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

Document AVC-440
January 1993

SOURCE : UK
TITLE : Specification of upsampling filters for hierarchical coding in H.26x/MPEG-2
PURPOSE : Discussion

The topic of upsampling filters for spatial scalability in H.26x/MPEG-2 was discussed at the Ipswich meeting in October 1992. Two approaches were identified:

1. The definition and specification of a (modest) limited number of individual upsampling filters, each tailored to a specific combination of scanning formats.

2. The specification of a means for downloading filter characteristics, for example FIR tap weights.

In a private communication to Mr. Morrison, some concerns over both these approaches were raised by Mr. Klenke (CLI). These included the lack of future extensibility of the former and possible implementation consequences of the latter.

He queried whether it is strictly necessary for the upsampling filters at coder and decoder to be identical. Such an assumption seems to be implicit in the above two approaches.

The tracking of these filters at coder and decoder has some similarity with that of the IDCT in H.261 and MPEG. Perhaps it is possible to define "theoretical target" filters and real implementations approximate to them. This allows choice of performance versus cost and continuous improvement of this ratio with advances in technology and innovative designs.

It is pointed out that although the upsampling filter at the decoder is indeed inside the decoding process, it is not within a recursive loop. Any mistracking error is therefore introduced only once, unlike the IDCT case where successive differences accumulate until intra mode is used. Further, errors will coincide with wanted picture detail which may well reduce their visibility.

There are several aspects which need further study before a decision can be made on the suitability of this approach. One is the response of the target filters. An ideal low pass is probably not a good candidate not only for VLSI implementation reasons but from a picture coding one too. Fast edges which are either present in the input image and well coded or which are introduced by coding distortion (eg block edges) in the layer being upsampled will cause an ideal low pass filter to ring appreciably. Such rings will be transferred into the next layer. Unlike normal single layer schemes where some ring on the decoded output may enhance the subjective quality, in multi layer schemes such rings are well within the total passband and undesirable. Thus it will be necessary to expend bits in the higher layer to encode an inverted version of the ring so that it can be cancelled.

Further consideration and study of this approach is encouraged.