Issues in Video Transmission Over Broadband ATM Networks

Tony L. Mitchell  Steven L. Blake

Center for Communications and Signal Processing
Department of Electrical and Computer Engineering
North Carolina State University
Raleigh, NC  27695-7914

Abstract

Broadband ATM networks at multi-megabit data rates will provide the necessary bandwidth for high-quality interactive video applications. Performance requirements for high quality video services are introduced. Results of statistical simulations of broadband networks transmitting video signals are shown to demonstrate that VBR coding provides gains in multiplexed channel utilization and in access link capacity. Issues relating to protocol design for ATM networks, particularly congestion control and error control are introduced. Finally various techniques for handling cell loss are compared.

1 Introduction

Current proposals for the Broadband Integrated Services Digital Network (BISDN) specify a user-network interface at ~ 155 Mbps, using the Asynchronous Transfer Mode (ATM) as the packet switching protocol. The key feature of the ATM protocol is the formatting of all user data into fixed size packets, named cells, which are 53 octets long (5 octets of header and 48 octets of payload). An advantage of the ATM is that the user can dynamically allocate bandwidth in cell increments to various applications communicating over the BISDN link. The BISDN will be designed to guarantee the order of cell delivery, to provide low end-to-end delay, and to provide a high grade of service (low probability of bit error and cell loss) [1, 2, 3].

The introduction of the BISDN as a ubiquitous wide area network (WAN) service will facilitate the development of advanced broadband applications such as high definition television (HDTV) distribution, videoconferencing, and multimedia database access. These video applications each impose service requirements which are difficult or impossible to satisfy with current local area networking (LAN) technologies (i.e.: Ethernet, FDDI). The desire to support broadband applications at the desktop has led to recent proposals for utilizing ATM technology in next-generation LAN architectures [4, 5]. An ATM-based LAN could support video applications which seamlessly integrate with the public-switched BISDN network.

The choice of the ATM protocol has an impact on the design of video coding algorithms for use over broadband networks. Advances in VLSI video compression processors promise to allow a single BISDN link to handle many simultaneous video connections at high quality levels (HDTV equivalent). The possibility of cell loss due to network congestion or misdelivery can have a severe impact on video services. The real-time requirements of video transmission also have an impact on the network protocol design. In this paper the impact of the ATM protocol on video transmission is discussed. In particular, the advantages of variable bit rate (VBR) coding are described. Also, techniques for preventing network congestion and minimizing the impact of cell loss on video service quality are discussed.

2 Network Performance Requirements

An examination of the coding algorithms currently being proposed to the FCC for broadcast transmission of digitally encoded HDTV will highlight some of the important issues faced in the digital transmission of video signals. High compression ratios are required to provide mega-pel resolution at large signal-to-noise ratios (SNR), due to the narrow bandwidth of the allocated broadcast channel (6 MHz) [6]. The bit rate of each proposal is approximately 20 Mbps. The constant bit rate of the broadcast channel requires coding algorithms that produce a data stream of constant bit
rate (CBR). This is not inherent in any of the hybrid coding algorithms proposed. In fact, restricting the output bit rate to a constant level can lead to severe distortion on frames with significant motion, and can lead to wasted data capacity in regions of little motion. This results in a video image with time varying SNR, where the SNR decreases significantly for frames with large motion. This problem is partially masked by the fact that the human perception of spatial resolution decreases in regions of temporal motion, but it is a noticeable problem for coders with high compression ratios. This problem is usually addressed by adding an elastic rate buffer at the interface between the coder and the channel interface. The goal is to absorb variations in the coder's output bit rate. These buffers are included in the HDTV coder proposals. The benefits of rate buffers fail when they overflow due to sustained motion. This problem is often addressed by adaptively reducing the number of quantizer levels in the coding algorithm as the buffer fills up.

To quantify the performance characteristics that must be provided by a broadband ATM network to support high quality video applications, it is first important to estimate the necessary bandwidth that must be supplied for each service. The current HDTV coder proposals should be suitable for high resolution video distribution applications and will require ~20 Mbps average bandwidth. The MPEG video coding algorithm targets multimedia data applications at a bit rate of ~1.5 Mbps [7]. Current videoconferencing systems utilize bit rates ranging from 56 Kbps to 768 Kbps. Videotelephony systems are being developed at data rates as low as 64 Kbps [8].

Most video applications today operate at frame rates of 30 frames/sec (33 msec/frame). To guarantee jitter-free delivery of the video signal at the receiving node the cell jitter in the ATM switching fabric should be bounded by 33 msec (this is anticipated in current network proposals). In addition, a dejittering buffer should be provided at the receiving node to account for cell jitter and for the variable number of bits/frame in VBR coding systems [9]. The transmitting node must be responsible for specifying in the bit stream when to refresh the frame. This signal, along with other synchronization signals, should be delivered by the network with high priority.

3 Advantages of VBR Coding

Current techniques for network video transmission rely on CBR coding, due to the nature of existing data networks which do not permit fine grained bandwidth allocation. This leads to a costly tradeoff. If a consistently high quality video signal is to be provided, then a bit rate much higher than the required average rate of a fixed SNR coder must be reserved, due to the variable nature of the image entropy and video delay constraints which limit the amount of buffering that can be introduced. From a users point of view it would be desirable to view the broadband network as a transparent bit pipe that can receive bits as they are generated with minimal buffering. This would allow the elimination of the elastic rate buffer from the transmitting coder [10]. Significant reductions in the average link load can be achieved by using VBR video coding [9]. The tradeoff can now be cast as constant bit rate vs constant SNR video transmission [11]. Obviously, the user would like to maintain a consistently high quality video service while minimizing the average bit rate utilized.

Because the ATM network allocates bandwidth in cell increments, it is very easy for the user to dynamically allocate bandwidth between multiple applications. This permits statistical multiplexing on the broadband link, where N bursty sources whose summed mean bit rate is less than the link rate but whose summed peak rate exceeds the link rate can be carried. This statistical multiplexing at the link level facilitates increased throughput over the user-network interface. The current ATM standards define four classes of network service: VBR video coding falls into class 2 (variable bit rate, connection-oriented, and preserved timing between source and destination) [12]. Statistical multiplexing gains can also be observed at the network multiplexer level. Statistical multiplexing gain for video sources would best be defined as the ratio of CBR video sources to subjectively equivalent VBR sources that could be carried within a given bandwidth. Because this measure is difficult to evaluate, a more common definition is the ratio of source peak rate to source equivalent load, where equivalent load is defined as the link bandwidth divided by the number of statistically equivalent sources multiplexed [13]. In general, when statistically multiplexing bursty sources, the network cannot operate at 100% utilization (bandwidth in excess of the source average rate must be allocated per source). Network congestion caused by simultaneous bandwidth peaks of many video channels at the link or multiplexer level will generally lead to lost cells in the network.

In [14] the authors developed a hardware test rig for measuring the statistics of television images coded using Huffman and run-length coding of interframe difference signals. Five television images
were observed simultaneously over 1 minute intervals. They determined the average image bit rate to be 370 Kbps/frame, with a standard deviation of 63 Kbps/frame. For the cases of 5 and 25 video sources multiplexed on a single channel, they recorded standard deviations of 76 and 27 Kbps/frame, versus 140 and 700 Kbps/frame standard deviations which would be obtained through variance addition. They determined that ~1150 Kbps/frame would have to be allocated per a single video source to achieve a probability of blocking of $10^{-3}$, while for the 25 signal multiplex case only ~450 Kbps/frame would have to be allocated to achieve the same blocking probability, yielding a statistical multiplexing gain of 2.5. They reported that the maximum observed bit rate for the multiplexed channel was about 1.6 times the mean bit rate (utilization of 53%).

In [9, 13] the authors studied statistical multiplexing gains for a variety of video sources using a DPCM coder. Assuming that $N$ independent video sources are multiplexed and the network switch queues are neglected, then the probability density function (pdf) of the multiplexed signal is the $N$-fold convolution of the individual source pdf's. Negative cumulative distributions were computed to determine the probability that a given bit rate would be exceeded by the multiplexed signal. For a 64 signal multiplexer, it was determined that with a bandwidth allocation of 7 Mbps per source for sources with a peak bit rate of 14 Mbps, the probability of exceeding the allocated bandwidth of the multiplexed channel was ~ $10^{-8}$, on the order of magnitude of the anticipated cell loss rate. This yields a statistical multiplexing gain of 2. In a continuation of their study they calculated the mean, peak, and equivalent multiplexed bit rates for various video sources. While peak/mean ratios of greater than 4:1 were observed for videophone signals, the ratio of equivalent load/mean bit rates on a multiplexed channel was typically ~1.3 (utilization of 77%).

From these studies it was shown that when multiplexing $\geq 25$ independent VBR video sources bandwidth gains of ~2 can be achieved over allocating the peak bit rate for each channel. If the typical ratio of peak rate/equivalent load is assumed equal to 2 for the general class of video sources, then data capacity of 50% of the peak bit rate can be freed for other uses on a broadband link if statistical multiplexing is used instead of allocating the peak bit rate. Peak bandwidth allocation for VBR video sources leads to extremely low network utilization rates. These coding gains can fail if the image sources are correlated (multiple instances of the same source). Because video signals are bursty, care must be taken by the network congestion control protocol to prevent cell losses.

4 Congestion Control

Utilization gains from statistical multiplexing can only be achieved if the network has some control over the peak bit rate generated by each video source. Bursty video sources can quickly fill network switch queues. Normally bandwidth allocation (traffic control) is implemented in two ways: reactive control and/or preventive control [15]. Reactive controls rely on feedback cells from the network's ATM switches relaying data on current traffic load. Due to the high data rates and bursty nature of VBR video, reactive controls are not suitable for WANs, since the network's outstanding load and the rate of video cell delivery can increase rapidly within the delay time of network responses. Preventive control is more appropriate for video applications. A policing function is performed by the transmitting node to ensure that its traffic conforms to the statistical parameters negotiated between the node and the network at the initiation of the video connection [9]. In [15], a proposed technique for negotiating grade-of-service requirements with the network is described based on the source's data statistics and real-time requirements.

It was shown in [13] that the period in which a video signal is generating at peak rate can span several seconds. This will have an impact on network congestion unless the transmitting node can buffer its output cells to smooth its traffic distribution, or can selectively drop cells before they enter the network. The use of a large smoothing buffer is equivalent to a rate buffer in a CBR coder, and is therefore not a desirable solution for VBR coders. A more appropriate solution is to allow the transmitting node to police its traffic by assigning priorities to the video data cells and to drop the low priority cells when the traffic exceeds negotiated thresholds. This is facilitated in the ATM protocol by the inclusion of a cell loss priority bit in the cell header. Certain CBR coding algorithms maintain a constant data rate by dropping or coarse quantizing high frequency components in the spatial and/or temporal domain. In a VBR coder, these signals could be assigned to low priority cells and would be transmitted if network traffic levels permit [16, 17]. In [18] congestion control was managed by negotiating a fixed bit rate for high priority image information and transmitting the lower priority information as traffic permits.
One particular method for policing node traffic that is receiving attention in the literature is the leaky bucket method [19]. Tokens are provided to a token pool buffer at the mean bit rate of the source. Cells either pick up a token and are transmitted, or are dropped at the transmitting node. Assuming that some cell priority scheme can be defined based on video signal quality criterion, the leaky bucket method can be modified to utilize cell priorities. A cell buffer can be placed in front of the token pool. If a low priority cell reaches the output of the buffer when there is no token available, it is immediately dropped. It may be necessary to drop low priority cells within the cell buffer as high priority cells enter if a high priority cell is stalled at the output of the buffer and the system wants to maintain its negotiated traffic levels. The advantage of dropping cells at the transmitter end is that it reduces congestion throughout the network. The cell priorities can be used by ATM switches to determine which cells to drop if switch buffers fill up. An alternative approach is to allow all video cells to leave the transmitting node, and allow the network to drop them downstream if congestion develops. This approach will reduce the statistical multiplexing gains at the access link level.

5 Error Control

The common technique used by data link control (DLC) protocols to handle lost packets is to generate an automatic repeat request (ARQ). This technique is suitable for nonreal-time applications, especially since in data transfers the goal is to transfer data without loss or error. The ARQ technique is not suitable as an error control protocol for video transmission over broadband ATM networks, for several reasons [12]. The requirement for a frame refresh rate of \( \sim 30 \) frames/sec means that the end-to-end delay between users must be less than 16 ms, or more than one frame’s worth of data must be buffer at the receiving node, if cell retransmissions are to be used (this end-to-end delay requirement could be achieved in a LAN environment, but not in a large WAN). Also, the mechanism that leads to most cell losses, network congestion, is likely to produce more than one consecutive lost cell. Requests for retransmission (demand refreshment) will only tend to increase the network’s congestion [14]. Since the probability of multiple cell losses occurring consecutively is high, error corrective coding (ECC) techniques would probably prove to be ineffective since the potential number of lost bits is large. To be effective, large ECC codes would have to be applied over many cells. The best technique for dealing with this problem for video transmission is to design the video coding algorithm to be robust in the presence of cell loss.

A megapixel HDTV signal coded for transmission on an ATM network at a mean bit rate of 20 Mbps at 30 frames/sec will be coded on average with \( \sim 1700 \) cells/frame and \( \sim 52000 \) cells/sec. Assuming a cell loss rate of \( 10^{-7} \), one cell loss will occur on average every 3 minutes. If cell loss errors are not concealed, this error rate would be unacceptable [13]. Since cell loss will be encountered at some rate, techniques for error concealment are required to deliver high quality video services.

The most efficient technique for reducing the impact of cell loss errors on the quality of the video signal is to restrict the losses to cells that are carrying signal information of low priority (enhancement data). The two available cell priorities in the ATM protocol allow the coder designer to allocate signal information of high subjective importance to high priority cells. This type of coder is known as a layered or hierarchical coder. One simple method for dividing signal data into priority levels is bit-plane separation [12]. Here some fraction of the most significant bits of coefficient values are transmitted at the high priority, while the lower significant bits are transmitted at the low priority. An alternative to this method is frequency-domain separation. Here, information from low frequency bands produced by transform or sub-band coders is transmitted at the high priority, while the high frequency bands, which are less perceptible, are transmitted at the low priority. The objective is to reduce the perceptibility of cell loss errors by making them occur only for data that is subjectively less important.

One problem that can significantly affect the observability of errors is their propagation throughout the image due to differential (DPCM) intraframe and/or interframe coding. Periodic command refreshment is often required [12]. In addition, high frequency image components that are assigned to low priority cells are often excluded from interframe predictor loops [16, 17]. A more direct form of error concealment is to replace the data in the frame from the cell in error with corresponding data from the previous frame. This technique will work well on images with low motion. Another alternative is to use interpolation and/or filtering on the frame’s region of error, using temporally and/or spatially adjacent pel values as a basis. The problem with this technique is that it may be computationally complex and difficult.
to implement in real-time.

In [18] the authors implemented a complex sub-band coder utilizing 11 sub-band regions in both the spatial and temporal dimensions. The information in band 1 (low spatial, low temporal frequency), was determined to be the data of most perceptual importance. This data was DPCM coded, with periodic command refreshment of the predictor. The other bands showed reduced variance, and were thus PCM coded with coarse quantizers. Experiments were performed with packet losses in subbands 2,5,9,10,11 (low priority bands). The packets used in this experiment were not ATM-sized cells but were 1024 octets long. The effect of packet loss in the low priority sub-bands was the introduction of ringing and noise artifacts into the image. The errors were not perceptually objectionable. Errors were also simulated for band 1. Errors in this high-priority sub-band led to significant image blur during image motion.

In [20] the author implemented a two layer coding algorithm. The high priority layer consisted of a DCT coder operating on 8 x 8 pel blocks. A motion compensator was included to support interframe prediction. Compression was achieved in this layer by using progressively fewer quantization levels for high frequency coefficients. The low priority layer was implemented as an enhancement channel; the difference between the original frame and the output of the DCT coder was DPCM coded and transmitted. The lowest frequency coefficient value for each block was transmitted as a PCM value. This scheme should restrict errors on a single block to its region. Experiments were performed for packet loss on both the high and low priority signals. For 10% packet loss for a still image, error artifacts were not visible. For 10% packet loss after an image change, block errors were easily noticeable, though the error disappeared after one frame. Errors due to the loss of most of the enhancement (low priority) packets resulted in visible noise artifacts across the whole image.

6 Conclusions

The design of the ATM protocol appears to accommodate the implementation of high quality video services. The low cell jitter specification, as well as the inclusion of cell priorities, are features with direct benefit to video coder design. More work needs to be done on both image statistical modeling and congestion control protocols before the widespread implementation of broadband video can begin.

Variable bit rate coding appears to be the most suitable method for implementing broadband video services. The potential gains in bandwidth that can be achieved through statistical multiplexing and the careful reservation of bandwidth by the video user promise to allow high channel utilizations within the broadband network. The flexibility of the ATM protocol will allow video compression services that can deliver consistent, user selectable image quality.

Current standardization efforts for CBR video coding algorithms (H.261, MPEG, digital HDTV) will eventually lead to economical VLSI codec implementations. Any successful VBR coding algorithm should be able to interoperate with standard codecs and should utilize the same core hardware components. Techniques for extending the standard coding algorithms for VBR operation in a broadband ATM environment are currently being investigated by the authors.

References


