Rate Scalable Image and Video Compression Techniques

Edward J. Delp, Paul Salama, Eduardo Asbun, Martha Saenz, and Ke Shen
Video and Image Processing Laboratory (*VIPER*)
School of Electrical and Computer Engineering
Purdue University
West Lafayette, Indiana 47907-1285

***Abstract*— In this talk we will describe embedded image and video compression techniques. We describe an embedded zero tree-like approach that exploits the interdependency between color components that is known as Color Embedded Zero Tree Wavelet (*CEZW*). We will also present a video compression technique, Scalable Adaptive Motion Compensated Wavelet (*SAMCoW*) compression, that uses the *CEZW* data structure described above. We show that in addition to providing a wide range of rate scalability, *SAMCoW* achieves comparable performance to the more traditional hybrid video coders.**

I. Introduction

The transmission of video or images over heterogeneous data networks for multimedia applications, has recently become an area of active research. Meeting bandwidth requirements and maintaining acceptable image quality simultaneously is a challenge. Continuous rate scalable applications can prove valuable in scenarios where the channel is unable to provide a constant bandwidth to the application. Such decoders are particularly attractive because of their flexibility in allowing only one image or sequence to be stored in the database, avoiding the overhead of maintaining several coded images or sequences at different data rates. Hence, rate scalability allows one to "encode once and decode on any platform fed by any data pipe."

A specific coding strategy known as *embedded rate scalable coding* is well suited for continuous rate scalable applications [1]. In embedded coding, all the compressed data is embedded in a single bit stream and can be decoded at different data rates. The decompression algorithm receives the compressed data from the beginning of the bit stream up to a point where a chosen data rate requirement is achieved. A decompressed image at that data rate can then be reconstructed and the visual quality corresponding to this data rate can be achieved.

Several approaches have been proposed to achieve rate scalability in still image compression. Shapiro's zerotree [1] exploits the interdependence between the subbands of a wavelet decomposed image. Said and Pearlman [2] investigated different tree structures and coding strategies that improved the quality of the decomposition in some applications. These approaches have been extended to video compression. For example, the separable 3-D multiresolution coding algorithm of Taubman and Zakhor [3], [4] employs several levels of the wavelet transform for the temporal dimension.

In this paper, we describe wavelet based rate scalable image and video compression algorithms that allow the data rate to be dynamically changed during decoding. We shall refer to the image compression algorithm as **Color Embedded Zerotree Wavelet** (*CEZW*), and to the video compression technique as the **Scalable Adaptive Motion COmpensated Wavelet** (***SAMCoW***) algorithm [5]–[11].

II. Embedded Coding of Color Images: *CEZW*

A. Basic Algorithm

Rate scalable algorithms, such as EZW [1] and SPIHT [2], were originally developed for grayscale images. To code a color image, the color components are usually treated as three individual grayscale images and the same coding scheme is used for each component. The interdependence between the color components is not explicitly exploited.

To exploit the interdependence between color components, one may decorrelate the color components. In Said and Pearlman's algorithm [2], the Karhunen-Loeve (KL) transform is used [12]. This technique shall be referred to here as CSPIHT. An alternative technique we described in [7], [8] uses a spatial orientation tree (SOT) that spans the color space.

The color space we initially used is the YUV space which consists of a luminance component and two chrominance (color difference) components. This space is popular because the luminance signal can be used to generate

This work was supported by grants from the AT&T Foundation, the Rockwell Foundation, Lucent Technologies, and Texas Instruments. Address all correspondence relative to this paper to E. J. Delp, ace@ecn.purdue.edu.

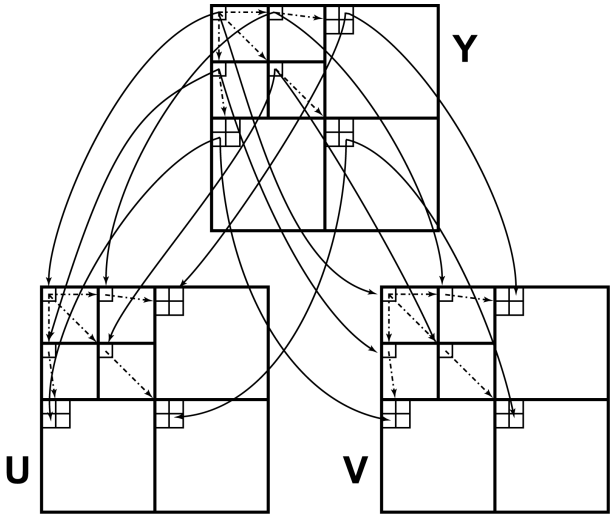


Fig. 1. Diagram of the parent-descendent relationships in *CEZW*. This tree is developed on the basis of the tree structure in Shapiro's algorithm. The YUV color space is used.

a grayscale image, which is compatible with monochrome systems.

Experimental evidence has shown that at the spatial locations where chrominance signals have large transitions, the luminance signal also has large transitions [8]. Transitions in an image usually correspond to wavelet coefficients with large magnitudes in high frequency bands. Thus, if a transform coefficient in a high frequency band of the luminance signal has small magnitude, the transform coefficient of the chrominance components at the corresponding spatial location and frequency band should also have small magnitude [13], [7], [8], [10]. In embedded zerotree coding, if a zerotree occurs in the luminance component, a zerotree at the same location in the chrominance components is highly likely to occur. This interdependence of the transform coefficients signals between the color components is incorporated into the SOT we developed to achieve a higher degree of compression [7], [8], [10]. A diagram of the SOT is shown in Figure 1. The luminance component is scanned first in a dominant pass. The two chrominance components are alternately scanned after the luminance component, in a subordinate pass, which is essentially the same as that in Shapiro's [1] algorithm [7], [8], [10], [11]. This still image compression technique is known as the Color Embedded Zerotree Wavelet (*CEZW*).

B. Extensions

In [7], [8], [10] the color space used was the 4:1:1 YUV color space the algorithm has since been extended for 4:4:4 YUV images in [11]. This algorithm shall be referred to as Color Zerotree Wavelet (*CZW*). Furthermore, the coding technique was modified such that the SOT used was that proposed by Said and Pearlman [2], whereas the coding strategy was that described in [7], [8]. This compression scheme is referred to as *SP-CZW* in [11].

Fig. 2. Performance of KLT & *CZW*, YUV & *CZW*, YUV & SPIHT, KLT & *SP-CZW*, and CSPIHT at 0.25 bpp.

In [11] we show that the use of SOTs that exploit color interdependencies is more efficient than coding each color component independently. In addition, using the YUV color space resulted in images that visually appeared to be better than those that were transformed using the KL transform [11].

A combination of color transformations, spatial orientation trees and embedded encoding algorithms were evaluated and are described in more detail in [11]. The performance of some of the schemes at 0.25 bits per pixel (bpp) is shown in Figure 2.

III. *SAMCoW*

A. Basic Algorithm

SAMCoW uses an embedded rate scalable coding strategy to obtain continuous rate scalability. Its main features are adaptive block-based motion compensation in the spatial domain to reduce temporal redundancy and improve image quality at low data rates, and *CEZW* on the intracoded and predictive difference frames.

B. Adaptive Motion Compensation

In a scalable codec, the decoded frames have different distortions at different data rates, making it impossible for the encoder to generate the exact reference frames as in the decoder for all the possible data rates. One solution is to have the encoder lock on to a fixed data rate (usually the highest data rate) and let the decoder run freely. The codec will work exactly as the non-scalable codec when decoding at the highest data rate. However, when decoding at a low data rate, the quality of the decoded reference frames will deviate from that at the encoder [5], [6], [8], [10]. Hence, both the motion prediction and the decoding of the predicted error frames contribute to the increase in distortion of the decoded video sequence. This distortion also propagates from one frame to the next within a group of pictures (GOP).

SAMCoW uses a technique known as *adaptive motion compensation* (AMC) [6], [5], [8], [10], in which a feedback loop is added in the decoder. The decoded reference frames at both the encoder and the decoder are then locked to the same data rate. This makes the reference frames at the encoder and the decoder identical, hence eliminating error propagation. *SAMCoW* also uses overlapped block motion compensation (OBMC) for the Y component to reduce the blockiness in the prediction error image [8], [10].

C. Advanced coding options

Results presented in [8], [10] indicate that perceived image quality in *SAMCoW* degrades at rates less than 32 kbps. Several reasons exist for this behavior. First, the wavelet transform is obtained for the entire image and the algorithm cannot allocate extra bits to areas with high activity. Second, *SAMCoW* uses only I and P frames. The MPEG standards have demonstrated that the use of B-frames can improve the performance of the coding scheme without significantly increasing the data rate. Furthermore, B frames can be discarded when needed because they are not used to predict future frames [9].

SAMCoW was modified to include B frames, rate control, half-pixel accuracy search techniques for motion vectors, and unrestricted motion vectors [9]. The advantage of B frames is that they can be quantized coarser than I or P frames or discarded completely. This allows for rate control whereby the intracoded (I) frames are given preference over P and B frames during bit allocation [9]. Using unrestricted motion vectors permits for motion vector searches that exceed image boundaries. In this case each dimension of each reference image is increased by a particular number of rows/columns. When a pixel referenced by a motion vector lies outside the original reference image boundaries, the value used is that of the closest edge pixel.

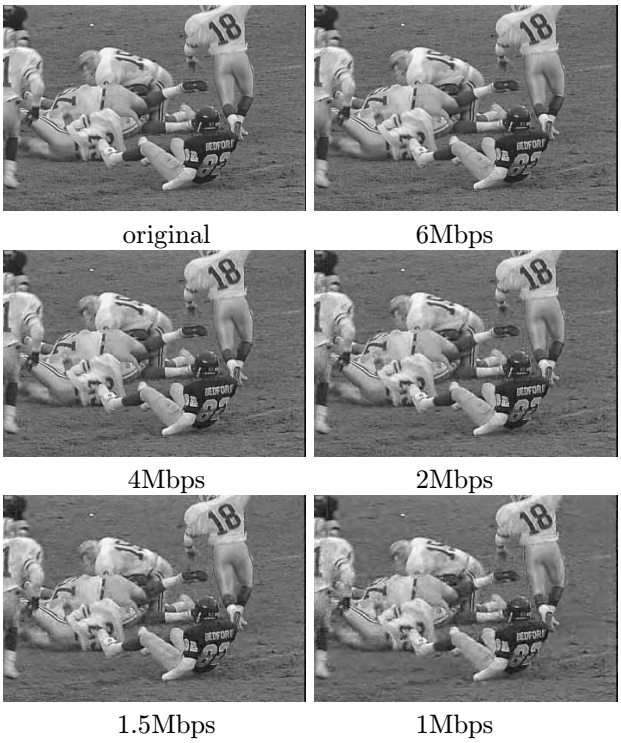
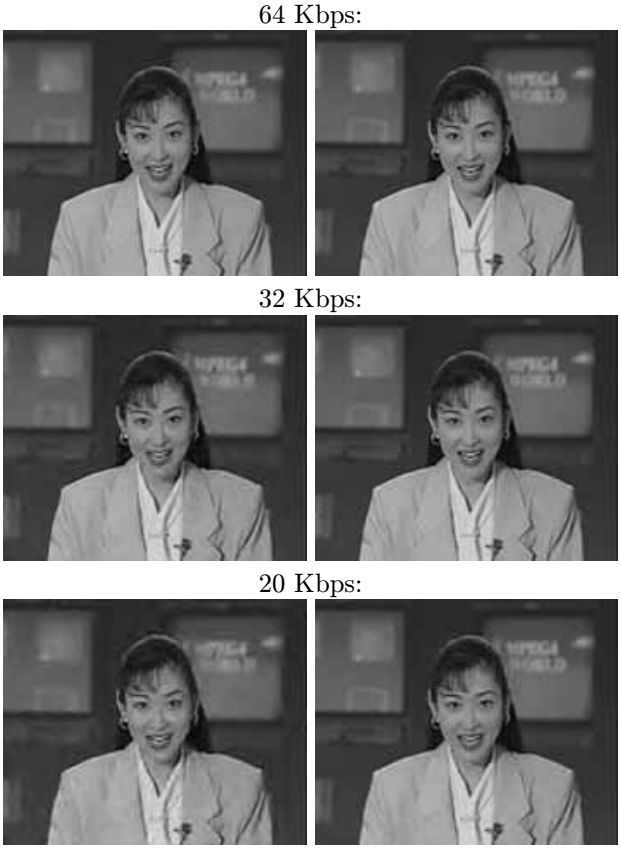


Fig. 3. Frames of the *football* sequence, decoded at different data rates using *SAMCoW* (CIF, 30 frames per second).

Dynamic bit allocation between frames, post processing, and the selective coding of spatial orientation trees have also been used to enhance the performance of *SAMCoW* [14].

Two types of video sequences were used for the experiments. One type is a CIF (352x240) sequence with 30 frames per second. The other is a QCIF (176x144) sequence with 10 frames per second or 15 frames per second. The CIF sequences were decompressed using *SAMCoW* at data rates of 1 megabits per second (Mbps), 1.5 Mbps, 2 Mbps, 4 Mbps and 6 Mbps. A representative frame decoded at the above rates is shown in Figure 3. Subjective experiments have shown that our algorithm produces decoded video with comparable visual quality as MPEG-1 at every tested data rate [10].

The QCIF sequences were compressed and decompressed using *SAMCoW* at data rates of 20 kilobits per



Akiyo (*SAMCoW*) *Akiyo* (H.263)

Fig. 4. Frames of the *Akiyo* sequence, decoded at different data rates (QCIF, 10 frames per second).

second (Kbps), 32 Kbps, 64 Kbps, 128 Kbps, and 256 Kbps. The same set of sequences were compressed using the H.263 algorithm at the above data rates. Our experiments indicate that *SAMCoW* performs comparably to H.263 at all tested data rates [10].

IV. Conclusion

In this paper we have provided an overview of new scalable image and video compression techniques. Papers describing *CEZW* and *SAMCoW* in more detail can be found at `http://www.ece.purdue.edu/~ace`

References

[1] J. M. Shapiro, "Embedded image coding using zerotrees of wavelet coefficients," *IEEE Transactions on Signal Processing*, vol. 41, pp. 3445–3462, December 1993.

[2] A. Said and W. A. Pearlman, "A new, fast, and efficient image codec based on set partitioning in hierarchical trees," *IEEE Transactions on Circuits and Systems for Video Technology*, vol. 6, no. 3, pp. 243–250, June 1996.

[3] D. Taubman and A. Zakhor, "Multirate 3-D subband coding of video," *IEEE Transactions on Image Processing*, vol. 3, no. 5, pp. 572–588, September 1994.

[4] D. Taubman and A. Zakhor, "A common framework for rate and distortion based scaling of highly scalable compressed video," *IEEE Transactions on Circuits and Systems for Video Technology*, vol. 6, no. 4, pp. 329–354, August 1996.

[5] M. L. Comer, K. Shen, and E. J. Delp, "Rate-scalable video coding using a zerotree wavelet approach," in *Proceedings of the Ninth Image and Multidimensional Digital Signal Processing Workshop*, Belize City, Belize, March 3–6 1996, pp. 162–163.

[6] K. Shen and E. J. Delp, "A control scheme for a data rate scalable video codec," in *Proceedings of the IEEE International Conference on Image Process ing*, Lausanne, Switzerland, September 16–19 1996, vol. II, pp. 69–72.

[7] K. Shen and E. J. Delp, "Color image compression using an embedded rate scalable approach," in *Proceedings of the IEEE International Conference on Image Process ing*, Santa Barbara, California, October 26–29 1997, vol. III, pp. 34–37.

[8] K. Shen, *A Study of Real Time and Rate Scalable Image and Video Compression*, Ph.D. thesis, School of Electrical and Computer Engineering, Purdue University, December 1997.

[9] E. Asbun, P. Salama, K. Shen, and E. J. Delp, "Very low bit rate wavelet-based scalable video compression," in *Proceedings of the IEEE International Conference on Image Processing*, Chicago, Illinois, October 1998, pp. 948–952.

[10] K. Shen and E. J. Delp, "Wavelet based rate scalable video compression," *IEEE Transactions on Circuits and Systems for Video Technology*, vol. 9, no. 1, pp. 109–122, February 1999.

[11] M. Saenz, P. Salama, K. Shen, and E. J. Delp, "An evaluation of color embedded wavelet image compression techniques," in *SPIE Conference on Visual Communications and Image Processing '99*, San Jose, California, January 1999, pp. 282–293.

[12] H. Hotelling, "Analysis of a complex of statistical variables into principal components," *Journal of Educational Psychology*, vol. 24, pp. 417–441 and 498–520, 1933.

[13] Q. Wang and M. Ghanbari, "Scalable coding of very high resolution video using the virtual zerotree," *IEEE Transactions on Circuits and Systems for Video Technology*, vol. 7, no. 5, pp. 719–727, October 1997.

[14] E. Asbun, P. Salama, and E. J. Delp, "Encoding of predictive error frames in rate scalable video codecs using wavelet shrinkage," in *Proceedings of the IEEE International Conference on Image Processing*, Kobe, Japan, October 1999.