Advanced Telecommunications and Signal Processing Program

Academic and Research Staff
Prof. Jae S. Lim

Research Affiliate
Carlos Kennedy

Graduate Students
Brian Heng, Ken Schutte, James Thornbrue, Wade Wan, Chuohao Yeo

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 making 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.

1. Multiple Description Video Coding for Streaming Video Delivery

Sponsor
Advanced Telecommunications Research Program

Project Staff
Brian Heng

Multiple description video coding is used to encode a video sequence into two or more complimentary streams. This is done in such a way that each stream is independently decodable with acceptable quality, but together the streams combine to provide the highest quality. Using this type of multiple description coding can help to offset the losses and delays caused by lossy packet networks such as the Internet.

It has been shown that using multiple description video coding can lead to significant benefits in terms of received video quality. We are investigating new source coding techniques for providing multiple description streams in order to determine methods which provide greater gains in resiliency with fewer losses in coding efficiency.

2. Energy Distributions in Clear Speech and Implications for Enhancement of Undegraded Speech

Sponsors
Advanced Telecommunications Research Program

Project Staff
Ken Schutte

Clear speech refers to the speaking style one often uses when confronted with a difficult speaking situation such as being in a noisy environment or speaking to the hearing impaired. Tests have shown that clear speech is consistently more intelligible than typical speaking style (referred to as conversational speech), leading to a 17% increase in the identification of key words. Many applications could benefit from a signal processing scheme which could transform a given conversational speech signal into some approximation of its clear speech counterpart.

Recent research in RLE's Sensory Communication Group has attempted to isolate the acoustic properties of clear speech which provide its high intelligibility. Some interesting results have been found, but they have not yet led to a successful technique for intelligibility enhancement. This work is another attempt at analyzing clear and conversational speech to determine (a) what causes the intelligibility improvement and (b) how and if this could be applied as a general speech enhancement scheme. This work differs slightly from previous studies by looking more closely at a few samples and taking into account different measurements in the statistical analysis.

3. Adaptive Format Conversion Information as Enhancement Data for the HDTV Migration Path

Sponsors
Advanced Telecommunications Research Program

Project Staff
James Thornbrue

Although the high-definition television (HDTV) standard has significant improvements over its NTSC predecessor, there are still limitations on the resolution that it supports. Specifically, the target resolution of 1080x1920 pixels, progressively scanned, at 60 frames per second (1080P), is not permitted in the HDTV standard. 1080P requires a sample rate of approximately 125 Mpixels/s, which exceeds the maximum sample rate of 62.2 Mpixels/s allowed by the MPEG-2 video compression portion of the HDTV standard. In addition, 1080P can not be compressed into a single 6 MHz broadcast channel without significant loss of picture quality for difficult scenes. The question of how to add support for 1080P and other higher-resolution video formats while dealing with these two issues—backward compatibility and limited bandwidth—is what is known as the migration path problem for high-definition television.

Previous research suggests that a scalable video codec using adaptive format conversion (AFC) information as a low-bitrate enhancement layer may be an ideal solution to the migration path problem. AFC information tells the decoder which of several predefined interpolation methods to use, on a block-by-block basis, in order to best reconstruct the higher-resolution video sequence. The low-bitrate AFC enhancement layer is sent in addition to an independent base layer which conforms to the MPEG-2 sample rate constraint.

Previous research, using an interlaced base layer and four deinterlacing modes, showed that adaptive format conversion significantly improves video quality using a very low enhancement layer bitrate. However, the implementation ignored issues relating to encoder optimization and the specific bandwidth constraints of high-definition television. Current research explores these issues by determining, first, the performance that can be expected for typical HDTV bitstreams, and second, when (and if) adaptive format conversion outperforms nonadaptive format conversion when the total bandwidth (base plus enhancement layer) is fixed.

4. Adaptive Format Conversion as Enhancement Information in Scalable Video Coding

Sponsors
Advanced Telecommunications Research Program

Project Staff
Wade Wan

Many video broadcasting applications must provide service to a multicast environment. In this environment, multiple clients require different types of service, also known as service levels, due to variations in their available bandwidth, processing power and memory resources. It is desirable for a video server to provide different resolutions and/or qualities of the same video sequence to satisfy each client type while minimizing the cost of reaching the audience. From the video coding point of view, this cost is the total bandwidth transmitted from the server. Scalable coding techniques are often used to efficiently provide multicast video service and involve transmitting a single independently coded base layer and one or more dependently coded enhancement layers. Clients can decode the base layer bitstream and none, some or all of the enhancement layer bitstreams to obtain video quality commensurate with their available resources. Enhancement layers are dependent on the decoded base layer and the efficient coding of information in enhancement layers has been a topic of substantial research interest. A well-known type of enhancement information is residual coding, which is used in scalable coding schemes such as the spatial scalability profiles in the MPEG-2 and MPEG-4 multimedia standards. There is another type of information that can be transmitted instead of (or in addition to) residual coding. Since the encoder has access to the original sequence, it can utilize adaptive format conversion to generate the enhancement layer and transmit the different format conversion methods as enhancement data. The use of adaptive format conversion for scalable coding has not been studied in detail, but recent research has shown that it can significantly improve video scalability.

This research examines the use of adaptive format conversion information as enhancement data in scalable video coding. The main focus is to determine the conditions where adaptive format conversion can be used instead of or in addition to residual coding to improve video scalability. Simulation results are performed for a wide range of base layer qualities and enhancement bitrates to determine when adaptive format conversion can improve video scalability. Since the parameters needed for adaptive format conversion are small compared to residual coding, adaptive format conversion can provide video scalability at low enhancement layer bitrates that are not possible with residual coding. In addition, adaptive format conversion can also be used in addition to residual coding to improve video scalability at higher enhancement layer bitrates. Many scalable applications may benefit from the use of adaptive format conversion. An application that adaptive format conversion is well-suited for is the migration path for digital television where it can provide immediate video scalability as well as assist future migrations.

5. An Investigation of Methods for HDTV Format Conversions

Sponsors
Advanced Telecommunications Research Program
Public Service Commission (PCS) of Singapore

Project Staff
Chuohao Yeo

For the last few years, broadcasters have been in transition from the old NTSC standard to the new ATSC digital television standard. A key feature of the new standard is the plethora of supported video formats. While this gives broadcasters greater flexibility in their operations, it has also necessitated format conversion at the transmitter-level and the receiver-level. It is possible for current production equipment to produce video at a resolution of 720x1280 progressively scanned (720p) at 60 frames per second or at a resolution of 1080x1920 interlaced (1080i) at 60 fields per second, but broadcasters can choose to transmit in many other video formats. In addition, a typical consumer HDTV set has only one display format, therefore other video formats that are received must be first converted to its native display format. The ATSC digital television standard also allows for multicasting, or the broadcasting of multiple programs on the same channel. Therefore, broadcasters may be interested in broadcasting programs in a video format that has lower resolution than production formats. While computation time may be an issue in commercial systems, the only criterion used here to evaluate different methods is the output quality. A number of conversion techniques used for de-interlacing and image resizing have been investigated to determine which combination of methods offers the best performance for a small subset of possible format conversions.

Publications

Journal Articles Published

E. Reed and J. Lim, “Optimal Multi-Dimensional Bit Rate Control for Video Communication,” *IEEE Transactions on Image Processing,* Vol. 11, No.8, pp. 873-885, August, 2002.

Lim, J.S., “High Definition Television (HDTV),” of *Wiley Encyclopedia of Telecommunications*, scheduled for publication in December, 2002.

Theses

Cheng, Christopher, “The Resolution Enhancement of Real-Time Captured Images,” March, 2002.

Yeo, Chuohao, “An Investigation of Methods for Digital Television Format Conversions,” May, 2002.

Wan, Wade K., Adaptive Format Conversion Information as Enhancement Data for Scalable Video Coding,” June 2002.