ISO/IEC JTC1/SC29/WG11
MPEG93/004
January 1993

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

Document AVC-409
21-22 January 1993

Source: UK
Title: Compatible coding
Purpose: Discussion

1. Introduction

This document describes how a compatible approach to video coding provides many advantages to manufacturers, service providers and consumers. It then suggests that the concept of compatible coding can be applied to the whole range of applications to achieve maximum benefit.

Different approaches to compatible coding using a layered approach are discussed. It is proposed that a layered approach with loose coupling between the layers provides the optimum performance from a service point of view, providing more efficient bandwidth utilisation than the un-coupled simulcast approach and more flexibility in the choice of source formats and algorithms than more tightly coupled approaches.

A demonstration of a four layer loosely-coupled system using the MPEG Test Model 3 algorithm is given. This is shown to provide significantly better quality than the corresponding simulcast system.

2. Advantages of compatible coding

There are many advantages to manufacturers, service providers and consumers of adopting a compatible approach to video coding. These are summarised below.

Compatible coding protects investments of consumers and manufactures of hardware and software against developments in video compression technology. It allows manufacturers to develop a range of different products with different performances that will all work to the same coded material. Compatible coding enables conditional access of different picture qualities which allows program generators to charge differently for these different picture qualities. It also allows interworking of terminals connected to networks of different capabilities, for example, in the multipoint videoconference application, all terminals can transmit at their peak capabilities rather than being forced to the highest common standard.

3. Methods of achieving compatible coding

There are many ways to achieve compatible coding, but all are based on the concept of coding layers: some bits are decoded to generate a low quality sequence and additional or other bits are decoded to generate a higher quality sequence. This does not imply that the layered bitstreams are transmitted separately or even that they can be separated.

This section describes three forms of layered coding and discusses their advantages and disadvantages.

3.1. Simulcast coding

The simplest method of achieving compatible coding is the simulcast approach. In this scheme, all layers are coded independently. This has the advantage that the coding schemes used for each layer can be chosen and optimised independently but is inefficient in bandwidth. It is indicative of the lowest quality that can be achieved by a layered scheme.

3.2. Tightly-coupled layered coding
Another approach is to have tight coupling between the layers, that is, the coding scheme used in each layer is related. In these approaches the up and down-sampling of source formats is actually part of the coding scheme. Examples of such schemes are the so-called 'frequency scalable' and 'dct split' schemes. This is more efficient in use of bandwidth than simulcast but places many constraints on the layers, both in terms of the coding algorithms that can be used and the resolutions and picture rates that can be coded. This scheme can be made resilient to bitstream errors by protecting the layers appropriately.

3.3. Loosely-coupled layered coding
A third approach is to have loose coupling between the layers, that is, the coding algorithms used to code the layers can be chosen independently, but use is made of the reconstructed pictures produced by lower coding layers. Examples of such schemes are the so-called 'spatial scalable' and 'pel split' schemes. This is more efficient in use of bandwidth than simulcast and more flexible in its possible uses than tightly coupled systems. The coding scheme used for the different layers can be chosen independently, as can the resolutions for the layers, and the particular methods of up and down-sampling. This scheme can be made resilient to bitstream errors by protecting the layers appropriately. Experimental results to date have shown that loosely-coupled layered coding uses bandwidth as efficiently as tightly-coupled systems.

4. The vision for layered coding
It is proposed here that all future video coding schemes should be developed with layering in mind. Our vision is that this would allow compatibility to be achieved over the whole range of video services from PSTN videophone to broadcast HDTV. This does not imply that receiving a low quality version of a HDTV broadcast on a mobile video terminal is imagined to be a highly desirable service, but rather that compatibility is desirable between applications using neighbouring segments of the above bandwidth range. For example, compatibility between PSTN/mobile videophone and ISDN videophone, between ISDN video phone and video conference, between video conference and business TV, and between conventional broadcast TV and HDTV may be extremely desirable. To achieve this, compatibility is required over the whole range of bandwidth.

It is proposed here that the best method for achieving this vision for layered compatible coding is the loosely-coupled layered coding scheme described above. This is flexible in that it places no restriction on the definition of any coding algorithm except that it may take advantage of 'prediction' from a reconstructed picture of a lower layer when it is beneficial to do so. This allows greater efficiency of use of bandwidth than simulcast, with almost as much flexibility. In the extreme case of not choosing to use the reconstructed lower layer picture, this scheme becomes the simulcast scheme.

As stated above, there is no restriction on the coding scheme used in each layer, except that it may use prediction from a reconstructed lower layer picture whenever it is beneficial to do so. In the future, schemes based on knowledge based algorithms, vector quantization, or fractals, etc. could be developed and achieve compatibility with existing schemes.

5. Demonstration
This section describes a simulation of four layered loosely-coupled coding to demonstrate the principles outlined above. The aim is to demonstrate the possibilities of layered coding rather than the state of the art of video coding at any of the bit rates and resolutions shown.

5.1. Source formats
The original sequence used was the CCIR 601 resolution Susie sequence. This is 704 pels by 576 lines, interlaced, at 25Hz.

This was down-sampled to a progressive SIF format by use of a three field aperture vertical-temporal filter. This is 352 pels by 288 lines, progressive, at 25Hz.

This progressive SIF was filtered and subsampled by a factor of two vertically and horizontally and subsampled by a factor of two temporally to produce a QSIF format. This was 176 pels by 144 lines, progressive, at 12.5Hz.

This progressive QSIF was filtered and subsampled by a factor of two vertically and horizontally and subsampled by a factor of two temporally to produce a QQSIF format. This was 88 pels by 72 lines, progressive, at 6.25Hz.

5.2. Encoder description

The encoder software used for this demonstration was based on MPEG Test Model 3. The value of N was 124

The lowest resolution layer, the QQSIF, was coded with the Test Model running without the extended syntax, that is, in accordance with the MPEG-1 standard. The bit rate was 32kbit/s. All frames were coded. Motion estimation was done as a full search on coded pictures over a +/-7.5 search range. Only predicted pictures were coded.

The next layer, the QSIF, was coded with the Test Model running with the extended syntax. Two runs were performed: one compatible run using prediction from the up-sampled coded QQSIF and another as simulcast. The QQSIF was up-sampled by horizontal and vertical interpolation and frame repeating; this format is shown on the demonstration tape. The bit rate was 96 kbit/s for this layer. All frames were coded. Motion estimation was done as a full search on coded pictures over a +/-7.5 search range. Predicted and interpolated pictures were coded with M=2.

The next layer, the SIF, was coded with the Test Model running with the extended syntax. Two runs were performed: one compatible run using prediction from the up-sampled compatibly coded QSIF and another as simulcast. The QSIF was up-sampled by horizontal and vertical interpolation and frame repeating; this format is shown on the demonstration tape. The bit rate was 384 kbit/s for this layer. All frames were coded. Motion estimation was done as a full search on coded pictures over a +/-4.5 search range around vectors calculated on source pictures using a full search over +/-16 integer pels. Predicted and interpolated pictures were coded with M=4.

The top layer, the CCIR 601, was coded with the Test Model running with the extended syntax. Two runs were performed: one compatible run using prediction from the up-sampled compatibly coded SIF and another as simulcast. The SIF was up-sampled by horizontal interpolation and vertical-temporal interpolation; this format is shown on the demonstration tape. The bit rate was 1536 kbit/s for this layer. All frames were coded. Motion estimation was done as a full search on coded pictures over a +/-4.5 search range around vectors calculated on source pictures using a full search over +/-16 integer pels. Predicted and interpolated pictures were coded with M=3. Adaptive rate control was used. The coded pictures are shown at actual resolution on the demonstration tape.

5.3. Demonstration Tape and Results

All the coded layers except the highest are shown at double their coded horizontal, vertical and temporal resolutions to avoid the need to view excessively small pictures while keeping the size on the display in proportion to the actual resolution.

As stated before, the aim of this demonstration is to show the flexibility and benefits of layered coding, rather than state of the art picture quality. The algorithms used are unlikely to be optimum, especially at the lower bit rates where the amount of overhead is high. Consequently the picture quality that can be observed is not the best obtainable.

There is noticeable difference in picture quality between the compatible coding simulations and the simulcast simulations, especially in the middle of the sequence where the motion is rapid and difficult to predict. A summary of SNR figures is given in tables 1 to 4.

6. Conclusions

This document has explained how a compatible approach to video coding provides many advantages to manufacturers, service providers and consumers and that these advantages can be maximised by extending the concept of compatible coding to the whole range of video coding applications.

Different approaches to compatible coding using a layered approach were discussed. It was proposed that a layered approach with loose coupling between the layers provides the optimum performance from a service point of view, providing more efficient bandwidth utilisation than the un-coupled simulcast approach and more flexibility in the choice of source formats and algorithms than more tightly coupled approaches.

A demonstration of a four layer loosely-coupled system using the MPEG Test Model 3 algorithm was given. It was shown that this can provide significantly better quality than the corresponding simulcast system.

7. Tables of results

	Compatible		Simulcast	
	Predicted	Interpolated	Predicted	Interpolated
Number of pictures	30	-	30	-
SNR for Y	36.36	-	36.36	-
SNR for U	41.08	-	41.08	-
SNR for V	40.80	-	40.80	-
Mean value of QP	5.21	-	5.21	-

Table 1. Coding statistics for QQSIF. Bit rate 32 kbit/s.

	Compatible		Simulcast	
	Predicted	Interpolated	Predicted	Interpolated
Number of pictures	30	30	30	30
SNR for Y	37.31	35.49	36.93	35.05
SNR for U	43.30	42.92	43.08	42.90
SNR for V	43.15	42.86	43.09	42.99
Mean value of QP	5.01	12.99	5.58	14.66

Table 2. Coding statistics for QSIF. Bit rate 32 + 96 kbit/s.

	Compatible		Simulcast	
	Predicted	Interpolated	Predicted	Interpolated
Number of pictures	30	90	30	90
SNR for Y	39.86	38.20	39.35	37.66
SNR for U	45.23	44.79	44.85	44.57
SNR for V	45.21	44.78	44.87	44.62
Mean value of QP	4.51	12.27	5.44	14.82

Table 3. Coding statistics for SIF. Bit rate 32 + 96 + 384 kbit/s.

	Compatible		Simulcast	
	Predicted	Interpolated	Predicted	Interpolated
Number of pictures	40	80	40	80
SNR for Y	40.26	39.30	39.12	38.38
SNR for U	45.58	45.07	44.89	44.60
SNR for V	45.70	45.14	44.93	44.64
Mean value of QP	4.00	10.93	6.63	16.91

Table 4. Coding statistics for CCIR 601. Bit rate 32 + 96 + 384+ 1536 kbit/s.