Compressed Video Transmission over Digital Networks: Analysis and Design

Qing Tang[1] and Jesse S. Jin[1,2]
[1]School of Information Technologies
University of Sydney, NSW 2006, Australia
[2]School of Computer Science & Engineering
University of New South Wales, NSW 2052, Australia
tqtracy@it.usyd.edu.au

Abstract

In the past two decades, new technologies in digital video compression have paved the way for video transmission. This is providing an environment for the widespread availability of a wide range of real video networking services.

The purpose of this paper is to introduce fundamental principles and important technologies used in design and analysis of compressed video transmission systems over digital networks.

We start with the overview of video transmission and video compression techniques and present gradually towards important issues of video transmission systems. The proposed error resilience approach will be addressed to some extent in last section.

Keywords: Video Transmission, MPEG, Packetizing, Multiplexing, Synchronizing, Interoperability, Transcoding and Error Resilience.

1 Introduction

Digital video compression is a field in which fundamental technologies were motivated and driven by practical applications so that they often lead to many useful advances. Especially, MPEG, the most recognized standard for digital video compression, has enabled many successful digital-video applications. These applications range from digital-video disk (DVD) and multimedia CDs on a desktop computer, interactive digital cable television, to digital satellite networks. Nowadays, video compression technologies are being used in almost all modern digital video systems and networks. Not only is video compression equipment being implemented to increase the bandwidth efficiency of communication systems, but video compression also provides innovative solutions to many related video-networking problems.

2 MPEG Video: An Overview

MPEG video is now an integral part of most digital video transmission and storage systems. Comparing with analog video, the use of MPEG video provides lower costs in video distribution, increases the quality and security of video, and allows for interactivity.

There exist various compression techniques that are in part competitive and in part complementary. Many of these techniques are already applied in industries, while other methods are still undergoing development or are only partly realized. In the following sections the most relevant work in the standardization bodies concerning video coding is outlined.

2.1 MPEG-1

MPEG-1 is a generic standard for coding of moving pictures and associated audio for digital storage media at up to about 1.5Mbits/s. The video compression technique developed in MPEG-1 covers many applications from interactive VCD to the delivery of video over telecommunications networks.

The MPEG-1 video algorithm has been developed with respect to the JPEG and H.261 activities. It was intended to retain a large degree of commonalty with the H.261 standard so that implementations supporting both standards were plausible. However, MPEG-1 was primarily targeted for multimedia CD-ROM applications, requiring additional functionality supported by both encoder and decoder. Important features provided by MPEG-1 include picture based random access of video, fast forward/fast reverse *(FF/FR)* searches through compressed bit streams, reverse playback of video and editing ability of the compressed bit stream. The basic MPEG-1 video compression technique is based on a Macroblock structure, motion compensation and the conditional replenishment of Macroblocks. As outlined in Fig. 1 the MPEG-1 coding algorithm encodes the first picture in a video sequence in Intra-picture coding mode (I-picture). Each subsequent picture is coded using Inter-picture prediction (P-pictures) – only data from the nearest previously coded I- or P-picture is used for prediction. The MPEG-1 algorithm processes the pictures of a video sequence block-based.

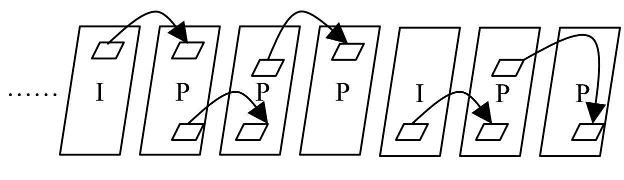

Figure 1. An example of MPEG-1 coding structure

2.2 MPEG-2

Basically MPEG-2 can be seen as a superset of the MPEG-1 coding standard and was designed to be backward compatible to MPEG-1 – every MPEG-2 compatible decoder can decode a valid MPEG-1 bit stream. New coding features were added by MPEG-2 to achieve sufficient functionality and quality, thus prediction modes were developed to support efficient coding of interlaced video. In addition scalable video coding extensions were introduced to provide additional functionality, such as embedded coding of digital TV and HDTV, and graceful quality degradation in the presence of transmission errors.

2.2.1 Field and Frame Pictures

MPEG-2 has introduced the concept of *frame pictures* and *field pictures* along with particular *frame prediction* and *field prediction* modes to accommodate coding of progressive and interlaced video. For interlaced sequences it is assumed that the coder input consists of a series of odd (top) and even (bottom) fields that are separated in time by a field period. Two fields of a Frame may be coded separately. In this case each field is separated into adjacent non-overlapping Macroblocks and the DCT is applied on a field basis. Alternatively two fields may be coded together as a frame (frame pictures) similar to conventional coding of progressive video sequences. Here, consecutive lines of top and bottom fields are simply merged to form a frame. Notice, that both frame pictures and field pictures can be used in a single video sequence.

2.2.2 Field and Frame Prediction

New motion compensated field prediction modes were introduced by MPEG-2 to efficiently encode field pictures and frame pictures. In filed prediction, predictions are made independently for each field by using data from one or more previously decoded field, i.e. for a top field a prediction may be obtained from either a previously decoded top field (using motion compensated prediction) or from the previously decoded bottom field belonging to the same frame. Generally the Inter-field prediction from the decoded field in the same frame is preferred if no motion occurs between fields. An indication which reference field is used for prediction is transmitted with the bit stream. Within a field picture all predictions are field predictions.

2.3 MPEG-4

Compared to MPEG-1 and MPEG-2, the MPEG-4 standard brings a new paradigm as it treats a scene to be coded as consisting of individual objects; thus each object in the scene can be coded individually and the decoded objects can be composed in a scene.

MPEG-4 has made several improvements in coding of intra macroblocks (INTRA) as compared to H.263, MPEG-1/2. In particular it supports the following:

- DPCM prediction of the DC coefficient,
- DPCM prediction of a subset of AC coefficients,
- Specialized coefficient scanning based on the coefficient prediction,
- Huffman table selection,
- Non-Linear inverse DC Quantization.

3 Functions of Video Transmission Systems

Just as there are techniques on how best to compress digital video, there are also efficient methods to mange, transmit, store and retrieve compressed digital video. The general term, video transmission, involves *packetizing*, *multiplexing*, *synchronizing* and *extracting* of video signals.

The functionality and format of video transmission systems can be summarized as follows:

- To provide a mechanism for packetizing video data with functionalities such as packet synchronization and identification, error handling, conditional access, random entry into the compressed bit stream and synchronization of the decoding and presentation process for the applications running at a receiver,
- To schedule and multiplex the packetized data from multiple programs for transmission,
- To specify protocols for triggering functional responses in the transport decoder, and
- To ensure the video bit stream level interoperability between communication systems.

The issues of transporting compressed digital video in modern communication systems and networks are addressed in the following sections.

3.1 The Packetization Approach

The fixed length packet usually has a format shown in Fig. 2. The so-called "link" header contains fields for packet synchronization and identification, error indication, and conditional access. The adaptation header carries synchronization and timing information for decoding and presentation process. It can also provide indicators for random access points of compressed bit streams and for "local" program insertion. The payload could be any multimedia data including compressed video and audio streams.

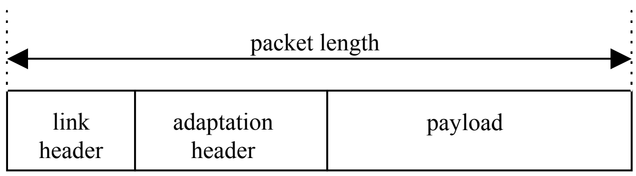

Figure 2. The fixed-length packet format

The choice of this packet size is motivated by a few key factors at the time. The packets need to be large enough so that the overhead of the transport headers does not become a significant portion of the total data being carried. The packet size should not be too large that the probability of packet error becomes significant under standard operating conditions (due to inefficient error correction). Another motive for the particular packet length selection is interoperability with the ATM packet.

3.2 Multiplexing Functionality

The overall multiplexing approach can be described as a combination of multiplexing at two different layers. In the first layer, a single-program transport stream is formed by multiplexing one or more elementary bit streams at the transport layer, and in the second layer the multiple program transport streams are combined (using asynchronous packet multiplexing) to form the overall system. The functional layer in the system that contains both this program and system level information that is going to be described is called the Program Specific Information (PSI).

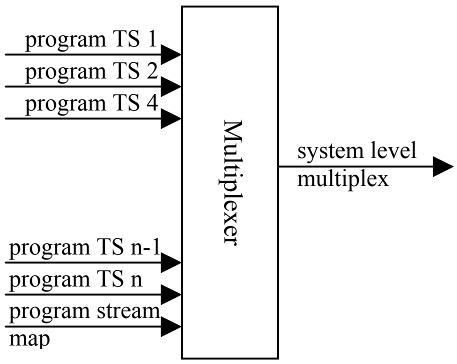

Figure 3. A functional diagram of stream multiplexer

3.3 Synchronizing Functionality

Synchronization and timing recovery process specified in the transport system involves the sampling of the analog signals, encoding, encoder buffering, transmission, reception, decoder buffering, decoding, and presentation of digital audio and video in combination.

Synchronization of the decoding and presentation process for the application running at a receiver is a particularly important aspect of real time digital data delivery systems. Since received packets are processed at a particular rate (to match the rate at which it is generated and transmitted), loss of synchronization leads to either buffer overflow or underflow at the decoder, and as a consequence, loss of presentation / display synchronization.

One solution to this issue in a transport system is to transmit timing information in the header of selected packets, to serve as a reference for timing comparison at the decoder. This can be done by transmitting a sample the system clock in the specified field, which indicates the expected time at the completion of the reading of that field from the bit stream at the transport decoder. The phase of the local system clock running at the decoder is compared to the sampled value in the bit stream at the instant at which it is obtained, to determine whether the decoding process is synchronized. In general, the sampled clock value in the bit stream does not directly change the phase of the local clock but only serves as an input to adjust the clock rate.

3.4 Interoperability and Transcoding

The problem has been raised frequently about the bit stream level interoperability of the transport system. There are two sides to this issue. One is whether a transport bit stream for one system can be carried on other communication systems, and the other is the ability of the transport system to carry bit streams generated from other communication systems.

In order to transmit the compressed video over the networks with different characteristics and bandwidth, video transport packets have to be able to adapt the changes in the video elementary stream. In the case where only one user is connected to the source, or independent transmission paths exist for different users, the bandwidth required by the compressed video should be adjusted by the source in order to match the available bandwidth of the most stringent link used in the connection. In the case where several users are simultaneously connected to the source and receiving the same coded video, as happen in video on demand (VoD) services, CATV services and Internet video, the existence of links with different capacities poses a serious problem. In order to deliver the same compressed video to all users, the source has to comply with the sub-network that has the lowest available capacity. This unfairly penalizes those users that have wider bandwidth in their own access links. By using transcoders in communication links, this problem can be resolved.

4 Proposed Error Resilience Approach

Digital video, when used in networked multimedia applications, suffers from data losses / errors. This is a serious problem in the case of digital networks. Currently, there are several ways to recover from these losses or errors.

The MPEG video standard supports error recovery by inserting resynchronization markers into the bit stream. To avoid long chains of interdependent macroblocks, MPEG-4 creates self-containing video packets (VP) comparable to the group of blocks (GOB) structure in H.261/H.263 and the definition of slices in MPEG-1/MPEG-2.

Our future work will also comprise extensive tests of the efficiency of error resilience tools.

5 References

Ralf Schafer and Thomas Sikora, “Digital video coding standards and their role in video communications”, Proceeding of IEEE, Vol. 83, No. 6, pp. 907-924, June 1995.

N. S. Jayant and P. Noll, Digital Coding of Waveform, Englewood, Cliffs, NJ: Prentice-Hall, 1984.

N.Ahmed, T.Natrajan and K.R.Rao, “Discrete Cosine Transform”, IEEE Trans. on Computers, Vol. C-23, No. 1, pp. 90-93, December 1984.

K. Jack, Video Demystified, 3rd ed., San Diego: HighText Interactive, 2000.

ISO/IEC 11172-2, “Information Technology – Coding of Moving Pictures and Associated Audio for Digital Storage Media at up to about 1.5Mbits/s – Video”, Geneva, 1993.

J. W. Woods (ed.), Subband Image Coding, Boston: Kluwer Academic Publishers, 1991.

ISO/IEC 13818, “Information Technology – Generic Coding of Moving Pictures and Associated Audio, Recommendation H.262”, International Standard, Paris, 25 March 1995.

ISO/IEC 14496-2, Information Technology – Generic coding of audio-visual objects – Part 2: Visual, Atlantic City, Nov. 1998.

Robert O. Eifrig, Xuemin Chen, and Ajay Luthra, “Intra-macroblock DC and AC coefficient prediction for interlaced digital video”, US Patent Number 5974184, Assignee: General Instrument Corporation, Oct. 26, 1999.