ITU Telecommunication Standardization Sector
Study Group 15
Q.2&3/15 Rapporteur Meeting
Eibsee, 24-27 September 1996

Document AVC-1038
Version 1
18 September, 1996

SOURCE : BT
TITLE : Variable bit rate transmission of audiovisual conversational services on ATM Networks
PURPOSE : Discussion

1. Introduction

This contribution overviews the basic principles of the H-series video encoding algorithms, H.261, H.262 and H.263, explains why the use of constant bit rate transmission limits the quality of service that can be provided to audiovisual conversational applications, and then explains how these limitations can be overcome by the use of variable bit rate transmission if the bit rate can be varied over a sufficiently long period of time.

Two possible means to support variable bit rate transmission, and the consequences of each, are described; and suggestions are made for the work that SG15 should do in order to produce a recommendation for variable bit rate audiovisual conversational services on ATM networks.

2. Overview of H-series video encoding algorithms

The basic principles of the H-series video coding algorithms, H.261, H.262 and H.263, are illustrated in figure 1. The algorithms are hybrids of motion compensated temporal prediction and discrete cosine transform (DCT) based coding. Temporal redundancy, present in stationary areas and regions with linear motion, is removed by subtracting from the input picture a prediction made from offset parts of previously coded pictures. The resulting prediction error is transformed to the frequency domain using the DCT. Although the DCT does not of itself give any data reduction, the values from it are better suited to quantisation, and hence compression. The resulting quantised data, together with the displacement vectors used in the prediction process and various overheads, are variable-length coded for transmission.

3. The limitations of constant bit rate transmission for audiovisual conversational services

The H-series algorithms, as shown in figure 1, inherently produce data at a non-constant rate: still areas produce few bits, moving detail produce the most. When constant bit rate transmission is used, a smoothing buffer is required between the coding kernel and the transmission line.

The size of the smoothing buffer determines the period of time for which the encoder can run 'open loop', that is, without changing the quantisation parameter and consequently changing the picture quality. A large buffer could provide less sudden variations in picture quality, but would add significantly to the end-to-end delay. For conversational applications, the buffer size is typically such as to cause a delay of about 150 ms.

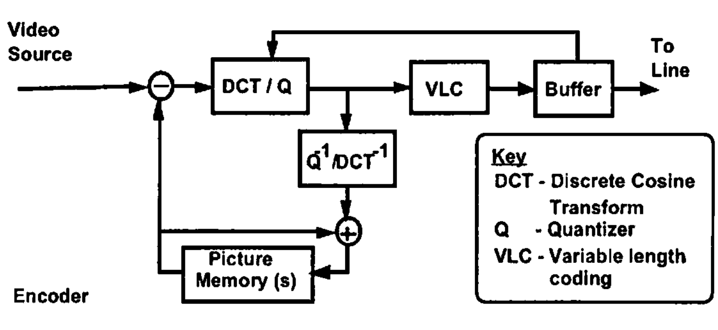

Figure 1. The basic principles of the video coding algorithm of the H-series recommendations.

4. The potential benefits of variable bit rate transmission for audiovisual conversational services

If variable bit rate transmission is used, then theoretically the buffer could be removed, and the coding kernel connected directly to the network.

The improvement in quality of service perceived by the user depends on the amount of variability allowed for the transmission bit rate. Variations in bit rate within a short time period, for example variations such that about the same number of bits would be transmitted in each non-overlapping period of 33 or 40 ms, would not allow any noticeable improvements in quality of service to be achieved: either the smoothing buffer is retained and performance is exactly the same as for constant bit rate transmission, or the buffer and smoothing delay are eliminated but the capability to avoid sudden changes in picture quality is lost.

The quality of service perceived by the user could be improved if the bit rate could vary within non-overlapping periods of about 150 ms, even if the total number of bits transmitted in each of these periods were about the same. This would allow the same bits to be transmitted as in the case of constant bit rate transmission, and the same pictures to be reconstructed by the receiver, but the smoothing delay of 150 ms would be eliminated.

While allowing bit rate variations within a window of 150 ms would allow most of the end-to-end delay to be eliminated, the picture quality would still vary in time: the quality would be good when there is little activity in the scene (still pictures or ones with little detail), but the quality would deteriorate when there is more activity (fast moving pictures or ones with detail). The picture quality in this case would be exactly the same as in the case of constant bit rate transmission, but the end-to-end delay would be reduced.

Constant picture quality could be achieved if bit rate variations were possible on the timescale of scenes in video sequences, that is, if bit rate variations could take place over tens of seconds or even minutes and hours.

5. Possible means to support variable bit rate transmission

One means to support variable bit rate transmission in the network is to allocate terminals their peak bandwidth requirement. This would guarantee a quality of service to the application at least as good as for constant bit rate transmission, but would be wasteful of bandwidth, and therefore expensive, unless the terminal could make use of the bandwidth that had been allocated, but not used by the video encoder, for other applications.

Another means to support variable bit rate transmission in the network is to allocate terminals less than their peak bandwidth requirement, and achieve statistical multiplexing in the network. This would make efficient use of bandwidth but may lead to reduced quality of service. There will be problems if a large number of terminals, connected to the same network link, simultaneously transmit at or near their peak rate. Network nodes contain cell buffers to smooth short term variations in bit rate, but may occasionally overflow with the consequent loss of one or more cells. Although the use of large cell buffers could reduce the amount of cell loss, they would cause delay that would be unsuitable for conversational services. A related problem is that of increased cell delay variation (CDV). Depending on the fullness of the network buffers, cells will experience varying waiting times at the network nodes, and hence have varying transmission delay through the network, the range of these delays being known as CDV.

6. Work required to produce a recommendation for variable bit rate audiovisual conversational services

Only a minimal amount of work would need to be done to produce a recommendation if peak bandwidth is allocated: additional DSS2 signalling codepoints would be required, as well as additional codepoints in H.245.

The remainder of this section relates to the case of allocating less than peak bandwidth.

6.1. Traffic characteristics, Service class and Quality of Service

It will be necessary to understand the typical Quality of Service that would be provided by the network when less than peak bandwidth is allocated. In particular, it will be necessary to know the amount and nature of cell loss and cell delay variation.

It should be possible to achieve Quality of Service estimates by simulation. The traffic characteristics of variable bit rate encoded video could be obtained, from simulation or real-time hardware, and input to network loading models together with a mix of other traffic, such as constant bit rate services, to produce estimates of cell loss ratio and cell delay variation.

When the video traffic characteristics are known, it should be possible to select a suitable service class. For example, real-time VBR or ABR (available bit rate), or perhaps a new class may need to be defined.

6.2. Network adaptation

When the achievable Quality of Service is known, appropriate network adaptation should be defined. This includes the selection of an AAL, and the mapping of audiovisual information into the AAL-PDU.

Network adaptation functionality may be needed in addition to that provided by the AAL in order to provide a satisfactory Quality of Service to the application. This could include functionality to detect and/or correct cell loss, and to reduce the effect of cell delay variation.

The Network adaptation could be defined to make use of more than one ATM VC for audiovisual information. For example, information requiring a high quality of service, such as audio and timing information, could be transmitted on one VC, perhaps at constant bit rate, while other information, such as video, could be transmitted on a VC that provided lower quality of service.

6.3. DSS2 Signalling and H.245 negotiation

When the service class has been selected and the achievable Quality of Service is known, signalling parameters should be defined. This would probably require additional DSS2 signalling codepoints, as well as additional codepoints in H.245.

END