Chapter 3. Advanced Television Research Program

Academic and Research Staff

Professor Jae S. Lim, Professor William F. Schreiber

Graduate Students

John G. Apostolopoulos, Babak Ayazifar, Matthew M. Bace, David M. Baylon, Warren H. Chou, Ibrahim A. Hajjahmad, David D. Kuo, Peter A. Monta, Aradhana Narula, Julien J. Nicolas, Julien Piot, Ashok C. Popat, Paul X. Shen, Lon E. Sunshine, Adam S. Tom

Technical and Support Staff

Debra L. Harring, Cynthia LeBlanc

3.1 Introduction

The present television system was designed nearly 35 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 techology in developing future television systems, Japan and Europe established large laboratories funded by government or industry-wide consortia. The lack of this type of organization in the United States was considered detrimental to the broadcasting and equipment manufacturing industries, and, in 1983, the Advanced Television Research Program (ATRP) was established at MIT by a consortium of U.S. companies.

Currently, consortium members include ABC, Ampex, General Instrument, Kodak, Motorola, NBC, NBC Affiliates, PBS, Tektronix, and Zenith.

The major objectives of the ATRP are:

1. To develop the theoretical and empirical basis for the improvement of existing television systems, as well as the design of future television systems;

2. To educate students through television-related research and development and to motovate 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 to which problems and proposals can be brought for discussion and detailed study; and

5. To transfer the technology developed from this program to the sponsoring companies.

The research areas of the program include (1) the design of channel-compatible advanced television (ATV) system, (2) receiver-compatible ATV system and digital ATV system, and (3) development of transcoding methods. Significant advances have already been made in some of these research areas. A channel-compatible ATV system has been designed and is scheduled to be tested in 1991 by the FCC for possible adoption as the U.S. HDTV standard for terrestrial broadcasting.

3.2 ATRP Facilities

The ATRP Laboratory's computer facilities are currently based on a network of seven Sun-4 workstations. There are approximately 13.8 GB of disk space, distributed among the various machines. Attached to one Sun-4 is a DATARAM Wide Word Storage system with 320 MB of RAM. The high speed interface to the Wide Word system drives a three-dimensional interpolator that was constructed by graduate students in the Lab. The three-dimensional interpolator has the ability to perform separable spatio-temporal interpolation. The output of the interpolator feeds a custom-built data concentrator which drives a Sony 2k by 2k monitor, running at 60 frames/second.

In addition to displaying high resolution real time sequences, the ATRP facilities include a 512 x 512 Rastertek frame buffer and an NTSC encoder. The Rastertek distributes static images to nearly a dozen monitors around the Lab and offices. The NTSC encoder allows image sequences to be recorded onto either three-quarter inch or VHS videotape. For hard copy output, the Lab uses an Autokon 8400 graphics printer for generating high resolution black and white images directly onto photographic paper.

Other peripherals include an Exabyte 8 mm tape drive, capable of storing 2.3 GB on an 8 mm cassette, a 16-bit digital audio interface with two channels and sampling rates to 48 kHz per channel, and an "audio workstation" with power amplifier, speakers, CD player, tape deck, etc.

For preparing presentations, the ATRP facilities also include a Macintosh SE30 microcomputer, a Mac IIx, and an Apple LaserWriter.

To support the growing computation needs of the group, several additional Sun-4 workstations will be installed in the near future. They will have 24-bit color displays, local disk storage, and DSP boards to assist with computation-intensive image processing. Some of the existing machines may also be supplemented with such DSP boards.

A fast network (FDDI) is under consideration to augment the current 10 Mbps Ethernet. The new network would enable much faster data transfer to display devices such as the Dataram and support large NFS transfers more easily.

3.3 Coding of the Motion Compensated Residual for an All-Digital HDTV System

Sponsor

Advanced Television Research Program

Project Staff

John Apostolopoulos, Professor Jae S. Lim

An All-Digital High Definition Television system is being developed at MIT to transmit higher quality video and audio information in the same channel bandwidth as today's conventional television. To achieve this goal, the system must eliminate redundant information which exists because of the high correlation inherent to video and audio. For normal television broadcasts, the image from one frame to the next is very similar. In order to reduce this redundancy in the temporal direction, we implement a motion estimation/motion compensation algorithm in the transmitter and the receiver. This algorithm generates motion vectors which are used by the receiver to predict the next frame from the current frame. The transmitter will transmit these motion vectors as well as the error between the predicted and the actual next frame. The receiver will apply these motion vectors to create the predicted next frame, and add to it the predic-

tion error to produce the next displayed frame. The prediction error, or residual, as it is commonly called, also contains much redundancy which can be reduced by further coding. The purpose of this work is to find the optimal method to code this motion compensated residual to achieve the "best" image quality at the receiver.

Methods for coding the residual that will be analyzed include the blocked DCT and various QMF subband filtering schemes. The primary performance criteria for comparing the various methods are (1) visual image quality and (2) the mean square error between the image at the receiver and the original image. Other issues to be considered include the effects of energy compaction of the residual, quantization of the coefficients and pixel-adaptive selection of the transmitted coefficients.

3.4 Motion-Compensated Vertico-Temporal and Spatial Interpolation

Sponsor

Advanced Television Research Program

Project Staff

Babak Ayazifar, Professor Jae S. Lim

In this project, we examine the application of a motion-estimation algorithm to the field and line-rate conversion issues which exist in the process of converting video signals from the European to the Amercian standards and vice-versa.

Topics explored are simultaneous temporal and vertical interpolation of image sequences and strictly spatial interpolation of individual frames (e.g., line and column doubling) using a novel generalized form of the well known "spatio-temporal" constraint equation-based motion estimation algorithms.

3.5 Receiver-Compatible Adaptive Modulation for Television

Sponsors

Project Staff

Matthew M. Bace, Professor Jae S. Lim

There have been numerous proposals for developing methods to improve upon the quality of the current NTSC television picture. Most of these proposals have concentrated on methods that increase either the spatial or the temporal resolution of the television picture. While these proposals promise significant improvements in picture quality, the fact remains that until an effective scheme for combatting channel noise has been introduced, these improvements can not be fully realized. Degradations such as random noise ("snow"), echo, and intersymbol interference (channel crosstalk) are still the greatest barriers to high-quality television.

This research developed a receiver-compatible scheme to reduce the effects of channel imperfections on the received television picture. In particular, the method of adaptive modulation was employed in an attempt to make more efficient use of the currently underutilized bandwidth and dynamic range of the NTSC signal. By concentrating more power in the higher spatial frequencies and using digital modulation to send additional information in the vertical and horizontal blanking periods, the existing television signal can be altered so that it is more robust in the presence of high frequency disturbances. Furthermore, it is possible to adjust the parameters of this scheme so that the modulated signal may be received intelligibly even on a standard receiver (although an improved receiver will be required to realize the full benefits of adaptive modulation).

Before we concluded which adaptive modulation scheme was optimal, many details were considered. Among the parameters which may be varied were: the control over the adaptation and compression factors; the form of the input low-pass filters; the interpolation scheme to be used at both the transmitter and receiver; and the encoding of the digital data. These parameters were adjusted to optimize the performance of the modulation scheme with repect to two fundamental performance criteria: the degree to which the channel degradations are removed when the signal is received on an improved receiver, and the degree to which the signal is distorted when received on a standard receiver.

This research was completed in June 1990.

3.6 Adaptive Amplitude Modulation for Transform Coefficients

Sponsor

Advanced Television Research Program

Project Staff

David M. Baylon, Professor Jae S. Lim

Adaptive amplitude modulation/demodulation (AM/DM) has been shown to be an effective noise reduction technique. However, the adaptation factors must be transmitted as side information. It is important to minimize the required side information in systems that have limited transmission bandwidth. This research focused on representing the adaptation factors by a few parameters exploiting properties of the signal in the transform (frequency) domain.

Previous investigations of adaptive amplitude modulation have been based upon time domain methods. Specifically, in two-dimensional subband filtering, an image is decomposed into a set of spatial frequency subbands that are adaptively modulated. Similarities among subbands are exploited in reducing the number of adaptation factors to about one-sixth the number of data points. Nevertheless, further reduction in the amount of side information is desirable.

This research took a different approach to reducing the amount of side information by adaptively modulating the transform of the signal. Transform coefficients of typical images tend to decrease in energy away from DC. By exploiting this property, the transform coefficients and the adaptation factors can be modeled with a few parameters (for example, an exponential model). Consequently, the amount of side information can be significantly reduced compared with the amount required by previous approaches.

Research focused on determining the best way to model the adaptation factors with a few parameters in systems that are bandwidth and peak power constrained. Performance criteria of the various AM/DM schemes included signal-to-noise ratios and overall image quality. Among the many ways of obtaining the coefficients (such as using subband filtering or the lapped orthogonal transformation (LOT)), the discrete cosine transformation (DCT) was used because of its many desirable properties, including coefficient uncorrelation, energy compaction, and efficient computation using the fast Fourier transform (FFT). Issues that were addressed included choosing the appropriate block size and determining the best AM/DM

method with an adaptive coefficient selection scheme (such as used in image coding systems). Both two-dimensional images and three-dimensional video were studied.

This research was completed in June 1990.

3.7 Transform Coding for High Definition Television

Sponsor

Advanced Television Research Program

Project Staff

Ibrahim A. Hajjahmad, Professor Jae S. Lim

Field of image coding is usefully applied in many areas. For instance, one of the most prominent of these areas is compressing channel bandwidth for image transmission systems; this is helpful for applications in HDTV, video conferencing, and facsimile. Another important area is reducing storage requirements for tasks such as digital video recording.

Image coding can be divided into a number of classes, depending on which aspects of the image are being coded. In the transform image coder,[1] an image is transformed from the spatial domain to a different domain more suitable for coding. Then, the transform coefficients are quantized and coded. When received, the coded coefficients are decoded and then inverse transformed to obtain the reconstructed image.

To perform transform coding, one must select an appropriate transform. In particular, the Discrete Cosine Transform (DCT) is very useful because of two important properties.[2] The first is the energy compaction property, which states that a large amount of energy is concentrated in a small fraction of the transform coefficients (typically the low frequency components). Because of this property, only a small fraction of the transform coefficients need to be coded, while little is sacrificed in terms of quality and intelligibility of the coded images. The second property is the correlation reduction property in which the high correlation among pixels intensities in the spatial domain is reduced. In effect, the redundant spatial information is not coded.

Current research is investigating the use of the DCT for bandwidth compression. In addition, new adaptive techniques for quantization and bit allocation are being studied to further reduce the bit rate without sacrificing image quality or intelligibility.

3.8 Adaptive Spatio-temporal Filtering

Sponsors

Advanced Television Research Program
Kodak Fellowship

Project Staff

David D. Kuo, Professor Jae S. Lim

The current NTSC television standard specifies a frame rate of 60 fields/second throughout the transmission chain. This frame rate was chosen to minimize the visibility of annoying flicker at the display. However, to eliminate flicker, only the display must operate at the high frame rate; there is no need to constrain the channel to 60 frames/second. It is widely accepted that there exists a great deal of correlation between neighboring frames of an image sequence. As such, a high frame rate through the channel seems bandwidth inefficient.

One way to take advantage of the correlation between neighboring frames is to transmit only a temporally subsampled version of the original sequence and to rely on the receiver to recover the inbetween frames. However, prior work along these lines suggested that the receiver must have more information than simply the subsampled frames. This research focused on using motion vectors as part of the image sequence representation.

There were three main areas of focus in this research. First, it considered the use of adaptive spatio-temporal prefiltering as a means of reducing the aliasing that arises from temporal subsampling. Secondly, the characteristics of the motion vectors were explored. Finally, our research considered how to use multiple frames of data to improve the motion estimation process.

This research was completed in June 1990.

---

[1] J.S. Lim, *Two-Dimensional Signal and Image Processing* (Englewood Cliffs, N.J.: Prentice Hall, 1990); R.J. Clarke, *Transform Coding of Images* (London: Academic Press, 1985).

[2] N. Ahmed, T. Natarajan, and K.R. Rao, "Discrete Cosine Transform," *IEEE Trans. Comput.* C-23: 90-93 (1974).

3.9 Signal Processing for Advanced Television Systems

Sponsor

Advanced Television Research Program

Project Staff

Peter A. Monta, Professor Jae S. Lim

Digital signal processing (DSP) will play a large role in future advanced television systems. Source coding to reduce the channel capacity necessary to transmit a television signal and display processing such as spatial and temporal interpolation are its major applications. Present-day television standards will also benefit significantly from signal processing designed to remove transmission and display artifacts. This research will focus on algorithms and signal models designed to enhance current standards (both compatibly and with some degree of cooperative processing at both transmitter and receiver) and to improve proposed HDTV systems.

Given a receiver with a high-quality display and significant computation and memory, the American television standard, NTSC, can be improved in a number of ways. Interlace artifacts, such as line visibility and flicker, can be removed by converting the signal to a progressive format prior to display. Color cross-effects can be greatly reduced with accurate color demodulators implemented with DSP. If the original source material is film, an advanced receiver can recover a much improved image by exploiting structure in the film-NTSC transcoding process; such an algorithm has been implemented and tested.

Similar ideas apply to HDTV systems. For example, film will be a major source material well into the next century, and HDTV source coders should recognize film as a special case, trading off the inherent reduced temporal bandwidth for better spatial resolution. The MIT Channel Compatible (MIT-CC) HDTV system will adapt to film in this way.

3.10 MIT Channel Compatible System

Sponsor

Advanced Television Research Program

Project Staff

Julien J. Nicolas, Professor Jae S. Lim

The MIT Channel Compatible system was developed by ATRP over the course of the last three years as a new high-efficiency alternative to existing television systems and as a candidate for the FCC/ATTC competition. Key features of this system include (1) decomposition of the signal into a set of spatio-temporal frequency subbands, (2) a hybrid analog/digital representation of the signal for both source and channel coding; and (3) use of a technique called adaptive modulation to reduce the effects of the channel noise and other impairments in the areas where they are most noticeable.

Recent investigations have been aimed at finding efficient ways of representing the hybrid information and especially at reducing the amount of digital data required by this system. In the MIT-CC system, digital data is used to represent the lowest spatio-temporal subbands to code the location of the largest subband coefficients and the adaptive modulation coefficients. New methods to reduce selection information are being studied currently. These methods will be used in conjunction with effective adaptive modulation techniques previously developed and a hybrid digital/analog transmission scheme to meet the bandwidth constraints for terrestrial broadcasting and to optimize the information transfer in a large portion of the service area.

Future work will involve comparing the spatio-temporal subband decomposition approach with techniques based on coding the motion-compensated prediction errors, as is commonly done in digital systems. This research is aimed at gaining a better understanding of the advantages and limitations of the different types of signal representations pertaining to motion picture coding.

3.11 Subband Coding for Channel-Compatible Transmission Of High-Definition Television

Sponsor

Advanced Television Research Program

Project Staff

Ashok C. Popat, Professor William F. Schreiber

In recent years, subband coding has received considerable attention from the image coding community as a simple and effective means of efficiently

representing image and image-sequence data.[3] A three-dimensional (horizontal, verticle, and temporal) subband coding technique has been proposed by the Advanced Television Research Program at MIT for application to a 6-MHz channel-compatible high-definition television (HDTV) distribution system.[4] Although preliminary "proof-of-principal" tests have demonstrated that the technique is effective, these tests have also shown that considerable improvement could be achieved by adjusting various parameters in the system (such as the degree of data compression; the number of subbands in each dimension; the type and length of the subband analysis/synthesis filters; and the means of selecting subband pixels to be retained and transmitted). A high degree of interdependency among many of the system parameters has been observed; this interdependency has complicated the process of identifying the particular combination of parameters that is best suited to the present application. In particular, the strong interdependency seems to eliminate the possibility of finding the best choice for each parameter separately. A major objective of our research is to search through the vast parameters separately by judicious choice of parameters for computer simulation and by objective and subjective evaluation of the simulation results.

One of the more critical set of system parameters is the set of coefficients used in the subband analysis/synthesis filter banks. A novel approach to designing such filters based on time-domain numerical search has been developed; the approach is fairly general and has resulted in critically-sampled filter banks that are extremely well-suited to image coding applications.[5]

A seemingly basic principle of image subband/transform coding has emerged from the present study. In particular, it is evident that the best choice for the length of the analysis/synthesis filters depends only slightly on the number of sub-bands and depends more strongly on the spatial extent over which the image can be well-modeled as stationary. Thus, the nonstationarity of images inevitably leads to an uncertainty-principle based

tradeoff in the selection of the number of sub-bands and lengths of filters.

We found that it is extremely important that the allocation of channel capacity was allowed to be spatially varying; that is, it is essential to be able to increase the number and/or fidelity of samples used in representing action regions of the image at the expense of more poorly representing inactive regions. A fixed-rate, practicable means of exploiting this principle was devised.

This research was completed in March 1990.

3.12 Hybrid Analog/Digital Representation of Analog Signals

Sponsor

Advanced Television Research Program

Project Staff

Lon E. Sunshine, Professor Jae S. Lim

Transform coding has been shown to be an effective way to represent images, allowing for a significant amount of data-compression while still enabling high-quality reproduction of the original picture. One result of the transform coding of images is that at a given signal-to-noise ratio, the (spatially) low-frequency components are much more sensitive to additive noise than the high-frequency components.

In the MIT-CC television system, we must employ a noise reduction technique which is sufficient to eliminate the effects of additive noise in the low frequencies. We have chosen to do this is by representing these analog (continuous-amplitude) coefficients by a hybrid analog/digital signal. This representation consists of a new analog value plus a discrete-valued piece of side information. The advantage of using this hybrid format is that we can reduce the noise added to a particular coefficient by at least 6 dB for each bit used in the side information.

3 J.W. Woods and S.D. Oneil, "Subband Coding of Images," *IEEE Trans. ASSP* 34: 1278-1288 (1986); H. Gharavy and A. Tabatabai, "Subband Coding of Monochrome and Color Images," *IEEE Trans. Circuits Syst.* 35: 207-214 (1988).

4 W.F. Schreiber, et al., *Channel-Compatible 6-MHzHDTV Distribution Systems*, CIPG Technical Report ATRP-T-79, MIT, January 1988.

5 A.C. Popat, "Time-Domain Numerical Design of Critically Sampled Filter Banks," presentation viewgraphs, Advanced Television Research Program, MIT, October 1988; A.C. Popat, "A Note of QMF Design," unpublished memo, Advanced Television Research Program, MIT, December 1988.

This research considers the task of determining the "best" hybrid representation for an image. Here, "best" is characterized by a tradeoff between sufficient noise reduction, simplicity in implementation, and minimization of necessary side information.

3.13 Channel Equalization and Interference Reduction Using Adaptive Amplitude Modulation and Scrambling

Sponsor

Advanced Television Research Program

Project Staff

Adam S. Tom, Professor William F. Schreiber

Terrestrial broadcast channels and cable channels are imperfect. Random noise, multipath (ghosts), adjacent and co-channel interference, and an imperfect frequency response degrade these transmission channels so that the quality of the signal at the receiver is significantly below that at the transmitter. In order to appreciate the increased resolution of high definition images, a means of reducing the degradation due to channel defects needs to be employed. Conventional methods of channel equalization use adaptive filters. These methods are limited by convergence time, length of filters, and computational complexity. We are doing research on a new method of channel equalization and interference reduction based upon the ideas of adaptive amplitude modulation and pseudo-random scanning (scrambling). This new method is not bound by the above limitations; however, it is limited by the energy of the channel degradations produced.

Adaptive modulation is a noise reduction technique. It is applied only to the high frequency components of the signal. Prior to transmission, a set of adaptation factors are multiplied with the input signal. The net effect of this is to raise the amplitude of the signal according to the strength of the signal. At the receiver, the signal is divided by these same adaptation factors. In this manner, the random noise added in the channel is reduced by a factor equal to the adaptation factor. The noise is reduced more in the blank areas relative to the busy areas.

Scrambling is a technique to reduce the effects of multipath, adjacent and co-channel interference, and an imperfect frequency response. Prior to transmission, the input sequence is pseudo-randomly scanned; thus, in essence scrambling the signal so that it appears as random noise. This scrambled signal is transmitted through the channel, and the reverse of the scrambling is performed at the receiver. Consequently, any degradations in the channel are themselves scrambled at the receiver and, thus, have the appearance of pseudo-random noise in the decoded signal while the desired signal remains sharp and in full bandwidth.

Since the resultant signal in the receiver now has a noisy appearance, we apply the noise reduction technique of adaptive modulation to the input signal. In our scheme we first apply adaptive modulation to the input signal and then scrambling. The coded signal is transmitted through the imperfect channel and decoded at the receiver. The decoding consists of doing the reverse of the scrambling and then the reverse of the adaptive modulation. Scrambling causes any channel degradations to have a noiselike appearance, and Adaptive Modulation reduces the appearance of this pseudo-random noise. In this manner the degradations to a transmitted signal are reduced, and the channel is equalized.

This research was completed in June 1990.

Professor Donald E. Troxel has made significant contributions to image processing research in RLE. His primary focus is now in the area of computer-aided fabrication of integrated circuits.