch dynamische I-Bild Erzeugung vermeiden. Abhängig von einem bestimmten Kriterium, wie z. B. der Größe der Differenzbilder entscheidet der Encoder selbständig eine neue GOP mit einem neuen I-Referenzbild aufzusetzen. Man erhält Datenströme variabler GOP-Länge.

3.7 Dynamische Auflösungssteuerung

CCTV Encoder lassen die schnelle Umschaltung zwischen verschiedenen Auflösungsstufen bei der Kompression zu. Neben der Einsparung von Speicherkapazität beschreibt Abschnitt 2.3 den Einsatz dieses Werkzeuges für die Optimierung des Präsentationsdurchsatzes. Der Einsatz dieses Werkzeuges in einer Standard-Konformen Weise ist problematisch. Bei MPEG-2 ist die Bildauflösung z. B. eine globale Eigenschaft des Streams. Sie wird dem Decoder einmalig mitgeteilt. Wechselt die Auflösung während der Wiedergabe hängt es von der Robustheit des Decoders ab, ob die Wiedergabe möglich ist oder nicht. Sollen die Exporte eines CCTV Systems mit dynamischer Auflösung mit beliebigen Decodern wiedergebar sein, so muss es beim Export ein zeitaufwändiges Umskalieren auf eine einheitliche Auflösung geben.

3.8 Paritätsseparation

Dieses Kompressionswerkzeug von CCTV Encodern hat mehrere positive Effekte:

- Halbierung der Bitrate und damit der Speicherkosten gegenüber 4CIF Kompression
- Erzeugung unabhängig komprimierbarer Streams für verschiedene Aufgaben
- höhere Qualität als 4CIF Bilder bei starken Zeilensprungartefakten bei Bewegungen

Dem steht als Nachteil der Verlust an vertikaler Auflösung gegenüber. Dieser spielt jedoch oft nur bei statischen Bildern ohne Bewegung eine Rolle, was den Kompromiss in den vielen Anwendungsfällen akzeptabel erscheinen lässt. Wie bei Anwendung der anderen Werkzeuge sollte ein CCTV Codec dieses spezielle Kompressionswerkzeug alternativ zur 4CIF-Vollbildverarbeitung anbieten, um eine optimale Adaption an die Projektverhältnisse zu erlauben. Mittels spezieller Implementationstechniken lassen sich aus den um eine Parität beraubten Kompressionsdaten nichtsdestotrotz MPEG konforme Streams erzeugen, die mit Standard Decodern wiedergegeben werden können.

3.9 Weitere Kompressionswerkzeuge

3.9.1 Multi-Referenz Bewegungskompensation

Dieses Werkzeug ist Bestandteil des H.264-Standards, der auch als MPEG-4/AVC oder MPEG-4 Part 10 bekannt ist. Die Bewegungskompensation vorangegangener MPEG-Standards (Abschnitt 3.3) geht von einem oder zwei zeitlich unmittelbar benachbarten Referenzbildern für die Kompression von P- und B-Bildern aus. In H.264 wird dies auf mehrere Referenzbilder erweitert. Im Falle von P-Bildern liegen diese alle zeitlich vor dem zu komprimierenden P-Bild. Im Falle von B-Bildern können die Referenzen zeitlich sowohl vor als auch nach dem Referenzbild liegen. In [3] wird das Verfahren beschrieben.

Der rechentechnische Anspruch an die Codecs übersteigt bei weitem den ohnehin schon gewaltigen Anspruch der „einfachen“ Bewegungskompensation aus Abschnitt 3.3. Der Hardware-Aufwand steigt durch die Notwendigkeit der Pufferung einer Vielzahl von Referenzbildern im Encoder und Decoder.

Eine DSP Implementation dieses H.264-Werkzeuges ist mit der gegenwärtigen Schaltkreisgeneration nur mit hohen Kosten möglich. Da Rechenleistung aber bislang immer nur eine Frage der Zeit war, werden zukünftige DSP-Generationen sicher auch die Leistung zur Beherrschung dieser Werkzeuge kostengünstig bereitstellen. Allerdings ergibt sich mit den Erläuterungen der vorhergehenden Abschnitte die Frage ob man derartig aufwändige Werkzeuge im CCTV-Bereich überhaupt braucht, bzw. ob man nicht besser in die im Vorhergehenden dargestellten alternativen CCTV-spezifischen Werkzeuge investieren sollte.

Verwendet man B-Bilder, die aus Referenzen berechnet werden, die weit in der Zukunft liegen, läuft die Multi-Referenz Bewegungskompensation den Zielen der meisten CCTV-Projekte zuwider. Die in Abschnitt 3.2 beschriebenen Verzögerungseffekte nehmen hier stark zu, so dass

eine Brauchbarkeit für Echtzeitanwendungen meist nicht gegeben ist. Mehrfachreferenzen auf zeitlich zurück liegend Bilder erhöhen diese Latenzen nicht, verkomplizieren aber den Zugriff auf das Einzelbild sehr stark.

3.9.2 Global Motion Compensation (GMC)

Das GMC-Werkzeug ist Bestandteil des MPEG-4 Standards im Rahmen des Advanced Simple Profiles (MPEG-4/ASP). Die GMC soll globale Veränderungen des Kamerabildes, welche durch Translation, Rotation und Zoom entstehen, kompensieren. Mit diesem Ziel wäre das Werkzeug für die Anwendung auf Domekameras gut geeignet. Wie bei der „normalen“ Bewegungskompensation legt der Standard nicht die Algorithmen fest welche ein MPEG-4/ASP Encoder für die Berechnung der GMC verwenden soll. Es wird lediglich die Syntax definiert, welche den Decoder steuert, um die Bewegung zu kompensieren. Wie auch die folgende Quarter Pixel Motion Compensation ist das Werkzeug optionaler Bestandteil des Standards. Die Angabe von MPEG-4/ASP in einem Datenblatt heißt damit nicht, dass ein Encoder mit diesem Werkzeug arbeitet.

Der rechen-technische Anspruch an die GMC ist wie bei allen Verfahren der Bewegungskompensation riesig. Deshalb ist dieses Werkzeug in heutigen MPEG-4 Encodern noch kaum in Gebrauch. Obwohl der Standard von MPEG-4/ASP Decodern eigentlich verlangt, dass sie mit GMC-Datenmaterial umgehen können, sind selbst die verfügbaren Decoder Implementationen meist nicht in der Lage derartige MPEG-4 Streams wiederzugeben.

3.9.3 Quarter Pixel Motion Compensation

Dieses Werkzeug ist ebenfalls Bestandteil der MPEG-4/ASP Spezifikation. Es soll der Verfeinerung der pixelbasierten Bewegungskompensation dienen. Die Suche nach passenden Bildblöcken im Referenzbild erfolgt auf einem feineren Raster als dem groben Pixelraster. Da dieses Raster physikalisch natürlich nicht existiert, müssen dessen Werte aus den Werten der wirklich verfügbaren Pixel interpoliert werden. Da Bewegungen der Szene sich nicht an Pixel-Abständen orientieren, bringt das Werkzeug eine Verbesserung der Kompressionsrate dadurch mit sich, dass die im Referenzbild gefundenen Blöcke genauer auf die zugehörigen Blöcke des aktuell komprimierten Bildes passen. Damit werden die Differenzen kleiner und die Kompression wird bei gleicher Qualität besser.

Für den Rechenaufwand und die aktuelle Verbreitung dieses Werkzeuges in verfügbaren Codec-Implementationen gilt die gleiche Aussage wie bei der GMC.

4. Was ist MPEG4CCTV

MPEG4CCTV steht als Markenname für ein digitales CCTV System der Firma Geutebrück, welches mit verschiedenen MPEG-Codecs (Multistandardsystem) operieren kann. Dabei hat der Anwender die Freiheit sich zwischen der Flexibilität einer DSP-basierten MPEG-Implementation und einer Multimedia-Chip basierten starren MPEG-Implementation frei zu entscheiden. So können die Vorteile, die beide Ansätze bieten, optimal entsprechend der Projektgegebenheiten kombiniert werden. Durch die offene MPEG4CCTV-Aufzeichnungsschnittstelle (Bild 11) kann das System auch MPEG-Encoder integrieren, die z. B. nach HDTV Norm arbeiten. So trägt die Architektur der wachsenden Forderung nach höheren Auflösungen Rechnung.

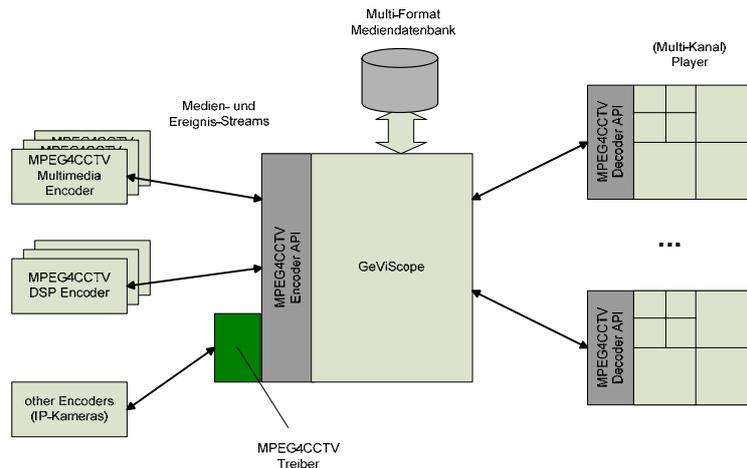


Bild 11:

Der DSP-basierte CCTV-Codec stellt heute schon eine Reihe der in den vorhergehenden Abschnitten beschriebenen CCTV-spezifischen Kompressionswerkzeuge in einem MPEG-konformen Rahmen bereit. Außerdem unterstützt er die Echtzeitsensorik aus Abschnitt 2.8 und die Bildverarbeitungsverfahren aus Abschnitt 2.10. Der DSP-basierte MPEG4CCTV Codec arbeitet mit den folgenden Kompressionswerkzeugen:

- DCT
- Differenzkompression
- Unidirektionale vorwärts prädizierte B-Bilder garantieren die Latenz eines Einzelbildverfahrens und optimale Zugriffe auf das Einzelbild
- optionales Paritätssplittung erlaubt zwei voneinander unabhängige Kompressionskanäle pro Kamera
- dynamische Qualitäts- und Auflösungssteuerung mit sehr geringer Umschaltverzögerung durch Steuerung der Referenzbilderzeugung
- variable Bildrate (VFR)
- dynamische GOP-Längen und Struktur-Steuerung
- intelligente Kompressionsdynamik durch latenzfreie Umschaltung von Qualitätsprofilen in Abhängigkeit der Aktivitätssensorik des Encoders
- Audio Kompression, umschaltbare synchronisierte und unsynchronisierte Live-Wiedergabe
- Constant Quality mit VBR

Neben der reinen Medien-Kompression stellt der DSP-basierte MPEG4CCTV Encoder auch Werkzeuge für Bild- und Audio-Sensorik sowie Bildverarbeitung bereit. Bereits verfügbar sind:

- Aktivitätsdetektion als Sensor für die Qualitätssteuerung des Encoders und einfache Alarmquelle
- Audio-Schwellwert-Sensor
- Privatbereiche

Weitere in den vorangegangenen Abschnitte gezeigte Werkzeuge wie z. B. Qualitätsprofile, Wasserzeichen, Videobewegungsmelder und neuartige Bildsensoren können wegen der Flexibilität des DSP Ansatzes einfach in Form von Software-Updates ohne Austausch der Hardware bereitgestellt werden. Der Codec kann also den wachsenden Ansprüchen an die Bildverarbeitung im Lebenszyklus eines CCTV-Systems gerecht werden.

Der hochoptimierte Decoder erlaubt die Dekompression von bis zu 400 Bildern auf aktueller PC-Standard-Technik, was einer simultanen Darstellung von 16 Videokanälen mit voller Bildrate pro Arbeitsplatz entspricht.

Der Multimedia-Chip basierte MPEG4CCTV Codec demgegenüber kann in allen Anwendungen zum Einsatz kommen bei denen es vordergründig um:

- volle Bildraten von 25 fps und 4CIF Auflösungen
- gute Kompressionseffizienz
- CBR für deterministische Bandbreiten
- kostengünstige Lösungen mit wenig Parametrieraufwand

geht. Ein bedeutendes Szenario ist die Überwachung von Kasinos. Wegen der hohen Geschwindigkeit der an den Spieltischen ablaufenden Vorgänge wird fast immer eine Permanentaufzeichnung mit maximaler Qualität und Bildrate gefordert. CCTV-typische Werkzeuge z. B. VFR, wie sie im DSP-basierten MPEG4CCTV Codec implementiert sind, können hier kaum genutzt werden, um die Kompression zu verbessern.

5. Zusammenfassung

Multimedia Chip basierte Video-Encoder mit starrer Implementation der Kompressionswerkzeuge können nur kleine Teilbereiche der sehr vielfältigen Anforderungen an CCTV Systeme abdecken. Haupteinsatzgebiet von darauf basierenden CCTV Systemen sind Videoaufzeichnungen und Übertragungen mit voller Bildrate bei guter Qualität. Derartige Anwendungsfälle sind jedoch die Ausnahme in CCTV. CCTV Projekte, die nur wegen der Inflexibilität der Encoder zu Aufzeichnungen und Übertragungen mit voller Bildrate gezwungen sind, sind unökonomisch und profitieren nur unzureichend von der speziellen CCTV-Situation, um ein besseres Nutzen/Kostenverhältnis zu erzielen.

Völlig untauglich sind Multimedia Encoder basierte CCTV Systeme oft wenn es neben reiner Aufzeichnung und Übertragung auch um Themen wie Echtzeit-Sensorik und -Image Processing geht. Problematisch sind auch die oftmals hohen Latenzen für die Live-Übertragung bzw. Kompressionssteuerung derartiger Chips. Eine Optimierungsmöglichkeit durch Beeinflussung von Encoder-Algorithmen besteht nicht.

Demgegenüber bieten DSP basierte CCTV Encoder mit auf den Einsatzfall CCTV zugeschnittener Firmware die Möglichkeit sich individuell an das jeweilige Szenario anzupassen und Vorteil aus den Eigenschaften des jeweils vorliegenden Bildmaterials zu schöpfen. Damit wird in einer Vielzahl von Anwendungsfällen ein erheblich besseres Nutzen/Kostenverhältnis als bei Multimedia-Encodern erzielt. Eine Vielzahl von Problemstellungen von CCTV sind mit Multimedia-Encodern prinzipiell nicht zugänglich, während sie bei DSP basierten Systemen im Rahmen von Firmware-Entwicklungen und über Software-Updates realisierbar sind. Nachteile von DSPs wie die niedrigere Performance im Vergleich zu fest programmierten Schaltkreisen werden zunehmend in den Hintergrund gedrängt. Aktuelle DSP Generationen sind heute auch bei niedrigen Schaltkreiskosten in der Lage zumindest mit PAL/NTSC basierendem Videomaterial in Echtzeit problemlos umzugehen.

Insgesamt bieten DSP basierte CCTV Systeme eine erheblich höhere Flexibilität als Multimedia Chip basierte. Damit erhöhen sie die Investitionssicherheit von CCTV-Systemen und tragen den im Laufe des Lebenszyklus einer Anlage wachsenden Anforderungen Rechnung.

Der Beitrag demonstrierte an typischen MPEG-Kompressionswerkzeugen wie diese im Rahmen einer DSP-basierten Encoder-Implementation unter Ausnutzung von Standard-Freiheitsgraden genutzt werden können um die Funktionalität von CCTV Produkten besser den individuellen Projektgegebenheiten anpassen zu können. Es zeigt sich, dass die Antwort auf die simple Frage nach einem optimalen CCTV Codec extrem kompliziert ist und generalisiert nicht gegeben werden kann. Jedes CCTV-Projekt braucht individuelle und idealerweise von den Eigenschaften der Einzelkanäle abhängige Kompromisse. Diese kann mit DSP basierten Encoder-Implementationen am besten erreicht werden.

Quellen

- [1] Matthias G. Döring: Digitale CCTV-Systeme. Economica Verlag - Heidelberg, 2004
- [2] Iain E. G. Richardson: H.264 and MPEG-4. WILEY, 2003
- [3] Girod, Bernd; Flierl, Markus: Videocodierung mit mehreren Referenzbildern. it 5/2003. Seiten 280-284

Dr. Matthias Döring
Februar 2006