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

Document AVC - 479

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/SG29/WG11
MPEG 93/314
March 1993

Title: **Evaluation of Scalable Coding Schemes**

Source: Thomas Sikora

on behalf of Australian "Universal Video Communication" Consortium

Purpose: Information

## 1. Introduction:

The attempt of this contribution is to provide some insight into the functionality and restrictions associated with the two Scalability approaches, Frequency Scalabilitity and Spatial Scalability, currently discussed in MPEG. Both methods havē been developed within MPEG to provide layered coding with resolution scalability, yet both methods serve different purposes and in part support different application profiles. In this contribution we compare the two scalability approaches under various functionality requirements and service dependent aspects.

Although this draft does not attempt to serve as a comprehensive comparison we believe that it already provides a guide to the ability of either method to serve certain applications. It appears that both methods can essentially provide layered coding with resolution scalability. But the different emphasis of the methods, on flexibility and compatibility on one hand (Spatial Scalability), and efficiency and simplicity on the other hand (Frequency Scalability) clearly indicate, that the two methods mostly don't serve as alternatives to one and another, but rather suit different service profiles.

## Comparative Results and Conclusion:

The Table shown below allows some conclusions to be drawn. The Simulcast case is referenced as a benchmark for comparison purposes.

a.) Spatial Scalability:

Spatial Scalability has been developed to serve as a very flexible approach which incorporates features like compatibility with existing and possible future standards and can easily provide different input formats for services (e.g. interlaced and progressive layers can be supported on different layers).

With the incorporation of many different service functionalities in one scalable approach sacrifices have to be made in terms of coding efficiency and hardware complexity. A severe decrease of coding efficiency and a substantial increase in hardware complexity overhead, both compared to a MPEG2 stand-alone coder (single layer coder), seem to be inherent in the approach. Spatial Scalability layering schemes generally achieve a coding efficiency comparable or better than the Simulcast approach. The hardware complexity of the highest resolution decoder is substantial and more complex than would be required for a Simulcast solution in most applications. A multi-loop decoder is required for the reconstruction of the highest resolution video to retain the flexibility of the approach for features like panning and zooming.

b.) Frequency Scalability:

Frequency Scalability as described in the context of MPEG has been developed to provide scalable features with a coding efficiency and hardware complexity comparable or close to that of a MPEG2 stand-alone implementation. The general philosophy is that, as an extension to a MPEG2 stand-alone coder, scalable coding should not sacrifice significantly the quality of the highest resolution video nor place any significant hardware overhead burden to the MPEG2 decoder (in cases this can not be tolerated).

The encoder complexity can be tailored to the needs of the services requirements - scalability can be achieved by using one encoding loop or multiple encoding loops to encompass desired quality at lower scales. A single loop decoder is sufficient for the reconstruction of the highest resolution video for either single-loop or multiple-loop encoding option.

Features like compatibility with existing standards (like H.261 and MPEG1) have not been targeted to retain simplicity and are thus not supported. Scalability features which may be desirable for some broadcast applications, like panning and zooming for lower resolution video, are possible but are not implemented. Also very low bit rates at lower resolutions are not easily achieved. The method seems especially well suited for a number of computer video applications with progressive sources, i.e. Desktop Video Conferencing, Software decoding of lower resolution video etc..

Both Spatial Scalability as well as Frequency Scalability methods have shown good results for cell loss resilience for packet video applications. Again, Frequency Scalability offers a low complexity and highly efficient solution with scalable features and does not sacrifice the quality of the highest resolution video. Spatial Scalability can offer more flexibility with the possibility to encode lower layers compatible to existing standards. This may be an appropriate choice if quality reduction of the highest resolution video is acceptable and complexity is not an issue.

c.) Selection Criteria:

From the above discussion it is apparent that the functionality required by a particular service as well as the particular constraints have to be considered when selecting the appropriate layering scheme. It is also clear that some application profiles can only be satisfied by one or the other methods exclusively. There are also overlapping application profile areas. This is outlined in Figure 1 below:

I.e. if complete flexibility and compatibility are key issues Spatial Scalability and Simulcast are the only possible options (provided that implementation complexity and coding efficiency are considered less important). If high coding efficiency and low implementation complexity are of highest priority only Frequency Scalability can be the choice (provided that compatibility with existing standards and complete flexibility are not considered important). The possibility to implement software decoding again seems to be an option which may be much easier achieved using Frequency Scalability techniques.

Typical applications in the overlapping area could be i.e. Desktop Video Conferencing and HDTV/TV interworking where both methods are applicable.

| | Frequency | Spatial | Simulcast |
|---|---|---|---|
| **Interworking Capabilities** | | | |
| Flexible distribution of bit rates between layers: | yes (but constrained) | yes | yes[1] |
| Ability to drop all layers and code single layer: | yes | yes | yes |
| Ability to code in interlace-interlace mode: | yes | yes | yes |
| Ability to code in interlace-progressive mode: | possible[2] (not attempted) | yes | yes |
| Ability to code in progressive-progressive mode: | yes (can be very efficient) | possible (not attempted) | yes |
| **Coding Performance** | | | |
| Better than simulcast efficiency: | yes (for some applications) | yes (for almost all applications) | |
| Approaching single layer efficiency: | yes (if simplest scheme) | no (does not seem possible) | |
| **Network Issues** | | | |
| Can provide robustness against cell loss: | yes (simple scheme possible) | yes | possible |
| Ability to provide constant bit rate on all layers: | yes | yes | yes |
| **Compatibility with existing standards** | | | |
| Compatible with H.261 or MPEG1 on lower layer: | no | yes | possible |
| Compatibile with MPEG2 on lower layer: | no | yes | possible |
| **Hardware complexity** | | | |
| Additional hardware required to facilitate layering: | yes (can be very small) | yes (very large at this stage) | yes |
| Additional encoder cost/complexity: | small-high | high | high |
| Highest layer decoder cost/complexity[3]: | small | high | small |
| Lowest layer decoder cost/complexity: | can be very small | small | small |
| Lowest layer bitstream software decodeable: | seems possible[4] | does not seem possible | does not seem possible |

**Table I:** Comparison Between Frequency and Spatial Scalability Coding Schemes

---

1 Free distribution of bitrate is constraint in Simulcast. If all bits are allocated to the lowest layer no video is transmitted in the upper layers unlike in the other layering approaches..

2 Technically feasable but currently not provided by the synatx.

3 As compared to a MPEG2 stand-alone implementation (single layer).

4 This is reslated to computing power available in a workstation environment today. Future technology developments may also enable more complex schemes, i.e. Spatial Scalablility approaches, to be decoded in real time using software only.

| | Frequency | Spatial | Simulcast |
|---|---|---|---|
| **Functionality** | | | |
| Short encoding delay possible: | yes | yes | yes |
| Short decod.ng delay: | yes | yes | yes |
| Panning: | possible | yes | yes |
| Zooming: | possible (but constrained) | yes | No |
| Possibility to have minimum complexity at encoder: | yes | no | |
| Possibility to have a single loop decoder: | yes | possible (but not attempted) | yes |
| Fleibility to have complexity/quality tradeoff at encoder and decoder:[5] | yes | ? | possible |
| Does the family of schemes suport downward compatibility:[6] | yes | ? | ? |
| **Applications** | | | |
| Suitable for Desktop video-conferencing: | yes | yes | yes |
| Suitable for HDTV-TV interworking: | yes | yes | yes |
| Is it possible to incorporate future services: | constrained[7] | yes | yes |

**Table I cont.:** Comparison Between Frequency and Spatial Scalability Coding Schemes

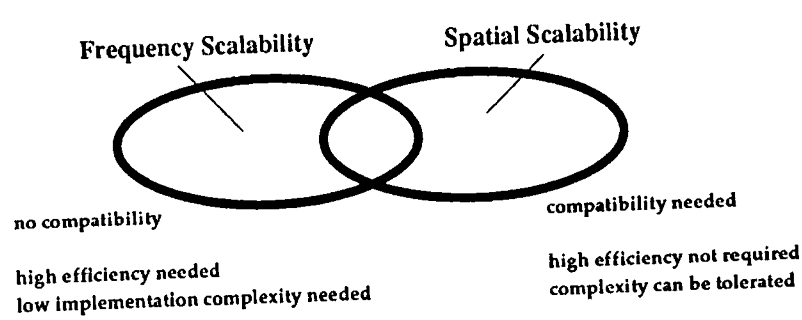


**Figure 1:** Virtual applications areas for Frequency and Spatial Scalability Coding Schemes.

---

5 We assume that a syntax defines a family of coding schemes which differ in implementation complexity (as FrequencyScalability does). This functionality may allow that quality of a service can be traded off with implementation complexity.

6 Downward compatibility as defined in this context is the possibility to achieve full downward compatibility of coders with one family. Decoders with higher implementation complexity are downward compatible to schemes with lower complexity

7 Future higher resolution services would require a DCT size of higher dimension which may have implications for motion compensation as well as DCT implementation complexity. More work is needed!