TechTalk 85

Digital-ATV – Understanding the latency of a MPEG-2 transmission

by Ken Konechy W6HHC
& Hans Hass DC8UE

As already discussed in the earlier OCARC TechTalk83 article, there will be a signal delay (called latency) between the camera and the receiving display during a DATV transmission.

During an analog and uncompressed ATV-broadcast, the pictures of the video camera can be transmitted and then received immediately (really “real-time”, no delays). When transmitting Digital-ATV signals, you will see that the received television picture shows a delay on the displayed signal compared with an analog transmission. This TechTalk will try to explain the reason for this DATV delay time.

Analog Video Signal Preparations

When you plan digital (compressed) video broadcasting, the picture signal first has to run through a complicated digitalization process to prepare it for MPEG-2 compression. If the camera delivers an analog composite signal (CCVS) as a source signal, this must be disassembled in the first treatment step into the components. This will generate the component signals Y (brightness signal), R-Y (colour difference signal “red minus brightness”) and B-Y (colour difference signal “blue minus Y”). Then these signals have to be Analog-to-Digital converted and transferred into a digital data stream resulting in a bit-rate of 216 Mbit/sec.

These 216 Mbit/sec bit data-rate occurs during the video digitizing process with 8-bit resolution. The 5MHz analog Y-signal will be Analog-to-Digital converted with a sampling rate of 13.5 MHz and generates therefore 13.5 x 8 = 108 Mbps for the brightness signal. The colour difference signals R-Y and B-Y will be digitized with half of the luminance sampling frequency of 6.75 MHz and produce 6.75 x 8 = 54 Mbps per colour component signal.

Hence, the whole digitized video signal includes a data-bit-rate of 108 Mbps (Y) + 54 Mbps (R-Y) + 54 Mbps (B-Y), producing a total data-bit rate of 216 Mbps. In a professional TV studio, using a 10-bit digitization (after CCIR 601), the data-rate will increase to 270 Mbps.

For the input to an MPEG-2 encoder, a conversion must be done, depending on the incoming input signal. There are several input signals possible, like CCVS, Y/C, Y/U/V, or CCIR601 (parallel digital). These must be converted into a component signal with the chosen video resolution D1, HD1 or SIF. In a software solution the signal must be put into the memory in a first step.

Then a processor can calculate the source format and afterwards the converted picture must be read out of memory again. Already these two process steps creates some delays, because simply transferring the picture to the memory consumes time and therefore adds a small amount of delay. In addition to this comes the time required for the calculation and the signal transport. A hardware solution would be quicker, but is neither so flexible nor adaptable and is more expensive.

MPEG-2 Compression

The MPEG-2 compression uses the fact, that a television picture contains enough redundancy from frame-to-frame, to make a permanent repetition of the whole picture content unnecessary. In order to compress, the MPEG encoder generates three different kinds of encoded picture frames:

I-Frame

This type of video frame contains all the necessary information to display a whole picture frame (low JPEG-like compression).

B-Frame

This type of compressed video frame contains only information to describe the changes from the I-Frame and the P-Frame.

P-Frame

This type of compressed frame contains only changes since the preceding I-Frame or preceding P-Frame.

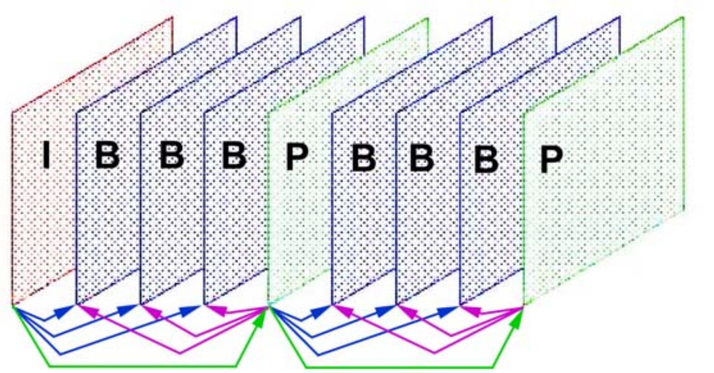

Figure 1 – A Sequence of Compressed Frames (MPEG-2 using GOP Mode called IBBBP)

green arrow: forward prediction of P-Frames
blue arrow: forward prediction of B-Frames
red arrow: backward prediction of B-Frames

B-frames obtain much compression by comparing differences to I-frames and P-frames. The I frame only uses "run-length" compression, similar to JPEG compression. Some details on the compression of a B-frame based on comparing pixel differences between a preceding I-frame and a following P-frame are shown in **Fig 2.**

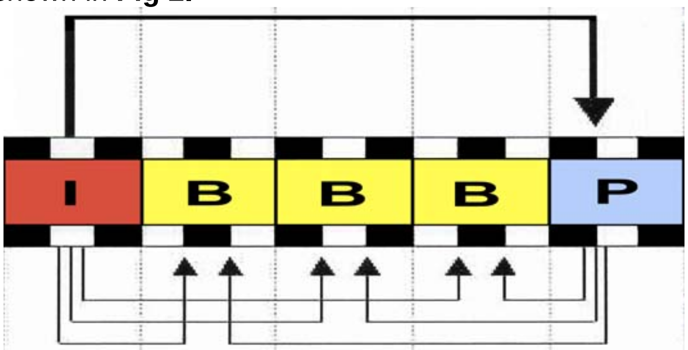

Figure 2 - Details of prediction order of I, B and P frames

The receiver can generate an entire picture frame by using only the I-frame information. To reconstruct a complete picture frame from a B-frame, information from both the I-frame and the following P-frame must be obtained and processed. To guarantee that there are no observed delays between frames while actually viewing the video, the frames are not sent in continuous order. Therefore the P-frame follows immediately after the I-frame. In this way the P-frame will be already available, when the receiver needs it to decode the incoming B-frames.

As **Fig 4** shows, the video frame sequence of the exit of the encoder is changed compared with the encoder exit sequence is changed from the encoder's "natural" camera-based frame sequence. The signal at the output of the encoder must be delayed until the necessary information is available and has been calculated.

To create the "broadcast order" sequence shown in **Fig 4**, the encoder must wait with the transmission, until the next P-Frame is digitized and the difference information is calculated. Because a frame (using NTSC) lasts 33 ms, an essential time factor delay is created by the waiting. In **Fig 4**, the wait is a minimum of four frame-periods, plus other processing delays. The absolute duration of the delay depends on the GOP (Group of Pictures) number of the pictures between two I-Frames (called the GOP Number) and the frame type sequence. In MPEG-2, a GOP Number of 12 is typical for ham radio transmissions.

Group of Picture (GOP) Settings

However, other orders of B-and P-frames can be also transmitted and for that reason the delay times inside the encoder can change. The term for the sequence of I, B, and P frames used is called the GOP Mode. A sample of different GOP Modes that can be used are:

- I mode
- IP mode
- IBBP mode (see example in Fig 5)
- IBBBP mode (see example in Fig 1)

Fig 1 showed an example of the encoded frames produced by the IBBBP GOP Mode. **Fig 5** shows the sequence of encoded frames produced by the GOP

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

Figure 3 - Encoder camera-based sequence of compressed frames

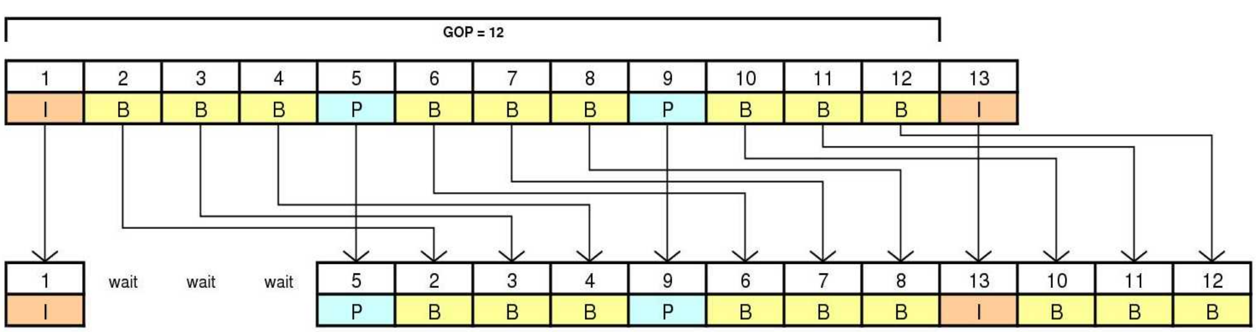

Figure 4 - **Encoder exit frame sequence in broadcasting order**

Mode, IBBP. Typically, the GOP mode of IBBP is the default setting used in SR-System DATV boards.

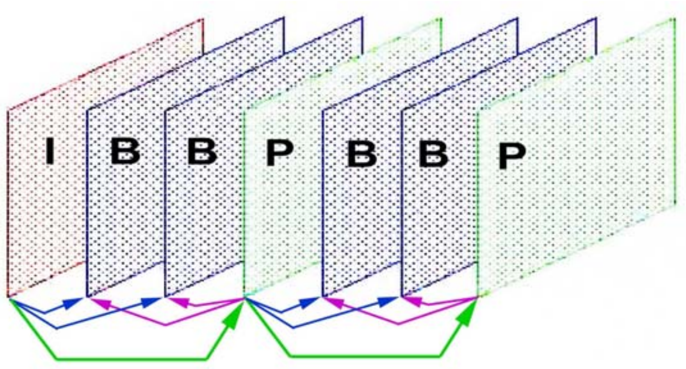

Figure 5 - Sequence of Encoded Frames using GOP Mode IBBP

The minimum delay originates if exclusively I-frames (that is: I-frame only) are transmitted. Unfortunately, only a very low compression is available with this GOP mode and the needed RF bandwidth increases considerably. As Stefan DG8FAC of SR-Systems explains: “if you enable the GOP Mode "I", then you will have the lowest delay…but you create a payload net-data-bit-rate of around 10Mbps in D1 resolution (at least four times larger data rate and therefore wider RF bandwidth than with the IBBP mode)”!! Because the B-frames contain the smallest amount of data (greatest amount of compression), a large number of B-frames reduces the bandwidth considerably, however, raises also the waiting period up to the next P-frame. Next, the separation of the signals in the I- and Q-portion and then the calculation of the error protection (FEC) require two more processing steps. The final QPSK modulation then runs very fast.

Especially with professional broadcast-encoder, there are different operation modes available. You can choose between "low delay", "very low delay", "normal" and different "seamless modes". While in the "low-delay" operation mode no B-frames are processed, the speed is very fast, at the expenses of the data rate and image quality. In the "seamless"-modes a free of interruption change of the data rates is possible, indeed, only by the evaluation of several successive I-Frames and therefore the transmission can delay around several seconds. Merely in the "normal" mode, the delay depends to the necessary minimum by the frame order.

The following parameters and settings will have effect to the encoder delay:

- what kind of the input signal
- choice of the video-resolution
- GOP Mode (mix of frame types)
- GOP Number

The mix of frame types and GOP (Group Of Pictures) are the determining factors which are responsible for the delay from something about half a second inside the MPEG-2 encoder.

SetTopBox Decoder Delays

The delay on the receiver side (SetTopBox, aka STB,) is not so serious. Because the picture sequence will be received for the decoder already in suitable order, therefore the decoding can begin immediately with the treatment. However, good receivers will first store first the picture and display it only delayed. This has several reasons. So received and corrupted picture parts (usually done at the “pixel block” level….groupings of 8x8 pixels) can be replaced by suitable parts of the following picture and an apparently undisturbed picture will be generated. Stefan DG8FAC of SR-Systems has tested many STB receivers and reports that standard STBs exhibit around 7-14 frames of delay. In NTSC, this equates to a latency ranging from about 0.25 seconds to almost 0.5 seconds, in typical STB receivers.

During a digital (packetized) transmission the picture packets and sound packets will not arrive at the same time on the receiver site. To ensure, nevertheless, a lip-synchronous presentation, all packages will get so-called "timestamps” during the encoding process. In the receiver this information will be evaluated and the video output will be controlled and steered by the "presentation timestamp". Therefore, the output of the signals must be held back until all signals are available together.

If the video signal is transmitted with other resolution than the desired television standard, a format conversion must be done on the video output side. The video resolution conversion (e.g. from SIF to D1 or from 16:9 to 4:3) requires again a signal processing and calculating, which produces one more delay.

If the signal will be displayed e.g. on a PC screen, then another set of digitizing and conversion of the pixel numbers into the format of the desired display is necessary. If yet also an adaptation to the picture refresh frequency of the monitor becomes necessary (e.g., from 50 to 60 or 75 hertz), then this video resolution adaption requires some necessary interpolation processing and again adds considerable delay time.

Finally, one will discover that the transmission of TV signals using the MPEG-2 standard allows a reduction of the RF bandwidth compared with an analog transmission, however, this narrow bandwidth is bought dearly by a significant increase of the latency time. The important factor that creates latency is not the digitization, but rather is the signal compression (that is: the reduction of net-data-bit-rate needed).

So it will be necessary during digital MPEG-2 broadcasts, to choose whether a low-latency and wider bandwidth scenario is desired…or a larger-latency and narrow bandwidth is desired. This choice determines the parameter settings of the MPEG-2 compression process. Narrow bandwidth DATV transmissions can not be achieved without incurring any delays.

Avoiding Additional Repeater MPEG Latency

When implementing a DATV, normally an additional set of MPEG delays can be incurred at the repeater site itself. Here is a list of the location of MPEG delays in a "normal" DATV transmission through a DATV repeater:

1) Transmitting DATV home station – MPEG Encoder
2) Repeater site STB Receiver – MPEG Decoder
3) Repeater site re-Encoding for Transmitter – MPEG Encoder
4) Receiving DATV home station STB – MPEG Decoder

SR-Systems offers an interesting DATV product called a NIM Tuner board (see **Fig 6**). The NIM Tuner can be tuned to the typical DATV microwave frequencies (for example 1.2 GHz and 2.4 GHz) and outputs the Transport Stream (TS) directly in a ribbon cable that can be connected directly to the TS input-connector of a DATV exciter. The ribbon connector on the right hand side in **Fig 6** is the TS output signal. This approach avoids the performing the MPEG-2 decode step and the re-encode step at a DATV repeater location. See **Fig 7** that compares the delays between using a STB at the repeater and using a NIM tuner at a repeater – side-by-side.

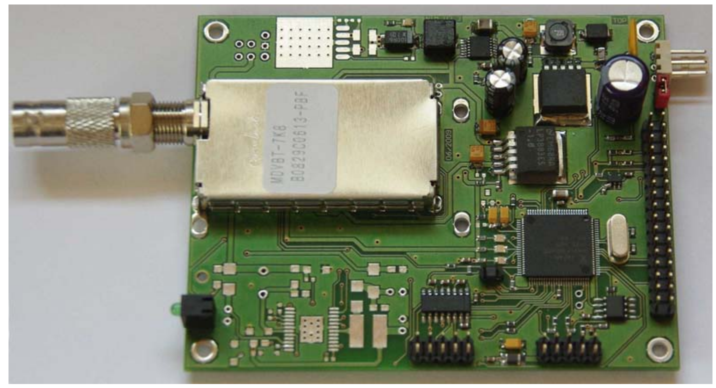

Figure 6 – NIM Tuner board with Transport Stream (TS) output.

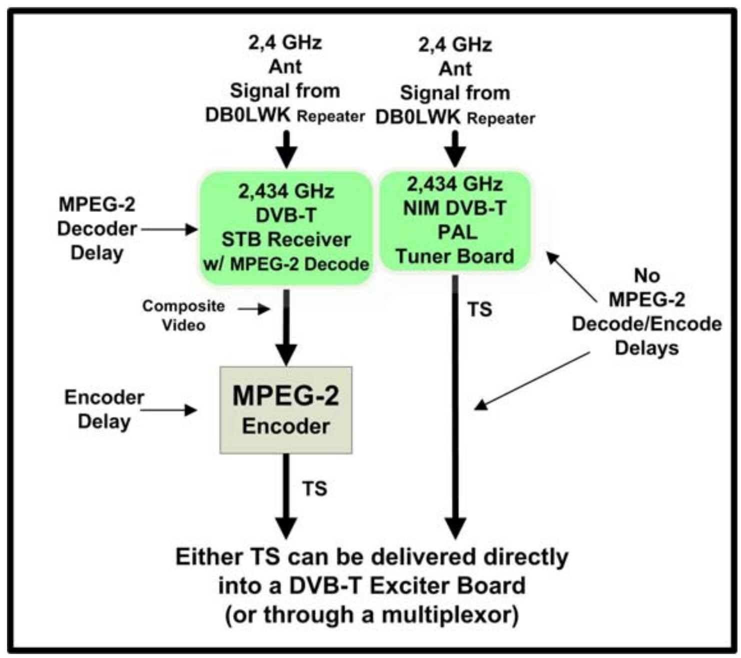

Figure 7 - Comparing NIM Tuner with STB.
NIM avoids any Decode/Encode delays being added.

Conclusion

It is our hope that the reader can now appreciate the processing and complexities involved with the MPEG-2 aspect of DATV. You should now also be able to recognize the purpose/function of some of the variable parameters used for MPEG-2 and the impact that can occur for various setting values…and why DATV narrow bandwidths incur latency.

(See Interesting DATV links on page 6)

Interesting DATV Links

- Wikipedia on MPEG-2 – see **http://en.wikipedia.org/wiki/MPEG-2**
- British ATV Club - Digital Forum – see **www.BATC.org.UK/forum/**
- German portal for DATV streaming repeaters and downloads – see **www.D-ATV.net** (in German)
- AGAF D-ATV components (Boards) – see **www.datv-agaf.de** and **www.AGAF.de**
- Lechner DATV Boards - **http://lechner-cctv.de/d-atv-dvb.151.de.html?mwdSID=9agn7phuiu46fvm2ok3aueltf3**
- Complete ready-to-go DATV transmitters – see **www.d-atv.org/D-ATV-Modulator.pdf**
- SR-Systems D-ATV components (Boards) – see **www.SR-systems.de**
- DGØVE microwave amps, up-converters, down-converters – see **www.DG0VE.de**
- Kuhne Electronics (DB6NT) RF Amplifiers – see **www.Kuhne-Electronic.de**
- Orange County ARC newsletter entire series of DATV articles – see **www.W6ZE.org/DATV/**