Chapter 2. Advanced Telecommunications and Signal Processing Program

Academic and Research Staff

Professor Jae S. Lim

Visiting Scientists and Research Affiliates

M. Carlos Kennedy

Graduate Students

John G. Apostolopoulos, David M. Baylon, Ruben E. Galarza, Raynard O. Hinds, Eric C. Reed, Lon E. Sunshine, Wade Wan

Technical and Support Staff

Cindy LeBlanc, Christopher R. O'Donnell, Denise M. Rossetti

2.1 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 Advanced Television Research Program (ATRP) was established at MIT in 1983 by a consortium of U.S. companies.

The major objectives of ATRP are:

1. To develop the theoretical and empirical basis for the improvement of existing television systems, as well as for the design of future television systems.
2. To educate students through television-related research and development and to motivate them to undertake careers in television-related industries.
3. To facilitate continuing education of scientists and engineers already working in the industry.
4. To establish a resource center where problems and proposals can be brought for discussion and detailed study.
5. To transfer the technology developed from this program to the television industries.

The research areas of this 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 ATRP group represented MIT in MIT's participation in the Grand Alliance, this alliance 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 addition to research on issues related to the design of digital television system, our research program also includes research on (1) signal processing for telecommunications applications and (2) speech enhancement.

2.1.1 Patents

Lim, J.S. "Method and Apparatus for Encoding, Decoding, and Compression of Audio-Type Data." Patent #5,625,746. April 29, 1997.

Lim, J.S. "Encoding, Decoding and Compression of Audio-Type Data Using Reference Coefficients Located Within a Band of Coefficients." Patent #5,640,486. June 17, 1997.

2.2 Signal Processing for Signals with Arbitrarily Shaped Regions of Support

Sponsors

Advanced Telecommunications Research Program
AT&T Fellowship

Project Staff

John G. Apostolopoulos

Many applications today and in the near future will entail efficient processing of multidimensional (MD) signals with arbitrarily shaped (nonrectangular) regions of support. For example, most high level representations of images or video incorporate 2D or 3D models which decompose the scene into arbitrarily shaped objects or regions. Medical imaging often results in 2D or 3D imagery where the relevant information is localized over an arbitrarily shaped region. Furthermore, many areas of scientific research involve problems defined over arbitrarily shaped domains. High quality and computationally efficient processing of these *arbitrarily shaped signals* is important for the success of these applications.

Transform/subband representations form a basic building block for many signal processing algorithms and applications. Most of the effort has focused on developing representations for infinite-length signals, with simple extensions to finite-length 1D and rectangular support 2D signals. However, the important problem of representing 2D or general MD signals with arbitrarily shaped regions of support has received relatively little attention. We have developed a novel framework for creating critically sampled perfect reconstruction transform/subband representations for discrete 1D, 2D, and general MD signals defined over arbitrarily shaped regions of support. Our method selects an appropriate subset of vectors from an (easily obtained) basis for a larger (superset) signal space in order to form a basis for the arbitrarily shaped signal. In particular, we have developed a number of promising wavelet representations for arbitrary-length 1D signals and arbitrarily shaped 2D/MD signals that provide high performance with low complexity.

2.2.1 Publications

Apostolopoulos, J.A., and J.S. Lim. "Transform/Subband Representation for Signals with Arbitrarily Shaped Regions of Support." *Proceedings of ICASSP 97*, Munich, Germany, April 21-24, 1997, pp. 2097-2100.

2.3 Source Multiplexing for Variable Bit Rate Video with Partial or Complete Information

Sponsor

Advanced Telecommunications Research Program

Project Staff

David M. Baylon

Video compression algorithms for fixed bandwidth systems typically operate at fixed targeted compression ratios. A problem with fixed bit rate systems is that video quality can vary greatly within an image and across an image sequence. In order to achieve more constant quality video, variable bit rate coding can be performed, where bit allocation is allowed to vary more widely as the video statistics change.

Although it may be inefficient to transmit a single variable rate encoded source over a constant rate channel, it becomes more appropriate if several variable rate sources are combined to share a given fixed bandwidth. In source multiplexing, bit-rate fluctuations of each source are effectively averaged, making the aggregate bitstream more suitable for a constant rate channel. Applications for source multiplexing include video transmission over terrestrial or cable channels, where it may be desirable to transmit several video programs simultaneously over a fixed rate channel.

Traditional video compression and video multiplexing systems for bandwidth limited channels employ causal processing of the video data. This is necessary for delivery of real-time video programs. However, many video programs such as movies are prerecorded and can be pre-processed before performing the actual compression and multiplexing for transmission. By exploiting information about each of the entire video sequences before the compression and multiplexing, better video quality can be delivered compared to causal compression and multiplexing systems. The goal of this research is to characterize what gains in video quality can be achieved by having noncausal information about the video programs being compressed and multiplexed. This work focuses on MPEG-2 compression, and it is assumed that the video encoder and decoder buffer sizes are constrained, and that the transmission duration for a video program over a bandwidth limited channel is the same as the video program duration. Preliminary results with noncausal processing of a single video program show improved and more constant video quality compared to causal processing.

2.4 Speech Enhancement

Sponsors

GEM Fellowship
U.S. Federal Bureau of Investigation

Project Staff

Ruben E. Galarza

The problem of enhancing speech degraded by additive background noise has received considerable attention over the years. Various techniques have been developed to solve this problem, each capitalizing on specific characteristics or constraints. To date, model based enhancement systems (enhancement systems based on parameter estimation) have produced the best results in reconstructing speech degraded by additive noise. However, these techniques are ultimately limited by the model in which they are based.

In general, the underlying assumption of all model based algorithms for enhancing speech is that each speech segment is assumed stationary for a fixed window. This assumption is false, since the duration of stationarity varies for different classes of sounds and speakers. A viable solution is to find the desired stationary regions before the model based enhancement takes place.

Following this train of thought, an algorithm was developed for the segmentation of speech into (nearly) stationary regions by means of M-band decomposition and adaptive windowing. Noise reduction is then achieved by applying a modified Wiener filter based on selective linear prediction, taking into consideration the local signal to noise ratio of the region being enhanced. The main purpose of this research is to study alternatives to improve this technique and make it more computationally efficient for real- time applications.

2.5 Real-Time Video on the Internet

Sponsors

Advanced Telecommunications Research Program
Lucent Technologies Fellowship

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 can not tolerate excessive delay. Packets arriving after their scheduled playback point at the receiver are discarded; however, 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. 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 to minimize the distortion for a given bit-rate constraint. Future research must be done to find a solution when motion compensation is allowed.

2.6 Digital Processing of Underwater Images

Sponsors

Advanced Telecommunications Research Program
Charles S. Draper Laboratory
U.S. Navy - Office of Naval Research
NDSEG Graduate Fellowship

Project Staff

Eric C. Reed

This project focuses on compression of underwater images taken from an unmanned undersea vehicle (UUV) system. The channel that is available for communication between a UUV system and the remote location may support a bit-rate capacity as low as 10 kbps. However, the video sequence of underwater images may have in the order of a few million pixels/sec at a resolution of 8 bits/pixel. Therefore the bit-rate reduction required for video transmission of underwater images is by a factor of several thousand. To achieve such a high rate of compression, a video compression system must fully adapt to the characteristics of underwater images.

Underwater images are often blurred and have low contrast because of low light illumination, scattering of light, water turbulence, and sedimentation. In addition, since a UUV system typically does not move at a very fast speed relative to the frame rate, the correlation that exists in the temporal domain is very high. We are exploring methods to exploit the temporal redundancy. For example, we are considering temporal subsampling of the source prior to encoding and then reconstructing the skipped frames at the decoder by motion-compensated temporal interpolation. We are also considering global parametric motion models which require negligible overhead but have enough degree of freedom to describe the 2D image motion induced from the 3D motion of the camera attached to the UUV.

Furthermore, since an image is formed by recording the light reflected from an object that is illuminated by a light source attached to the UUV, an image can be modeled as a product of illumination and reflectance components. Since the illumination component is approximately stationary over long periods of time, it may represent redundant information and therefore only needs to be transmitted once over long periods of time.

The goal of this project is to compress the underwater images to a bit rate of 10 kbps while preserving image detail and overall video quality for a human operator to observe. This technology may have a significant impact on a variety of future underwater applications.

2.6.1 Publication

Reed E., and J.S. Lim. "Efficient Coding of DCT Coefficients by Joint Position-dependent Encoding." *Proceedings of ICASSP*, 1998.

2.7 HDTV Transmission Format Conversion and the HDTV Migration Path

Sponsors

Advanced Telecommunications Research Program

Project Staff

Lon E. Sunshine

The current proposal for terrestrial HDTV broadcasting allows for several possible transmission formats. Because production and display formats may differ, it will be necessary to convert between formats in an effective way. A key to this process is the de-interlacing process. Since HDTV will presumably move toward progressive display systems, it will be necessary to de-interlace nonprogressive source material. The research will consider topics relating to conversion among the six formats being proposed for the US HDTV standard.

As HDTV evolves, it is possible (and likely) that more transmission formats will be allowed. Furthermore, additional bandwidth may be allocated for some channels (terrestrial and/or cable). This research will consider the issues related to the migration of HDTV to higher resolutions. Backward compatibility and image compression and coding issues will be addressed.

2.8 Study on Migration to a Higher Resolution Digital Television System

Sponsors

Advanced Telecommunications Research Program

Project Staff

Wade Wan

High-definition television will begin broadcasting over the next few years. The need to broadcast at higher resolutions than those currently allowed has already been recognized. A generally accepted resolution goal for terrestrial high-definition broadcasting is more than 1000 lines with progressive scanning at 60 frames per second within one single 6 MHz channel. Current transmission and compression technologies cannot support such a high-video resolution within a 6 MHz channel.

We are currently investigating methods to display video at the desired higher resolution formats using video enhancement bits. This approach would consist of sending two sets of information within the bandwidth allocated for HDTV transmission. Standard video bits would be transmitted at a lower resolution format that current technologies can support. An advanced HDTV receiver could receive the standard video bits and then convert the video to a higher resolution format with the assistance of the video enhancement bits. Standard HDTV receivers could ignore the enhancement bits and display video at the lower resolution format.

This backward-compatible approach is highly desirable tonot make standard HDTV receivers obsolete. Previous research in our laboratory has shown that this approach is feasible for transmitting 1080 line interlaced scanned images at 60 fields per second and using the enhancement bits to assist the deinterlacer performed at the receiver. We will be investigating the use of different transmission formats to migrate toward higher resolution formats using this enhancement data approach. For example, progressively scanned video may be transmitted with 720 lines at 60 frames per second and then spatially upsampled at the receiver. This research will allow us to determine the best transmission format and signal processing schemes for optimum picture quality.