Advanced Telecommunications and Signal Processing Program

Academic and Research Staff
Prof. Jae S. Lim

Research Affiliate
Carlos Kennedy

Visiting Scientists
Dr. Cheul-hee Hahm
Mr. Gi-Byoung Park

Graduate Students
David M. Baylon
Rubén E. Galarza
Theresa Huang
Raynard O. Hinds
Eric C. Reed
Wade Wan

Support Staff
Cindy LeBlanc

Introduction

The present television system was designed nearly 50 years ago. Since then, there have been significant developments in technology, which are highly relevant to the television industries. For example, advances in the very large scale integration (VLSI) technology and signal processing theories make it feasible to incorporate frame-store memory and sophisticated signal processing capabilities in a television receiver at a reasonable cost.
To exploit this new technology in developing future television systems, the research areas of the program focused on a number of issues related to digital television design. As a result of this effort, significant advances have already been made and these advances have been included in the U.S. digital television standard. Specifically, the ATSP group represented MIT in MIT's participation in the Grand Alliance, which consisted of MIT, AT&T, Zenith Electronics Corporation, General Instrument Corporation, David Sarnoff Research Center, Philips Laboratories, and Thomson Consumer Electronics. The Grand Alliance digital television system served as the basis for the U.S. Digital Television (DTV) standard, which was formally adopted by the U.S. Federal Communications Commission in December 1996. The standard imposes substantial constraints on the way the digital television signal is transmitted and received. The standard also leaves considerable room for future improvements through technological advances. Current research focuses on make these future improvements.

In addition to research on issues related to the design of digital television system, the research program also includes research on signal processing for telecommunications applications and research on speech enhancement.

Video Compression with Complete Information for Pre-Recorded Sources

Sponsor: Advanced Telecommunications Research Program

Project Staff: David M. Baylon

Traditional video compression algorithms focus on causal processing of the video data. The causal constraint is necessary for real-time video programs where delay is an important consideration. However, many video programs such as movies are pre-recorded and can be processed offline prior to compression. This study proposes using pre-processing for pre-recorded video programs to obtain useful information for compression. By exploiting information obtained about a program prior to compression, better video quality can be achieved.

An application of this general approach to a buffer-constrained quantization problem using MPEG-2 intraframe video compression is studied. Intraframe video compression is useful for noisy channels or when lower complexity is desired, although the approach can be applied to interframe compression as well. The issues addressed include what a priori information should be extracted, how this information should be used, and what performance gains can be achieved. In particular, rate-distortion curves are proposed as a set of a priori sufficient statistics for intraframe compression, and a noncausally computed target quality level is introduced into the distortion function. A new noncausal iterative algorithm is developed to reduce the number of quantization changes, and it is shown to be optimal under certain conditions. It is demonstrated that by using this algorithm, improved and more constant video quality can be delivered using the noncausal approach relative to a traditional causal approach. Gains of up to 1.0 dB in PSNR were observed in multiscene sequences, corresponding to savings of up to 10% in bit rate.

Since there exist many applications involving the transmission and storage of pre-recorded video programs, such as video-on-demand, digital versatile disk (DVD), etc., there are many potential applications of the proposed approach. In addition, the approach can be applied to processing and compression of other sources known a priori such as images and audio.

Speech Enhancement

Sponsors: GEM Fellowship, U.S. Federal Bureau of Investigation, the Innovative Development and Enterprise Advancements Program, and the Advanced Telecommunications Research Program

Project Staff: Rubén E. Galarza

Enhancement of speech degraded by additive noise has been an active area of research for the past 20 years. Many algorithms have been developed to try and solve this problem, each with varying degrees of success. This project focuses on one such technique, which achieves enhancement by segmenting speech into (nearly) stationary regions in the time-frequency plane.

The speech signal is divided into stationary regions through filtering and adaptive windowing. A filter bank is used to separate the signal into frequency channels. Then, an adaptive length window is applied to each channel. The length of the window is determined by a similarity measure based on the cross-correlation of spectra in adjacent time segments. This scheme allows the windowing stage to adapt following specific speech characteristics at different regions in frequency. After the stationary regions have been selected, each of them is enhanced using linear prediction modeling and Wiener filtering. The parameters of the linear prediction model and Wiener filter are adjusted according to the region's estimated signal to noise ratio (SNR).

Although preliminary testing using this system showed promising results, a simpler implementation with higher quality output was desired. To this end, several stages of the original system were modified. For example, the adaptive windowing stage was changed to update parameters according to SNR. Also, the enhancement stage was modified so that low SNR regions at higher frequencies are less affected by Wiener filtering than high SNR regions at lower frequencies. These adjustments simplified the algorithm significantly while moderately improving speech intelligibility. More testing is needed to increase intelligibility in the enhanced speech.

Study on Duplicate Video System

Sponsors: Samsung Electronics
Advanced Telecommunications Research Program
Project Staff: Cheul-hee Hahm

In video transmission with high quality and reliability, such as contribution and primary distribution of video for Olympic broadcasting, a duplicate redundant system based on private channels is often set up in order to prevent transmission failure. The basic idea of the duplicate redundant system is to decrease the risk of transmission failure by distributing the video signal into several streams. Generally, a regular channel and a backup channel are used in the duplicate video system. The regular channel is used for the normal periods without any transmission failure, and when it has some troubles, a backup channel is used. The requirements for the duplicate video system are to transmit the encoded video stream through separate channels and to reconstruct the picture from any single stream.

To satisfy these requirements, a simulcast scheme and a flat multi-scalable scheme using a shifted picture have been used. In the simulcast scheme, the same video bitstream is transmitted through each channel and the picture is reconstructed from each stream. Because the regular channel is used for most of time, the capacity of the backup channel of this scheme is usually wasted. In the flat, multi-scalable scheme using a shifted picture, the preprocessor shifts the pixel positions of the picture for one channel before encoding and the decoder restores the original positions of the pixels in one channel by inverse shifting and averages the two decoded pictures. However, the utilization of the backup channel of this scheme is low. So, to increase the utilization of the backup channel for failure-free video transmission, an alternate temporal sub-sampling scheme is currently being examined. In this scheme, the encoded even pictures and the encoded odd pictures are transmitted through separate channels and the pictures at the decoder are chosen from each decoded picture of each channel alternatively. By this method, we can use the full backup channel for failure-free video transmission.

Real-Time Video on the Internet

Sponsors: Lucent Technologies Fellowship
Advanced Telecommunications Research Program
Project Staff: Raynard O. Hinds

The Internet has become sufficient at transmitting data over packet-switched networks. This is because data has no inherent delay constraints and can handle the delay jitter that occurs due to variable queuing delays across the network as well as the excess delay that occurs from retransmission of lost packets. This research has looked at transmitting real-time video over this same network. Real-time video cannot tolerate excessive delay. Packets arriving after their scheduled playback point at the receiver are discarded. Video sequences are capable of tolerating loss. Block-based video coders, which rely on motion-compensated block prediction for more data compression, have been used to code video over networks. With the resulting packet loss that occurs on congested networks, coding mode selection for each macroblock is significant in determining the overall distortion on the decoded video sequence. In the past, a methodology for optimal mode selection in the presence of potential macroblock loss was developed for a restricted set of video coders. For a given channel erasure description and error concealment method, macroblock modes are selected for block-based video coders with zero-motion compensation only to minimize the distortion for a given bit-rate constraint. This algorithm was extended to find the optimal mode selection when motion compensation is allowed. This has led to a more efficient robust coder in the presence of loss.

Compression of Underwater Video Sequences using Quantizer/Position Dependent-Encoding

Sponsors: Charles Stark Draper Laboratory,
Advanced Telecommunications Research Program

Project Staff: Theresa Huang

Transmission of underwater video sequences obtained by an unmanned underwater vehicle to a mother ship requires a transmission channel with a bit rate capacity on the order of megabits per second. The transmission channel available has a bit rate capacity of 10 kilobits per second. In order to obtain such low bit rates, the video sequence is first preprocessed. Motion estimation is used to reduce the temporal redundancies, and an 8x8 block Discrete Cosine Transform (DCT) is applied to minimize the spatial redundancies. Each 8x8 block of DCT coefficients is then encoded into a bit stream. Typically, one standard codebook is used to assign codewords to each event that occurs. Position-Dependent Encoding (PDE) further reduces the bit rate by exploiting the varying statistical properties of the runlength and the amplitude as a function of the starting position. Coefficients in the lower frequencies are most likely to have a large amplitude and a short runlength, whereas coefficients in the higher frequencies are most likely to have zero amplitude and a long runlength. Further compression is achieved by using Quantizer/Position-Dependent Encoding. Not only do the statistics vary with position, they also change as a function of the quantizer. When a very fine quantizer is used, very little data is thrown away. However, when a coarse quantizer is used, most of the DCT coefficients become zero. As a result, Quantizer/Position-Dependent codebooks significantly reduce the bit rate.

Multi-Dimensional Bit Rate Control for Video Communication

Sponsors: Charles Stark Draper Laboratory
Advanced Telecommunications Research Program

Project Staff: Eric Reed

This thesis focuses on bit rate control for applications involving the real-time transmission of video signals at very low bit rates. In conventional bit rate control schemes, the buffer level is controlled by adjusting the quantizer parameter. The emphasis has been on how to choose quantizers that will maximize video quality under a delay constraint. During the video coding process, the coding frame rate and spatial resolution are typically fixed and chosen independently of the coding process. At very low bit rates, this particular scenario often leads to poor video quality and loss of important image details.

To overcome this problem, we investigate a more general Multi-Dimensional (M-D) bit rate control where a vector of coding parameters, rather than a single quantization parameter, is jointly adapted to control the bit rate. In our approach, temporal and spatial subsampling parameters are adapted along with the quantizer parameter throughout the coding process. The advantage of this approach is that each parameter including the quantizer can vary at a much slower rate when compared with a scheme that adjusts only one parameter while under the same delay constraint. This gives the encoder more control over the quantizer decisions that can be made. The added flexibility of M-D bit rate control may be useful for a variety of applications requiring very low bit rate video transmission especially in environments where the channel and the source characteristics can change abruptly.
As a case study, we experiment with the compression of underwater images taken from an unmanned undersea vehicle (UUV) down to below 10 kbps. The goal is to transmit these images to the mother ship at the ocean surface.

Publication

Reed, E., and F. Dufaux. “Constrained Bit Rate Control for Very Low Bit Rate Streaming Video Applications.” Submitted to IEEE Transactions on Circuits and Systems for Video Technology, October 1999.

Migration to Higher Resolution Digital Television Systems

Sponsors: Advanced Telecommunications Research Program

Project Staff: Wade Wan

The new U.S. digital television standard that recently went into commercial service incorporates many technological advances made over the past few decades. As a result, digital television systems are significantly better than their analog counterparts. Despite the substantial improvements, the digital television standard has a significant limitation in its video resolution. The 1080P format (1080 lines with progressive scanning at 60 frames per second) is not allowed in the standard. The need to migrate to this resolution in the future has already been recognized and desired by terrestrial broadcasters. This research is focused on developing methods to migrate to the 1080P format using two sets of bitstreams. Standard video bits would be transmitted at a resolution format allowed by the current standard. In addition, enhancement video bits would also be transmitted. An advanced receiver would receive the standard bits and convert the standard video to the desired 1080P format with the assistance of the enhancement bits. Standard receivers would ignore the enhancement bits. This backward-compatible approach is highly desirable so as not to render earlier receivers obsolete.

Three issues with this backward-compatible approach to migrate to the desired 1080P format are currently being investigated. The first issue is the selection of the resolution format that will be used to generate standard video bits. This selection will significantly affect the enhancement bits. For example, 720P video (720 lines with progressive scanning at 60 frames per second) may be transmitted and then spatially upsampled at the receiver. Alternatively, 1080I video (1080 lines with interlaced scanning at 60 fields per second) may be transmitted and then deinterlaced at the receiver. The second issue is the division of the channel bandwidth between the standard and enhancement bitstreams. A 6 MHz terrestrial channel can currently support approximately 18 megabits per second (Mbps). One migration scheme may involve using 17 Mbps for standard video and 1 Mbps for enhancement data. Another scheme may utilize 15 Mbps for standard video and 3 Mbps for enhancement data. The third issue is what information should be encoded as the enhancement data to efficiently aid the migration from the standard format to the desired 1080P format. Understanding these issues will allow the development of efficient migration schemes for the future and may influence the current debate between 720P and 1080I as the preferred digital format for broadcasters.