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

Document AVC-467
Version 1
April, 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
MPEG 93/253
Sydney, March, 93

Title: **Inter working and complexity requirements on scalable video coding algorithms**

Source: **PTT Research (NL)**

Purpose: **Discussion**

1 Introduction

Transcoding inflicts unnecessary hardware costs and impairs the signal quality. Therefore, *it is important that transcoding is in principle not necessary in any AV-chain.* In this paper, we asses the consequences of this requirement, not only from the codec equipment but also from the medium (network, recording media) point of view.

I - Transcoding must be kept to an absolute minimum in any AV-chain
IIa - The codec equipment cost should be roughly proportional to the image size.
IIb - The data rate should be roughly proportional to the image size.

2 Factors that can obstruct inter working in an AV-chain and necessitate transcoding

Interoperatability

1. the image size or frequency differs along the chain (e.g. simultaneous broadcast to HDTV, EDTV, and LDTV television sets, or 50 Hz and 60 Hz CCIR 601),
2. incompatible standards are used (e.g. MPEG1, MPEG2, CCITT H.261),
3. incompatible subsets of the standard have been implemented (profiles),

Interconnect ability

4. the bitrates in one part of the chain are too high for a subsequent part (network bottleneck, recording of e.g. HDTV).

ad 1

At the European Seminar on Digital Television, the *European launching group for digital television*, supported by Philips, Thomson and many European broadcasters, put forward that it is considering to implement a digital television system that broadcasts at three different resolution levels,

- HDTV for fixed receivers, e.g. with a roof top antenna,
- EDTV for fixed receivers, with a antenna located on the television set,
- SDTV mobile receivers with an internal antenna.

ad 3

)

AVC-467

The number of profiles is must be kept to a minimum, and is preferably 1. If it is not feasible to limit the number of profiles to 1; profiles should be defined in a *hierarchical* fashion, i.e. the profiles can be ordered such that "larger" profiles contain all "smaller" profiles.

ad 4

A coded AV-bit stream should preferably contain an AV-sub stream *of much lower data rate* which allows the sub stream to be

- much heavier protected against transmission errors,
- forwarded through media of limited data transport capacity.
- This property, we denote as *bit stream scalability*. For the aforementioned applications of bit stream scalability it is crucial that the bitrate of the sub stream is of a much lower data rate, and should be for most applications of a fixed, controllable data rate.
- encryption of higher layers for pay-tv like services, with the lowest resolution as free preview channel

3 FOUR INTER WORKING REQUIREMENTS

From the inter working Requirement I it can be concluded that MPEG2 should unite the following properties,

Ia 3-level scalability,
Ib compatibility between MPEG2 and MPEG1 syntax (CCITT H.261 syntax),
Ic hierarchical profiles,
Id bit stream scalability.

The fourth point oftentimes does not get the attention it deserves, but is of great importance to network operators and for recording purposes. Requirement Id ties in with requirement IIb to an important constraint on the system.

The inter working requirements (Ia-d) and the complexity requirements (IIa,b) are to be understood as measuring sticks for compatible/scalable proposals for MPEG2.

4 JUDGEMENT OF PEL-SPLIT AND DCT-SPLIT SCHEMES BY THE REQUIREMENTS

Within MPEG, pel-split and DCT-split schemes have been developed as means to ease inter working. We use the terms pel-split and DCT-split to avoid confusion between properties and names of schemes ('spatial scalable'). The requirements Ia-d, IIa,b formulated above should play an important role in a fair comparison between pel-split and DCT-split schemes. So called pel-split schemes, allow for a side information video sequence from which blocks can serve as an alternative prediction to the motion estimated prediction (reference: several British Telecom, BBC, PTT Research contributions; TM4.2). DCT-split schemes make that the lowest 4x4 out of 8x8 DCT coefficients correspond to an image of half the resolution, and the lowest 2x2 coefficients correspond to an image of a quarter the resolution. Table 1 gives an overview, which requirements are satisfied by both families of schemes.

Requirement Scheme	Pel-split	DCT-split
Inter working		
Ia 3-level scalability	+	+
Ib Compatibility	+	-
Ic Few profiles	+	+
Id Bitrate scalability	+	0
Ie Code interlaced HDTV-interlaced TV	+	0[1]
Complexity		
IIa Proportional hardware	+	+
IIb Proportional bitrate	+	0
Flexibility		
IIIa Coding with side information (e.g. 3DTV)	+	-
IIIb Flexible image hierarchy	+	-
IIIc Modular hardware	+	-
Visual quality		
IIIIa Lower level image	+	-
IIIIb Better than simulcast efficiency	+	0

Table I Judgement of proposals by the requirements

DCT-split does not easily allow for compatibility with e.g. CCITT H.261 (Ib). For DCT-split schemes, the bitrate of the middle and lower resolution is difficult to adapt (Id, IIb), and for these schemes to work efficiently, the bitrate of the lower resolutions must be a large fraction of the total bitrate (IIb). Properties IIIa-c are discussed in the next paragraph. A point not noted in the table is that pel-split schemes can offer better quality of the down conversion than DCT-split can (image quality).

5 ADDITIONAL FLEXIBILITY OF PEL-SPLIT SCHEMES

Feature IIIa

Pel-split schemes can use the upscaled low resolution images as side information (viz. an alternative prediction) at the encoding of the next higher layer. This structure allows in fact for *any* video sequence to be used as side information. An important example currently studied within RACE2-DISTIMA is coding of stereoscopic television. Images from the left video sequence are used as side information for the coding of the right video sequence. In fact, a compensation for the 'disparity' of the location of objects as seen by the left and right eye is carried out on a block basis to improve the prediction value of the left video sequence. A single MPEG2 stream can carry multiple video windows. Pel-split schemes allow for these kinds of future applications of the MPEG2 standard, DCT-split schemes do not.

1 Low resolution image quality is poor.

Feature IIIb
Because any video sequence can be used as side information, there is no restriction on the ratio of the resolutions of the different levels, or the image frequencies. Therefore, pel-split has more flexibility with regard to extension to future image formats. In addition, pel-split schemes can offer a better down conversion than DCT-split schemes, for which the down conversion is done implicitly by the DCT.

Feature IIIc
Pel-split encoders/decoders, have an extra input for the side information images. As a result, an N-level encoder/decoder consists of N functionally identical blocks that operate at different SNR's, resolutions and/or image sizes. This is a modular structure.

DCT-split encoders/decoders require a more integrated design, that cannot be composed of standard blocks.

Feature IIIIa
The lower level image quality is for DCT-split schemes bound to the efficiency of a downconversion in the frequency domain.

Feature IIIIb
For all applications the Pel-split scheme works better then simulcast. The DCT-split scheme can not give this performance, see also "Evaluation of Scalable Coding Schemes" by Thomas Sikora.

6 Conclusion

DCT-split schemes
- are more difficult to regulate in bitrate at all image levels simultaneously,
- achieve efficient operation if the fractional bitrates of the lower layers is large with respect to the ratio (e.g. 1/4) of the image areas.

Pel-split based schemes
- satisfy all of the inter working and complexity requirements considered here,
- allow for *'coding with side information images'* at the encoder and decoder, which is a very powerful concept indeed (e.g. application to 3DTV, multi-viewpoint television),
- put no restrictions on the image formats in the hierarchy,
- implicitly have a modular encoder/decoder structure for the various levels.

In addition, pel-split schemes allow for better down conversions than those based on the DCT.