

CCITT SGXV
Working Party XV/1
Experts Group for ATM Video Coding
(Rapporteur's Group on Part of Q.3/XV)

Document AVC- 411
Jan. 21, 1993

SOURCE: Japan
TITLE: Simulation results on compatibility
PURPOSE: Proposal/Information

1. Introduction

Two sets of simulations were carried out on compatibility. One was a comparison between embedded and simulcast in the interlace-interlace compatible framework. An improvement gained by introducing the embedded method was compared with that for the interlace-progressive compatible framework.

The other was concerning the improvement by a spatio-temporal weighting method. According to the result shown at the London meeting, the improvement by spatio-temporal weighting seems to be small. Simulations were carried out to confirm this point both for the interlace-progressive and interlace-interlace cases.

2. Simulation condition

The simulation program was basically the same as that described in AVC-367 (MPEG92/586).

TM2 (rev.1) M=3, N=15 Fr structure, Fr/Fi prediction
rate control: including step 3
spatio-temporal weighting: weighting 1/2 only
Combination of the base layer and bi-directional prediction in the B picture was not allowed (as agreed at London)
sequences: Football (new), Flower Garden (2 seconds)
Motion vector search range
base layer ± 7.5 pel/fr (FG), ± 15 pel/fr (FB)
upper layer ± 15 pel/fr (FG), ± 30 pel/fr (FB)

(1) Interlace-progressive case

bit rate MPEG1 layer : SM3 1.15 Mbps Total bit rate : 4 Mbps
base layer format: SIF odd

(2) Interlace-interlace case

bit rate MPEG1 layer : TM2 1.15 Mbps Total bit rate : 4 Mbps
base layer format: SIF interlace (intra-field down sampled using a filter described in TM)
up sampling: intra-field coefficients for vertical filter odd: 1,2,1 even: 1,3,3,1

3. Simulation for interlace-interlace compatibility

Embedded and simulcast were compared in a case that both the upper layer and the base layer were interlace. Table 1 shows the simulation result. This result shows that a gain in coding efficiency by embedding was small compared with the interlace-progressive case. Furthermore, a lower SNR than in the simulcast case was obtained for "Flower Garden". The reasons which occurred to our mind were as follows.

- (1) An SIF interlace format made by intra-field up-down sampling has only one half resolution of an SIF odd format in the vertical direction and has an aliasing component. This makes the signal fit less for the Rec.601 format as a prediction.
- (2) In embedded coding, excessive information is included compared with simulcast (compatible type).
- (3) Intra macroblocks are differently quantized for embedded and simulcast.

To cope with the first reason, up-down sampling using a BT/BBC filter (MPEG92/758) was applied. To cope with the third reason, the same quantizing as Intra macroblock was applied for compatible macroblocks in the I pictures. The results are shown in Table 2. These results show the following.

Concerning the vertical resolution, intra-field up-down sampling and BT/BBC up-down sampling didn't show a large difference. This seems to be because although the BT/BBC filter prepares a prediction signal with a higher vertical resolution than an intra-field up-sampling filter in regions with a small motion, a temporal prediction also fits well and the SIF interlace cannot contribute to the prediction very much in such a region. If this is true, the BT/BBC filter will have an advantage in the I picture for pictures with a small motion.

On the other hand, a filter with a good performance is desirable for down sampling because the picture quality of the base layer is determined by the performance of the down sampling filter.

Concerning quantization, Intra quantization for compatible macroblocks in the I picture gave a higher SNR. This should also be true for an interlace-progressive case.

As a result, the gain in coding efficiency by embedding seems to be small in the interlace-interlace case compared with the interlace-progressive case.

4. Problem of currently proposed up-sampling methods

The up-sampling method proposed by BT/BBC utilizes 3 successive fields to transform interlace to progressive. So, the previous and the next fields to the current field are necessary. This leads to at least 1 field delay increase. Especially, for cases using the B picture in base layer coding, the filter must wait until the necessary fields will become available by frame re-ordering. This will lead to a large delay.

For the same reason, an up sampling method proposed by AT&T (MPEG92/509) also increases the delay for the same B picture cases.

5. Spatio-temporal weighting

A simulation was carried out to investigate the performance of spatio-temporal weighting compared with the prediction from the base layer both in the interlace-progressive and interlace-interlace cases. The results are shown below.

- (1) Interlace-progressive case

For the interlace-progressive case, the result shows that the gain from the prediction from the base layer was 0.17 dB for "Football" and 0.23 dB for "Flower Garden". (The result was slightly different from that of AVC-367 because of removing the combined prediction of the bi-directional and base layer in the B picture.)

(2) Interlace-interlace case

For the interlace-interlace case, the simulation condition was to use a BT/BBC filter for up-down sampling and with quantization modification. The result showed that the gain from the prediction from the base layer was 0.13 dB for "Football" and 0.02 dB for "Flower Garden".

From these results, it has been confirmed that the spatio-temporal weighting method is not very effective compared with the prediction from the base layer. This seems to be because the improvement of the prediction signal is canceled by an increase in excessive motion vector information, as was pointed out at London. This concludes that the spatio-temporal method is not necessary considering implementation complexity.

6. Conclusion

(1) In the interlace-inter case, a gain in coding efficiency by embedding was smaller than in the interlace-progressive case.

(2) The up-sampling filtering method proposed by BT/BBC doesn't seem to contribute to prediction very much even though it gives a better base layer picture.

On the other hand, a good down sampling filter is desirable because the picture quality of the base layer is determined by the performance of the down sampling filter.

(3) The up sampling filter proposed by BT/BBC has a problem that it increases the delay, especially for cases using the B picture in base layer coding. The up-sampling method proposed by AT&T also increases the delay for the same B picture cases.

(4) The improvement obtained by spatio-temporal weighting seems to be small both for interlace-interlace and interlace-progressive cases. Therefore, we propose not to include spatio-temporal weighting, which makes the hardware more complex than a simple prediction from the base layer.

END

Table 1 Interlace - Interlace

Football

	SNR (I; P; B)	Compatible mode (%)
Simulcast (2.85 Mbps)	33.45 (34.42; 33.42; 33.38)	
Embedded(spatio-temporal)	34.66 (34.68; 35.09; 34.49)	100; 80.3; 43.0

Flower Garden

	SNR (I; P; B)	Compatible mode (%)
Simulcast (2.85 Mbps)	27.92 (28.38; 27.65; 27.99)	
Embedded(spatio-temporal)	27.83 (27.42; 27.69; 27.94)	99.6; 36.7; 12.4

Table 2 Interlace-Interlace
Modified embedded (including spatio-temporal)

Football

	SNR (I; P; B)	Compatible mode (%)
BT/BBC filter	34.69 (34.74; 35.13; 34.52)	100; 82.0; 43.0
Intra quantization	34.84 (35.81; 35.37; 34.57)	100; 78.2; 41.4
BT/BBC + Intra Q	34.88 (35.85; 35.42; 34.60)	100; 79.9; 41.5

Flower Garden

	SNR (I; P; B)	Compatible mode (%)
BT/BBC filter	27.90 (27.52; 27.78; 28.00)	99.8; 40.6; 13.1
Intra quantization	28.13 (28.64; 27.90; 28.18)	99.6; 33.0; 10.6
BT/BBC + Intra Q	28.20 (28.73; 27.99; 28.23)	99.8; 36.6; 11.4

Table 3 Spatio-temporal vs pred. from base layer
(Interlace - Progressive)

Football

	SNR (I; P; B)	Compatible mode (%)
Simulcast (2.85 Mbps)	33.45 (34.42; 33.42; 33.38)	
Prediction from base layer	35.31 (35.76; 35.82; 35.08)	100; 76.9; 41.0
Spatio-temporal weighting	35.48 (35.82; 36.05; 35.24)	100; 91.8; 55.9

Flower Garden

	SNR (I; P; B)	Compatible mode (%)
Simulcast (2.85 Mbps)	27.92 (28.38; 27.65; 27.99)	
prediction from base layer	28.79 (28.55; 28.72; 28.85)	99.9; 52.5; 15.2
spatio-temporal weighting	29.02 (28.68; 29.08; 29.02)	99.9; 77.2; 26.4

Table 4 Spatio-temporal vs pred. from base layer
(Interlace - Interlace)

Football

	SNR (I; P; B)	Compatible mode (%)
Simulcast (2.85 Mbps)	33.45 (34.42; 33.42; 33.38)	
Prediction from base layer	34.75 (35.58; 35.40; 34.44)	100; 62.4; 23.7
Spatio-temporal weighting	34.88 (35.85; 35.42; 34.60)	100; 79.9; 41.5

Flower Garden

	SNR (I; P; B)	Compatible mode (%)
Simulcast (2.85 Mbps)	27.92 (28.38; 27.65; 27.99)	
Prediction from base layer	28.18 (28.70; 27.95; 28.22)	99.8; 15.2; 5.4
Spatio-temporal weighting	28.20 (28.73; 27.99; 28.23)	99.8; 36.6; 11.4