

TechTalk 85

Digital-ATV – Signal-Verzögerung bei der MPEG-2 Übertragung

von Ken Konechy W6HHC
und Hans Hass DC8UE

In TechTalk 83 wurde unter anderem über die Signal-Verzögerung zwischen Kamera und Wiedergabe-Display bei einer digitalen ATV-Übertragung berichtet.

Während einer analogen unkomprimierten ATV-Sendung können die Bilder der Video-Kamera sofort gesendet und wieder empfangen werden (wirklich "in Echt-Zeit", ohne Verzögerungen). Bei der Übertragung von digitalen ATV-Signalen fällt dagegen sofort auf, dass das wiedergegebene Fernsehbild gegenüber einer analogen Übermittlung eine deutliche Verzögerung aufweist. Es soll hier der Versuch unternommen werden, die Ursache dieser Latenzzeit zu ergründen.

Vorbereitungen des analogen Video-Signals

Während bei einer analogen (unkomprimierten) Übertragung das aufgenommene Kamerabild sofort gesendet werden kann, muss bei einer digitalen (komprimierten) Sendung das Bildsignal zunächst einen aufwendigen Bearbeitungsprozess durchlaufen. Wenn die Kamera als Ausgangs-Signal ein analoges Composite-Signal (FBAS) abliefern muss, muss dieses in einem ersten Bearbeitungs-Schritt in seine Komponenten zerlegt werden. Hierbei entstehen die Komponenten-Signale Y (Helligkeits-Signal), R-Y (Farbdifferenz-Signal Rot) und B-Y (Farbdifferenz-Signal Blau). Danach werden diese Signale abgetastet und damit in einen digitalen Datenstrom mit 216 Mbit/sec überführt.

Diese 216 Mbit entstehen bei einer Digitalisierung mit 8 Bit Auflösung. Dabei wird das 5 MHz breite Y-Signal mit einer Abtastfrequenz von 13,5 MHz abgetastet und erzeugt somit $13,5 \times 8 = 108$ Mbit/ für das Helligkeits-Signal. Die Farbdifferenz-Signale R-Y und B-Y werden mit der halben Luminanz-Abtastfrequenz von 6,75 MHz gesampelt und ergeben daher je $6,75 \times 8 = 54$ Mbit/s pro Farbart-Signal.

Das gesamte Bildsignal umfasst daher 108 Mbit/s (Y) + 54 Mbit/s (R-Y) + 54 Mbit/s (B-Y), also eine Gesamtdatenrate von 216 Mbit/s. Bei der 10Bit-Digitalisierung (nach CCIR 601) in einem professionellen Studio entstehen sogar 270 Mbit/s..

Im Eingang eines MPEG-Encoders muss also eine Wandlung abhängig vom angelieferten Eingangssignal vorgenommen werden. Als Eingangssignal sind dabei FBAS, Y/C, Y/U/V, oder CCIR601 (Parallel Digital) möglich. Diese müssen in ein Component-Signal mit der gewählten Auflösung D1, HD1 oder SIF umgewandelt werden. Bei einer Software-Lösung muss dazu das Signal zunächst in den Speicher eingelesen werden.

Dann kann der Prozessor-Teil das Ausgangs-Format berechnen und anschließend muss das umgerechnete Bild wieder ausgegeben werden. Schon dieser Bearbeitungsschritt verbraucht Zeit, da schon das Einlesen eines Bildes nur in Echtzeit erfolgen kann. Dazu kommt die anschließende Rechenzeit und der Signaltransport zu den nächsten Bearbeitungs-Stufen. Eine Hardware-Lösung ist dabei schneller, aber nicht so flexibel und auch aufwendiger und teurer.

MPEG-2 Kompression

Bei der (MPEG-) Komprimierung macht man sich die Tatsache zu nutze, dass in einem Fernsehbild genügend Redundanzen vorhanden sind, die eine regelmäßige Wiederholung des gesamten Bildinhaltes unnötig machen. Der MPEG-Encoder erzeugt dabei drei unterschiedliche Arten von Bildern:

I-Frame

Dieses Typ eines Video-Frames enthält alle zur Darstellung des Bildes notwendigen Informationen, um das gesamte Bild darzustellen (JPEG-ähnliche Kompression).

B-Frame

Dieser Kompressions-Typ enthält nur Informationen, um die Änderungen zwischen dem umgebenden I-Frame und dem P-Frame zu beschreiben.

P-Frame

Dieser Kompressions-Typ enthält nur Änderungen bezogen auf das vorangegangene I-Frame oder vorangegangenen P-Frame.

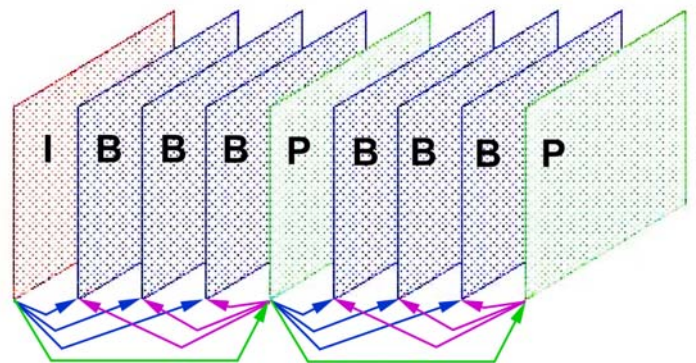


Bild 1 – Reihenfolge der komprimierten Frames (im IBBBP GOP Mode)

grüner Pfeil: vorwärts Vorhersage eines P-Frames
blauer Pfeil: vorwärts Vorhersage eines B-Frames
roter Pfeil: rückwärts Vorhersage eines B-Frames

B-Frames erreichen ihre hohe Kompression durch die Verwendung der Veränderungs-Informationen der umgebenden I-Frames und P-Frames. Dabei verwendet der I-Frame-only-Mode eine "Lauf-Längen"-Kompression, ähnlich wie sie bei der JPEG Kompression benutzt wird. Einige Details zur B-Frame-Kompression, basierend auf den Pixel-Veränderungen zwischen vorangehendem I-Frame und folgendem P-Frame, zeigt **Bild 2**.

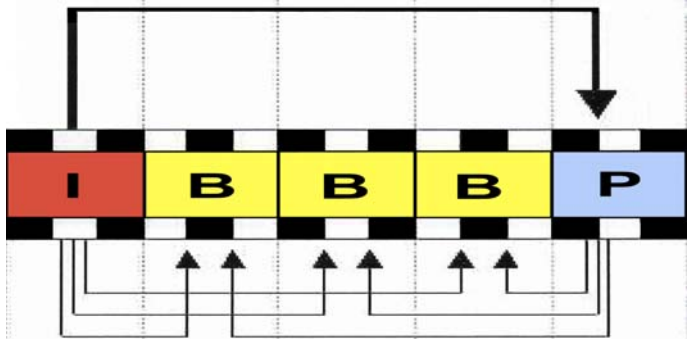


Bild 2 - Details der Vorhersage-Verknüpfung von I, B und P Frames

Der Empfänger kann dabei aus einem I-Frame ein vollständiges Bild erzeugen. Um allerdings aus einem B-Frame ein vollständiges Bild zu rekonstruieren, müssen beide Informationen aus I-Frame und nachfolgendem P-Frame vorliegen. Um dies zu gewährleisten, werden die Frames nicht in kontinuierlicher Reihenfolge gesendet. So folgt das P-Frame unmittelbar auf das I-Frame, damit es im Empfänger bereits vorliegt, wenn das B-Frame übertragen wird.

Wie **Bild 4** zeigt, wird dabei die Signalfolge am Encoder-Ausgang im Vergleich mit dem Encoder-

Eingang und der natürlichen Kamera-basierten Reihenfolge vertauscht. Daher muss der Encoder das Signal solange verzögern, bis die notwendigen Informationen vorliegen und berechnet sind.

Damit diese, in **Bild 4** gezeigte "Sende-Reihenfolge" möglich wird, muss der Encoder mit der Aussendung also solange warten, bis das nächste P-Frame digitalisiert und die Differenz-Information berechnet ist. Da ein Frame (bei PAL) aber jeweils 40 ms dauert, ist damit ein wesentlicher Zeitfaktor alleine durch das Warten begründet. Die Wartezeit in **Bild 4** beträgt minimal vier Frames zuzüglich den anderen Verarbeitungs-Verzögerungen. Die absolute Größe hängt dabei von der GOP (Group of Pictures = Anzahl der Bilder zwischen zwei I-Frames) und der Frametyp-Reihenfolge ab. Bei MPEG-2 ist eine GOP von 12 typisch.

Group of Picture (GOP) Einstellungen

Es können aber auch andere Reihenfolgen von B- und P-Frames gesendet und damit die Verzögerungszeiten verändert werden. Die Anzahl und Verteilung der I, B, und P Frames wird Group of Picture (GOP) genannt. Ein Beispiel der verschiedenen verwendbaren GOP Modes sind:

- I Mode
- IP Mode
- IBBP Mode (siehe Beispiel in Bild 5)
- IBBBP Mode (siehe Beispiel in Bild 1)

Bild 1 zeigt ein Beispiel der Reihenfolge einer im IBBBP-Mode encodierten Bilderfolge. **Bild 5** zeigt die GOP-Folge einer kodierten Sequenz im IBBP-Mode.

1	2	3	4	5	6	7	8	9	10	11	12	13
I	B	B	B	P	B	B	B	P	B	B	B	I

Bild 3 – Encoder in der Kamera-basierte Reihenfolge der komprimierten Bilder

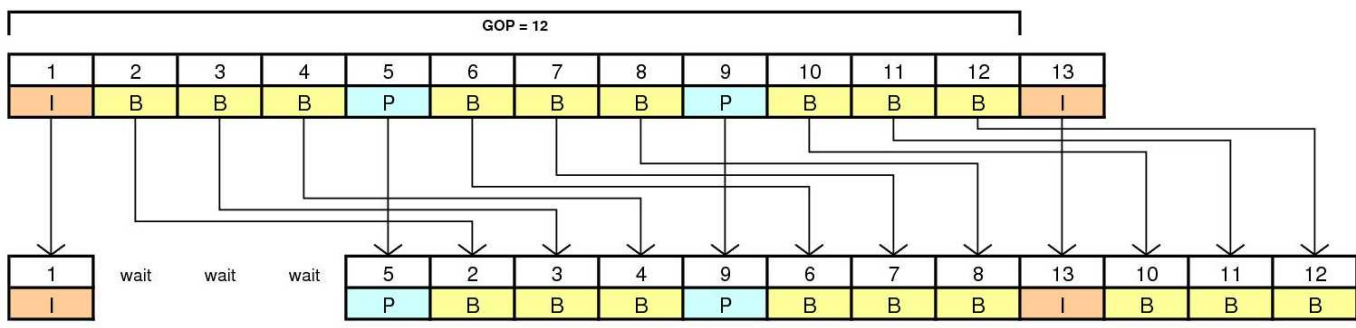


Bild 4 – Encoder-Ausgang mit den Bildern in der Sendereihenfolge

Die GOP-Betriebs-Einstellung IBBP ist die typische Grundeinstellung in den SR-System D-ATV Boards.

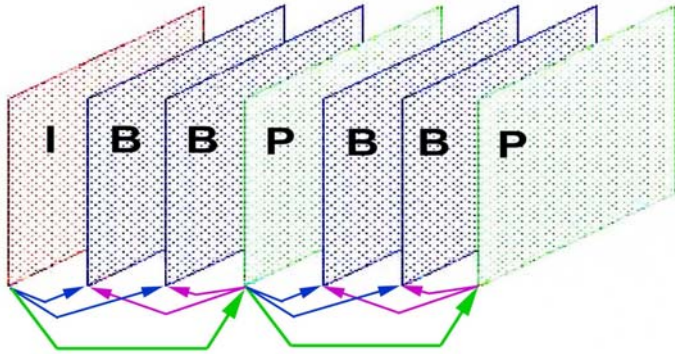


Bild 5 – Verknüpfung der encodierten Bilder im GOP Mode IBBP

Die geringste Verzögerung entsteht, wenn ausschließlich I-Frames (I-Frame only) gesendet werden. Leider ist damit auch nur eine geringere Kompression erreichbar und der Bandbreiten-Bedarf steigt erheblich. Stefan (DG8FAC) von SR-Systems erklärte dazu: "wenn die Betriebsweise GOP Mode "I" gewählt wird, hat man zwar die geringste Verzögerungszeit... aber man erzeugt bei einer D1-Auflösung eine Nutzdatenrate von rund 10Mbps (eine mindestens vier mal größere Datenrate und damit auch breitere HF-Bandbreite als im IBBP-Mode)"!! Da B-Frames die geringste Datenmenge beinhalten, reduziert eine große B-Frame-Anzahl die Bandbreite erheblich, erhöht aber auch die Wartezeit bis zum nächsten P-Frame. Die anschließende Aufspaltung der Signale in den I- und Q-Anteil und die Berechnung des Fehlerschutzes FEC erfordern einen weiteren Bearbeitungs-Schritt. Die abschließende QPSK-Modulation gehen dann sehr schnell.

Speziell bei professionellen Broadcast-Encodern gibt es verschiedenen Betriebs-Modi. Man kann wählen zwischen „Low Delay“, „Very Low Delay“, „Normal“ und verschiedenen „Seamless Modes“. Da in den „Low-Delay“-Betriebsarten keine B-Frames verarbeitet werden, ist die Geschwindigkeit auf Kosten von Datenrate und Bildqualität sehr schnell. In den „Seamless“-Modes ist eine unterbrechungsfrei Veränderung der Datenraten möglich, allerdings nur durch die Auswertung mehrerer aufeinanderfolgender I-Frames und das kann die Aussendung um mehrere Sekunden verzögern. Lediglich im „Normal“-Mode wird die Verzögerung auf das von der Frame-Folge abhängige Minimum reduziert.

Für die Verzögerung sind daher im Encoder folgende Einstellungen verantwortlich:

- Art des Eingangs-Signals
- Wahl der Video-Auflösung
- GOP (Anzahl der Bilder)
- GOP Frametyp-Folge

Dabei sind die Frametyp-Folge und GOP (Group of Pictures) die entscheidenden Faktoren, die bei Standard-Einstellung für eine Verzögerung von etwas über einer halben Sekunde verantwortlich sind.

SetTopBox Dekoder Verzögerung

Die Verzögerung auf der Empfängerseite (STB, SetTopBox) ist nicht ganz so gravierend. Da die Bildfolge für den Decoder schon in passender Reihenfolge empfangen wird, kann dieser sofort mit der Bearbeitung beginnen. Gute Empfänger speichern allerdings zunächst das Bild und geben es erst verzögert aus. Dies hat mehrere Gründe. So können gestört empfangene Bildteile (üblicherweise Pixel-Blöcke, gruppiert in Form von 8x8 Pixeln) durch entsprechende Teile des nachfolgenden Bildes ersetzt werden und somit ein scheinbar ungestörtes Bild ausgegeben werden. Stefan (DG8FAC) von SR-Systems hat viele STB-Empfänger getestet und berichtet, dass Standard SetTopBoxen eine typische Verzögerung von 7 – 14 Frames aufweisen. Dies ergibt bei PAL eine Zeitspanne von 0,28 bis 0,56 Sekunden für einen typischen STB-Empfänger.

Bei einer digitalen (paketweisen) Übertragung kommen die Bild- und Ton-Pakete nicht gleichzeitig sondern nacheinander beim Empfänger an. Um trotzdem eine lippensynchrone Darstellung zu ermöglichen, werden alle Pakete sendeseitig mit sogenannten „Timestamps“ versehen. Im Empfänger werden diese Informationen ausgewertet und durch den „Presentation-Timestamp“ gesteuert ausgegeben. Deshalb muss die Ausgabe der Signale solange zurückgehalten und zwischengespeichert werden, bis alle Signale zur gemeinsamen Ausgabe verfügbar sind.

Wird das Bildsignal mit anderer Auflösung als der gewünschte Fernsehstandard übertragen, muss Ausgangsseitig noch eine Format-Konvertierung vorgenommen werden. Sowohl die Umrechnung (von z.B. SIF auf D1 oder von 16:9 auf 4:3) erfordert wieder einen Bearbeitungs-Schritt, der mit einer zeitlichen Verzögerung verbunden ist.

Wird das Signal dann noch z.B. über ein PC-Display wiedergegeben, so ist neben der erneuten Digitalisierung eine weiter Umrechnung auf die Pixelzahl der gewünschten Display-Ausschnitts erforderlich. Wenn jetzt noch eine Anpassung an die Bild-Wiederholfrequenz des Monitors notwendig wird (z.B. von 50 auf 60 oder 75 Hz), erfordert dies durch die notwendige Interpolation nochmals eine erhebliche Bearbeitungszeit.

Abschließend muss man feststellen, dass die Übertragung von TV-Signalen im MPG-2-Standard zwar eine Verringerung der HF-Bandbreite gegenüber einer analogen Ausstrahlung ermöglicht, dies aber durch eine deutliche Zunahme der Latenz-Zeit erkauft wird. Der maßgebliche Faktor ist hierbei nicht die Digitalisierung sondern die Signal-Kompression (auf die benötigte oder gewünschte Netto-Daten-Rate).

Es ist auch bei digitalen MPEG-2-Sendungen ständig abzuwägen, ob eine schnelle breitbandige oder schmalbandige und langsame Übertragung gewünscht wird. Das bestimmt die Wahl und die Einstellung des Kompressions-Verfahrens. Es gibt keine schmalbandige D-ATV-Übertragung ohne Verzögerung.

Vermeidung zusätzlicher Relais MPEG Laufzeiten

Mit der Einführung von D-ATV wird möglicherweise eine weiter Verzögerung im Relais selbst erzeugt. Hier ist eine Liste der möglichen MPEG-Verzögerungen auf dem Weg von einer Heim-Station über das Relais bis zur anderen Heim-Station:

- 1) Sender in der D-ATV Heim-Station – MPEG Encoder
- 2) Relais-seitiger STB - Empfänger – MPEG Decoder
- 3) Relais- Re-Encoding für den Sender– MPEG Encoder
- 4) STB-Empfänger in der Heim-Station – MPEG Decoder

Bei SR-Systemen verfügt man über ein interessantes D-ATV-Produkt, genannt NIM Tuner Board (siehe **Bild 6**). Der NIM Tuner kann auf die typischen D-ATV Mikrowellen-Frequenzen (zum Beispiel 1,2 GHz und 2,4 GHz) abgestimmt werden und liefert am Ausgang einen Transport-Strom (TS) direkt an ein Flachbandkabel, das sofort mit der TS-Stream-Eingangsbuchse des Modulators verbunden werden kann. Der Flachkabel-Anschluss auf der rechten Seite von **Bild 6** liefert das TS-Ausgangs-Signal. Dieser Ansatz vermeidet den MPEG-2-Decodierungs-Schritt und die erneute Re-Codierung am Umsetzer-Standort.

In einer Gegenüberstellung vergleicht **Bild 7** die Verzögerung bei der Verwendung einer STB im Vergleich zu einem NIM-Tuner am Relais-Standort.

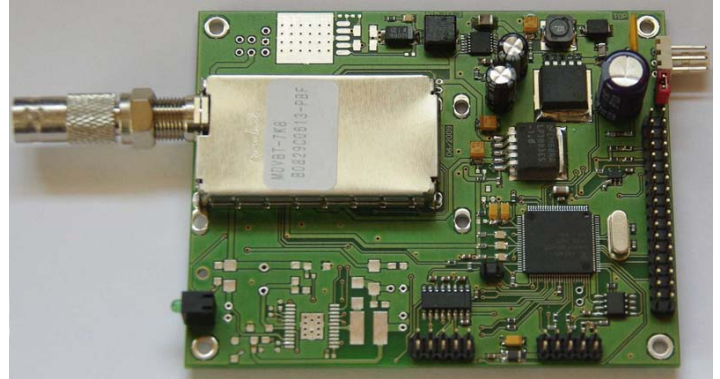


Bild 6 – NIM Tuner- Board mit TransportStream (TS) -Ausgang.

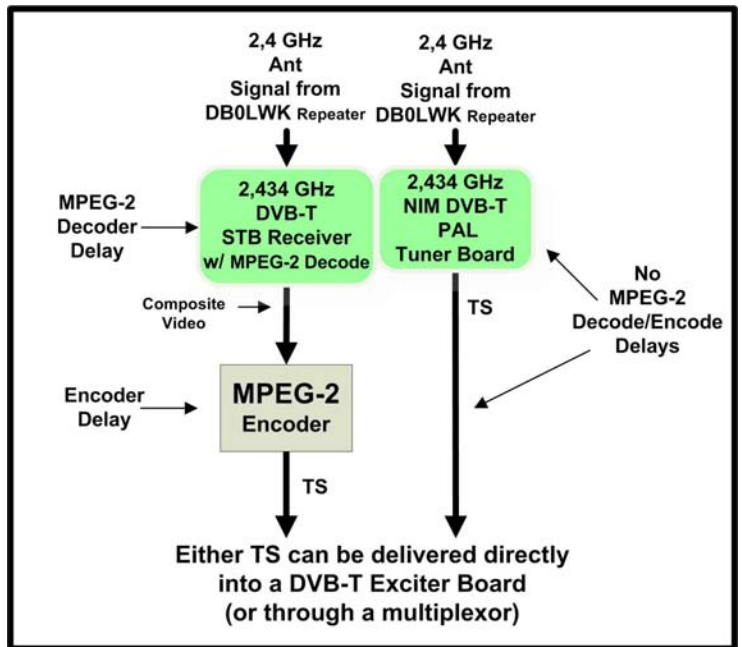


Bild 7 - Vergleich NIM Tuner mit einer STB. NIM vermeidet eine Addition der Decoder/Encoder Verzögerung.

Zusammenfassung

Wir hegen die Hoffnung, dass der Leser nun ein wenig die Komplexität des MPEG-2-Bearbeitungsprozesses bei einer D-ATV-Übertragung sowie den Funktion und Zweck einiger variabler Parameter sowie deren Auswirkung bei veränderter Einstellung abschätzen kann. ...und warum sich bei schmalbandigen D-ATV-Übertragungen Verzögerungszeiten ergeben.

(siehe nützliche D-ATV Links auf der nächsten Seite)

Nützliche D-ATV Links

- Wikipedia / MPEG-2 – siehe <http://en.wikipedia.org/wiki/MPEG-2>
- British ATV Club - Digital Forum –siehe www.BATC.org.UK/forum/
- Portal für D-ATV Streaming Relais und Downloads – siehe www.D-ATV.net (in Deutsch)
- AGAF D-ATV Komponenten (Boards) – siehe www.datv-agaf.de und www.AGAF.de
- Lechner D-ATV Boards - <http://lechner-cctv.de/d-atv-dvb.151.de.html?mwdSID=9agn7phuiu46fvm2ok3aueltf3>
- Komplette betriebsfertige D-ATV-Sender – siehe www.d-atv.org/D-ATV-Modulator.pdf
- SR-Systems D-ATV Komponenten (Boards) – siehe www.SR-systems.de
- DGØVE Microwave-Verstärker, Up-Converter, Down-Converter – siehe www.DG0VE.de
- Kuhne Electronics (DB6NT) HF-Verstärker– siehe www.Kuhne-Electronic.de
- Orange County ARC Newsletter, umfangreiche D-ATV Artikel-Serie – siehe www.W6ZE.org/DATV/