In Microsystems technology for Multimedia Applications
Sheu et al. Eds. IEEE Press, May 1995

Chapter 1.1

Video transmission over ATM networks

Antonio Ortega

Signal and Image Processing Institute
Dept. of Electrical Engineering-Systems
University of Southern California

Abstract

In this chapter we will introduce the current state of the art, potential benefits and main challenges in establishing transmission of video over Asynchronous Transfer Mode (ATM) networks. ATM is likely to become the preferred transport technology within future Broadband Integrated Services Data Networks (B-ISDN), where ATM switching and multiplexing will be used to transport end-user traffic end to end through the network. We will emphasize the comparison between ATM transmission and currently available methods. We will concentrate on the more specific video coding aspects while just outlining some of the main issues pertaining to network management.

1.1.1 Introduction

Recent years have witnessed the increasing popularity of digital storage and transmission of video. The emergence of digital video has been driven by the availability of fast hardware at affordable prices, largely thanks to the economies of scale obtained through standardization of compression algorithms. The advantages of digital coding and transmission are well known. When combining a digital format with compression techniques one can provide quality comparable to that of current analog systems, while requiring a fraction of the bandwidth. Moreover, digital transmission systems have the advantage of guaranteeing constant received quality provided that the channel conditions are within certain bounds. As attested by the success of digital audio products such as the compact disk, this latter characteristic can be very attractive in consumer electronic products.

Digital video is being used, or will soon be used, in many applications ranging from videoconferencing (where two or more parties can carry out an interactive communication) to video on demand (where several users can access the video information stored at a central location). Each application has different requirements in terms of bit rate, end-to-end delay, delay jitter, etc. ATM technology is targeted to be used within emerging B-ISDN and promises to allow flexible and efficient delivery of multimedia data, accommodating many different delay and bit rate requirements.

In this chapter we will describe the main advantages of the ATM concept, from the point of view of both video quality and network efficiency. While specification of most aspects of the ATM technology is well underway [1, 2] the definition of the mechanisms to be used for video transmission is still an open question. We will survey some of the key issues in designing video transmission systems over ATM networks, emphasizing those aspects that are specific of video.

This chapter is organized as follows. Section 1.1.2 introduces variable bit rate source coding and rate control as well as various techniques for network transmission of video. In Section 1.1.3 we outline the main issues involved in ATM network design. Section 1.1.4 describes the video coding design problems that are more specific of the ATM environment. Finally, Section 1.1.5 points out the relevance of the choice of an interface between video services and the network. Many issues presented here are a subject of current research. Our goal is not to present an exhaustive survey of all the issues and we choose to concentrate on the general ideas. Details can be found in the references.

1.1.2 Variable bit rate coding and transmission

1.1.2.1 Video compression: variable bit rate coding

International standards for image compression (JPEG[3, 4]), VCR quality video (MPEG-1[5]) and broadcast quality TV/HDTV video (MPEG-2[6]) have been established in the past few years. The first all-digital TV satellite broadcasting system has already begun operations in the U.S. and it will be only a matter of time until these techniques are widely used in other transmission environments.

Video compression algorithms are designed to exploit spatial and temporal redundancies in the input sequence in order to reduce the number of bits needed to transmit a sequence. (For an introduction to image and video compression techniques see [7].) These algorithms are "lossy" in that the decoded images may be different from the original. However, the loss is such that the original and decoded image are almost indistinguishable. One can increase the compression ratio at the cost of increased perceptual degradation. The analysis of this trade-off can be formalized within the Rate-Distortion theory [8]. The amount of redundancy is variable and therefore *the output bit rate is also variable*, even if the quality of the output is maintained constant. For example, under a cosntant set of coding parameters it is clear that for a given scene an increase in motion will result in an increase in bit rate.

1.1.2.2 Rate control

Typically video applications involve real-time display of the decoded sequence. Transmission over a Constant Bit Rate (CBR) channel requires that there is a constant end-to-end delay between the time the encoder processes a frame and the time at which that same frame is available to the decoder [9]. Since the channel rate is constant it will be necessary to buffer the variable rate information generated by the video encoder, where the size of the buffer memory will depend on the acceptable end-to-end delay. It is thus necessary to provide a *Rate Control* mechanism which will ensure that the buffer does not overflow and thus all the information arrives at the decoder. The basic idea is to lower the video quality for scenes of higher "complexity" so as to avoid overflow.

The simplest approach to rate control relies on a deterministic mapping of each buffer occupancy level to a fixed coder mode of operation [10]. More recent works have proposed to use models of the coder behavior to set up coding rate predictions that are used to drive the buffer control [11]. In other cases ideas from control theory are used to devise the buffer controller [12]. In general, the buffer control is designed for a particular encoding scheme and thus scheme-dependent heuristics tend to be introduced [13, 14]. A survey of the algorithms proposed for rate control goes beyond the scope of this chapter, a more detailed review of the problem can be found in [15].

It is important to note that the goal of all these algorithms is to maximize the received quality given the available resources. For example, rate control algorithms increase the rate to avoid underflowing the buffer (a situation that occurs if the encoder is producing consistently fewer bits than the channel can transmit.) We will see how this design philosophy may have to change in the ATM case.

1.1.2.3 Circuit switching, packet switching and ATM

Consider two transmission environments that are currently used for video and exemplify the circuit switching and packet switching techniques: ISDN and computer networks, both local and wide area.

Whereas ISDN offers both circuit switched and packet switched channels, video transmission typically uses a circuit switched channel (this is the case for most videoconferencing products). In this scenario, the transmission capacity available to the end user is constant throughout the duration of the call. The main advantage of this approach is its reliability, since the channel capacity is guaranteed. However, this leads to inefficiency in that a fixed capacity is reserved regardless of the amount of information that needs to be sent.

In computer networks, on the other hand, video is manipulated just as any other type of data. Video data is packetized and routed through the network, sharing the transmission resources with other available services such as remote login, file transfer, etc which are also built on top of the same transport protocols. Such systems are being implemented over LANs [16] and WANs [17, 18], with both point-to-point and multipoint connections. This method is efficient in that network resources are used only when necessary but has the drawback of not providing guarantees on the performance. These systems are often referred to as *Best Effort* because they provide no guarantees on the end-to-end transmission delay and other parameters. As a result, in a best effort environment, the received video quality may change significantly over time.

ATM networks seek to provide "the best of both worlds" by allowing the flexibility and efficiency of computer networks while providing sufficient guarantees so as to permit reliable transmission of real time services as well. As in current computer networks, the information is split into small *cells* or packets, which are then routed through the network and eventually re-assembled at the destination point. Each cell has a size of 53 bytes and contains some overhead bits for routing, cell priority and other purposes, as well a "payload" of user information (48 bytes). See [1] for details on the cell format.

Using this technique allows flexible use of capacity, permitting dynamic routing and re-utilization of bandwidth. Because video compression algorithms produce a variable number of bits, the periods of low activity of one source can be re-used by other sources. For example if N sources each require CBR channels of rate R bits/s, it might be possible to transmit them together over a single channel with rate less than RN. This reduction in capacity is the so-called *Statistical Multiplexing Gain* (SMG). For example, experiments in [19] show gains of up to a factor of 2 when multiplexing 16 sources. While numerical results depend on many factors it is clear that the potential efficiency gains make ATM techniques very attractive. Note that similar gains can be achieved by combining several video sequences into a CBR bit stream [20]. Also note that the SMG has been used for years within the telephone network where the Time Assigned Speech Interpolation (TASI) technique takes advantage of silences in phone calls.

If all sources are using their maximum capacity simultaneously, packets might be lost and thus transmission can only be guaranteed "most of the time". However, contrary to the best effort characteristic of most computer networks, ATM networks are being designed so as to allow quality of service (QOS) parameters to be met, *at least statistically*. Whereas in a circuit switched network the capacity is deterministically available throughout the duration of a connection, the QOS parameters in ATM will only be defined statistically. A circuit switched network would guarantee R bits for as long as the connection is active, while an ATM network may guarantee a bound on the probability that a connection requiring on average R bits exceeds some end-to-end transmission delay. For example, one could negotiate that the maximum permissible delay for the information sent in a frame interval is only exceeded (and thus some information is lost) 0.01% of the time.

ATM networks aim to accommodate very heterogeneous sources, thus allowing a customized set of QOS parameters to be selected by each application. Indeed the ATM design should be able to support both the best effort network protocols and circuit switched connections. However, while the QOS parameters are meaningful for non real-time data (where the main concern is the total duration of the transmission) they are not the only factor to take into account for real time video transmission. For a video service, the appropriate measure of performance should be the perceptual quality achieved for the given bit rate. In that respect, video differs from other type of data in that acceptable transmission quality can be achieved *even if some of the data is lost*. The effect of packet losses, which in other applications result in retransmission of data, can be reduced in the video case with appropriate encoding strategies (see Section 1.1.4.1) combined with error concealment techniques (see Section 1.1.4.2).

1.1.3 Network Design Issues

We now examine some of the network design issues, emphasizing those aspects that directly affect video transmission. (See [21] for a recent survey of open issues.)

1.1.3.1 Admission control, contracts and traffic parameters

Admission control is the task of deciding whether a new connection with a given set of requested QOS parameters can be allowed into the network. In very simple terms, the connection should be admitted if it can be guaranteed to have the required QOS without degrading the QOS of other ongoing connections. Since video encoders produce variable rate, a key factor in the admission control problem is to find statistical models for the expected bit rate of video sources. A model characterizes the bit rate of a video connection at various time scales (bit rate per frame, distribution of rate within a frame), and will attempt to capture the short and long term dependencies in the bit rate as well. Providing good models for various types of video sources and encoders has thus been an active area of research in recent years (see for instance [22, 23] for recent results in this area). Typical models are autoregressive, where the number of bits for the current frame is correlated to the number of bits for the previous frame [24], or Markov chains [25]. While most models have focused on the short term behavior some recent work has also shown significant long term dependencies in the bit rate [26].

For a given model the performance of the network under different routing and queueing strategies can then be examined (see [21] for a review of the various queueing disciplines and models which have been considered) and the decision on whether to admit a call can be made based on the expected performance of the network.

Note that admission control is much simpler for circuit switched networks. Since the transmission resources are constant throughout the duration of the call, the only issue is to find out whether currently unused resources are sufficient to carry the additional call. If they are, the call can be completed, otherwise it will be rejected. In the ATM environment the resources needed by each of the sources change over time and thus the main problem is to estimate the likelihood that resources will be insufficient to guarantee QOS. Best effort networks do not perform explicit admission control, although insufficient QOS during high load conditions will drive users out or make them delay their connection.

Admission control is part of the negotiation process between user and network to set up a connection. The result of the negotiation is a *contract* which will specify the *traffic parameters* of the connection. Typical traffic parameters are peak cell rate and sustainable cell rate. These are operational measures of the offered bit rate (over, respectively, short and long term) and are implemented with counters.

1.1.3.2 Policing

The network will have the means to monitor usage by each of the sources after the connection has been established. This function, called Usage Parameter Control (UPC) or *Policing* mechanism, has the goal of preventing sources from maliciously or unwillingly exceeding the traffic parameters negotiated at call set up. Typically the network will look at policing methods that are directly linked to the negotiated traffic parameters. For instance if a certain peak cell rate has been agreed upon, then the policing mechanism may consist of a counter which tracks the peak rate and verifies it does not exceed the negotiated value.

One of the most popular policing mechanisms, due to its simplicity, is the so-called *leaky bucket* [27]. A leaky bucket is simply a counter incremented with each cell arrival and decremented at fixed intervals such that the decrement is equivalent to an average cell rate of R. The other parameter of the leaky bucket is the size of the bucket (i.e. the maximum allowable value for the counter). The network can detect violations by monitoring whether the maximum value of the counter is reached. Other policing mechanisms have been proposed and their relative performance can be judged based on criteria such as their simplicity and success in detecting violating sources without affecting compliant ones [1] (for a comparison of some of the proposed policing mechanism see [27]). If some cells are found to be violating the policing functions the network can decide to delete them or to just mark them for possible deletion in case of congestion.

As will be seen later the choice of policing function is important because it may determine the type of rate that the sources will transmit through the network. Indeed, given the agreed upon parameters, it is likely that the video sources will be able to adapt their transmission rates so as to not violate the constraints.

1.1.4 Video Coding Issues

Because of real time display, the information corresponding to a given video frame has to reach the decoder within a certain time in order to be useful. Therefore, strategies involving retransmission of packets that have been lost are not possible and video transmission schemes have to be designed under the assumption that some of the information will be lost. Thus, video encoding algorithms for ATM transmission need to be *robust to packet losses*. This can be achieved in part by using multiresolution encoding schemes along with different levels of priorities for the cells corresponding to each resolution. Additionally, error concealment techniques can be used to mask to some extent the perceptual effects in the decoded sequence of the loss information.

1.1.4.1 Multiresolution coding

Multiresolution encoding schemes separate the information into two or more layers or resolutions. The coarse resolution contains a rough approximation to the full resolution image or sequence. The enhancement, or detail, resolution provides the additional information needed to reconstruct at the decoder the full resolution sequence at the target quality. Typically the coarse resolution sequence is obtained by reducing the spatial or temporal resolution of the sequence (i.e. reducing the frame size or the frame rate, respectively) or by simply having images of lower quality (i.e. where encoding artifacts are more visible).

A survey of multiresolution encoding techniques can be found in [28]. The recent MPEG-2 standard [6] includes several so-called scalable modes which are specifically targeted at applications such as video over ATM. Since the main requirement for these scalable techniques is to provide robustness to packet losses, it will be more natural to resort to methods that preserve the spatial and temporal resolution. In these schemes the coarse layer has worse perceptual quality but the spatial and temporal resolution is the same as in the original. To take advantage of the multiresolution encoding the information is packetized into two classes of packets according to whether the priority bit provided by the ATM format [1] is set or not. The coarse resolution will be transmitted using high priority packets while the detail resolution will be sent with the low priority ones. It has been shown that using the priorities so that the packets with lower priority are discarded first in case of congestion is beneficial in terms of the end to end quality [29]. As an example, the interface proposed in [30] would send each of the video resolutions within a separate connection, each with different QOS service requirements.

1.1.4.2 Error concealment

A further advantage of using multiresolution coding schemes is that they enable efficient error concealment techniques. Essentially, the idea is to use the available information, i.e. packets that were not lost, to interpolate the missing information. When multiresolution coding is used the information decoded from only the lower resolution layer may be sufficiently good. Other approaches that have been proposed involve interleaving the information so that a cell loss causes minor perceptual degradation in several image blocks rather than severe degradation in just a few. Also losses can be compensated utilizing motion information by, for instance, replacing lost information with that corresponding to the previous frame. See [31, 32, 33] for examples of several approaches to error concealment.

1.1.5 Design of the video/network interface

1.1.5.1 Rate control and rate shaping

ATM transmission provides the possibility of transporting variable number of bits per frame and thus would seem to make unnecessary the use of rate control. However, this view is not realistic because each connection will be specified by a series of traffic parameters which will be monitored by the network. Transmission over the limits set by the traffic contract may result in lost packets and thus rate control is still necessary under this situation[9]. This is an important observation since most models that are proposed for video coders do not take into account rate control. A key issue is thus to choose an appropriate video/network interface as this will be determine both the achievable video quality and the "shape" of the video stream within the network.

We can make the distinction between rate control, which entails changing the rate produced by the encoder (and thus affects the quality of the encoded video) and rate shaping which only affects the times at which cells are sent to the network, but not the total amount of information transmitted. It has been shown (see for instance [34]) that smoothing or shaping the traffic can be beneficial since timing the transmission of the cells may prevent violation of peak rate constraints.

1.1.5.2 Choices in video/network interface

As was shown in [35] under policing constraints such as those imposed by a Leaky Bucket it is easy for the encoder to perform rate control so as to always utilize a bit rate close to the maximum allowed by the contract. In that scenario there will hardly be any SMG and thus all the benefits of packet transmission would be lost. Thus additional mechanisms have to be implemented to ensure that so-called greedy coding (i.e. encoders which will attempt to use as much rate as possible within the constraints) is avoided. One such mechanism is to implement tariffication based on the total amount of actual rate used, rather than on the connection parameters. Thus connections which use fewer bits than the maximum allowed will cost less. The main disadvantage of this approach is that the worst case scenario may preclude any substantial SMG within the network. Indeed, it is conceivable that sources would prefer to use the full allowable rate, and pay the extra amount for it, while using the corresponding bandwidth in part to transmit other information such as data.

Another approach is to assume that users will always attempt to use as high a rate as possible and consequently to provide policing mechanisms for which this worst case scenario is less "damaging" to the network. Some such approaches including rate constrained to a set of discrete levels as in [36], defined by a histogram as in [37] or constrained by multiple Leaky buckets [38] have been proposed. As outlined in Section 1.1.3 the network performance analyses that are used in the admission control process rely on the accuracy of the source models. However the bit rate that each source will produce depends on the traffic contract (just as in the CBR case, the bit rate produced by the source depends on the buffer size) and thus on the specific traffic parameters and constraints that are used within the video/network interface. Thus, the video/network interface design will be one of the key issues in bringing about video over ATM.

1.1.6 Conclusions

We have presented an overview of some of the key research issues in designing systems for video transmission over ATM networks. While some of the problems (e.g. Multiresolution video coding and error concealment) affect only the video quality, we have motivated that the definition of the video/network interface is key in simultaneously obtaining good video quality and efficient network utilization.

REFERENCES

[1] ATM Forum, *ATM User-Network Interface Specification, Version 3.0.* Prentice-Hall, 1993.

[2] "ATM Forum WWW Home Page." URL: `http://www.atmforum.com/`.

[3] G. K. Wallace, "The JPEG still picture compression standard," *Communications of the ACM*, vol. 34, pp. 30–44, Apr. 1991.

[4] W. B. Pennebaker and J. Mitchell, *JPEG Still Image Data Compression Standard.* New York, NY: Van Nostrand Reinhold, 1993.

[5] D. LeGall, "MPEG: a video compression standard for multimedia applications," *Communications of the ACM*, vol. 34, pp. 46–58, Apr. 1991.

[6] Inform. Technology - Generic Coding of Moving Pictures and Associated Audio, ITU Draft Rec. H.262, ISO/IEC 13818-2, Mar. 1994.

[7] A. N. Netravali and B. G. Haskell, *Digital Pictures: Representation and Compression.* New York: Plenum Press, 1988.

[8] T. M. Cover and J. A. Thomas, *Elements of Information Theory.* New York: Wiley, 1991.

[9] A. R. Reibman and B. G. Haskell, "Constraints on variable bit-rate video for ATM networks," *IEEE Trans. on CAS for video tech.*, vol. 2, pp. 361–372, Dec. 1992.

[10] B. G. Haskell, "Buffer and channel sharing by several interframe picturephone coders," *Bell Systems Tech. J.*, vol. 51, Jan. 1972.

[11] J. Zdepsky, D. Raychaudhuri, and K. Joseph, "Statistically based buffer control policies for constant rate transmission of compressed digital video," *IEEE Trans. on Comm.*, vol. 39, pp. 947–957, June 1991.

[12] J.-P. Leduc and S. D'Agostino, "Universal VBR videocodecs for ATM networks in the Belgian broadband experiment," *Image Communication*, vol. 3, pp. 157–165, June 1991.

[13] C.-T. Chen and A. Wong, "A self-governing rate buffer control strategy for pseudoconstant bit rate video coding," *IEEE Trans. on Image Proc.*, vol. 2, pp. 50–59, Jan. 1993.

[14] K.-H. Tzou, "An intrafield DCT-based HDTV coding for ATM networks," *IEEE Trans. on Circ. and Sys. for Video Tech.*, vol. 1, pp. 184–196, Jun. 1991.

[15] A. Ortega, K. Ramchandran, and M. Vetterli, "Optimal trellis-based buffered compression and fast approximations," *IEEE Trans. on Image Proc.*, vol. 3, pp. 26–40, Jan. 1994.

[16] A. Eleftheriadis, S. Pejhan, and D. Anastassiou, "Algorithms and performance evaluation of the Xphone multimedia communication system," in *Proc. of the ACM Multimedia 93 Conf.*, (Anaheim, CA), pp. 311–320, Aug. 1993.

[17] J.-C. Bolot and T. Turletti, "A rate control mechanism for packet video in the internet," in *Proc. of Infocom'94*, (Toronto), pp. 1216–1223, Jun. 1994.

[18] M. Macedonia and D. Brutzman, "MBone provides audio and video across the Internet," *Computer*, vol. 27, pp. 30–36, Apr. 1994.

[19] W. Verbiest, L. Pinnoo, and B. Voeten, "The impact of the ATM concept on video coding," *IEEE J. on Sel. Areas in Comm.*, vol. 6, pp. 1623–1632, Dec. 1988.

[20] G. Keesman and D. Elias, "Analysis of joint bit-rate control in multi-program image coding," in *Proc. VCIP'94*, (Chicago, IL), 1994.

[21] J. Kurose, "Open issues and challenges in providing quality of service guarantees in high-speed networks," *ACM Comp. Comm. Review*, vol. 23, pp. 6–15, Jan. 1993.

[22] "Special issue on packet video." IEEE Trans. on Circ. and Sys. for Video Tech., Jun. 1993.

[23] "Proc. of the 6th Intl. Workshop on Packet Video, (Portland, OR)," Sept. 1994.

[24] B. Maglaris, D. Anastassiou, P. Sen, G. Karlsson, and J. Robbins, "Performance models of statistical multiplexing in packet video communications," *IEEE Trans. on Comm.*, vol. 36, pp. 834–843, July 1988.

[25] D. P. Heyman, A. Tabatabai, and T. Lakshman, "Statistical analysis and simulation study of video teleconferencing traffic in ATM networks," *IEEE Trans. on Circ. and Sys. for Video Tech.*, vol. 2, pp. 49–59, Mar. 1992.

[26] M. W. Garrett and W. Willinger, "Analysis, modeling and generation of self-similar VBR video traffic," in *Proc. of ACM SigComm*, (London), Sept. 1994.

[27] E. P. Rathgeb, "Modeling and performance comparison of policing mechanisms for ATM networks," *IEEE J. on Sel. Areas in Comm.*, vol. 9, pp. 325–334, April 1991.

[28] M. Vetterli and K. M. Uz, "Multiresolution coding techniques for digital television: A review," *Special Issue on Multidimensional Processing of Video Signals, Multidimensional Systems and Signal Processing*, pp. 161–187, Mar. 1992.

[29] M. W. Garrett and M. Vetterli, "Joint source/channel coding of statistically multiplexed real time services on packet networks," *IEEE/ACM Trans. on Networking*, vol. 1, pp. 71–80, Feb. 1993.

[30] P. Haskell and D. Messerschmitt, "Open network architecture for continuous-media services: The Medley gateway," *IEEE/ACM Trans. on Networking*, 1994. Submitted.

[31] K. Ramchandran, A. Ortega, K. M. Uz, and M. Vetterli, "Multiresolution broadcast for digital HDTV using joint source-channel coding," *IEEE J. on Sel. Areas in Comm.*, vol. 11, pp. 6–23, Jan. 1993.

[32] Q.-F. Zhu, Y. Wang, and L. Shaw, "Coding and cell-loss recovery in dct-based packet video," *IEEE Trans. on Circ. and Sys. for Video Tech.*, vol. 3, pp. 248–258, Jun. 1993.

[33] M. Ghanbari and V. Seferidis, "Cell-loss concealment in ATM video codecs," *IEEE Trans. on Circ. and Sys. for Video Tech.*, vol. 3, pp. 238–247, Jun. 1993.

[34] K. Joseph and D. Reininger, "Source traffic smoothing for VBR video encoders," in *Proc. of 6th Intl. Workshop on Packet Video*, (Portland, OR), Sept. 1994.

[35] A. Ortega, M. W. Garrett, and M. Vetterli, "Toward joint optimization of VBR video coding and packet network traffic control," in *Proc. of the 5th Packet Video Workshop*, (Berlin), March 1993.

[36] H. Heeke, "A traffic-control algorithm for ATM networks," *IEEE Trans. on Circ. and Sys. for Video Tech.*, vol. 3, pp. 182–189, Jun. 1993.

[37] P. Skelly, M. Schwartz, and S. Dixit, "A histogram-based model for video traffic behavior in an ATM multiplexer," *IEEE/ACM Trans. on Networking*, vol. 1, pp. 446–459, Aug. 1993.

[38] A. Ortega and M. Vetterli, "Multiple leaky buckets for increased statistical multiplexing of ATM video," in *Proc. of the 6th Packet Video Workshop*, (Portland, OR), Sep. 1994.