13

Development of a MPEG Data Stream Characterization for Use with ATM Networks

Olen L. Stokes
IBM Corporation
P.O. Box 12195, Research Triangle Park, N.C., 27709
(919) 467-5200, olstokes@vnet.ibm.com

Arne A. Nilsson
North Carolina State University
Department of Electrical and Computer Engineering
Raleigh, N.C., 27695, (919) 515-5130, fax (919) 515-5523
nilsson@eos.ncsu.edu

Abstract

The various parameters which affect MPEG data streams are discussed. An AR model is developed to characterize both constant bit rate and variable bit rate videos. Three videos with different bit rate modes, audio content, frame type mixture, and group of pictures sizes are analyzed. The accuracy of the AR model in predicting the size of the next frame based on previous frames is examined. Possible enhancements to the AR model are also introduced. The role of the AR model in the development of ATM networks is discussed.

Keywords

ATM networks, MPEG, AR model

1 INTRODUCTION

"We are moving into an age of Information Networking, with users anticipating increasing freedom to communicate with people and retrieve information anytime, anywhere, and in multiple media" (Albanese, 1991). These expectations translate into expanding markets and profits when desirable new services can be provided economically. The key to these services and their revenue growth is video. In fact, "digital video is likely to become the dominant B-ISDN bandwidth driver, particularly for on-demand entertainment services" (Terry, 1992). Providing this information will require higher speed services with more flexibility than is possible with today's networks. Broadband ISDN (B-ISDN) networks with Asynchronous Transfer Mode (ATM) services have been proposed to provide the needed bandwidth and versatility.

"The study of the statistical properties of packet video streams, to model video sources, is a required step in the process of designing B-ISDN networks to handle heterogeneous traffic" (Pancha, 1994). For network control, the ability to predict data rates could enable effective congestion control measures. For a constant bit rate (CBR) connection, the amount of bandwidth required and the degree of multiplexing possible are easily determined at connection setup. The task becomes to insure that this guaranteed traffic arrives with an acceptable delay and jitter. However, for a variable bit rate (VBR) video connection, the required bandwidth and resulting multiplexing must be adjusted over time. By predicting the upcoming requirements, the network can adjust its allocations to meet the new conditions. Likewise, an encoder could request a new allocation based on its predictions (Pancha, 1993).

2 MPEG VIDEO

In order to effectively utilize the available bandwidth and to provide the desired picture quality, compressed video encoding/decoding methods are being developed. One such standard which is receiving much attention has been developed by the Moving Pictures Expert Group (MPEG). The current standard is commonly known as MPEG-1 and is directed at providing image quality comparable to VHS video and sound quality similar to audio CDs. Although it is targeted for digital storage media, MPEG "is flexible enough to be used in a variety of video applications" (Pancha, 1993) including transmission through B-ISDN ATM networks. Also, an MPEG-2 standard is being developed for higher resolution images and correspondingly higher data rates.

The data stream from an MPEG video is inherently variable. This comes in part from the three different algorithms for encoding a video picture. An intra-coded picture, or I-frame, is an encoding of the picture based entirely on the information in that frame. A predictive-coded picture, or P-frame, is based on motion compensated prediction between that frame and the previous reference frame (I- or P-frame). A bidirectionally predictive-coded picture, or B-frame, is based on motion compensated prediction between that frame and the previous or next reference frame (I- or P-frame). The size of the resulting frame varies significantly between frame types. I-frames are the largest while B-frames are the smallest.

Further, within each frame type, the size of the resulting encoded information varies. The size of an I-frame varies based on picture content. P- and B-frames vary depending on the motion present in the scene as well as picture content. Also, the quantizer-scale parameter (q) can be varied by the encoder to change the size and corresponding quality of each frame. A small q value produces a higher quality picture that requires a larger data rate. A large q value generates a smaller data rate at the expense of a lower quality picture. The value(s) chosen by the encoder are determined by the desire to either provide a constant picture quality or maintain a particular data bit rate.

An MPEG data stream which contains audio as well as image data also includes system level control information. The control information defines the mixing of data packets which contain either audio or image data. This interleaving of audio and video packets is not necessarily constant, nor are the audio packets always equally sized. Therefore, the audio data rate also varies over certain time scales.

An MPEG encoder chooses between two basic service types: constant bit rate (CBR) and variable bit rate (VBR). Note that bit rate refers to the rate at which data is written to or taken from storage media (such as a CD-ROM) or transmitted over a network. For VBR video, the data rate is allowed to vary over time while the picture quality potentially remains constant. For CBR video, the picture quality is varied to insure that the resulting data stream can be written, read, or transmitted at a constant rate. To accomplish this, the data rate and picture quality are manipulated during encoding by adjusting q.

This CBR data stream does not imply a constant frame size. Rather, it means that the output of an encoder buffer can be read at a constant bit rate. By using a sufficiently large buffer which is partially filled prior to removing the first video information, and by varying q so that the buffer never overruns or underruns, the encoder can produce a data stream which appears to be a CBR source to the remainder of the system. The encoder system shown in Figure 1 can operate in such a CBR manner or in a VBR manner. The CBR versus VBR decision is made at encoding time.

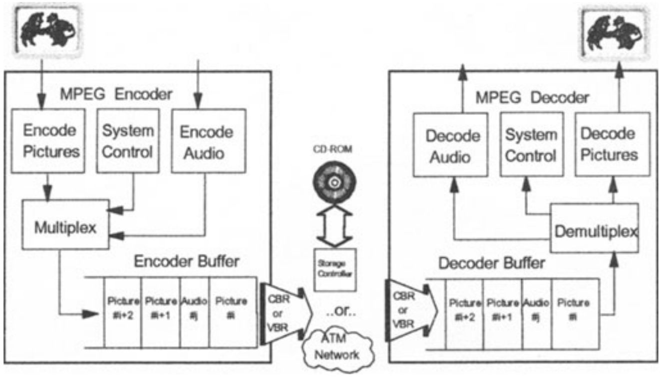

Figure 1 MPEG encoding and decoding systems.

The MPEG data shown in Figure 1 is part of a system level data stream as discussed above. The control information is not shown, nor are any padding data packets which might be required to obtain a CBR stream. The system and pad data rates are significantly smaller than the video and audio data rates. Note that many of the MPEG data streams available for analysis contain only image or audio information, but not both.

The decoder system shown in Figure 1 can also operate under CBR or VBR conditions. The decode buffer is initially filled to a level specified by the data stream before any actual picture and/or audio information is removed from the buffer and decoded. Thereafter, the buffer is filled on one end by the arriving data and emptied periodically on the other by the decoders. As long as the buffer does not overrun or underrun, the video can be displayed properly. A generic MPEG decoder, the system target decoder (STD), is defined in the MPEG standard (ISO, 1993). The encoder must insure that the data stream it produces can be played properly by a directly connected STD. Therefore, the encoder must define the decoder minimum buffer size, initial delay, and audio and video data rates. Note that any delay jitter introduced by transmission over a network is not necessarily anticipated by the encoder.

The initial delay to allow the decode buffer to partially fill normally causes no concern when playback is from a digital storage medium. However, when the playback is part of an interactive or real-time display, this delay may create problems. For example, an initial delay in the decode buffer of 250 msec will significantly increase the probability that the round-trip delay will have a noticeable affect on a conversation. In these cases, a VBR mode of operation, where a frame's data is transmitted as soon as it is encoded, may be required. The encoder can either maintain a constant picture quality or continue to use the CBR algorithm without buffering. This latter mode would allow the data at the remote site to be both displayed as well as buffered for storage and future CBR playback.

The frame type (I, P, or B) selected for a particular frame is based on a repeating sequence of frames called a Group Of Pictures (GOP). Two typical groups are shown in Figure 2. The sequence used is chosen by the encoder. The MPEG standard does not completely specify the encoding process, but rather defines a syntax from which decode by the STD is possible. This allows numerous encoding options which increases the difficulties in handling and controlling MPEG traffic.

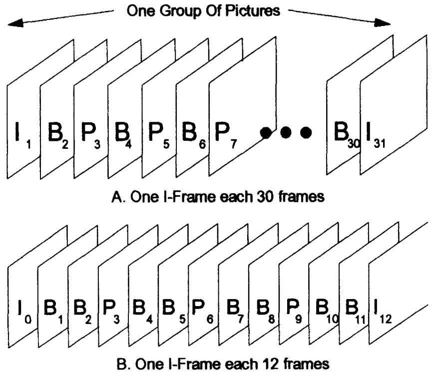

Figure 2 MPEG frame display sequence.

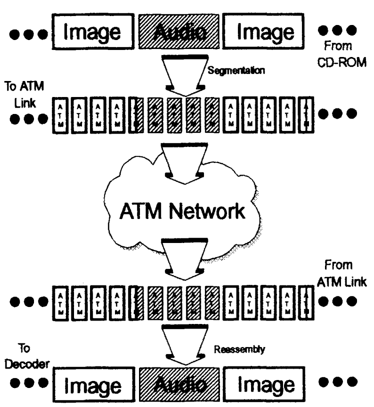

Figure 3 MPEG-1 data stream to/from ATM cells.

Once the data stream is encoded, it can be used locally or transmitted through a network. It can be displayed immediately or written to a digital storage medium for future display. Whenever the data stream is transmitted through an ATM network, the data must be divided into blocks which match the size of the ATM cell payload. This means 44- to 48-octet blocks depending on the ATM Adaptation Layer (AAL) used. At the receiving end, the data from the cell payloads is reassembled into the original MPEG stream. This process is shown in Figure 3. Should the data not be interleaved, different data types would travel on separate ATM virtual connections.

3 MPEG CHARACTERIZATION

"In spite of all efforts, so far no general consensus has been reached for how to model data rates generated by typical video codecs..." (Heeke, 1993). The derivation of an MPEG-1 mathematical characterization is challenging. As discussed above, MPEG-1 data streams are affected by numerous factors including picture content, object motion, audio content, and encoder options. The result is a random sequence of frame sizes which is difficult to analyze. Figure 4 shows three examples of MPEG-1 data stream frame sizes. The sizes are in terms of ATM cells with 44-octet payloads.

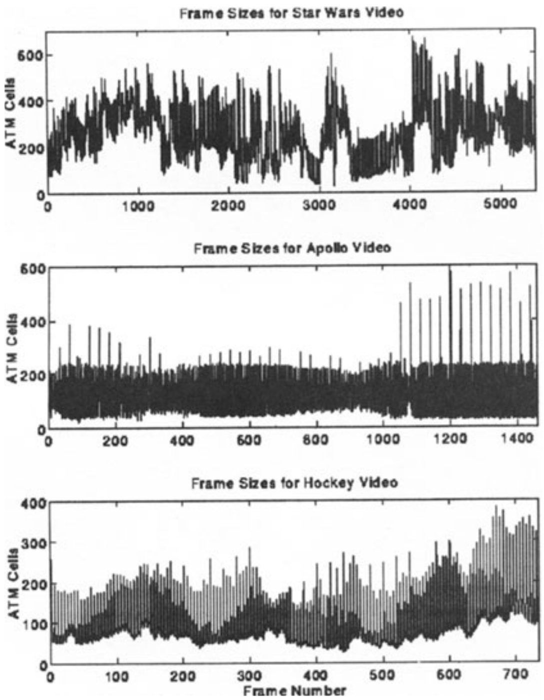

Figure 4 Frame sizes for sample video sequences.

These MPEG-1 data streams represent three different sets of encoder parameters. The *Star Wars* sequence is based on the image-only VBR data stream described in (Pancha, 1993). It contains only I- and P-frames and utilizes a constant q. The *Apollo* sequence is a system level CBR stream with interleaved audio and image (Aris). It includes all three frame types. The *Hockey* sequence is also a VBR image-only stream (North Valley Research). However, this encoding varies q by frame type and uses B-frames as well as I- and P-frames. All of these videos contain camera shot changes as well as changes in movement and camera panning.

Using these frame sizes as a sequence of random numbers, the average and standard deviation can be calculated as well as the autocorrelation functions (ACFs). These first two statistics are

DOCKET ALARM

Explore Litigation Insights

Docket Alarm provides insights to develop a more informed litigation strategy and the peace of mind of knowing you're on top of things.

Real-Time Litigation Alerts

Keep your litigation team up-to-date with **real-time alerts** and advanced team management tools built for the enterprise, all while greatly reducing PACER spend.

Our comprehensive service means we can handle Federal, State, and Administrative courts across the country.

Advanced Docket Research

With over 230 million records, Docket Alarm's cloud-native docket research platform finds what other services can't. Coverage includes Federal, State, plus PTAB, TTAB, ITC and NLRB decisions, all in one place.

Identify arguments that have been successful in the past with full text, pinpoint searching. Link to case law cited within any court document via Fastcase.

Analytics At Your Fingertips

Learn what happened the last time a particular judge, opposing counsel or company faced cases similar to yours.

Advanced out-of-the-box PTAB and TTAB analytics are always at your fingertips.

API

Docket Alarm offers a powerful API (application programming interface) to developers that want to integrate case filings into their apps.

LAW FIRMS

Build custom dashboards for your attorneys and clients with live data direct from the court.

Automate many repetitive legal tasks like conflict checks, document management, and marketing.

FINANCIAL INSTITUTIONS

Litigation and bankruptcy checks for companies and debtors.

E-DISCOVERY AND LEGAL VENDORS

Sync your system to PACER to automate legal marketing.

WHAT WILL YOU BUILD? | **sales@docketalarm.com** | **1-866-77-FASTCASE**

fastcase®
Smarter legal research.