Telecommunication Standardization Sector
Study Group 15
Experts Group for ATM Video Coding
(Rapporteur's Group on Part of Q.2/15)

Document AVC-544
July 1993

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/MPEG93/641

Source: I. Parke, M.E.Nilsson, BT Labs
Title: Frequency adapted spatio-temporal weighting
Purpose: Information and proposal

It should be possible to improve the performance of spatio-temporal weighting by adapting the weighting according to spatial frequency. The spatial prediction can only predict low frequencies while the temporal prediction can predict both low and high frequencies. This document describes an experiment done to investigate the effect of forming a prediction from the temporal prediction and the low frequency component of the spatial prediction.

Experiment

The calendar and cheerleader sequences were used. The lower layer was coded with MPEG-1 at 1.5Mbit/s, and the upper layer with an additional 2.5Mbit/s

Eight modes were available for each macroblock:

1)0, 0: All temporal prediction;
2)1, 0: Spatial prediction for field 0 and temporal prediction for field 1;
3)0, 1: Temporal prediction for field 0 and spatial prediction for field 1;
4)1, 1: All spatial prediction;
5)0.5, 0.5: Averaged temporal and spatial prediction;
6)f, 0: Frequency adapted weighting of predictions for field 0 and temporal prediction for field 1;
7)0, f: Frequency adapted weighting of predictions for field 1 and temporal prediction for field 0;
8)f, f: Frequency adapted weighting of predictions for both fields.

Only two bits of overhead were counted for indicating the choice of mode. This allowed comparison with previous results spatio-temporal switching. The mode with the lowest MSE was chosen.

Frequency weighted predictions were made as follows. A 1 2 1 filter was applied in the horizontal direction to produce low pass spatial and temporal predictions, pels at the left and right end of blocks being copied. High pass components were derived from the difference of the original and the low pass component. The frequency weighted prediction was made from the average of the two low pass components added to the temporal high pass component.

Results

The luminance SNR figures are compared in table 1 below with previous results where simple switching was used.

	Calendar	Cheerleader
Switching	28.35	29.09
Frequency weighted	28.48	29.25

Table 1. Luminance SNR figures.

The frequency adapted pictures were just noticeably better than the normally coded pictures. As some of the modes were chosen infrequently, it may be possible to retain the picture quality, but with the same number of overhead bits as normal.

The number of times that each of the modes was chosen is given below in table 2.

	Predicted pictures		Interpolated pictures	
Mode	Calendar	Cheerleader	Calendar	Cheerleader
0, 0	19727	10245	45863	34594
1, 0	1301	6212	2234	17683
0, 1	86	301	216	713
1, 1	4105	17894	1995	11922
0.5, 0.5	2696	4575	5471	12123
f, 0	15838	3565	17340	16286
0, f	1592	1440	3281	5745
f, f	3759	4608	4659	12364

Table 2. frequencies of each of the modes.

Conclusion

The simple method of frequency adapted weighting of spatial and temporal predictions used in this experiment has produced a small but noticeable improvement in picture quality.

It is proposed to define a CORE experiment to investigate this matter further. Additional lines of investigation include:

- different horizontal filter;
- the use of vertical filtering;
- making the prediction from the spatial low pass component and the temporal high pass component and
- relating the band split filter to the bi-linear upsampling filter as proposed in companion documents for use in the upsampling of the coded lower layer picture for spatial prediction.

End of document