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

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 Through Adaptive Segmentation

Sponsor
Advanced Telecommunications Research Program

Project Staff
Brian Heng

The Internet is a best-effort service characterized by variable bandwidths, packet losses and delays. This environment is inhospitable to real-time applications which require a minimum quality of service in order to maintain synchronization between the sender and receiver. In some cases, these problems may be reduced to some extent by the use of playback buffers which store data at the receiver in the event that the network throughput drops temporarily. However, this type of buffering introduces a significant delay into the system given the time necessary to initially fill the buffer. For some applications, like video conferencing, the amount of delay necessary to prevent buffer underflow may be unacceptable for maintaining the feeling of an interactive experience. In addition, the amount of buffering necessary in any situation is unknown ahead of time due to the time-varying properties of the network. For these reasons, today's solutions to streaming video either require large delays and/or suffer from severe glitches each time the network becomes congested.

This research focuses on improving streaming video applications through the use of multiple description coding. A multiple description coder segments a stream into two or more separately decodable streams and transmits these independently over the network. The quality of the received video improves with each received description, but the loss of any one of these streams does not cause complete failure. Thus video playback can continue, at a slight reduction in quality, without waiting for rebuffering or retransmission. Each approach to multiple description coding consists of a tradeoff between compression efficiency and robustness. How efficiently each method achieves this tradeoff depends on the level of quality and robustness desired and on the characteristics of the video itself. Previous approaches to multiple description coding have made the assumption that a single segmentation method would be used for an entire sequence. Yet, the optimal method of segmentation can vary depending on the goals of the system, it can change over time, and it can vary within a frame. This work introduces a unique approach to multiple description coding through the use of adaptive segmentation. By selecting from a set of segmentation methods, the system adapts to the local characteristics of the video as well as to rate-distortion goals of the system.

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 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.

Publications

Theses

Schutte, Kenneth T., “An Investigation of Clearly Spoken Speech and Possibilities of Intelligibility Enhancement by Redistribution of Energy,” May, 2003.

Thornbrue, James, “Adaptive Format Conversion Information as Enhancement Data for the High-definition Television Migration Path,” June, 2003.