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

Document AVC-339
Sept. 28, 1992

INTERNATIONAL ORGANIZATION FOR STANDARDIZATION
ORGANISATION INTERNATIONALE DE NORMALISATION
ISO/IEC JTC1/SC29/WG11
CODING OF MOVING PICTURES AND ASSOCIATED AUDIO

ISO/IEC JTC1/SC29/WG11
MPEG92/
29 Sept. 1992

Source: Toshiba Corporation
Title: Simulation results on prediction core experiment (Dual-prime, SFAMC) in TM2
Purpose: Proposal(MPEG), Information(CCITT)

1. Introduction

In this document, the performance of the dual-prime prediction mode has been evaluated, and it has been compared with that of SFAMC in the following four prediction structures, including low delay modes.

(1-a) Frame structure M=3
(1-b) Frame structure M=1
(2-a) Field structure M=1
(2-b) Field structure M=2

Observations of the selected prediction modes has also been carried out in the Annex, to confirm the effect of the dual-prime prediction mode as increasing vertical resolution and spatiotemporal loop filtering.

2. Simulation condition

•TM2
•GOP N=15
  M=1 IPPPPPPPPPPPPPP
  M=2 IBPBPBPBPBPBPBP
  M=3 BBIBBPBBPBBPBBP
•Sequences: Flower Garden, Mobile&Calendar, Bicycle (2 seconds)
•Rate: 4Mbps
•Rate control: Including step 3
•Motion vector search
  15.5 pixel/frame
  Telescopic search based on frame motion vectors
  Half pixel search: Using original picture
  Second stage search for Dual-prime prediction: Using local decoded picture
•Reference fields of the second(even) fields of the P-picture in the field structure (M=2)
  Dual: The last two P(I)-fields are used.
  Dual', SFAMC: The last two contiguous P(I)-fields are used.

3. Simulation results

The luminance SNRs with various combination of prediction modes including Dual' or SFAMC are summarized in Table 1.

Table 1 Luminance SNR[dB] in several prediction modes

(1-a) Frame Structure M=3

Prediction mode	Flower	Mobile	Bicycle
Fi./Fr.	29.86(0.00)	27.88(0.00)	27.45(0.00)
Dual'	29.36(-0.50)	26.64(-1.24)	27.81(+0.36)
Fi./Dual'	30.33(+0.47)	28.03(+0.15)	27.70(+0.25)
Fr./Dual'	30.27(+0.41)	28.10(+0.22)	27.90(+0.45)
Fi./Fr./Dual'	30.48(+0.62)	28.22(+0.34)	27.77(+0.32)
SFAMC	29.92(+0.06)	27.76(-0.12)	27.55(+0.10)
Fi./SFAMC	30.24(+0.38)	27.95(+0.07)	27.52(+0.07)
Fr./SFAMC	30.36(+0.50)	28.09(+0.21)	27.71(+0.26)
Fi./Fr./SFAMC	30.41(+0.55)	28.16(+0.28)	27.63(+0.18)

(1-b) Frame Structure M=1

Prediction mode	Flower	Mobile	Bicycle
Fi./Fr.	28.29(0.00)	25.71(0.00)	27.43(0.00)
Dual'	28.84(+0.55)	25.41(-0.30)	27.93(+0.50)
Fi./Dual'	29.55(+1.26)	26.25(+0.54)	27.90(+0.47)
Fr./Dual'	29.66(+1.37)	26.37(+0.66)	27.98(+0.55)
Fi./Fr./Dual'	29.66(+1.37)	26.38(+0.67)	27.94(+0.51)
SFAMC	29.37(+1.08)	25.92(+0.21)	27.60(+0.17)
Fi./SFAMC	29.32(+1.03)	25.88(+0.17)	27.68(+0.25)
Fr./SFAMC	29.39(+1.10)	25.95(+0.24)	27.71(+0.28)
Fi./Fr./SFAMC	29.44(+1.15)	26.02(+0.31)	27.75(+0.32)

(2-a) Field Structure M=1

Prediction mode	Flower	Mobile	Bicycle
Fi.	27.55(-0.42)	24.65(-0.17)	27.48(-0.22)
Fi./Dual	27.97(0.00)	24.82(0.00)	27.70(0.00)
Fi./Dual'	28.63(+0.66)	25.25(+0.43)	27.87(+0.17)
Fi./SFAMC	28.82(+0.85)	25.09(+0.27)	27.86(+0.16)

(2-b) Field Structure M=2

Prediction mode	Flower	Mobile	Bicycle
Fi.	29.12(-0.25)	26.20(-0.16)	27.85(-0.12)
Fi./Dual	29.37(0.00)	26.36(0.00)	27.97(0.00)
Fi./Dual'	29.76(+0.39)	26.68(+0.32)	28.05(+0.08)
Fi./SFAMC	29.79(+0.42)	26.65(+0.29)	27.99(+0.02)

3.1 Results on frame structure (M=3, 1)

• Both Dual' and SFAMC gave increasing in coding efficiency in combination with Frame and/or Field. The amount of increasing in coding efficiency was especially large, in case of M=1.
• Fi./Fr./Dual' was slightly better than Fi./Fr./SFAMC.
• Comparing Fi.(Fr.)/Dual' with Fi.(Fr.)/SFAMC, Fr./SFAMC outperformed Fr./Dual' in Flower with M=3. In other cases, Fi.(Fr.)/Dual' was better than Fi.(Fr.)/SFAMC.
• Using Dual' alone, its performance was worse than Fi./Fr. in Flower and Mobile. The reason for this result seems to be the following two points.

The first point is that the interpolation mode is inhibited in B-picture like SFAMC. The second point is that prediction signal is made by averaging two to four scanning line signals. (Refer to appendix A.2) So, it is not possible to make prediction signal by referencing one scanning line signal directly.

Therefore, it is necessary that Dual' is used in combination with Field and/or Frame.

3.2 Results on field structure (M=1, 2)

• Both Fi./Dual' and Fi./SFAMC outperformed Fi./Dual.
• Comparing Fi./Dual' with Fi./SFAMC, Fi./SFAMC was slightly better than Fi./Dual in Flower(M=1), and Fi./Dual' was slightly better than Fi./SFAMC in Mobile(M=1). In other cases, their performances were almost the same.

4. Effect of using local decoded images for ME and approximation of vector extension/shortening

In the dual-prime prediction, local decoded images are used in the second stage of motion estimation (ME) process, although the original images are used in the entire process of ME in SFAMC. Table 2 shows the performance of the dual-prime prediction when the original images are used in the whole ME process in the frame structure (M=3). The table shows that the amount of performance improvement using local decoded images was at most 0.04dB, in the frame structure (M=3).

In the current description of the dual-prime prediction in TM2, the motion vector extension/shortening operation is carried out by the integer division operator "/". Table 3 show the performance when it is aproximated as follows;

(recon * dist_$i) / dist_$j = (recon * (dist_$j + (dist_$i - dist_$j))) / dist_$j --> (recon * (2^L1 + (dist_$i -dist_$j) * N)) / 2^L1
where L1 = 4, N = (2^L1 // dist_$j); N has been calculated in advance.

The effect of approximation is small (less than 0.03 dB), as it has been already depicted in MPEG92/259, because the truncation process to the half pixel precision point cancels the error caused by this approximation in almost all cases.

Table 2 Luminance SNR[dB] in Dual' using original images in ME process

Frame structure m=3			
Prediction mode	Flower	Mobile	Bicycle
Fr./fi./Dual'	30.45(-0.03)	28.21(-0.01)	27.73(-0.04)

Table 3 Luminance SNR[dB] in Dual' using vector extension/shortening approximation with 4 bits

Frame structure m=3			
Prediction mode	Flower	Mobile	Bicycle
Fr./Fi./Dual'	30.45(-0.03)	28.21(-0.01)	27.76(-0.01)

5. Conclusion

The performance of the dual-prime prediction was evaluated, and it was compared with that of SFAMC. The simulation results showed that Both Dual' and SFAMC gave increasing in coding efficiency in combination with Frame and/or Field including low delay modes, and that their performance were about the same.

Although the precision of pel by pel basis interpolation in Dual-prime is lower than that in SFAMC, high performance can be achieved because of its ability to adapt to both increasing in vertical resolution like SFAMC and spatiotemporal loop filtering. (Refer to appendix A.2)

As for its implementation, Dual-prime has the following advantage.

1. Look up table is not necessary.
2. Pel by pel basis multiplication is not necessary.

We propose the dual-prime prediction mode, for its high prediction efficiency and its simplicity to the hardware implementation.

Appendix

A.1 Comparison between two prediction methods in field structure (M=2)

When Dual-prime prediction is applied to the field structure with m>=2, there are two candidates for the reference fields of the even field of the P-picture, as shown bellow.

(1) Reference field for "Pe1": "Po1", "Pe0"
dist_OtoE = 1;
dist_EtoE = 4;
(2) Reference field for "Pe1": "Po0", "Pe0"
dist_OtoE = 5;
dist_EtoE = 4;

Figure A.1 Field structure (M=2)

Odd fields:	Po0	Bo	Po1
Even fields:	Pe0	Be	Pe1

Table A.1 shows the simulation results using the two alternatives. The table shows that the prediction structure (2) was better than (1) in Flower and Bicycle. However, if the field structure is used in the compatible mode, the prediction structure (1) may be more appropriate than (2).

Table A.1 Luminance SNR[dB] using two prediction methods in field structure (M=2)

Reference	Flower	Mobile	Bicycle
(1) Po1, Pe0	29.61(+0.24)	26.54(+0.18)	28.07(+0.10)
(2) Po0, Pe0	29.76(+0.39)	26.68(+0.32)	28.05(+0.08)

A. 2 Effect of Dual-prime as increasing vertical resolution and spatiotemporal loop filtering

The operation of the dual-prime prediction is considered to be controlling balance between the effect of increasing vertical resolution and that of spatial-temporal loop filtering, by the differential motion vector "DMV".

If the averaging operation of the two fields is acted as increasing vertical resolution, both of the vertical components of the two motion vectors from the two fields to be averaged, are integer pixel precision (0, ±1, ±2,..).

On the other hand, If the averaging operation of the two fields is acted as spatial-temporal loop filtering, half pixel precision points (±0.5, ±1.5,..) are selected for the vertical components of the motion vectors.

Table A.2 shows the percentage of the selected prediction mode, in both Fr./Fi. and Fr./Fi./Dual', for the first P-picture with fixed Q-scales (QI,QP=9,QB=12). The dual-prime mode is further classified into the following three sub-modes, D-2, D-3 and D-4, for both even and odd fields to be coded.

In D-2 mode, both of the vertical component of the two motion vectors from the two fields to be averaged, are integer pixel precision. So, the prediction signal is made by averaging two scanning line signals. This mode seems to include the operation as increasing the vertical resolution like FAMC.

In D-3 mode, one of them is half pixel precision and the other is integer pixel precision. The prediction signal is made by averaging three scanning line signals with weighting.

In D-4 mode, both of them are half pixel precision. The prediction signal is made by averaging four scanning line signals. In this mode, the effect of spatial filtering seems to be maximized.

Table A.2 shows that large number of macroblocks are changed from Fr. or Fi. mode to Dual' mode for all three sequences. Furthermore, the percentage for D-3 or D-4 mode in Bicycle is large, compared with those of Flower and Mobile. This seems to be because of the effectiveness of spatial filtering for fast moving pictures like Bicycle.

Figure A.2 shows the selected modes for individual macroblocks in the first P-picture in Flower. D-2 is selected in "Flower garden" part to increase vertical resolution, and D-3 or D-4, which act as spatiotemporal loop filtering, is selected in "sky" area in the upper left and "moving tree" in the center.

Field prediction mode is also selected in the part where motion is not smooth.

From these observations, it seems that Dual-prime is adaptively used for both increasing the vertical resolution and spatiotemporal loop filtering.

Table A.2 Percentage of each prediction mode (P-picture) (%)

(a) Flower Garden

Prediction mode	Fr.	Fi	Intr.	Dual'	D-2	D-3	D-4
Fr./Fi.	37.3	58.2	4.5	-----	-----	-----	-----
Fr./Fi./Dual'	7.5	14.9	4.1	73.6	57.7	29.9	12.4

(b) Mobile and Calender

Prediction mode	Fr.	Fi.	Intr.	Dual'	D-2	D-3	D-4
Fr./Fi.	44.3	51.4	4.3	-----	-----	-----	-----
Fr./Fi./Dual'	19.5	28.0	3.9	48.6	53.2	36.9	9.9

(c) Bicycle

Prediction mode	Fr.	Fi.	Intr.	Dual'	D-2	D-3	D-4
Fr./Fi.	12.0	77.5	10.4	-----	-----	-----	-----
Fr./Fi./Dual'	5.1	47.4	9.1	38.3	19.4	49.9	30.7

Figure A.2 Selected prediction modes for individual macroblocks