Advanced Telecommunications and Signal Processing Group

Academic and Research Staff
Professor Jae S. Lim

Research Affiliate
Carlos Kennedy

Graduate Students
Fatih Kamisli

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 TV service is scheduled to discontinue in the spring of 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 planning future research in this area.

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

1. Reduction of Blocking Artifacts Using Side Information

Sponsor: Advanced Telecommunications Research Program, Higher Education Council of Turkey

Project Staff: Fatih Kamisli

A method that is often used in image and video compression is transform coding based on the discrete cosine transform (DCT). In this method, an image (or the motion-compensated residual in the case of a video) is divided into non-overlapping blocks of pixels and the DCT coefficients of each block are quantized. When the compression rate is high, each block is not well represented and discontinuities along the block boundaries become visible. This project addresses the problem of reducing this artifact, known as the blocking artifact.

Many approaches have been proposed to reduce blocking artifacts. Among the most common are post-processing methods. These methods apply various post-processing techniques to the compressed image to reduce blocking artifacts. Post-processing methods include adaptive filtering [1-3], statistical estimation [4], and projection onto convex sets [5] (POCS). These methods do not require any modification to current encoders.

In methods based on adaptive filtering, block boundaries where blocking artifacts occur are low-pass filtered. Filters are adapted to the local characteristics of the image so that sufficient smoothing of artifacts is achieved in smooth regions, while blurring of details is avoided in detailed or edge regions. Methods based on statistical estimation use probabilistic models for the compressed and deblocked images. Estimators are then derived based on these models. In methods based on POCS, the desired properties of the deblocked image are captured by defining appropriate convex sets. Projecting the image with blocking artifacts onto these convex sets gives the deblocked image. Methods based on statistical estimation and POCS result, most often, in iterative deblocking algorithms, which may require substantial computations. In video coding, adaptive filtering techniques are often preferred over the other techniques, in part because the computational requirements of adaptive filtering methods are lower.

An approach different from post-processing is to replace the DCT with lapped orthogonal transforms [6] (LOT). LOTs have basis functions that overlap with adjacent blocks. This property reduces blocking artifacts but creates other artifacts, such as increased ring effects around edges. Even though image coders based on wavelets, which can be seen as generalizations of lapped transforms, have been standardized, many image and video coding standards are based on blocked-based DCT. In addition, transforms that have basic functions that overlap with adjacent blocks create difficulties in video compression where inter-frame and intra-frame processing are used together in the same frame. Therefore, it is important to explore ways of reducing blocking artifacts.

In our research, we studied an approach where blocking artifacts are reduced using side information transmitted from the encoder to the decoder. In this approach, the processed image can be compared directly with the original undegraded image to improve the performance. For example, we could process an image with different methods and transmit the most effective method as part of the side information. In addition, computations at the receiver are reduced, since the receiver does not need to determine which method to use. A major goal of this research was to demonstrate that this approach has significant benefits and should be explored further in the future research on deblocking.

The ability to compare the result of processing with the original image can be used to improve the performance. For example, one can process each block-boundary with different methods and transmit the most effective method in regions with high detail or a strong edge. Specifically, in regions with strong edges, ringing artifacts arise due to loss of significant high frequency information. Ringing artifacts appear as artificial fluctuations on both sides of the edge. These fluctuations can be interpreted as high frequency content or detail by existing methods (for example, adaptive filtering methods) and thus might not be reduced to maintain details. Using the original image, however, it is possible to determine that these fluctuations are artificial and do not exist in the original image, and therefore it is desirable to remove them.

Another benefit is the reduction of computations at the decoder. Post-processing methods need to process local regions of the compressed image in order to determine how to deblock these local regions. For example, post-processing methods based on adaptive filtering extract information about local regions in order to determine the filtering method to be used. Smooth regions are generally filtered with substantial smoothing whereas detailed regions are only slightly modified. In the approach based on side information, no computation to extract information about local regions is necessary.

Sending side information results in an overhead in the bit-rate. In our research, we demonstrated that the increase in the bit-rate can be compensated for by an increase in the performance for some range of bit-rates in image and video coding. We developed a system example based on the proposed approach.

We compared this system with a post-processing type of system. Based on this comparison, we demonstrated that the reduction of blocking artifacts using side information is worthwhile to explore further in the future research on the reduction of blocking artifacts. The results of this research were reported in *Applications of Digital Image Processing XXIX*, Edited by Tescher, Andrew G., *Proceedings of the SPIE*, Volume 6312 (2006).

References

[1] B. Ramamurthi and A. Gersho, "Nonlinear space-variant postprocessing of block-coded images," *IEEE Transactions on Acoustics, Speech, and Signal Processing* **34**, pp. 1258-1268, 1986.

[2] S. D. Kim, J. Yi, H. M. Kim, and J. B. Ra, "A deblocking filter with two separate modes in block-based vide coding," *IEEE Transactions on Circuits and Systems for Video Technology* **9**, pp. 156-160, 1999.

[3] S.-C. Tai, Y.-Y. Chen, and S.-F. Sheu, "Deblocking filter for low bit rate mpeg-4 video," *IEEE Transactions on Circuits and Systems for Video Technology* **15**, 2005.

[4] R. L. Stevenson, "Reduction of coding artifacts in transform image coding," *Proc. of the IEEE Intl. Conference on Acoustics, Speech, and Signal Processing* **5**, pp. 401-404, 1993.

[5] Y. Yang, N. Galatsanos, and A. Katsaggelos, "Projection-based spatially adaptive reconstruction of block-transform compressed images," *IEEE Transactions on Image Processing* **4**, pp. 896-908, 1995.

[6] H. S. Malvar and D. H. Staelin, "The lot: transform coding without blocking effects," *IEEE Transactions on Acoustics, Speech, and Signal Processing* **37**, pp. 553-559, 1989.

2. Multiple Transform/Subband Representations for Video Processing

Sponsor: Advanced Telecommunications Research Program

Project Staff: Jae S. Lim

A common encoding technique for video compression uses motion compensation, in which an estimate is made of the magnitude and direction of motion of an image from one frame to the next, to provide an estimate of the next frame, and only the difference between the estimate and the actual frame (the "motion-compensated residual") is encoded. The amplitude of the residual will generally be much smaller than the intensity of the image, itself, and so fewer bits are needed for accurate encoding. Motion compensation can be used for all or just a portion of a frame.

Another common encoding technique, often used in conjunction with motion compensation, is transforming the intensity of the image (or the amplitude of the residual) to, e.g., a spatial frequency domain, and then digitizing the transform coefficients. For the same number of bits, transform encoding generally produces a higher quality image.

When image intensity is transformed and digitized, it is called "intra-frame" coding, as only information from the same frame is used. When residual amplitude is transformed and coded it is called "inter-frame" coding, as information from the current and at least one prior frame is used.

There are various advantages in using intra-frame coding. For example, because intra-frame coding does not involve other frames, it can be decoded without decoding other frames. This can be useful when a viewer changes television channels. In the United States digital television standard, an entire frame, referred to as an *I*-frame, is periodically encoded using intra-frame coding. When a channel change occurs, the television receiver can wait for the *I*-frame and begin decoding from that frame. It is also useful for VCR or DVD type applications, wherein only *I*-frames may be decoded to provide images during fast-forwarding. Also, intra-frame coding reduces the effect of error propagation, because errors that occurred in other frames do not affect intra-frame coded regions.

Inter-frame encoding is very useful in reducing the bit rate for some image regions. As noted above, motion compensation takes advantage of the fact that scenes often do not change substantially from one frame to the next, and thus once the previous frame is decoded, it can be used to predict portions of the current frame. By encoding only image aspects that cannot be predicted, the bit rate used can be significantly reduced. Except for the *I*-Frame, in which the entire frame uses only intra-frame coding, some portions of a typical video frame are encoded using intra-frame coding (when motion prediction is not good) and other regions are encoded using inter-frame coding.

The transform that is used in video compression is typically the discrete cosine transform (DCT), which is a block-based transform. The image is divided into many non-overlapping blocks (typically the sizes are 8×8), and the DCT coefficients of the image intensity, in the case of intra-frame coding, and the motion-compensated residual, in the case of inter-frame coding, are quantized. An advantage of using a non-overlapping region transform (non-ORT) such as the DCT is that each block can be treated separately from other blocks, thus making it simple to mix intra-frame coded blocks and inter-frame coded blocks to form a complete frame.

A significant disadvantage of non-ORTs is the occurrence of blocking artifacts. Because the blocks are treated separately and the image is not perfectly reconstructed after compression, discontinuities can occur along the block boundaries. This becomes particularly evident when the bit rate is low (compression is high). Once the blocking effects occur, they can propagate to other frames as a result of inter-frame coding.

An approach to reducing the effects of blocking artifacts is to use an overlapping region transform (ORT) in which there is overlap in the regions transformed. This can increase the number of transform coefficients to encode, however, as the overlapped regions are represented more than once. But some ORTs e.g., the lapped orthogonal transform (LOT), utilizes overlapping regions without increasing the number of coefficients relative to DCT.

Transform representations such as the DCT and LOT are related to subband representation. For a subband representation, a signal such as an image or a residual is filtered by a set of filters, and the results are subsampled. The filtered and subsampled signals are the subband representation of the signal. The filters used are called analysis filters because they are used in analyzing the signal. A different set of filters, of course, results in different subband coefficients for the same signal. In an image or video compression system based on a subband representation, the subband coefficients are quantized and the quantized coefficients are transmitted in applications such as digital television or stored in a storage medium in applications such as DVD.

The quantized subband coefficients can be used to reconstruct an estimate of the original signal by interpolation and filtering with a set of filters. The filters used in this process are called the synthesis filters. If we choose an appropriate set of analysis and synthesis filters, and perform the appropriate subsampling and interpolation functions, it is possible to reconstruct the original signal exactly from the unquantized subband coefficients. Because of the quantization process which is necessary in a typical image or video compression application, the reconstructed signal is only an estimate (approximation) of the original signal.

The transform coefficients and subband coefficients of the same signal may be very simply related to each other. For example, the DCT coefficients of a signal can be simply related one-to-one to the subband coefficients by choosing an appropriate set of analysis filters. For this reason, the transform coefficients and subband coefficients will be collectively referred to as transform/subband coefficients. The DCT and LOT are examples of the transform/subband representations.

In our research, we have developed an approach where different parts of video are represented by different transform/subband representations. This approach has a number of advantages over conventional methods where the same transform/subband representation is used for the entire video. As a specific example to illustrate the advantages of this approach, we encode the *I*-frame using a subband representation such as LOT that overlaps blocks. This reduces the blocking artifacts. The *B* and *P*-frames which utilize inter-frame coding, however, can be encoded with the DCT. The use of block-based transforms such as the DCT simplifies the mix of intra-frame coding for some blocks and inter-frame coding for some other blocks within the same frame.

Publications

Conference Proceedings, Published

Kamisli, Fatih and Lim, Jae S., “Reduction of Blocking Artifacts Using Side Information,” Proceedings of the SPIE, Volume 6312, pp. 63120A-1-11, Applications of Digital Image Processing XXIX, San Diego, CA, August, 2006.

Patents

Lim, J.S., “Method and Apparatus for Encoding Decoding and Compression of Audio-Type Data Using Reference Coefficients Located within a Band of Coefficients,” Canadian Patent No. 2,128,216 issued June 6, 2006.