

INTERNATIONAL ORGANISATION FOR STANDARDISATION
ORGANISATION INTERNATIONALE DE NORMALISATION
ISO-IEC/JTC1/SC29/WG11
CODED REPRESENTATION OF PICTURE AND AUDIO INFORMATION

ISO-IEC/JTC1/SC29/WG11/MPEG93/___

Source: M.E.Nilsson, D.G.Morrison, BT Labs
Title: Clarification for the use of VBV for low delay applications
Purpose: Information

This document provides some information about the use and operation of the VBV in low delay applications with the intention of clarification of these issues.

1. The need for skipped pictures in low delay applications

When skipped pictures are not allowed, that is, all source pictures are coded, the minimum end to end delay through the coder and decoder buffers is determined by the largest picture size in bits divided by the bit transfer rate from coder to decoder. Without skipped pictures the buffering must be such that ALL pictures have this delay. For those individual frames where high compression cannot be achieved with acceptable quality and the bit rate is low, this delay becomes unacceptable for some applications.

When skipped pictures are allowed, the delay through the encoder and decoder buffers is NOT the same for all pictures. Consider the following sequence of pictures.

Source pictures	1	2	3	4	5	6	7	8
Encoded pictures	1	2	3	4	-	-	7	8
Decoded pictures	1	2	3	3	3	4	7	8

Picture 4 is coded with a large number of bits, probably because it is the first picture of a new scene, causing the encoder to skip pictures 5 and 6. The decoder repeats picture 3 twice, then displays picture 4 before continuing normal decoding. The end to end delay for all pictures except picture 4 can be minimal, while picture 4 incurs an additional delay of 2 picture periods.

Skipped pictures allow operation with a low steady state delay which increases temporarily when needed.

The usual cause of skipped pictures is a scene change. It is the last frame of the old scene which gets held and repeated. Since there is a "disturbance" in the visual content anyway the "temporal artefact" is not particularly visible.

Good encoders will try to avoid skipped pictures and may use large step sizes at scene changes. The human visual system does not see these large distortions for a short time after the scene change. This is another reason not to use B pictures and reordering in low delay operation - the coded picture with large step size may not be the one in which the scene cut is first displayed.

2. The definition VBV_delay

VBV_delay should be defined in such a way as to allow low delay applications.

When operating in low delay mode, the encoder will usually try to keep the encoder buffer near to empty. It will therefore get the coded bits into the channel very quickly. The encoder will not know at the time of coding a picture header whether it will be able to code the picture with a small number of bits. It will only discover that it is coding a so-called 'big picture' as it codes the picture. By the time it knows, the picture header could have been sent to the channel..

The encoder must therefore fill in the VBV-delay and PTS/fields without knowing in advance that a picture may be a big one. Whilst it is possible to envisage encoders which go back and modify those fields after they realise there was/needs to be a big picture, this may conflict with the purpose of low delay.

3. Removal of pictures from VBV

The VBV concept in MPEG was derived from the Hypothetical Reference Decoder (HRD) of H.261. For the purpose of this discussion, the H.261 coding process is the same as the MPEG low delay process. The HRD was defined so that a picture was instantaneously removed if one were present, otherwise wait until a picture period later and inspect again. The maximum bits per coded picture was limited.

If bits are removed as they become available for a 'big-picture', there is no limit on the number of bits per coded picture. This could cause storage problems for some implementations of decoder.

The current definition involving the removal of all or none of the bits should be retained.

4. Possible low delay implementations

There is a simple way for a decoder to know if it has all the bits for one picture without parsing the whole data: look for picture headers going into the decoder buffer. If the header for the next picture hasn't gone in, then all the data for a picture is not in the buffer.

With real decoders as opposed to the VBV there are effectively two decoding architectures. In the first the decoding of a picture is completed in a holding store before being passed on to the reordering/display processing. This type is amenable to pausing its decoding operation when its coded data buffers (there may be more than one distributed over various places) are about to underflow. The incomplete picture will cause the reordering/display process to repeat the last complete picture.

The second type of decoder is more challenging but has lower delay. It attempts to decode close to the theoretical minimum delay. Pels (more accurately rows of macroblocks) are decoded just before they are needed for reordering/display. Clearly, stopping decoding at the onset of buffer underflow will cause a displayed frame to contain information from two time instants: the bottom part of the picture will be a repeat of the last picture. It may be possible to avoid or reduce this by looking at the headers going into the buffer as described above. With this technique the decoding would be held up until all data for a picture was received. However, this may introduce up to one picture delay and may increase the peak processing power needed in the decoder.

5. Conclusion

This document has attempted to clarify the use and operation of the VBV in low delay applications.

End of document