

Title: **An issue regarding to the B-picture coding**

Status: Input Document

Purpose: Proposal

Contact: Sherman (Xuemin) Chen

16215 Alton Parkway,
Irvine, CA 92619

Tel: +1 949 585 6185

Email: schen@broadcom.com

Source: Broadcom Corp

1. Video Decoding and Presentation Timestamps

In a video transport system, the system clock of a video program is usually used to create timestamps that indicate the presentation and decoding timing of video, as well as to create timestamps that indicate the instantaneous values of the system clock itself at sampled intervals. The timestamps that indicate the presentation time of video are called Presentation Time Stamps (PTS) while those that indicate the decoding time are called Decoding Time Stamps (DTS). It is the presence of these timestamps and the correct use of the timestamps that provide the facility to synchronize properly the operation of the decoding.

In MPEG-2 Systems [2], a MPEG-2 video elementary stream is assembled into a packetized elementary stream (PES). Presentation Time Stamps (PTS) are carried in headers of the PES. Decoding Time Stamps (DTS) are also carried in PES headers that have the picture header of an I- or P-picture when bi-directional predictive coding (B-picture – which cannot be used as a reference picture) is enabled. The DTS field is never sent with a video PES stream that was generated with B-picture coding disabled. The value for a component of PTS (and DTS, if present) is derived from the 90 kHz portion of the PCR that is assigned to the service to which the component belongs.

Both PTS and DTS are determined in video encoder for coded video pictures. If a stream only consists of I- or P-pictures, these pictures need not be delayed in the reorder buffer and the PTS and DTS are identical. This is, so called, the **low-delay** mode. If B-pictures are present in the video stream, coded pictures (sometime also called video access units) do not arrive at the decoder in presentation order. In this case, some pictures in the stream must be stored in a reorder buffer in the decoder after being decoded until their correct presentation time. In particular, I-pictures or P-pictures carried before B-pictures will be delayed in the reorder buffer

after being decoded. Any I- or P-picture previously stored in the reorder buffer is presented before the next I- or P-picture is stored. While the I- or P-picture is stored in the reorder buffer, any subsequent B-picture(s) is (are) decoded and presented. This can be called the **non low-delay** mode.

For MPEG-2 video [3], DTS indicates the time when the associated video picture is to be decoded while PTS indicates the time when the presentation unit decoded from the associated video picture is to be presented on the display. Times indicated by PTS and DTS are evaluated with respect to the current System Time Clock value -- locked to Program Clock Reference (PCR). For **B-pictures, which are never reordered**, PTS is always equal to DTS since these pictures are decoded and displayed instantaneously. For I- or P-pictures (if B-pictures are present), PTS and DTS differ by the time that the picture is delayed in the reorder buffer, which will always be a multiple of the nominal picture period, except in film mode. If B-pictures are not present in the video stream, i.e., B-picture type is disabled, all I- and P-pictures arrive in presentation order at the decoder, and consequently their PTS and DTS values are identical – i.e. the **low-delay** mode. Note that if the PTS and DTS values are identical for a given access unit, only the PTS should be sent in the PES header.

Example 1.

Consider a MPEG-2 video in which three B-picture are sent between each I- or P-picture, and pictures arrive at the decoder in the following decoding order:

I₀ P₄ B₁ B₂ B₃ P₈ B₅ B₆ B₇ I₁₂ B₉ B₁₀ B₁₁.....

These pictures will be reordered in the following presentation order:

I₀ B₁ B₂ B₃ P₄ B₅ B₆ B₇ P₈ B₉ B₁₀ B₁₁ I₁₂

Each time that the picture sync is active, the following picture information are usually required for time stamping of the picture:

- Picture type: I-, P-, or B-picture.
- Temporal Reference: A 10-bit count of pictures in the presentation order.
- Picture Sync Time Stamp (PSTS): A 33-bit value of the 90 kHz portion of the PCR that was latched by the picture sync.

In the normal video mode, the DTS for a given picture is calculated by adding a fixed delay time, T_d , to the PSTS. For some pictures in the film mode, the DTS is generated by $(T_d + \text{PSTS} - \text{a field time})$. T_d is nominally the delay from the input of the MPEG-2 video encoder to the output of the MPEG-2 video decoder. This delay is also called the end-to-end delay. In real applications, the exact value of T_d is most likely determined during system integration testing.

The position of the current picture in the final display order is determined by using the picture type (I, P or B). The number of pictures (if any) for which the current picture is delayed before presentation is used to calculate the PTS from the DTS. **If the current picture is a B-picture, then it is not delayed in the reorder buffer and the PTS and DTS are identical.** In this case, the PTS is sent usually in the PES header that precedes the start of the picture. If the current

picture is instead an I- or P-picture and the B-processing mode is enabled, then the picture will be delayed in the reordering buffer by the total display time required by the subsequent B-picture(s).

Example 2.

Consider the video stream given in Example 1. Assume that all pictures in the stream are coded as frame-structure pictures and in the non-film mode. Then one can compute DTS and PTS as follows.

Decoding order Index i	Encoded Pictures with a display order index	$PSTS_i$	$\Delta PSTS_i$	DTS_i	PTS_i
0	I_0	$PSTS_0$	F	$PSTS_0 + T_d$	$DTS_0 + F$
1	P_4	$PSTS_1$	F	$PSTS_1 + T_d$	$DTS_1 + 4F$
2	B_1	$PSTS_2$	F	$PSTS_2 + T_d$	DTS_2
3	B_2	$PSTS_3$	F	$PSTS_3 + T_d$	DTS_3
4	B_3	$PSTS_4$	F	$PSTS_4 + T_d$	DTS_4
5	P_8	$PSTS_5$	F	$PSTS_5 + T_d$	$DTS_5 + 4F$
6	B_5	$PSTS_6$	F	$PSTS_6 + T_d$	DTS_6
7	B_6	$PSTS_7$	F	$PSTS_7 + T_d$	DTS_7
8	B_7	$PSTS_8$	F	$PSTS_8 + T_d$	DTS_8
9	I_{12}	$PSTS_9$	F	$PSTS_9 + T_d$	$DTS_9 + 4F$
10	B_9	$PSTS_{10}$	F	$PSTS_{10} + T_d$	DTS_{10}
11	B_{10}	$PSTS_{11}$	F	$PSTS_{11} + T_d$	DTS_{11}
12	B_{11}	$PSTS_{12}$	F	$PSTS_{12} + T_d$	DTS_{12}
...	...	...	...	...	...

Where time interval F between PTS_i and PTS_{i-1} is exactly equal to the nominal picture time in 90 KHz clock cycles (i.e. $F = \frac{90 \times 10^3}{29.97} = 3003$ for NTSC since the picture rate equals 29.97, and

$$F = \frac{90 \times 10^3}{25} = 3600 \text{ for PAL since the picture rate equals 25}).$$

2. Problem

H.264 Stream with B pictures

In H.264 [1], the use of B pictures is indicated by the nal_unit_type. One key difference between H.264 B-picture and B-pictures in other standards is that some B-pictures in a H.264 stream can be used as a reference picture.

Example 3.

Consider a H.264 stream which is similar to the stream given in Example 1.

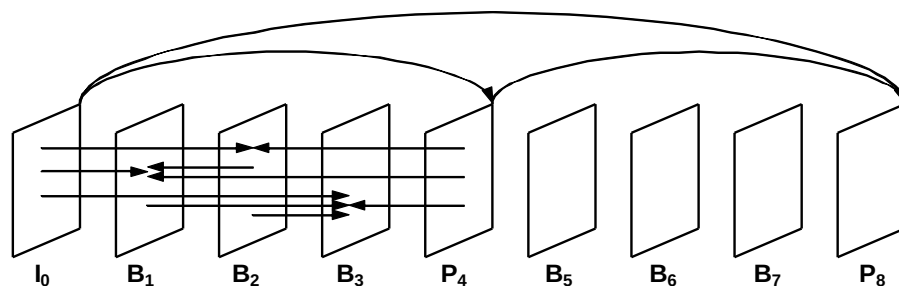


Figure 1 – Illustration of H.264 B picture concept

The location of pictures in the bitstream is in a data-dependence order rather than in temporal or display order. Pictures that are dependent on other pictures shall occur in the bitstream after the pictures on which they depend. Figure 1 shows a typical example, where three B pictures are inserted in display order between two successive P pictures. The P picture P₄ only depends on the first Intra picture I₀. The B picture B₂, which is temporally located between I₀ and P₄, depends on both of these pictures. The B picture B₁ depends on I₀, P₄, and B₂; the B picture B₃ additionally depends on B₁. The decoding order is given by

I₀ P₄ B₂ B₁ B₃

The display order for this example is given by

I₀ B₁ B₂ B₃ P₄

Example 3 shows that, beside I- and P- pictures, B-pictures in H.264 also need to be delayed in the reordering buffer, i.e. some B-pictures are also reordered. This implies that only use of the low-delay mode and the non low-delay mode is not sufficient for indicating such a delay and assisting calculation of time-stamp, e.g. PTS.

3. A Solution

Computation of PTS and DTS for H.264 Video

A delay parameter, which indicates the number of pictures being delayed, is needed for assisting computation of PTS. In H.264, both field and frame pictures are allowed and field pictures doesn't need to be transmitted in pairs. Thus, the unit of this delay parameter can be selected as a field. A frame picture is considered as a two-field unit. The delay parameter, namely **picture_delay_unit**, indicates maximum reordering delay in number of field-picture units. For example,

picture delay unit	configurations
0	Low delay
1	One field-picture delay
2	Two field-picture delay or one frame delay
.....	

To enhance coding efficiency, **picture_delay_unit** can be coded as a variable length code (VLC). The **picture_delay_unit** is up-bounded by number of short-term pictures (in field-picture unit), and also number of reference pictures, indicated by number_of_reference_pictures specified in H.264 Committee Draft, which defines the total number of short- and long-term picture buffers in the multi-frame buffer.

In order to make changes on **picture_delay_unit**, the coded video sequence needs to be terminated by a **sequence end** code.

Example 4

Now assume that all pictures in the stream discussed in Example 3 are frame-structured pictures. Then, we have **picture_delay_unit** = 4 and both DTS and PTS can be computed as follows.

Decoding order index i	Encoded pictures with a display order index	$PSTS_i$	$\Delta PSTS_i$	DTS_i	PTS_i
0	I_0	$PSTS_0$	F	$PSTS_0 + T_d$	$DTS_0 + 2F$
1	P_4	$PSTS_1$	F	$PSTS_1 + T_d$	$DTS_1 + 4F$
2	B_2	$PSTS_2$	F	$PSTS_2 + T_d$	$DTS_2 + 2F$
3	B_1	$PSTS_3$	F	$PSTS_3 + T_d$	DTS_3
4	B_3	$PSTS_4$	F	$PSTS_4 + T_d$	$DTS_4 + 2F$
...	...	...	...	...	...

Where F is the same as being defined in Example 2. This example is also shown in the following figure.

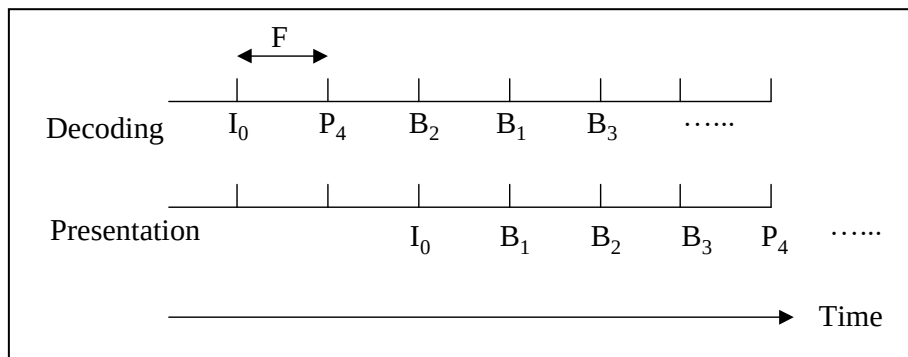


Figure 2 Time line for decoding and presentation for Example 4.

In general, at the beginning of the sequence, the PTS and DTS of the first picture are related as follows:

$$\text{PTS} = \text{DTS} + \text{picture_delay},$$

Where **picture_delay** = **picture_delay_unit** * (the field-picture period in 90 KHz clock cycles).

The following requirement or its equivalent should be included in the H.264 specification for clarification: “For two B-pictures with no data-dependence, transmission and decoding of these pictures shall follow temporal or display order.”

References

- [1] Committee Draft, JVT-C167, ITU-T Recommendation H.264 | ISO/IEC 14496-10 (MPEG-4 AVC), 2002.
- [2] ISO/IEC 13818-1 | ITU-T Recommendation H.222, (MPEG-2 Systems), 1995.
- [3] ISO/IEC 13818-2 | ITU-T Recommendation H.262, (MPEG-2 Video), 1995.

(Append for Proposal Documents)

JVT Patent Disclosure Form

International Telecommunication Union
Telecommunication Standardization Sector



International Organization for Standardization



International Electrotechnical Commission



Joint Video Coding Experts Group - *Patent Disclosure Form*

(Typically one per contribution and one per Standard | Recommendation)

Please send to:

JVT Rapporteur Gary Sullivan, Microsoft Corp., One Microsoft Way, Bldg. 9, Redmond WA 98052-6399, USA

Email (preferred): Gary.Sullivan@itu.int Fax: +1 425 706 7329 (+1 425 70MSFAX)

This form provides the ITU-T | ISO/IEC Joint Video Coding Experts Group (JVT) with information about the patent status of techniques used in or proposed for incorporation in a Recommendation | Standard. JVT requires that all technical contributions be accompanied with this form. *Anyone* with knowledge of any patent affecting the use of JVT work, of their own or of any other entity (“third parties”), is strongly encouraged to submit this form as well.

This information will be maintained in a “living list” by JVT during the progress of their work, on a best effort basis. If a given technical proposal is not incorporated in a Recommendation | Standard, the relevant patent information will be removed from the “living list”. The intent is that the JVT experts should know in advance of any patent issues with particular proposals or techniques, so that these may be addressed well before final approval.

This is not a binding legal document; it is provided to JVT for information only, on a best effort, good faith basis. Please submit corrected or updated forms if your knowledge or situation changes.

This form is *not* a substitute for the ITU ISO IEC Patent Statement and Licensing Declaration, which should be submitted by Patent Holders to the ITU TSB Director and ISO Secretary General before final approval.

Submitting Organization or Person:

Organization name **Broadcom Corp.**

16215 Alton Parkway, Irvine, CA 92619

Mailing address

Country

USA

Contact person

Sherman Chen

Telephone

+1 949 585 6185

Fax

+1 949 788 0972

Email	<u>schen@broadcom.com</u>
Place and date of submission	<u>Klagenfurt, Austria, July 22-26, 2002</u>
<u>Relevant Recommendation Standard and, if applicable, Contribution:</u>	
Name (ex: "JVT")	<u>JVT</u>
Title	<u>An issue regarding to the B-picture coding</u>
Contribution number	<u>JVT-D125</u>

(Form continues on next page)

Disclosure information – Submitting Organization/Person (choose one box)

☐

2.0 The submitter is not aware of having any granted, pending, or planned patents associated with the technical content of the Recommendation | Standard or Contribution.

or,

The submitter (Patent Holder) has granted, pending, or planned patents associated with the technical content of the Recommendation | Standard or Contribution. In which case,

☐

2.1 The Patent Holder is prepared to grant – on the basis of reciprocity for the above Recommendation | Standard – a free license to an unrestricted number of applicants on a worldwide, non-discriminatory basis to manufacture, use and/or sell implementations of the above Recommendation | Standard.

☐

2.2 The Patent Holder is prepared to grant – on the basis of reciprocity for the above Recommendation | Standard – a license to an unrestricted number of applicants on a worldwide, non-discriminatory basis and on reasonable terms and conditions to manufacture, use and/ or sell implementations of the above Recommendation | Standard.

Such negotiations are left to the parties concerned and are performed outside the ITU | ISO/IEC.

x

2.2.1 The same as box 2.2 above, but in addition the Patent Holder is prepared to grant a “royalty-free” license to anyone on condition that all other patent holders do the same.

☐

2.3 The Patent Holder is unwilling to grant licenses according to the provisions of either 2.1, 2.2, or 2.2.1 above. In this case, the following information must be provided as part of this declaration:

- patent registration/application number;
- an indication of which portions of the Recommendation | Standard are affected.
- a description of the patent claims covering the Recommendation | Standard;

*In the case of any box **other than 2.0** above, please provide the following:*

Patent
number(s)/status

Inventor(s)/Assignee(
s)

Relevance to JVT

Any other remarks: _____

(please provide attachments if more space is needed)

(form continues on next page)

Third party patent information – fill in based on your best knowledge of relevant patents granted, pending, or planned by other people or by organizations other than your own.

Disclosure information – Third Party Patents (choose one box)

☒ 3.1 The submitter is not aware of any granted, pending, or planned patents *held by third parties* associated with the technical content of the Recommendation | Standard or Contribution.

☐ 3.2 The submitter believes third parties may have granted, pending, or planned patents associated with the technical content of the Recommendation | Standard or Contribution.

For box 3.2, please provide as much information as is known (provide attachments if more space needed) - JVT will attempt to contact third parties to obtain more information:

3rd party name(s) _____

Mailing address _____

Country _____

Contact person _____

Telephone _____

Fax _____

Email _____

Patent number/status _____

Inventor/Assignee _____

Relevance to JVT _____

Any other comments or remarks: