Advanced Telecommunications and Signal Processing

Academic and Research Staff
Professor Jae S. Lim

Research Affiliate
Carlos Kennedy

Visiting Scientist
Heung-Nam Kim

Graduate Students
Zhenya Gu
Fatih Kamisli
Andy Lin

Support Staff
Cindy LeBlanc

Introduction

The present television system was designed nearly 60 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 digital TV system based on this standard has been deployed successfully. In 2006, digital television receiver sales exceeded analog television receivers in both number and dollar volume in the U.S. The analog terrestrial TV transmission was discontinued in the United States in June, 2009.

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. Future research will focus on making these improvements. The digital television system is a major improvement over the analog television system. The next major improvement over the digital television system is likely to be in the introduction of 3-D television. We are currently exploring methods for efficient transmission of 3-D TV signal.

The specific research topics where we made some progress are as follows:

1. Comparisons of Mobile TV Standards and Development of 3-D Mobile TV based on DMB

Sponsor
Advanced Telecommunications Research Program

Project Staff
Heung-Nam Kim

Mobile TV is the transmission of TV programs or video for the devices such as cellular phones, PDAs, and wireless multimedia devices. The mobile broadcast transmission can be serviced through terrestrial medium through terrestrial digital TV standards such as DVB-T in Europe, ATSC in U.S.A., and ISDB-T in Japan. However, these standards cannot be applied directly to mobile TV due to the tiny screen, the limited power consumption, and the characteristics of mobility. To resolve these problems, many standards have been developed in many countries such as DVB-H in Europe based on DVB-T, DMB in Korea and China based on DAB, OneSeg in Japan based on ISDB-T, and recently ATSC-M/H in U.S.A. based on ATSC. In addition, Qualcomm developed its proprietary solution, MediaFLO. In this research, we compare different mobile standards from the viewpoint of Audio/Video encoding, data encapsulation, transport format, channel coding, modulation technology, and advantages/disadvantages of each technology.

Since users have demanded for higher quality multimedia contents, reality is one of the major references in judging high quality. Providing mobility and increased reality in multimedia information services is thus a promising direction for the future. Specifically, 3-D mobile TV services are very attractive because (1) glassless 3-D viewing with small display is relatively easy to implement and more suitable to single user environments like DMB, (2) DMB is a new media and thus has more flexibility in adding new services on existing ones, and (3) 3-D multimedia handling capability of 3-D DMB terminals has lots of potential to generate new types of services if it is added with other components like built-in stereo camera. In this research, we propose attractive future mobile TV application, i.e., 3-D mobile TV which provides backward compatibility with DMB and better bit rates. ETRI (Electronics and Telecommunications Research Institute), developed DMB in 2004 and the world’s first broadcast mobile TV service was launched in 2005 by SK Telecom in Korea. The backward compatibility means that the existing DMB receivers can identify a 3-D DMB signal and recover the 2-D DMB audio-video signal without any problem.

In this research, we propose a cost-effective method of 3-D mobile TV services based on DMB system. One of the most important requirements of the 3-D mobile TV is the backward compatibility with the existing DMB system. To satisfy such a requirement, we define a new object descriptor (OD) of MPEG-4 Systems. Specifically we regard a pair of stereoscopic video signals and multi-channel 3-D audio signals as two objects. Then we consider that the two elementary streams of video for left and additional constitute the video object and similarly, the two elementary streams of audio for stereo and additional constitute the audio object. To verify the proposed scheme of 3-D mobile TV services over DMB, we implemented a prototype of 3-D mobile TV system. Through the experimental results, we show that the proposed system satisfies the required backward compatibility, provides acceptable depth impression even under the limited bandwidth of 3-D video (allocated less than or equal to 0.8Mbps) and a small-sized auto-stereoscopic 3-D LCD display.
In a practical situation, the bitrates of right-view images can be reduced much lower than that of left-view images because human perception is dominated by the high-quality component of a stereo pair. This can be achieved by separate rate control of each view. Hence it is necessary in the future work to study rate-control and test subjective quality for stereoscopic video coding in order to get better coding efficiency.

[1] John Arrnold, Michael Frater, Mark Pickering, “Digital Television: Technology and Standards,” Wiley-Interscience, U.S.A., 2007.

[2] Amitabh Kumar, “Mobile TV,” Focal Press, U.S.A., 2007.

[3] ITU-R Rec. BS. 775-1 (1994): Multichannel Stereophonic Sound System with and without accompaning Picture.

2. Transforms for Prediction Residuals in Video Coding

Sponsor: Advanced Telecommunications Research Program

Project Staff: Fatih Kamisli

To store or transmit the immense amount of data involved in visual media, digital image and video compression technologies are utilized. For example, to transmit the HDTV signal in a bandwidth originally designed for analog television broadcast, compression by a factor of 70 is performed. Compression is achieved by exploiting temporal and spatial redundancies present in visual media. Temporal redundancy refers to the observation that there are often very small changes from frame to frame within a video sequence. Spatial redundancy, present in a single frame or an image, refers to the similarity or slow variations of picture elements in a small neighborhood.

Spatial redundancy in images is eliminated by applying transforms on a small neighborhood of pixels. These transforms have the energy-compaction property, which means that a small number of transform coefficients are sufficient to capture the signal in that small neighborhood with adequate fidelity. In video sequences, first the temporal redundancy is eliminated and then the remaining spatial redundancy is eliminated. Temporal redundancy is eliminated using motion-compensated prediction techniques. Typically, a frame is divided into small blocks, and each block is predicted from previously transmitted frames by searching them for a good match for that block. The difference between the prediction and the frame to be coded is often called the motion-compensation residual (MC-residual). The MC-residual usually has some spatial redundancy left in some regions. To exploit this redundancy, standard video encoders use the same methods that image encoders use to exploit the spatial redundancies in images. Specifically, the same transform, the 2-D Discrete Cosine Transforms (DCT), is used. However, even though the MC-residual is intimately connected to images it has been obtained from, its spatial characteristics may differ from that of an image. This research focuses on developing transforms specifically designed for the MC-residual, as well as other residuals encountered in video coding, such as the resolution enhancement residual in scalable video coding and the disparity-compensation residual in multiview video coding.

The properties of the MC-residual have been studied by various researchers [1-3]. In [1], the auto-covariance of the MC-residual is modeled as a sum of a first-order Markov-process and independent white noise. This model reflects the relatively weaker correlation of the MC-residual compared to images in a simple and tractable way. In [2], the authors propose another compound model which fits the tails of the auto-covariance of MC-residuals better than the model in [1]. A more complicated analysis resulting in a more complicated model is provided in [3]. All these studies indicate that the statistical characteristics of the MC-residual may have some differences from the statistical characteristics of images. However, transforms accounting for these differences can often not be derived directly from such characterizations because most of these characterizations are rather complicated.

Recently, there has been a great deal of research on developing new direction-adaptive transforms for images [4-8]. These transforms take advantage of locally anisotropic features in images. Conventionally, the 2-D DCT or the 2-D Discrete Wavelet Transform (DWT) is carried out as a separable transform by cascading two 1-D transforms in the vertical and horizontal directions. This scheme does not take advantage of the locally anisotropic features present in images because it favors horizontal or vertical features over others. For example, the 2-D DWT has vanishing moments only in the horizontal and vertical directions. The main idea of these novel approaches is to adapt to local anisotropic features by performing the filtering along the direction where the image intensity variations are smaller. This is achieved by resampling the image intensities along such directions [6], by performing filtering and subsampling on oriented sublattices of the sampling grid [7], by directional lifting implementations of the wavelet transform [4], or by various other means. Even though most of the work is based on the wavelet transform, applications of similar ideas to DCT-based image compression have also been made [8]. However, it seems that the impacts of these ideas to modeling and compressing the MC-residual, as well as other prediction residuals, have not been investigated.

In this research, our goal is to develop transforms specifically designed for the MC-residual as well as other residuals encountered in video coding, such as the resolution enhancement residual in scalable video coding or the disparity compensation residual in multiview video coding. Using insights obtained from the research on direction-adaptive image transforms, we investigate how locally anisotropic features of images affect the MC-residual. We obtain an adaptive auto-covariance characterization of the MC-residual, which reveals some statistical differences between the MC-residual and the image. Based on this characterization, we propose a set of block transforms that can be used to compress the MC-residual. We also develop wavelet transforms/subband representations for prediction residuals based on the same insights. Experimental results indicate that these transforms, when combined with conventional transforms (2-D DCT or DWT), can provide important bitrate savings when compressing prediction residuals in video coding applications [9].

One aspect of future research is to design a new entropy coding method to encode the coefficients of our new block transforms. Another aspect is to investigate the applicability of our transforms to other prediction residuals, such as the resolution enhancement residual in scalable video coding and the disparity compensation residual in multiview coding. We also work on theoretical justifications of our transforms.

References

[1] C.-F. Chen and K.K. Pang. The optimal transform of motion-compensated frame difference images in a hybrid coder. *Circuits and Systems II: Analog and Digital Signal Processing, IEEE Transactions on [see also Circuits and Systems II:Express Briefs, IEEE Transactions on],* 40(6):393–397, Jun 1993.

[2] W. Niehsen and M. Brunig. Covariance analysis of motion-compensated frame differences. *Circuits and Systems for Video Technology, IEEE Transactions on,* 9(4):536–539, Jun 1999.

[3] K.-C. Hui and W.-C. Siu. Extended analysis of motion-compensated frame difference for block-based motion prediction error. *Image Processing, IEEE Transactions on,* 16(5):1232–1245, May 2007.

[4] C.-L. Chang and B. Girod. Direction-adaptive discrete wavelet transform for image compression. *Image Processing, IEEE Transactions on,* 16(5):1289–1302, May 2007.

[5] E. Le Pennec and S. Mallat. Sparse geometric image representations with bandelets. *Image Processing, IEEE Transactions on,* 14(4):423–438, April 2005.

[6] D. Taubman and A. Zakhor. Orientation adaptive subband coding of images. *Image Processing, IEEE Transactions on,* 3(4):421–437, Jul 1994.

[7] V. Velisavljevic, B. Beferull-Lozano, M. Vetterli, and P.L. Dragotti. Directionlets: anisotropic multidirectional representation with separable filtering. *Image Processing, IEEE Transactions on,* 15(7):1916–1933, July 2006.

[8] Bing Zeng and Jingjing Fu. Directional discrete cosine transforms for image coding. *Multimedia and Expo, 2006 IEEE International Conference on,* pages 721–724, 9-12 July 2006

[9] Kamisli, Fatih; Lim, Jae S., Transforms for the motion compensation residual, *Acoustics, Speech and Signal Processing, 2009. ICASSP 2009. IEEE International Conference on* , vol., no., pp.789-792, 19-24 April 2009

3. Using Enhancement Data to Deinterlace 1080i HDTV

Sponsor: Advanced Telecommunications Research Program

Project Staff: Andy Lin

Interlaced scan has always been an important aspect of broadcasted television. Established in the 1930's, interlaced scan was originally a solution to reduce flicker in televisions, without doubling the bandwidth or reducing the quality of video. Today, interlaced scan is outdated, as televisions no longer require interlacing to reduce flicker. Interlaced video leads to unsightly artifacts if not deinterlaced properly. Therefore, modern displays must deinterlace interlaced high definition television. However, high quality deinterlacing can be quite computationally intensive.

U.S. over-the-air (OTA) broadcasting changed to an all-digital format in June, 2009. The television industry is ready for the implementation of a system, which will improve high definition television (HDTV), in order for HDTV quality to migrate towards a new standard. This receiver compatible system will allow for an additional enhancement stream, which will enable new HDTV's to easily deinterlace 1080i streams. Older HDTV's will still be able to view HDTV by ignoring this stream and decoding the original 1080i stream.

The receiver compatible system is advantageous because it shifts computing costs from the receiver to the transmitter. There are far more receivers than transmitters, thus the net cost is reduced dramatically. Moreover, the receiver compatible system has access to the original video, allowing for the final result to have better quality than traditional deinterlacing systems.

For this project, we developed an algorithm to compute an enhancement stream that has a small enough bit-rate, yet still improves video quality significantly, compared to traditional systems. Using a small bit-rate stream allows us to propose using the receiver compatible deinterlacing system alongside mobile digital television. An implementation of an efficient small bit-rate receiver compatible enhancement stream is possible for 1080i HDTV streams. If decoded properly, this efficient small bit-rate stream yields results that are mathematically and visually superior to traditional deinterlacing algorithms.

References

[1] J. S. Lim, "A Migration Path to a Better Digital Television System", *SMPTE Journal*, vol. 103, No. 1, pp. 2-6, Jan. 1, 1994.

[2] J. Lim and L. Sunshine, "HDTV Transmission Formats and Migration Path," *International Journal of Imaging Systems and Technology,* vol. 5, pp. 286-291, Winter 1994.

[3] L. Sunshine, "HDTV Transmission Format Conversion and Migration Path," Ph.D. thesis, Massachusetts Institute of Technology, Cambridge, MA US, 1997.

[4] W. K. Wan, "Adaptive Format Conversion Information as Enhancement Data for Scalable Video Coding," Ph.D. thesis, Massachusetts Institute of Technology, Cambridge, MA US, 2002.

[5] W. Wan and J. Lim, "Adaptive Format Conversion For Scalable Video Coding," in *Proceedings of SPIE: Applications of Digital Image Processing XXIV,* vol. 4472, (San Diego, CA, USA), pp. 390-401, SPIE, July 2001.

4. Improving the H.264 Video Compression Standard with JPEG2000 Technology

Sponsor
Advanced Telecommunications Research Program

Project Staff
Zhenya Gu

In today's fast-paced world, digital media has become increasingly integrated into our everyday lives. In recent years, technological advances have allowed quality video streaming, vastly improved video conferencing and a proliferation of digital images on the Internet. Digital images and video however require large amounts of storage space and transmission bandwidth. Data compression is necessary to reduce data to a size within current storage and transmission capabilities. Many applications, such as digital television, require compression ratios of 50-70.

MPEG-4/H.264 AVC [1] is the most recently developed video compression standard and is widely used today. Video standards compress data by exploiting temporal and spatial redundancies. One commonly used technique is motion-compensated prediction. The next frame is predicted using motion vectors that indicate how objects in the scene change from one frame to the next. This prediction is subtracted from the original frame, and only the residual is transmitted. The amount of data transmitted is greatly reduced. Frames that are encoded with motion-compensated prediction are called inter-frames, or P and B-frames. Some frames are coded independent of other frames; these are referred to as intra-frames or I-frames. I-frames prevent an increasing drift between the encoder and decoder.

The most recent image compression standard is JPEG2000. JPEG2000 was finalized in late 2000, and it implements several new techniques to improve compression and image quality. One change is in the Fourier-related transform. JPEG 2000 uses the Discrete Wavelet Transform (DWT) instead of the DCT [2]-[3]. The DWT preserves color better and reduces blocking artifacts at low bit-rates. However, the DWT has more blurring artifacts. JPEG 2000 also uses a binary arithmetic encoding for entropy encoding.

In our research, we created a hybrid video compression system that uses JPEG2000 to encode the I-frames and H.264 to encode the P-frames. Because JPEG2000 and H.264 implement very different encoding techniques, using JPEG2000 to encode the still image frames might give improved performance. The hybrid system is evaluated against the H.264 system using a rate-distortion performance analysis and subjective analysis. While the rate-distortion performance between the two systems is similar, the subjective performance is very different because they exhibit very different artifacts at low bit-rates.

References

[1] ITU-T Recommendation H.264 and ISO/IEC 14496-10 MPEG-4 Part 10, Advanced Video Coding (AVC), 2003

[2] M. Antonini, M. Barlaud, P.Mathieu, and I. Daubechies, "Image coding using the wavelet transform," *IEEE Trans. Image Processing* 1: 205-220 (1992).

[3] D. J. Le Gall and A. Tabatabai, "Subband coding of digital images using symmetric short kernel filters and arithmetic coding techniques," in *Proc. IEEE Int. Conf. ASSP*, NY, 761-765 (1988).