CCITT SGXV
Working Party XV/1
Experts Group for ATM Video Coding

Document AVC-134
November 1991

SOURCE: Australia

TITLE: Picture Format Considerations

PURPOSE: Proposal

Abstract

In this document the implications of a flexible range of picture formats are discussed. A number of questions, raised in the previous meeting report [AVC-106R], are addressed. Flexible spatial resolutions[1], which lie within some maximum resolution limits, allow the coding system to be used in a wider range of applications and do not significantly impact on codec complexity. A limited subset of frame rates, consisting of those rates which are related to the maximum by integer factors, can also be dealt with reasonably easily. The ability to code sources which have frame rates not related to the 59.94 Hz coder rate, proposed for SCIF, by integer factors, is essential.

1. Introduction

The Experts Group are developing a coding algorithm for B-ISDN which should be useful in a wide range of applications. A large number of potential applications have already been identified [AVC-109] and it is recognised that others will emerge in the future. In a codec intended for a range of applications, flexibility in certain picture format parameters is essential.

The Experts Group have identified several issues regarding the use of a flexible set of formats, in section 4.2 and Annex 3 of AVC-106R. These are:

- *Fixed Decoder*. Implying that flexible formats will require a flexible decoder, which is potentially more complex.
- *MCU easier*. Implying that the MCU, which may be required to combine signals from separate sources into a single coded stream may be less complex for a fixed format.
- *Windows easier*. Implying that it will be easier to display multiple decoded video sequences on a single monitor if a single format is chosen.
- *Equal Burden*. Precisely what this means is difficult to understand. It is certainly not true that both regions have equal burden in converting to the SCIF format.
- *Display with only one conversion*. Implying that multiple conversions may be required to display signals with different formats on the same display.
- *Format negotiation is not required* (see section 4.2, AVC-106R, where this is listed as a potential disadvantage of the flexible formats option).

In the next section the argument for integrated terminal access is presented. It is argued that, to provide integrated access to a wide range of applications, requires a careful choice of picture formats which balances commonality and flexibility. Following this, the use of flexible and fixed formats is discussed separately for spatial resolutions and frame rates. This separation is important because the impact of introducing flexibility is different in these two cases.

2. Applications and Timeframe

The B-ISDN will be an integrated network which is used to carry the full range of communications services. This will include traditional telecommunication services, such as voice and video communications, broadcast services such as TV and HDTV and remote access services, such as video libraries. It makes little sense to develop an integrated network, without also integrating

[1] In this document spatial resolution is used to mean the number of horizontal and vertical pixels.

access to services so that a single terminal can access all services. Hence, in the future, when B-ISDN becomes a reality, services are unlikely to be accessed through purpose built terminals. In a working environment they are more likely to be accessed by a single desktop workstation, which is also used for other applications. This single workstation will provide point-point and multi-point communications as well as access to video broadcast information services, and remote video information databases.

This vision of an integrated terminal is already starting to become a reality with the introduction of ISDN. The likely timeframe for the introduction of the first fully capable B-ISDN equipment is 1995 and during this time the integrated terminal concept is likely to become well accepted and commonplace.

A crucial component in integrating the access to video services is the video coding algorithm. If different, incompatible coding algorithms are developed for each service, then integrated access will be expensive, and the potential of B-ISDN will be wasted. The first step in ensuring that integrated access is possible is to ensure some commonality in aspects of the picture formats which are coded. To reach the goal of wide application, it is also important to provide flexibility. The final choice requires a balance between commonality and flexibility to meet the dual goals of low terminal cost, and wide applicability. The Experts Group must bear in mind the implementation capabilities which are likely to be available in the 1995 timeframe when considering codec complexity issues.

3. Spatial Resolutions

In previous documents Australia has indicated it supports the concept of flexible spatial resolutions for coding. This capability is essential for dealing with a range of applications. A set of maximum resolutions should be defined for each service class. These resolutions represent the capabilities of the *decoder*. Coders can then use any resolution, from an appropriately defined subset, within this maximum. In addressing the points raised by the Experts Group (see introduction), as they impact on flexible spatial resolutions, Australia makes the following points:

- Decoders are assumed to be capable of decoding at the maximum resolution for the service class. Providing decoders with the capability to decode at some lower resolution will not increase complexity. Note that the MPEG 1 algorithm has this capability.
- The additional complexity in the MCU is only incurred if conversion from the coded format to the maximum of SCIF is used. This is unnecessary.
- The use of flexible spatial resolutions has no impact on the complexity of display in multiple windows on the same screen.
- The requirement to rescale pictures coded at a low resolution, to display them on a high resolution monitor, is not a part of the decoder, and would probably not be used since the decoded stream will be displayed in a window on the terminal. Resizing of windows, if implemented, is a terminal function and should not be subject to standardisation.
- The coder should signal the coding spatial resolution to the decoder. However, no negotiation is required since decoders will be capable of decoding at the maximum resolution. This functionality is already provided in the MPEG 1 algorithm.

In addition to these points, the question of suitable maxima and suitable subsets for the flexible formats option has been raised. Australia makes the following points:

- The maximum resolution chosen in the 601 class depends on whether wide-screen TV is included. The following maximum resolutions provide a square pixel format which also retains the maximum number of lines in CCIR Rec. 601:

  768x576 (4:3 aspect ratio)

  1024x576 (16:9 aspect ratio)

- The chrominance/luminance sampling patterns proposed for SCIF imply that an appropriate subset of spatial resolutions, less than the maximum, have dimensions which are multiples of 16.

4. Frame Rates

Flexibility in coded frame rate is also essential if the codec is to be suitable for a range of applications. For example, the quality requirements of video-phone services can be met by frame rates much lower than 59.94 Hz. There is a wide range of material, which will need to be coded, that is not available at a rate of 59.94 Hz, including film (24 Frames/sec) and European (and Australian) video sources. Again, it is important to stress that codecs should be capable of decoding at the maximum frame rate for their service class. However, a codec should be free to code at a frame rate, chosen from a defined set, which is suitable for the application in which it is used. The rate at which the sequence is displayed is also a matter of choice and should not be a part of the standard. In addressing the points raised by the Experts Group (see introduction), as they impact on flexible frame rates, Australia would like to make the following points:

- Within CCITT Rec. H.261, the frame rate defined for CIF is a maximum, rather than a fixed number. Coding at frame rates which, when multiplied by integer factors, give the CIF frame rate, is straight-forward and used extensively for low-rate video-phone coding. This type of flexibility must be a part of the new standard and its inclusion will not add significantly to decoder complexity. The need to decode and display at frame rates not related by an integer factor will introduce additional codec complexity if techniques more sophisticated than frame dropping or frame repeat are used for conversion.
- As above, the simple frame rate flexibility which is a part of H.261, should not effect MCU complexity significantly, while the use of more sophisticated conversion techniques will have an impact on MCU complexity. It should also be noted that the B-ISDN offers increased functionality for the provision of multipoint communications, and different functionality might be required of an MCU in this environment (see AVC-37).
- Displaying multiple windows on a single screen will require frame rate conversion, assuming a CRT display is used (LCD storage displays may be more flexible in this regard). Again, the additional complexity of this conversion depends on whether simple frame dropping/frame repeating or sophisticated motion compensated interpolation is used.
- Conversion from a range of different frame rates for display would not be particularly complex if frame dropping/frame repeating is used.
- Negotiation of a frame rate is not required if codecs are capable of decoding all allowable frame rates within the service class.

The obvious conclusion from these points is that simple frame rate conversion techniques can be used which have little impact on complexity. However, they will have a significant impact on the quality of the service. In considering service quality, Australia makes the following points:

- The conversion between different field/frame rates, without quality loss is an unsolved technical problem. Therefore, to meet the quality requirements of some applications (e.g. Broadcast TV, video retrieval) the coder must provide coding at sub-multiples of both 50 Hz and 59.94 Hz frame rates. Other frame rates, such as 24 Frames/sec (film), may be necessary to deal with a range of applications.
- The use of flexible frame rates will satisfy the quality requirements for intra-regional services. However, if unsophisticated frame rate conversion is used, inter-regional service quality will be compromised.
- For applications with less stringent quality requirements it is possible to code a 50 fields/sec (or frame/sec) sources using the flexible frame rate capabilities of CCITT Rec. H.261 and decode the sequence for display on either a 50 or 59.94 fields/sec (frames/sec) monitor. The conversion of frame rates can be achieved by simple frame dropping or frame repeats, and would not add complexity to codecs. It remains an open question whether this straight-forward flexible frame rate capability can meet the quality requirements of all intra and/or inter-regional services.

In considering the complexity which is introduced by using multiple frame rates, as opposed to a single frame rate, it is important to note that the use of multiple frame rates places equal conversion burden on both regions, since coders from both regions would employ conversion for inter-regional communications. The alternative of SCIF places the burden for conversion in codecs used in 50

fields/sec regions. An additional complexity burden would also be placed on MCU's which may need to be capable of conversion from multiple rates to a single output rate. This complexity burden will depend on the conversion quality which is implemented. As suggested above, simple techniques which do not introduce significant additional complexity can be employed. The alternative of a single SCIF frame rate places the burden of conversion on 50 fields/sec countries.

5. Conclusion

Australia supports the concept of flexible spatial resolutions and frame rates. These capabilities are seen as essential if the coding system developed is to be suitable for a wide range of applications. It has been stated that providing flexible spatial resolutions does not increase the complexity of codecs significantly. Providing the simple flexible frame rate capability used in CCITT Rec. H.261 must be a part of the new standard codec, particularly if backward/forward compatibility with this standard is required. Whether or not this capability alone can satisfy the requirements of all applications in both the regional source field/frame rates is an open question which needs further study.