

Im ersten Teil gingen wir auf die Systemstrukturen ein und verglichen die Techniken der analogen sowie digitalen Bildverarbeitung. In diesem Teil konzentrieren wir uns auf die Kodierung und Umsetzung im Übertragungsprotokoll.  
**Fortsetzung aus »de« 10/2010**

# Grundlagen der Videointegration (2)

## Zeitverhalten, Netzwerkanbindung, Protokolle und Datensicherheit

Die Bildverarbeitung und -übertragung hängt im starken Maße von der verwendeten Kodierung ab sowie von der Übertragungstechnik. Wir betrachten hier die Techniken mit verschiedenen gebräuchlichen Verfahren.

Bei der Planung und Realisierung digitaler Videosysteme ist die richtige Auswahl und Konfiguration des Netzwerks von entscheidender Bedeutung, denn eine Videoübertragung stellt hohe Anforderungen an das Netzwerk: große Datenmengen mit möglichst geringer Zeitverzögerung.

### Datenmengen und Zeitverhalten

Um es anschaulich zu machen: Ein DIN-A4-Seiten-Text benötigt bei digitaler Übertragung ein Datenvolumen von weniger als 5kByte, ein Videobild in PAL-Auflösung dagegen mehr als

1200kByte, also mehr als das 250-fache (Bild 5).

Diese große Datenmenge fällt bei einer Live-Übertragung 25 mal pro Sekunde an. Um die Netzwerke nicht bis an die Grenzen auszulasten, ist die Reduktion der zu übertragenden Datenmenge mittels geeigneter Video-Kompressionsverfahren unerlässlich. Dadurch erreicht man bei Inkaufnahme von mehr oder weniger starken Qualitätseinbußen (Artefakte, verringerte Auflösung oder Bildfrequenz) eine Datenreduktion von 50:1 bis zu 200:1. Der daraus resultierende Datenstrom von ca. 2MBit/s für einen

Videostream mit 25 PAL-Bildern pro Sekunde in guter Qualität ist immer noch anspruchsvoll für die Auslegung des Netzwerks, denn eine Livebild-Übertragung erfordert nicht nur große Datenmengen, sondern auch ein gutes Zeitverhalten.

### Kompromiss zwischen Auflösung und Zeitverhalten

Die Kompression und die paketorientierte Datenübertragung (bei IP-Netzen) bewirken prinzipbedingt immer einen Zeitversatz (Latenz) zwischen Sender und Empfänger. Dieser sollte idealerweise unter 150ms liegen, damit der Anwender keine Verzögerung bei der Steuerung von Kameras »spürt«. Solche Latenzzeiten erfordern jedoch einen hohen technischen Aufwand und bedingen damit hohe Kosten. Je nach Anwendungsfall sind daher verschiedene Lösungen wirtschaftlich optimal. Kritisch bezüglich Latenz ist beispielsweise die Verfolgung von Personen mittels Schwenk-Neige-Zoom-Steuerung, Audio bidirektional (Gegensprechen) oder das Betätigen von Schranken und Rolltoren aus der Ferne.

Wenn nicht genügend Übertragungsbandbreite zur Verfügung steht, muss man Abstriche an Bildqualität und Bildfrequenz machen. Das kann dann zu den Videoanlagen führen, von denen die Anwender sagen, die frühere Analogtechnik sei doch viel besser gewesen. Daher sollte man besser schon bei der Planung eines digitalen Videosystems auf die passende Dimensionierung des Netzwerks achten und über geeignete Konzepte zur Eindämmung der Datenflut nachdenken.

Eine Möglichkeit dazu bietet »Intelligenz« in den Bildquellen: Um die Netze nicht zu überlasten, überträgt

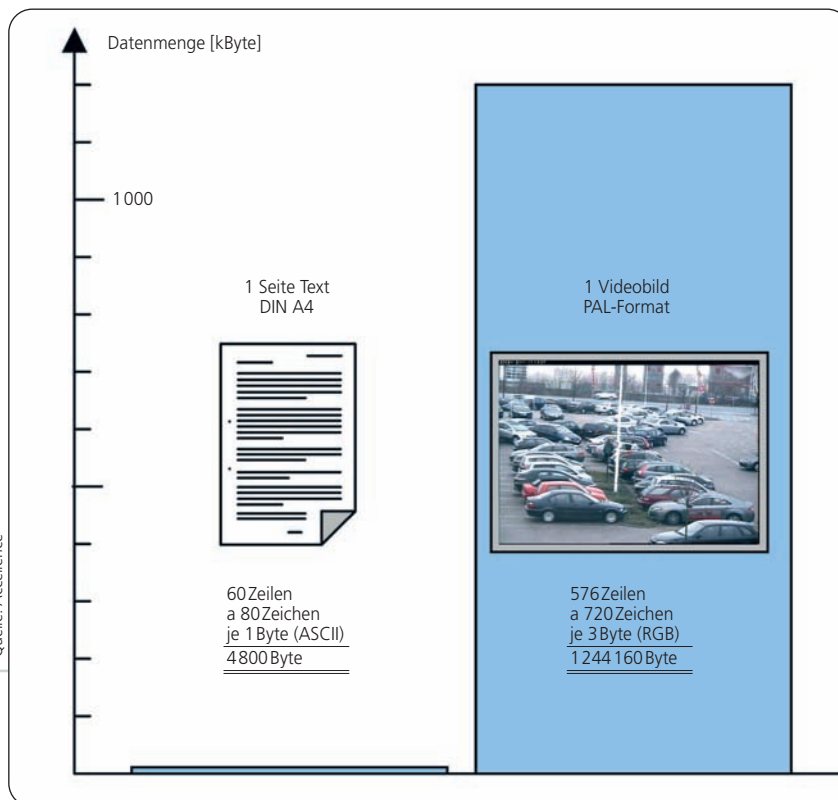


Bild 5: Vergleich der Datenmengen Text / Videobild

Quelle: Accellence

man nur dann Bilder, wenn sie relevante Informationen enthalten. Die Bewertung, wann Bilder relevant sind, stellt eine anspruchsvolle Aufgabe dar, die ausgefeilte Algorithmen und viel Rechenleistung erfordert. Verschiedene Hersteller arbeiten an solchen Lösungen; es ist je nach Anwendungsfall zu prüfen, welche dieser Ansätze passen und ob sie schon ausgereift genug sind, um in der Alltagspraxis einen echten Nutzen zu bringen.

Eine weitere Eindämmung der Datenflut gelingt durch immer ausgefeiltere Kompressionsverfahren, die aber prinzipbedingt auch unerwünschte Eigenschaften bezüglich des Zeitverhaltens oder der Bildqualität bei schnellen Bewegungen mit sich bringen können.

Eine aktuelle Entwicklung ist Scalable Video Coding (SVC): Hier überträgt man die Videodaten im Netzwerk in mehreren Teilstreams, jeder für sich dekodierbar. Durch Kombination der Teilstreams verbessert man die Bildqualität, je nach verfügbarer Bandbreite.

Es gibt Hersteller von Videosystemen, die eine besondere Kompetenz

**MEHR INFOS**

**Links**  
[www.bsi.de](http://www.bsi.de), [www.ebues.de](http://www.ebues.de)

**Noch Fragen?**  
 Sigurd Schobert  
 Telefon: (089) 12607-244  
[schobert@de-online.info](mailto:schobert@de-online.info)

bei der Übertragung von Video über schmalbandige Netze entwickelt haben (z.B. über ISDN mit 64kBit/s oder mit-Kanalbündelung, 128kBit/s), und das mit brauchbaren Resultaten. Andere Hersteller haben sich von vornherein auf breitbandige IP-Netze eingestellt, erzielen da mitunter bessere Ergebnisse, versagen jedoch, wenn die verfügbare Bandbreite knapp wird. All dies ist bei der Planung eines verteilten Videosystems zu beachten.

### Aspekte der Videointegration

Damit Videosysteme verschiedener Hersteller zusammen funktionieren, müssen sie entweder kompatibel sein oder man muss mittels einer geeigneten Software die Schnittstellen und

Datenformate konvertieren (Bild 6). Wenn im Videobereich von Kompatibilität die Rede ist, meint man oft nur Kompressionsverfahren wie MPEG oder H.264. Daneben gibt es aber eine Vielzahl weiterer Aspekte, die für den reibungslosen Betrieb eines heterogenen Videosystems wichtig sind:

- Netzwerkverbindung, Protokolle
- Datensicherheit, Authentifizierung
- Bildübertragung, Datenreduktion
- Zeitsynchronisation, Timecodes
- Metadaten
- Kamerasteuerung
- Fernwirken, Telemetrie
- Audio (Hören, Sprechen)
- Aufzeichnung, Recherche
- Ereignisübertragung
- Sonderfunktionen

### Netzwerkverbindung, Protokolle

Schon beim Netzwerkanschluss und den verwendeten Protokollen unterscheiden sich die angebotenen Videogeräte erheblich. Hier dominiert die Ethernet-schnittstelle mit 100MBit/s. Die Länge der Netzwerkkabel (Kupfer) beim 100Base-T-Anschluss endet nach 100m, Lichtwellenleiter (LWL) erlauben dagegen mehrere Kilometer.

**STICHWORT MEGAPIXEL**

Bei analogen Kameras orientiert sich die Bildauflösung (Anzahl der maximal unterscheidbaren Bildpunkte in horizontaler und vertikaler Richtung) meist an den Fernsehstandards PAL (in Deutschland) bzw. NTSC (in den USA). Rein digitale Systeme sind jedoch nicht mehr an diese Formate gebunden. Mit JPEG (genauer JFIF) können beispielsweise alle Bildauflösungen bis maximal 65536x65536 Bildpunkte übertragen und gespeichert werden. Einige Hersteller unterstützen Full-HD und HDMI, also Technologien, die aus dem Bereich der Unterhaltungselektronik kommen, andere Hersteller bieten nahezu beliebige Auflösungen, von QCIF bis hin zu aktuell 16 Megapixeln (Millionen Bildpunkte).

Welche Bildauflösung optimal ist, muss im Einzelfall geprüft werden, denn hohe Auflösungen erfordern viel Rechenleistung und belasten die Netzwerke – die Bilder kommen dann nicht mehr flüssig, und das Übertragen, Speichern und Nachbearbeiten wird unnötig teuer und langsam. Auch die Lichtempfindlichkeit ist bei Multi-Megapixel-Kameras ein kritisches Thema, weil jeder Bildpunkt (»Pixel«) bei gleicher Chipgröße mit einer kleineren Lichtmenge auskommen muss.

Außerdem ist eine hohe Megapixelzahl nur dann sinnvoll, wenn Optik und Mechanik darauf abgestimmt sind: Bei vibrierendem Kameramast, minderwertigen Linsen oder ungenauer Brennweite bringt eine hohe Megapixelzahl keinen Vorteil.

Im Gegensatz zum Anwendungsfall »Fernsehen«, wo viele Konsumenten im »Broadcast«-Verfahren das gleiche Bildmaterial im gleichen Format auf weitgehend standardisierten Geräten empfangen und es nur darauf ankommt, bei Anzeige des Gesamtbildes eine für das Auge angenehme Darstellung zu erreichen, ist es bei digitalen CCTV-Systemen möglich und je nach Anwendungsfall sinnvoll, mit unterschiedlichen Auflösungen zu arbeiten: Um vor dem Schließen eines Rollltores zu prüfen, ob da eventuell noch ein Kinderwagen steht, genügt eine Festkamera mit geringer Auflösung. Bei der Überwachung eines Kriminalitätsschwerpunktes (beispielsweise Bahnhofsvorplatz) kann dagegen eine hochauflösende Weitwinkelkamera durchaus sinnvoll sein, wenn dadurch nachträglich aus den aufgezeichneten Videobildern noch Gesichter oder Kennzeichen herausvergrößert werden können, um einen Täter zu identifizieren zu.

Die für solche Anwendungsfälle benötigten Auflösungen liegen deutlich über dem »HD-Format« aus dem Consumer-Bereich, weil hier nicht nur das Gesamtbild betrachtet wird, sondern Ausschnittsvergrößerungen zusätzliche Details liefern sollen. Außerdem bewirkt der technische Fortschritt, dass künftig immer höhere Auflösungen möglich sein werden: Die verfügbare Rechenleistung, der Speicherplatz und die Netzwerkkapazitäten werden voraussichtlich weiterhin exponentiell wachsen.

Aber auch dann müssen noch alte Bestandsgeräte mit entsprechend geringeren Auflösungen aufgeschaltet werden können. In der Praxis sind daher Systeme von Vorteil, die variable Bildauflösungen unterstützen und idealerweise dynamisch den jeweiligen Erfordernissen angepasst werden können, beispielsweise mit einer niedrigen Auflösung bei entsprechend hoher Bildfrequenz für schnelle Bewegungen und Kameraschwenks oder mit einer hohen Auflösung unter Inkaufnahme einer langsameren Bildfolge, wenn man Details erkennen will.

Daneben gibt es auch Geräte, die sich mittels eingebauter WLAN-Antenne drahtlos vernetzen lassen. Allerdings muss sich dann in Reichweite (abhängig von den Ausbreitungsbedingungen für die Funksignale etwa 5 ... 300m) ein passender WLAN-Access-Point befinden. Ab da geht es dann meistens mit Ethernet via Kabel weiter (siehe Kasten LAN – WAN).

Wieder andere Videogeräte arbeiten mit Mobilfunk (GSM, UMTS), um auch über größere Entfernungen und bei mobilen Anwendungen drahtlose Verbindungen zu ermöglichen. Allerdings ist bei diesen Übertragungstechnologien die verfügbare Netzwerkbandbreite vergleichsweise gering und schwankt, abhängig von der Anzahl der Teilnehmer in der Funkzelle und den Ausbreitungsbedingungen für die Funkwellen. Neben unabsichtlichen Störungen durch andere elektrische Geräte muss man beim Sicherheitskonzept auch mit absichtlichen rechnen: Beobachtete Personen wollen die Funkverbindungen verhindern. Hinzu kommen vergleichsweise hohe Kosten für den Datentransfer.

Ebenfalls mit kleinen Bandbreiten (64kBit/s bzw. 128kBit/s bei Kanalbündelung) müssen Videogeräte auskommen, die ihre Bilder über ISDN übertragen. ISDN wird zwar nach und nach abgelöst, spielt bei Bestandsanlagen aber immer noch eine Rolle. Man muss grundsätzlich zwei Übertragungsarten von Video über ISDN unterscheiden, und zwar:

#### IP über ISDN

Wie bei allen vorgenannten Verfahren nutzt man auch hier IP (Internet Protocol) als grundlegendes Protokoll. Damit lassen sich Daten in den höheren Protokollschichten immer in gleicher Weise übertragen, unabhängig von der Art



Quelle: Accellence

**Bild 6: Eine kleine Auswahl von Videogeräten, die typischerweise vom VMS integriert werden müssen**

der physikalischen Verbindung (Ethernet, WLAN, DSL, UMTS oder eben ISDN). Um IP-Datenpakete über ISDN zu übertragen, müssen sie ins PPP (Point to Point Protocol) »eingepackt« werden. Dies kann ein ISDN-Router erledigen; dafür müssen alle benötigten Verbindungen im Router konfiguriert werden.

Komfortabler geht es mit DFÜ-Verbindungen, die direkt auf dem PC konfiguriert werden können, auf dem auch das VMS läuft. Bei Anlagen mit mehreren Bedienplätzen setzt das einen ISDN-Server mit Remote-CAPI voraus. CAPI steht als Abkürzung für »Common Application Programming Interface« und stellt die Standard-Programmierschnittstelle für ISDN-Anwendungen dar. Aktuell ist die Version CAPI 2.0.

Mittels Remote Access Service (RAS) kann ein VMS die DFÜ-Verbindungen dynamisch verwalten, also je nach Bedarf anlegen, so dass die Verbindungsdaten zentral gepflegt werden

können. Soll dies auch für eingehende Verbindungen oder für mehrere Verbindungen von einem PC aus möglich sein, um z.B. gleichzeitig die Videobilder verschiedener ISDN-Gegenstellen auf einem Bedienplatz empfangen zu können, benötigt man einen CAPI-WAN-Treiber, der vom VMS konfiguriert und gesteuert werden kann und auf Anforderung mehrere PPP-Verbindungen über ISDN bereitstellt. All dies ist bei der Auswahl des VMS zu beachten.

#### Native ISDN

Durch direkten Zugriff auf die CAPI-Schnittstelle lässt sich der Datenfluss besser kontrollieren und erreicht eine geringere Latenz. Deshalb versenden einige Videogeräte die Bilder mittels proprietärer Protokolle direkt über ISDN. Der Preis dafür ist, dass die entsprechenden Schnittstellen und Protokolle im VMS verfügbar sein müssen. Dies ist vor dem Einsatz solcher Geräte zu prüfen.

#### Datensicherheit

Mit dem Thema Netzwerk geht der Begriff Datensicherheit eng einher: Sollen über WAN entfernte Bildquellen aufgeschaltet werden, sind Verbindungen nach außen nötig; damit besteht aber prinzipiell auch das Risiko von unberechtigten Zugriffen.

Planung, Installation und Administration eines Netzwerks erfordert daher ein hohes Maß an Kompetenz und Zuverlässigkeit. Mögliche Sicherheitslücken sollte man rechtzeitig erkennen und schließen. Grundlegende Regeln und Empfehlungen zur Sicherheit in digitalen Netzen bietet das Bundesamt für Sicherheit in der Informationstechnik (BSI) unter [www.bsi.de](http://www.bsi.de).

Eine erste Maßnahme zur Erhöhung der Datensicherheit kann sein, für die Videoüberwachung ein separates Netzwerk einzurichten. Dies erfordert nicht unbedingt ein eigenes Leitungsnetz, sondern es kann auch in Form eines Virtual Local Area Network (VLAN) realisiert werden: Video und Geschäftsdaten bleiben dann sauber getrennt, selbst bei Übertragung über die gleichen Leitungen.

Bei Verbindungen über das Internet empfiehlt sich ein Virtual Private Network (VPN): Eine Verschlüsselung schützt die Daten und sie lassen sich somit auf ihrem Weg durch das Internet nicht von Unbefugten abgreifen sowie unbemerkt manipulieren. An jedem Netzwerkzugang sollte eine Firewall selbstverständ-

### STICHWORT LAN – WAN

Von einem »Local Area Network« (LAN) spricht man, wenn es nur um kurze Entfernungen geht, in der Regel innerhalb eines Grundstücks.

Typische Übertragungstechnologien im LAN sind:

- Ethernet
- WLAN

»Wide Area Network« (WAN) steht für Datenetze, die große Entfernungen überbrücken. Das WAN wird in der Regel von »Providern« (z.B. der Deut-

schen Telekom AG) bereitgestellt. Technologien für WAN sind

- ISDN
- DSL
- GSM (GPRS, EDGE)
- UMTS (HSDPA, HSUPA).

Die Verbindung zwischen LAN und WAN stellt man über »Router« her: Kann eine Adresse innerhalb eines Netzes nicht gefunden werden, so vermittelt der Router die Verbindung automatisch in das nächste Netz.

lich sein, die nur solche Verbindungen zulässt, die vom Betreiber der Anlage ausdrücklich erwünscht sind.

Das VMS und die Bildquellen sollten Zugriffe nur nach Authentifizierung zulassen, also wenn ein gültiger Benutzername und das zugehörige Passwort eingegeben wurden. Unter Datenschutzgesichtspunkten kann auch das Vieraugenprinzip zum Tragen kommen, bei dem zwei verschiedene Passworte abgefragt werden, bevor man Zugang zu vertraulichen Daten erhält. Die Authentifizierungsverfahren der Bildquellen unterscheiden sich zum Teil sowohl im Ablauf als auch im Sicherheitsniveau erheblich.

### Bildübertragung, Datenreduktion

Wie bereits erläutert, erfordern die Videobilder zum effizienten Übertragen und Speichern eine Komprimierung. Eines der ältesten und am weitesten verbreiteten Kompressionsverfahren für Videobilder ist JPEG. Bilder in diesem Format lassen sich auf allen PCs und von sehr vielen Programmen direkt anzeigen und verarbeiten.

Im Laufe der Zeit wurde eine Vielzahl weiterer Kompressionsverfahren entwickelt, die vor allem eine höhere Datenreduktion zum Ziel hatten. Die Hersteller stehen im Wettbewerb darum, wer mit seinen Geräten bei begrenzter Bandbreite die bestmögliche Bildqualität und Bildfrequenz erzielt. Daher bringen sie immer neue Verfahren auf den Markt.

Es gibt nur wenige Video-Management-Systeme, die Bilder in allen marktüblichen Verfahren empfangen und weiterverarbeiten können. Denn es ist ein erheblicher Entwicklungsaufwand nötig, um alle diese Verfahren zu berücksichtigen und es kommen laufend neue hinzu.

Viele Hersteller übernehmen Kompressionsverfahren aus der Unterhaltungselektronik. Da bei der Videoüberwachung aber andere Anforderungen bestehen als etwa beim Fernsehen oder bei Spielfilmen, muss man diese Verfahren gegebenenfalls anpassen oder zumindest geeignet parametrieren (siehe Kasten CBR – VBR).

Bei begrenzter Netzwerkbandbreite ist es außerdem von Vorteil, wenn die Bildquelle je nach Anforderung verschiedene Bildfrequenzen und Bildauflösungen liefern kann: Bei Kameraschwenks oder schnellen Bewegungen wird auf eine reduzierte Auflösung

umgeschaltet, um eine schnellere Bildfolge zu erreichen; wenn es auf Details ankommt (Kennzeichen, Gesichter), wird zugunsten einer besseren Auflösung die Frequenz reduziert.

Viele der neuen Kompressionsverfahren geben nur einen Rahmen vor, innerhalb dessen unterschiedliche Umsetzungen (Implementierungen) möglich sind: So können sich, wo MPEG-4 draufsteht, in der Praxis durchaus je nach Hersteller abhängig von der ver-

fügbaren Rechenleistung und den konkret implementierten Algorithmen sehr unterschiedliche Eigenschaften hinsichtlich Bildqualität, Kompressionsrate und Bildfrequenz ergeben. Das macht es dem Errichter und Betreiber nicht gerade leicht, die jeweils passende Auswahl zu treffen (siehe Kasten GOP-Struktur).

(Fortsetzung folgt)

Hardo Naumann, Entwicklungsleiter der Accellence Technologies GmbH

## STICHWORT CBR – VBR

Wenn Kompressionsverfahren aus anderen Branchen für die Videoüberwachung genutzt werden sollen, sind einige Besonderheiten zu beachten:

Bei digitalem Fernsehen wird eine konstante Datenübertragungsrate (constant bit rate: CBR) angestrebt, da der Übertragungskanal (TV-Sender) eine fest definierte Bandbreite hat, die vollständig zur Übertragung der Bilder genutzt werden soll. Die Bilder des Fernsehprogramms ändern sich häufig, und die Zuschauer akzeptieren meist, dass bei besonders schnellen Bewegungen das Bild kurzzeitig etwas unschärfer wird.

Bei Videoüberwachung passiert dagegen in der meisten Zeit nichts; man könnte also zur Übertragung der sich kaum verändernden Bilder, die wenig neue Informationen enthalten, mit einer weit geringeren Übertragungsrate auskommen. Nur genau dann, wenn der Täter für kurze Zeit durch das Bild läuft, in der Szene also viel Bewegung stattfin-

det und die Informationsmenge, die für eine detaillierte Darstellung dieser Bilder nötig ist, sehr groß wird, soll die Übertragungsrate sprunghaft ansteigen: Die Bilder, auf denen sich der Täter durch das Objekt bewegt, sollen mit der größtmöglichen Detailtreue (also mit erhöhter Datenrate) übertragen werden. Bei Videoüberwachungsanlagen, bei denen sich die Videoströme mehrerer Kameras die verfügbare Bandbreite eines Netzes teilen, ist also nicht CBR, sondern VBR (variable bit rate) optimal.

Bei modernen Videokompressionsverfahren wie MPEG4 und H.264 kann konfiguriert werden, ob sie mit CBR oder VBR arbeiten. Der Errichter muss nur darauf achten, dass die von ihm verwendeten Geräte VBR unterstützen und dass er dies korrekt konfiguriert, weil die Anlage sonst ausgerechnet in den kritischen Momenten nur unscharfe Bilder liefert: dann gibt es kein brauchbares Fahndungsbild...

## STICHWORT GOP-STRUKTUR

Im Consumer-Bereich geht es darum, Spielfilme mit möglichst wenigen Daten zu speichern. Dazu wird ausgenutzt, dass zeitlich aufeinander folgende Bilder sich oft sehr ähnlich sind. Man kann viel Speicherplatz sparen, indem man nur die Differenz der Bildinformation zu vorangegangenen oder nachfolgenden Bildern speichert.

Die zu komprimierenden Einzelbilder können dazu in folgenden »Frame«-Typen gespeichert werden:

- I: enthält ein Bild vollständig
- P: enthält nur die Differenz zum vorangehenden I- oder P-Frame.
- B: Differenz zu vorangehendem und nachfolgendem I- oder P-Frame.

Die sich wiederholende Abfolge dieser Frames wird als GOP-Struktur bezeichnet. GOP steht als Abkürzung für englisch »group of pictures«. Sie bestimmt

maßgeblich die Eigenschaften eines Kompressionsverfahrens.

Bei JPEG gibt es nur I-Frames: Jeder Frame enthält die volle Bildinformation und kann auch dann korrekt dargestellt werden, wenn vorangegangene Frames gelöscht oder gestört wurden.

Bei Videokompression im Consumer-Bereich werden auch B-Frames intensiv genutzt, denn damit lassen sich die höchsten Kompressionsraten erzielen. Dies bewirkt allerdings auch eine prinzipielle Latenz, weil erst alle Bilder bis zum nächsten I- oder P-Frame abgewartet werden müssen, bevor die dazwischenliegenden B-Frames berechnet werden können.

Bei CCTV-Anwendungen mit Echtzeitanforderungen sollten daher keine B-Frames verwendet werden.