ISO/IEC JTC1/SC29/WG11
MPEG92/
July 1992

CCITT SGXV
Working Party XV/1
Experts Group for ATM Video Coding

Document AVC-275
June 23, 1992

SOURCE : JAPAN
TITLE : SCENE CHANGE HANDLING IN LOW DELAY MODE
Purpose: Dsicussion

----

1. Introduction

Low delay is one of the profiles for the generic coding to provide for communication applications. This document addresses how we should cope with scene changes which require large number of bits and consequently additional buffering delay.

2. Problem

As analyzed in AVC-179, inherent coding/decoding delay is caused by

- picture reordering due to the use of backward prediction, and
- buffering to obtain constant bit rate for transmission.

The current low delay mode defined in TM1 (Annex H to AVC-260) intends to solve the first item by using forward prediction only and the second item by distributing INTRA slices in multiple frames/fields.

When a scene change takes place, however, the first picture of a new scene requires much more bits than other pictures to maintain a certain quality as far as interframe prediction is used. If more bits are generated, more time is necessary to transmit them, and delay occurs in reproduced pictures if we want all pictures to be transmitted and displayed. We need some mechanism to cope with scene changes in the low delay mode.

Annexes 1 and 2 illustrate the above problem with TM experiments. Annex 1 contains some experimental results to show how SNR decreases at a scene change under constant bit rate control, and how information is generated under constant quality control. Annex 2 shows a performance comparison between the current TM1 low delay mode and other TM1 coding structures in terms of delay and SNR, suggesting that pure field coding with B-fields can also be considered for the low delay profile under the Test Model framework.

3. Simplified information generation model

For discussion purpose, we assume a simplified model as follows;

a. Coding: 30 Hz frame based
b. Transmission rate: 6 Mbit/s constant
c. Number of bits in the first frame in new scene: 800 kbits
d. Number of bits in all other frames: 200 kbits
e. Coder encodes an input signal as soon as it arrives and puts this encoded data continuously to a constant rate transmission channel.

f. Decoder buffer works as defined in VBV (see Annex C to AVC-260). As soon as the coded data of a frame are removed, the decoder starts to display the frame.

In this particular model, the VBV delay must be set to 4/30 second to accommodate the coded signal (see Figure 1). Generally, if X bits are generated at the first frame of scene change in a T bit/s transmission channel, delay of X/T second occurs.

4. Methods to reduce the delay due to scene change

We can cope with the scene change by one of the following methods:

A) All frames are transmitted with a nearly constant number of bits regardless of the picture content

This can be achieved by using a small size of transmission buffer. If 200 kbits are assigned even to the first frame of new scene in the above mentioned example (Frame G in Figure 1), there is no buffering delay.

One way will be to INTRA code the first frame with a large step size. Resultant pictures look like blurred, blocky and/or noisy for a while. Another way will be to INTRA code only a part of each of the first several frames of the new scene. In the above mentioned example every 1/4 of the frame can be INTRA coded for frames G, H, I and J. Resultant pictures look like being wiped from an old scene to a new scene.

This category requires no specification in the standard. All the tricks are covered by setting of the VBV delay parameter and coding control parameters in the encoding side.

B) Some frames are skipped

If the first frame of a new scene requires a large number of bits, some of the following frames may be skipped as shown in Figure 1 c) for applications where tranmsission and display of all the input frames are not mandatory. In the decoder, those skipped frames are not displayed and the frame preceding the first frame of new scene is displayed repeatedly. Visual effects will be some discontinuity of the picture appearance. It is expected that there may be masking effects due to a rapid change of scene. See Annex 3 for some experimental results.

This category requires definition of skipped frames in the syntax. It may be a frame with a picture header but without any coded data, or even the picture header can be dropped as in H.261. It should also be defined how to deal with these skipped frames in the decoder; to display the current frame if the next frame does not arrive, not to display those skipped frames, and to display the first available frame following those skipped frames.

5. Conclusion

Scene change handling in the low delay mode has been discussed. Some specifications will be required in the standard if frame dropping is to be allowed for this purpose. Visual effects of possible methods to cope with scene changes should be experimented.

According to the decision on this matter, clarification is necessary on how to process the first picture in the Test Model experiemtns.

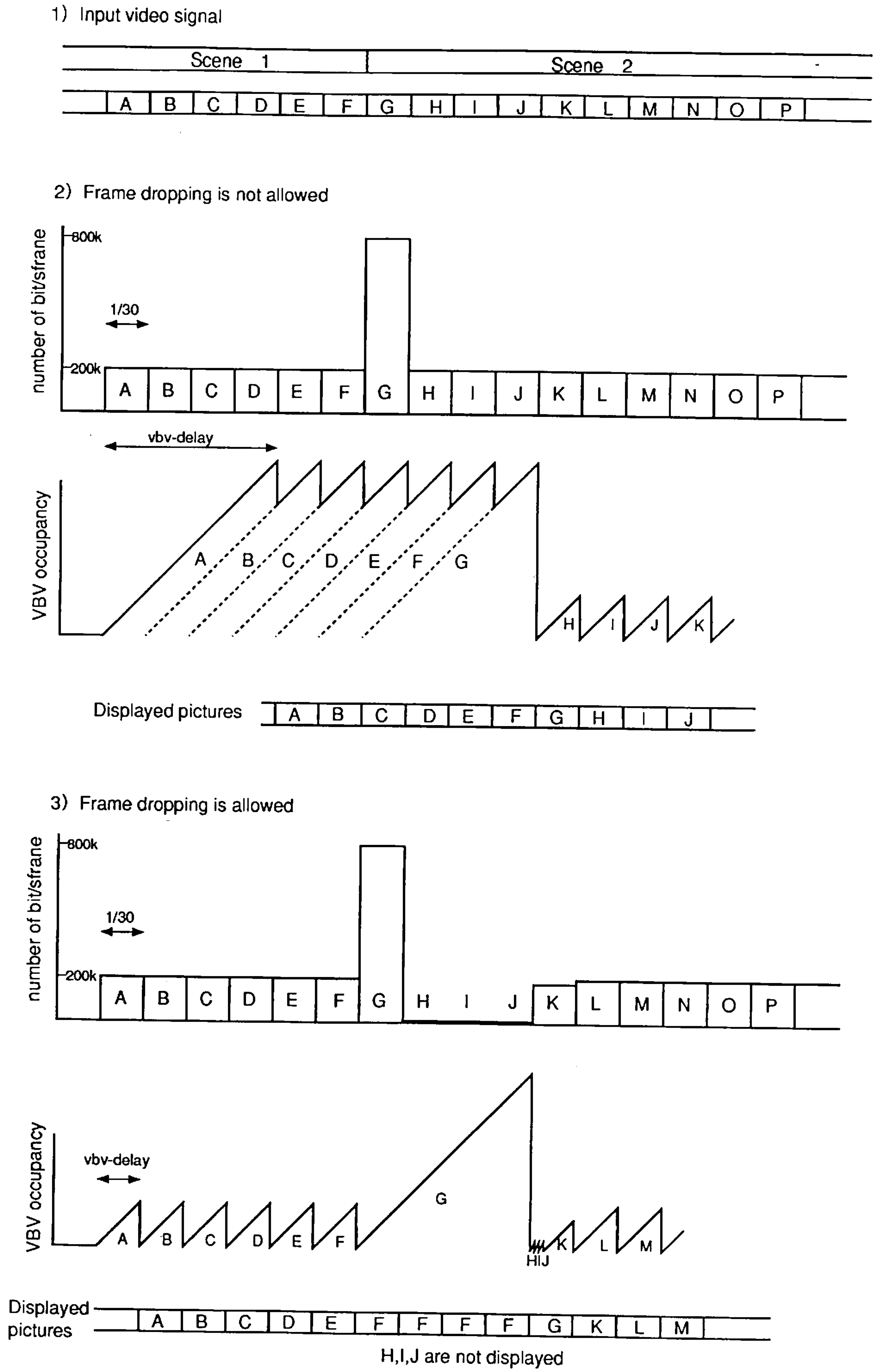


Figure1 Coding of a scene change

Annex1.

Estimation of S/N and VBV buffer transition in a scene change

1.Introduction

To achieve low delay, we must solve the problems of not only coding/decoding delay but also buffering delay. This annex estimates buffering delay using "Frame based coding" and "Field based coding", especialy in a scene change.

2.TEST conditions

Coding algorithms :	TM1 Fr based(Fr/Fi), pure-Fi
M :	M=1 INTRA SLICE (TM1 : APPENDIX H)
Rate control :	TM1 STEP2
Bitrate :	4Mbps
Sequence :	Jointed Sequence (FG 30frames + MC 30frames)

3.Results

Using M=1 INTRA SLICE in the low delay mode, when a scene change takes place, the number of generated bits increases. It is because there are many INTRA blocks and much MC Error in the pictures after the scene change.
Therefore it comes to the conclusion that

① When we attain the constant bitrate (CBR) using TM1 rate control (step2), the degradataion of about 4dB in the pictures after the scene is detected. What is more it takes about 1 second to recover the picture quality, not depending on whether it's "Fr based Fr/Fi" or "pure-Fi" algorithms. (see Figure 1 and 2)

② In order to keep good picture quality in the pictures after the scene change, we need ideal rate control algorithm such as the variable bitrate method(VBR). In this annex, VBR means the constant Q step (Q=12). In this case, large buffering delay occurs. (see Table1, Figure 7 and 8)

Fr/Fi : 181 (148[1]+33[2]) msec
pure-Fi : 82[3] msec

[1] [3] buffering delay in the scene change.(see Table1, Figure 7 and 8)

Table1:details

	Fr/Fi	pureFi
VBR	148 msec (592Kbits)	82 msec (327Kbits)
CBR (TM1)	76 msec (304Kbits)	40 msec (160Kbits)

[2] field merging delay.
There is no re-ordering delay in M=1. But Frame based algorithm needs delay of 2fields. (Enc:1field, dec:1field)

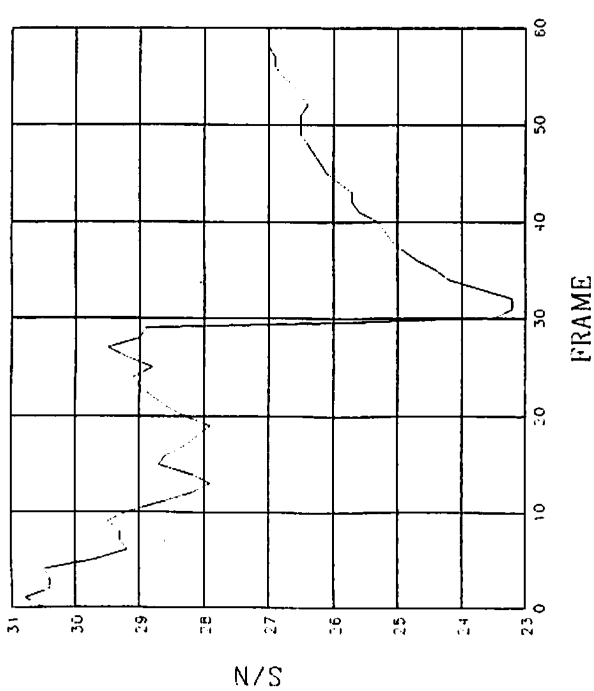


Figure1 : S/N Fr/Fi (CBR)

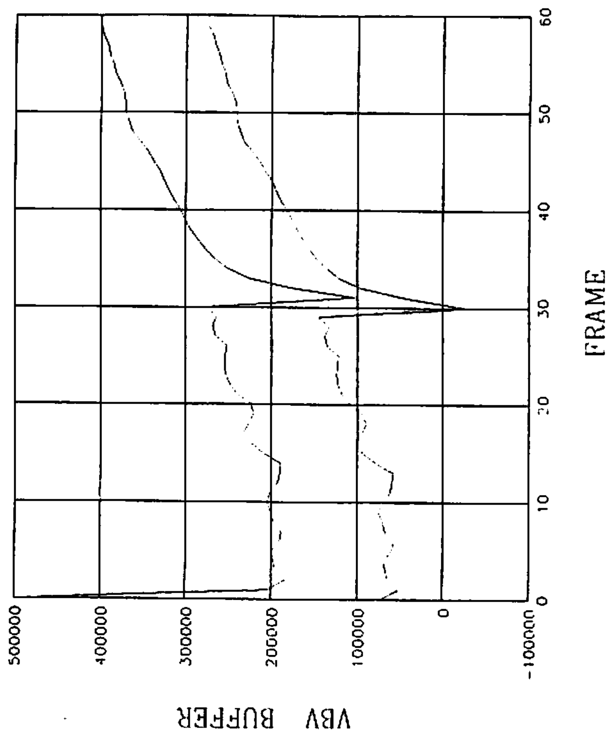


Figure3 : VBV Fr/Fi (CBR)

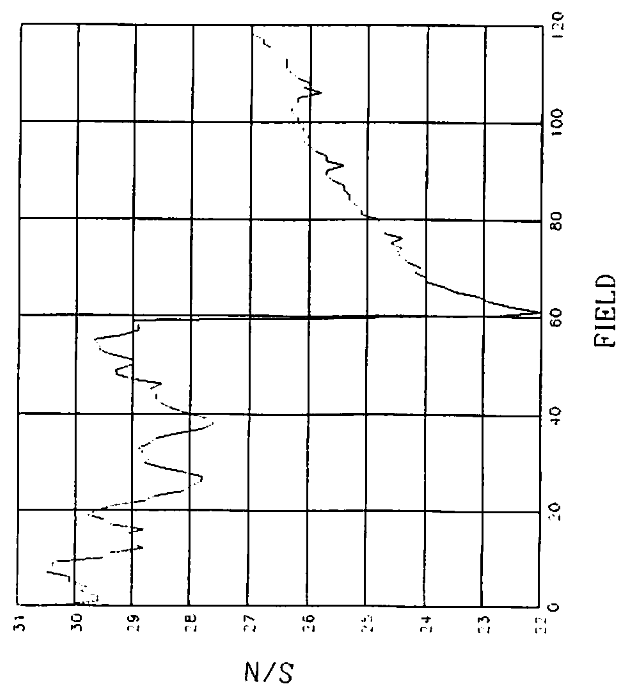


Figure2 : S/N pure-Fi (CBR)

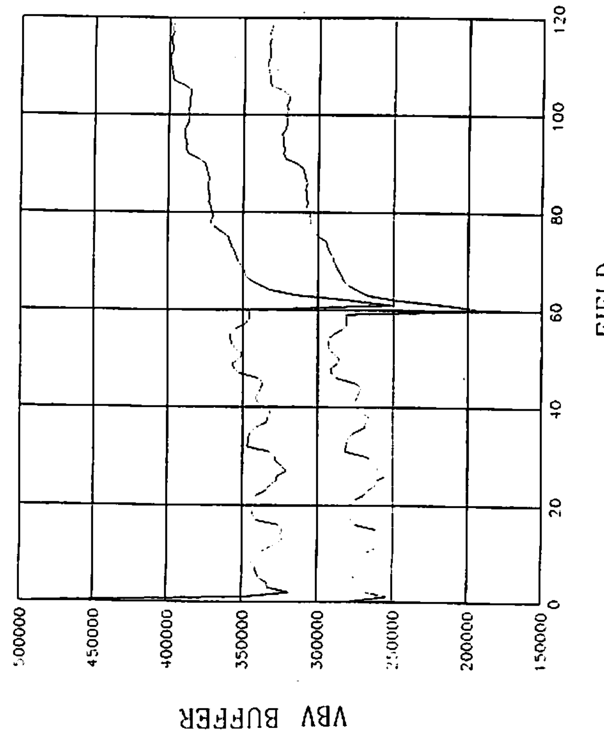


Figure4 : VBV pure-Fi (CBR)

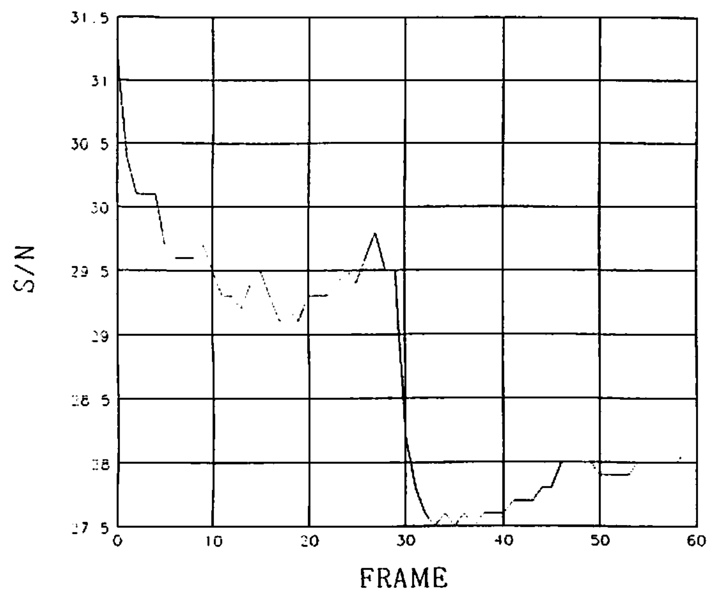


Figure5 : S/N Fr/Fi (VBR)

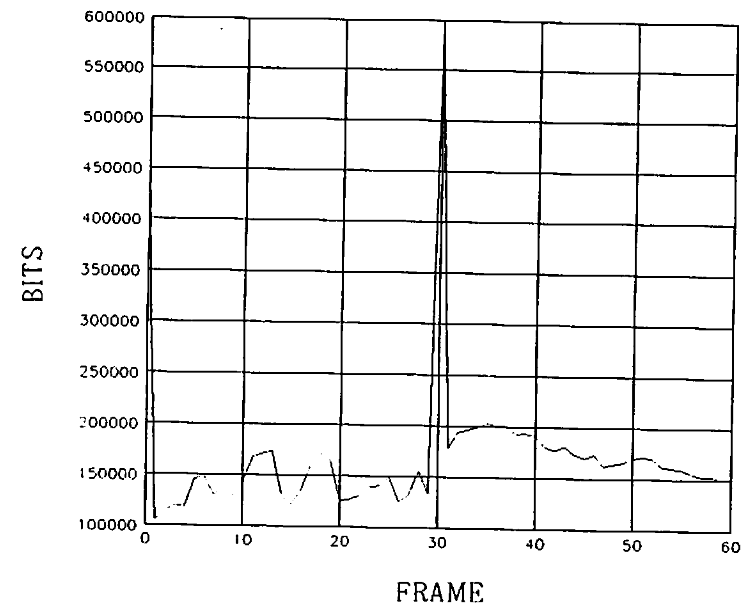


Figure7 : BITS Fr/Fi (VBR)

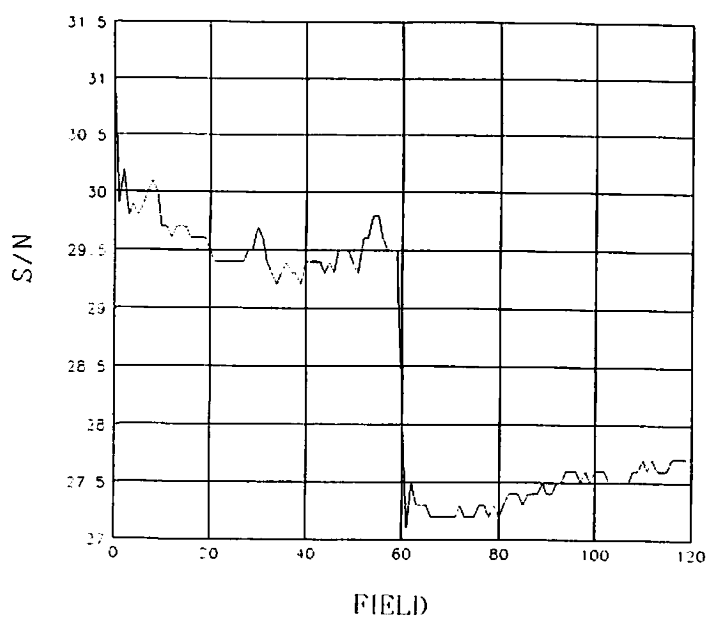


Figure6 : S/N pure-Fi (VBR)

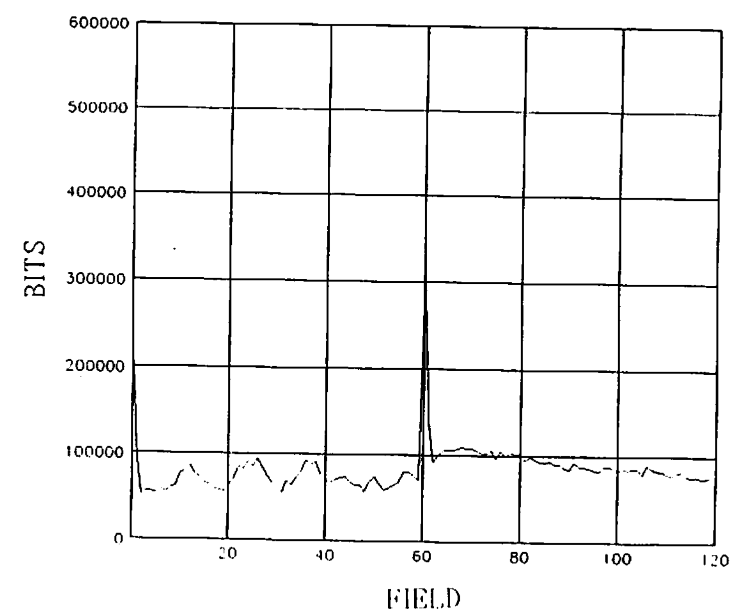


Figure8 : BITS pure-Fi (VBR)

ANNEX 2. An example for low delay coding with scene change robustness

1.Introduction

Low delay coding is one of the most important techniques for real time video communication. If the input sequence does not contain any scene changes, intra slice is very effective technique. However, a scene change causes picture quality degradation and/or picture skipping.

In this annex, an example for low delay coding with scene change robustness is introduced. Coding/decoding delay is estimated for this new scheme.

2.Low delay coding with scene change robustness

In this new coding, M=1 is assumed, and number of pictures in a GOP (N) varies according to scene change timing. When a scene change is detected, the first picture of the new scene is coded as I-picture. The following pictures are coded as P-picture until next scene change. In order to satisfy Video Buffering Verifier (VBV), rate control is performed as if N is equal to Nsc (e.g.150). For Nsc/30 seconds from scene change, the target bits for each picture is calculated according to the rate control described in TM1 (TM1 mode). In this period the VBV buffer occupancy in decoder increases up to the buffer fullness. When no scene change occurs for Nsc/30 seconds, the rate control scheme is switched to constant mode. In the constant mode target bits Tc is calculated as follows;

Tc = bit_rate / picture_rate

In the constant mode, the VBV buffer occupancy keeps the fullness. This means whole bits in the VBV buffer is allocated to next occasional I-picture when scene change interval is longer than Nsc/30 seconds.

If the scene change interval is shorter than Nsc/30 the allocated bits are shortened and picture quality degraded and/or pictures are skipped.

Fig.A2-1 illustrates the description of this rate control algorithm in Nsc==150, examples of long interval and short interval.

This rate control scheme is applicable to the codings with backward prediction (M=2,3...). In these cases, P-picture should be replaced by new I-picture. When scene change occurs at a B-picture, the B-picture and successive B-pictures should be skipped and the GOP must be terminated immediately.

3.Estimation of delay

The delay consists of basic delay and buffering delay. The basic delay is mainly caused by picture reordering for backward prediction. The buffering delay is caused by VBV buffering.

In this section, the both delays : basic delay and buffering delay of the new coding are estimated. And, the delays are compared to other scheme; pure field coding with 2 B-ref, pure field coding with 3 B-ref and conventional Field/Frame coding (M=3).

3.1.Basic Delay

The basic delays is estimated from prediction structure shown in Fig.A2-2. The delays are shown in Table A2-1.

Table A2-1 The basic delay

	coding base	M	basic delay	
New coding	frame	1	2 fields	33.3ms
pure field 2 B-ref.	field	3	3 fields	50.0ms
pure field 3 B-ref.	field	3	5 fields	83.3ms
Field/Frame(M=3)	frame	3	8 fields	133.3ms

3.Buffering Delay

The buffering delay is calculated with VBV buffer size. We assume that the VBV buffer size is defined as maximum number of bits for a picture : I-picture. With this assumption the buffering delay is estimated as follows:

buffering_delay = number_of_bits_for_INTRA_picture / bit_rate

We use initial value of Ti in TM1 rate control (TM1 p.61) for 'number_of_bits_for_INTRA_picture'. The buffering delays and the total delays are shown in Table A2-2. The total delay is calculated as basic delay + buffering delay. From Table A2-2, the new coding and two pure field coding satisfy the low delay requirement (150ms).

Table A2-2 Buffering delay (4Mbps : CBR)

	N	M	buffering delay		Total delay
New coding	150	1	350kbit	87.5ms	120.8ms
pure field 2 B-ref.	30	3	250kbit	62.5ms	112.5ms
pure field 3 B-ref.	30	3	250kbit	62.5ms	145.8ms
Field/Frame(M=3)	15	3	450kbit	112.5ms	245.8ms

4.Computer Simulation

The coding efficiency is measured with computer simulation for the new coding and two pure field coding which satisfy the low delay requirement. For the new coding, Nsc=150 is assumed. The specifications of the each model are shown Table A2-3.

Table A2-3 Specification of simulation model

	New coding	pure field 2 B-ref.	pure field 3 B-ref.
coding base	frame	field	field
N, M	(150,1)	(30,3)	(30,3)
algorithm	TM1 appendix H	TM1 appendix E	
prediction	Fr/Fi adapt.	2 ref. for B-picture	3 ref. for B-picture
motion estimation	15.5x15.5 (full search)		
rate control	modified TM1 step2	TM1 step2	
others	INTRA slice	16x16 MC	
bit rate	4Mbps		
sequence	Mobile&Calendar : 2sec		
	Football(old) : 2sec		

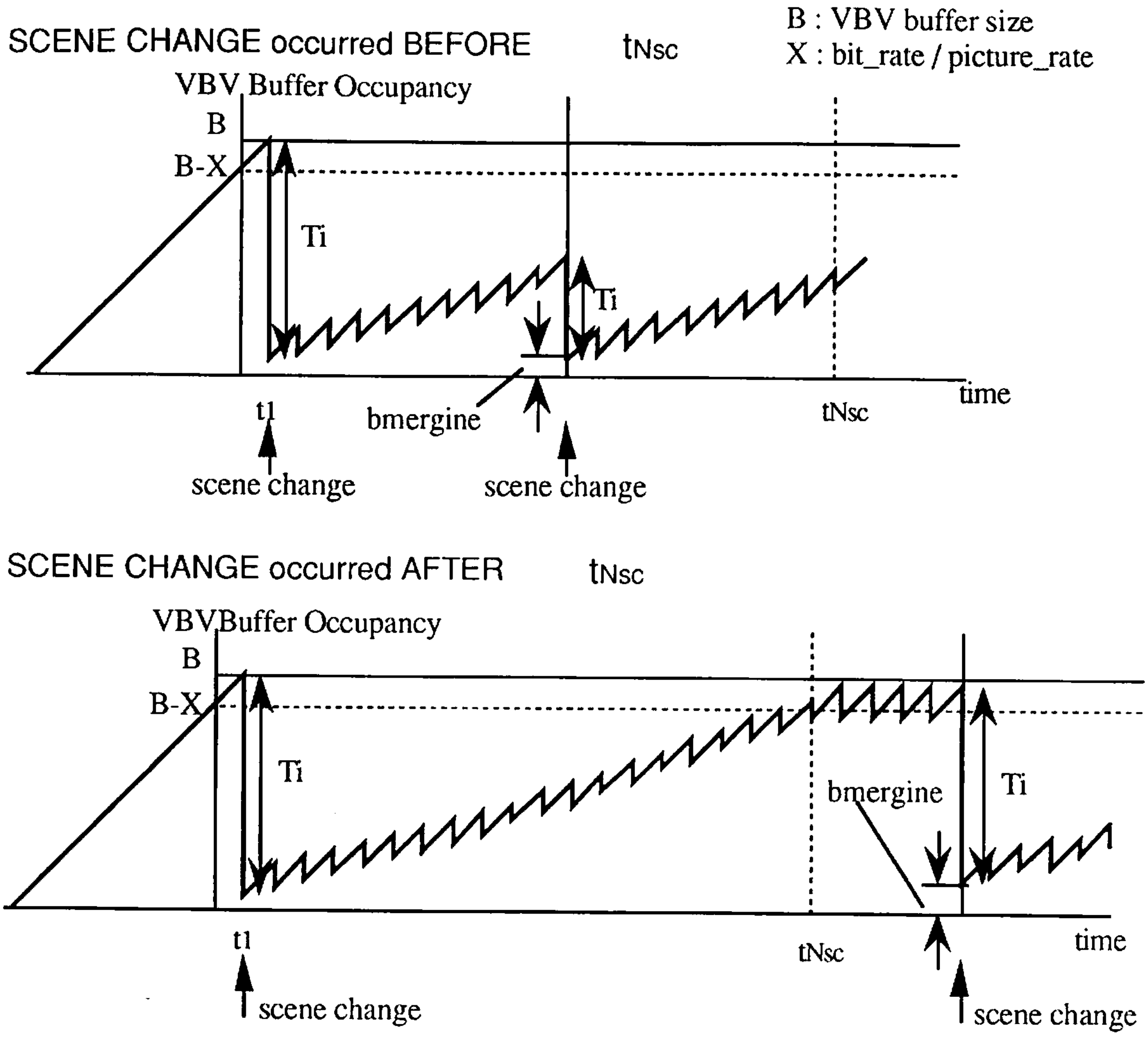
Table A2-4 shows the results of the simulation.

Table A2-4 SNR for Y signal

	mobile	football
New coding	26.7dB	31.0dB
pure field 2 B-ref.	27.4dB	32.1dB
pure field 3 B-ref.	28.1dB	32.3dB

5.Conclusion

The new low delay coding with scene change robustness satisfies the low delay requirement. However the pure field coding which also satisfies the requirement performs in higher picture quality. Further study should be needed for low delay coding with scene change robustness including the pure field coding.



Rate control algorithm

```
if ( scene_change_occurred ) {
   reset_GOP() ;  /* set number_of_picture_in_GOP = 1 */
   Ti =  VBV_buffer_occupancy - bmergine ;  /* bmergine == almost 0 */
   set_INTRA_coding_mode() ;
}else if ( number_of_picture_in_GOP <= Nsc ) {
   set Tp according to TM1 rate control.
}else {
   Tp =  bit_rate / picture_rate ;
}
```

Fig.A2-1 VBV buffer occupancy in New low delay coding

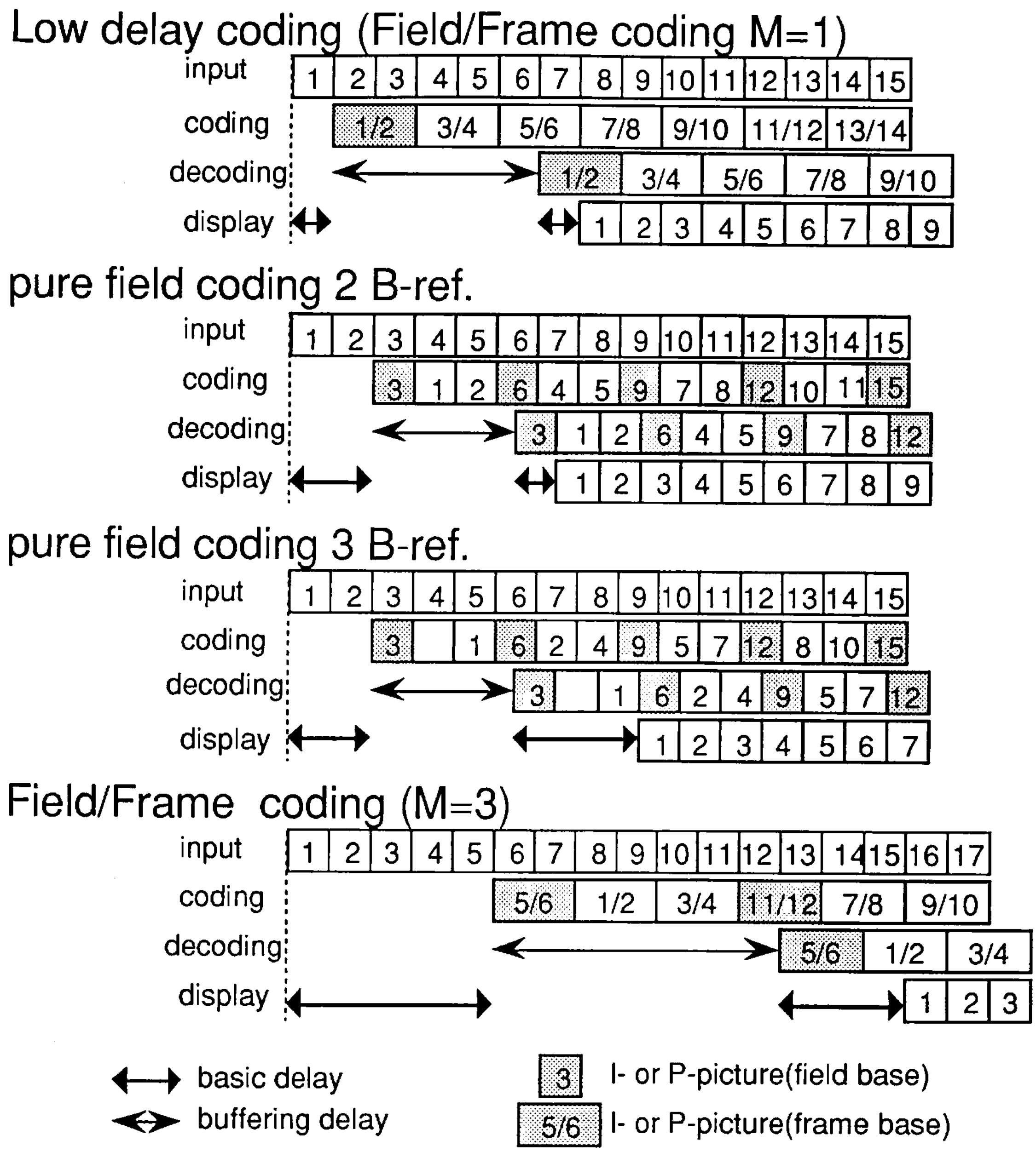


Fig.A2-2 Definition of coding/decoding delay

Video Demonstration

The purpose of this demonstration is to investigate the visual effects due to a rapid change of scene.

We compared the displayed patterns as shown in Figure 1. The Picture format is 525/30 CCIR-601. The picture sequences are shown in Figure 2. In Figure 1, each of A,B,...,N denotes a frame. The scene change occurs between the frame F and the frame G. The (a) is the case in which the frame dropping is not allowed. The (b-1), (c-1), (d-1) and (e-1) are the cases in which the frame dropping is allowed. The (b-2),(c-2), (e-2) and (e-3) are the case in which only a part of each of the first several frames of new scene is INTRA coded (wiped version). The (G/F) in the (b-2) denotes the upper 1/2 of the frame is INTRA coded for frame G and the lower 1/2 is not coded. In the same way, the (G/F/F) in the (c-2) denotes the upper 1/3 of the frame is INTRA coded for frame G and the other part is not coded, the (H/H/F) in the (c-2) denotes the upper 1/3 of the frame is INTER coded for frame H, the mid 1/3 is INTRA coded for frame H and the lower 1/3 is not coded, etc. Note that in this video demonstration each frame is original (not coded).

(a)	...A B C D E F	G	H	I	J	K L M N...
(b-1)	...A B C D E F	F	G	I	J	K L M N...
(c-1)	...A B C D E F	F	F	G	J	K L M N...
(d-1)	...A B C D E F	F	F	F	G	K L M N...
(e-1)	...A B C D E F	F	F	F	F	G L M N...
(b-2)	...A B C D E F	(G/F)	H	I	J	K L M N...
(c-2)	...A B C D E F	(G/F/F)	(H/H/F)	I	J	K L M N...
(e-2)	...A B C D E F	(G/F/F/F/F)	(H/H/F/F/F)	(I/I/I/F/F)	(J/J/J/J/F)	K L M N...
(e-3)	...A B C D E F	(F/F/G/F/F)	(F/H/H/F/F)	(F/I/I/I/F)	(J/J/J/J/F)	K L M N...

Figure A3-1 Displaying Frame Pattern

Sequence 1 :
```
---> FG(1.5sec) --> MC(1.5sec) --> FG(1.5sec) --> BL(1.0sec) --
|                                                              |
|______________________________________________________________|
```

Sequence 2 :
```
---> FB(1.5sec) --> MC(1.5sec) --> FG(1.5sec) --> BL(1.0sec) --
|                                                              |
|______________________________________________________________|
```

FG : FlowerGarden, MC : Mobile&Calendar
FB : Football, BL : black

Figure 2 Picture Sequences

The video demonstration shows skipping a few frames doesn't cause much visual degradation. For some applications, skipping a few frames in a scene change might be allowed.

END