

# ERROR RESILIENCE AND CONCEALMENT IN EMBEDDED ZEROTREE WAVELET CODECS

Paul Salama

Visual Communications Lab  
Dept. of Elect. and Comp. Engin.  
Indiana Univ-Purdue Univ Indianapolis  
Indianapolis, IN 46202  
USA

Ness Shroff, and Edward J. Delp

Video and Image Proc. Lab (VIPER)  
School of Elect. and Comp. Engin.  
Purdue University  
West Lafayette, Indiana 47907-1285  
USA

## ABSTRACT

In ATM networks cell loss or channel errors can cause data to be dropped in the channel. When digital images/video are transmitted over these networks one must be able to reconstruct the missing data so that the impact of the errors is minimized. In this paper we overview the problem of using EZW encoders in channels where data-loss is possible. We also describe an error resilience scheme based on unequal error protection and data interleaving that addresses the problem of using rate scalable encoders over ATM networks.

## 1. INTRODUCTION

Broadband computer networks must support a variety of applications involving high-resolution video and images. Asynchronous Transfer Mode (ATM) is one of the communication protocols used for these networks. To effectively transmit digital images or video over these networks, issues involved in designing coding schemes as well as packetizing the compressed data need to be studied. In particular the design of a robust coding schemes that will permit the reconstruction of missing data, due to ATM cell loss, by the decoder is necessary.

Of the many types of encoders that are attractive for transmission over broadband networks, are those that are *continuously scalable*. Continuous rate scalability provides the capability of arbitrarily selecting a specific decoder data rate from within a range of data rates.

A specific coding strategy known as *embedded rate scalable coding* [1, 2] is well suited for continuous rate scalable applications. In embedded coding, all the compressed data is embedded in a single bit stream and can be decoded at different data rates. To achieve best performance the bits that convey the most important information need to be embedded at the beginning of the compressed bit stream. Several embedded rate scalable coding schemes have been proposed [2, 3, 4, 5].

## 2. OVERVIEW OF EMBEDDED ZEROTREE WAVELET CODING

In the embedded zerotree wavelet (EZW) coding strategy, developed by Shapiro [1], a wavelet/subband decomposition of the image is performed. The wavelet coefficients/pixels are then grouped into Spatial Orientation Trees [2]. The magnitude of each wavelet

coefficient/pixel in a tree, starting with the root of the tree, is then compared to a particular threshold  $T$ . If the magnitudes of all the wavelet coefficients/pixels in the tree are smaller than  $T$ , the entire tree structure (that is the root and all its descendant nodes) is coded by one symbol, the zerotree symbol ZTR [1]. If however there exist significant wavelet coefficients/pixels, then the tree root is coded as being significant or insignificant (the symbol IZ, isolated zero, being used in this case), if its magnitude is larger than or smaller than  $T$ , respectively. The descendant nodes are then each examined in turn to determine whether each is the root of a possible subzerotree structure, or not. This process is carried out such that all the nodes in all the trees are examined for possible subzerotree structures. The significant wavelet coefficients/pixels in a tree are coded by one of two symbols, POS or NEG, depending on whether their actual values are positive or negative, respectively. The process of classifying the pixels as being ZTR, IZ, POS, or NEG is referred to as the dominant pass in [1]. This is then followed by the subordinate pass in which the significant wavelet coefficients/pixels in the image are refined by determining whether their magnitudes lie within the intervals  $[T, 3T/2)$  and  $[3T/2, 2T)$ . Those wavelet coefficients/pixels whose magnitudes lie in the interval  $[T, 3T/2)$  are represented by a 0 (LOW), whereas those with magnitudes lying in the interval  $[3T/2, 2T)$  are represented by a 1 (HIGH). Subsequent to the completion of both the dominant and subordinate passes, the threshold value  $T$  is reduced by a factor of 2, and the entire process repeated. This coding strategy, consisting of the dominant and subordinate passes followed by the reduction in the threshold value, is iterated until a target bit rate is achieved.

An EZW decoder reconstructs the image by progressively updating the values of each wavelet coefficient/pixel in a tree as it receives the data. The decoder's decisions are always synchronized to those of the encoder.

## 3. PROBLEMS WITH EZW

Although EZW is very efficient and effective in reducing the spatial redundancies in an image, it is very susceptible to transmission errors. The misinterpretation of a symbol by the decoder, as a result of transmission errors, is sufficient enough to result in loss of synchronization between the encoder and decoder. Consider the case of a tree structure being coded by the ZTR symbol, yet the data that the decoder receives indicates otherwise. In this case the decoder will be expecting extra data in order to reconstruct the wavelet coefficients/pixels belonging to the tree. Since there

is originally no data for those wavelet coefficients/pixels, the decoder will then commence to decode the data belonging to other trees and assign it to the current tree. In fact the misinterpretation of only one symbol is sufficient to derail the entire decoding process. The effect of incorrectly decoding a ZTR symbol as a POS symbol is observed in Figure 1b.

Our experiments have shown that of the four symbols used to encode an image, the most important and hence must be protected are the IZ (isolated zero) and ZTR (zero tree root) symbols. They comprise, on average, about 60 percent and 25 percent, respectively, of all the symbols used to code an image. Our experiments have also shown that the misinterpretation of the significant wavelet coefficients/pixels do not impact the subjective quality of the reconstructed image as severely as an error in decoding the IZ and ZTR symbols. Thus, the objective of any strategy used to pack the coded bitstream into ATM cells would be to protect the IZ and ZTR symbols. Thus, unequal error protection whereby sections of the bitstream are afforded higher protection than other sections is necessary. The packetization strategy should also permit the loss of the significant wavelet coefficients/pixels without any loss of synchronization between the encoder and decoder. One naive approach would be to pack all the ZTR symbols into high priority ATM cells, and interpret any missing symbol as being an IZ symbol. However, it has been observed that at least 40 percent of all the cells used to transmit the data have to be designated as being high priority cells.

#### 4. ERROR CONCEALMENT IN EZW

We present a brief overview of different approaches used by other authors [6, 7, 8, 9] to conceal the effect of transmission errors in EZW encoded streams.

In [6], all the zerotrees are grouped together into groups of  $N$  trees that are interleaved together to form  $M$  separate bitstreams. The  $M$  bitstreams are then transmitted separately. If an error is detected in a bitstream the decoder stops decoding and commences decoding another bitstream. A similar approach was adopted in [9], wherein the  $N$  interleaved trees are coded such that the resulting bitstream is exactly 48 bytes, an ATM cell user payload. If the size of the resulting bitstream is longer than 48 bytes, the trees are pruned. The disadvantage of both techniques is the loss of rate scalability.

In [8], the authors use Error-Resilient Entropy Coding (EREC) [10] to map blocks of variable length data into blocks of fixed length, called frames. A variation of the embedded zerotree wavelet codec developed by Said and Pearlman [2] was proposed in [7]. The new coder produces two bitstreams, denoted as the MAP and QUAN bitstreams. The QUAN bitstream consists of fixed length codes representing the quantizations of the significant wavelet coefficients/pixels. The coder constructs the MAP bitstream by producing a 1 whenever a coefficient/pixel or one of its descendants is significant relative to the threshold value at the current dominant pass, and a 0 whenever a coefficient/pixel and all its descendants are insignificant relative to the threshold value. Unequal error protection [11] is then employed to provide high protection to the MAP bitstream, and moderate and low protection to the QUAN bitstream.

It is to be emphasized that with the exception of [9], all authors test the robustness of their techniques by simulating binary symmetric channels, in which the data in the bitstream is corrupted. This is to be contrasted with the problem at hand in which entire

ATM blocks of data from the bitstream, 384 bits in particular, are lost.

Alternative protection of coded image data is described in [12] through the use Multiple Description Coding. The source encoder produces a number of correlated output sequences that are to be sent over separate channels with varying transmission reliabilities. Data corrupted in any particular group of channels is restored by means of the remaining uncorrupted data. The separate sequences are then used by the decoder to reconstruct the image.

#### 5. ALTERNATIVE APPROACH

Our approach to error concealment in EZW compressed images transmitted across ATM networks is based upon the use of unequal error protection and data interleaving. Unequal error protection is achieved through the use of a convolutional code, namely the (4,3,2) code [13], to code part of the data. This is done to protect the most significant data that is embedded at the beginning of the stream. This is in contrast to the technique described in [7], which relies upon the use of Rate Compatible Punctured Convolutional(RCPC) codes [11]. The disadvantage of using RCPC codes is that the effect of puncturing, used to achieve low rates, cannot be reversed by a decoder using Hard Decisions. The advantage, however, is a decrease in the encoder and decoder complexity. It is to be noted that other channel codes could have been utilized to achieve unequal error protection [14, 15, 16].

The compressed image data is segmented into two groups, the high priority group that will be encoded using the convolutional encoder, and the low priority group that will not be channel coded. The ezw source coder provides the channel encoder with the size of the high priority stream, which comprises 30% of the total size of the compressed image. The channel encoder segments the data into smaller blocks before coding. This is done to reduce decoding delay and preserve rate scalability, while providing small overhead. Subsequent to channel coding a 32 bit code is inserted between the two streams of data to delineate the beginning of the low priority data. The entire data stream is then divided into groups of 768 bytes (the size of 16 ATM cells). The data in each group is interleaved to disperse the effect of ATM cell loss over the entire coded block [17, 18]. Furthermore, the group of 16 ATM cells are organized and sent in the following order: 0, 4, 8, 12, 1, 5, 9, 13, 2, 6, 10, 14, 3, 7, 11, 15, where the numbers denote the cell number in every group of 16 ATM cells. To facilitate this reorganization of the ATM cells, the first 4 bits in every ATM cell payload are used as cell counters. These cell counters will be used by the receiver to reorganize the cells and determine which cells were lost during transmission. It was observed that the total added overhead with such a scheme was 15%.

The packetized data is sent across an ATM network. This was modeled as a queue with a fixed service rate  $\mu$  and a finite input buffer. In our experiments we used a buffer size of 500 ATM cells and changed the service rate  $\mu$  to attain different ATM cell loss rates. ATM cells arriving at the buffer were queued for service if the buffer capacity had not been exceeded, otherwise they were discarded. The arrival rate,  $\lambda$ , of incoming cells depended on the number cells per coded image. It is assumed that initially the input buffer is 95% full. This is to prevent the buffer of the ATM node from discarding cells carrying data from only the latter part of the coded data stream.

At the receiving end arriving ATM cells are reorganized, and their 4 bit counters are used to determine which cells were lost.

Missing cells are replaced by cells with payloads of zeros. The data is then de-interleaved, and the high priority data separated from the low priority data. The high priority data is subsequently decoded by a Viterbi decoder [13], and the output is prepended to the low priority data to form a stream that is decoded by an EZW decoder. When computing the Maximum Likelihood (ML) path through the trellis, it is assumed that bit errors are independent of future and past bit errors [13]. Furthermore, only hard decisions are made by the decoder.

## 6. RESULTS AND CONCLUSION

To test the performance of our scheme two images, we call them the *girls* and *bike* images were coded by an EZW encoder at a maximum data rate of 2 bits per pixel. They were subsequently channel encoded using our scheme, and subjected to 2%, 5%, and 10% ATM cell loss rates. At the receiver, the data was channel decoded and decompressed by an EZW decoder at data rates of 1, 0.5, 0.25 bits per pixel, respectively.

Figure 2 (a) shows the outcome of decoding the *girls* image at a data rate of 1 bits per pixel after having suffered a 10% ATM cell loss. Similarly, the outcome of decoding the image at a data rate of 0.25 bits per pixel after having suffered a 10% ATM cell loss is shown in Figure 2 (b). For comparison purposes, the uncorrupted images decoded at data rates of 1 and 0.25 bits per pixel are provided in Figures 2 (c) and (d), respectively.

Figure 3 (a) shows the outcome of decoding the *bike* image at a data rate of 1 bits per pixel after having suffered a 10% ATM cell loss. Similarly, the outcome of decoding the image at a data rate of 0.25 bits per pixel after having suffered a 10% ATM cell loss is shown in Figure 3 (b). For comparison purposes, the uncorrupted images decoded at data rates of 1 and 0.25 bits per pixel are provided in Figures 3 (c) and (d), respectively.

In Tables 1 and 2, the Peak Signal to Noise Ratio (PSNR) in dB, for the various reconstructed images at the different data rates and ATM cell loss rates are given. When obtaining PSNR values, the uncorrupted decoded images were used as references, and not the original image.

**Table 1.** PSNR values in dB for different data rates and cell loss rates for *girls*.

Data Rate	ATM Cell Loss rate		
	10%	5%	2%
1.0 bpp	34.92	40.27	70.7
0.5 bpp	38.11	58.12	67.72
0.25 bpp	58.10	61.86	73.04

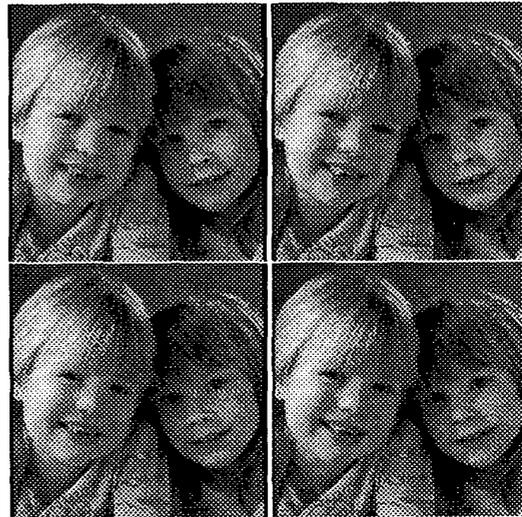
**Table 2.** PSNR values in dB for different data rates and cell loss rates for *bike*.

Data Rate	ATM Cell Loss rate		
	10%	5%	2%
1.0 bpp	22.65	27.36	43.55
0.5 bpp	24.68	53.50	56.23
0.25 bpp	45.31	52.75	57.73

## 7. 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] C. S. Barreto and G. Mendonca, "Enhanced zerotree wavelet transform image coding exploiting similarities inside subbands," in *Proceedings of the IEEE International Conference on Image Processing*, Lausanne, Switzerland, September 16-19 1996, vol. II, pp. 549-552.
- [4] K. Shen and E. J. Delp, "Color image compression using an embedded rate scalable approach," in *Proceedings of IEEE International Conference on Image Processing*, Santa Barbara, California, October 26-29 1997, vol. III, pp. 69-72.
- [5] 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.
- [6] C. D. Creusere, "A new method of robust image compression based on the embedded zerotree wavelet algorithm," *IEEE Transactions on Image Processing*, vol. 6, no. 10, pp. 1436-1442, October 1997.
- [7] H. Man, F. Kossentini, and M. J. T. Smith, "A class of EZW image coders for noisy channels," in *Proceedings of the International Conference on Image Processing*, Santa Barbara, California, October 26-29 1997, vol. III, pp. 90-93.
- [8] S. Thillainathan, D. Bull, and N. Canagarajah, "Robust embedded zerotree wavelet coding algorithm," in *Proceedings of the SPIE Conference on Visual Communications and Image Processing*, San Jose, California, January 28-30 1998, vol. 3309, pp. 58-99.
- [9] J. K. Rogers and P. C. Cosman, "Robust wavelet zerotree image compression with fixed length packetization," in *Proceedings of the IEEE Data Compression Conference*, Snowbird, Utah, March 30 - April 1 1998, pp. 418-427.
- [10] D. W. Redmill and N. G. Kingsbury, "The EREC: An error-resilient technique for coding variable-length blocks of data," *IEEE Transactions on Image Processing*, vol. 5, no. 4, pp. 565-574, April 1996.
- [11] J. Hagenauer, "Rate compatible punctured convolutional codes (RCPC Codes) and their applications," *IEEE Transactions on Communications*, vol. 36, no. 4, pp. 389-400, April 1988.
- [12] V. K. Goyal and J. Kovačević, "Optimal multiple description coding of gaussian vectors," in *Proceedings of the IEEE Data Compression Conference*, Snowbird, Utah, March 30 - April 01 1998, pp. 388-397.
- [13] S. Lin and D. J. Costello, *Error Control Coding: Fundamentals and Applications*, Prentice Hall, 1983.
- [14] R. H. Morelos-Zaragoza and S. Lin, "QPSK block-modulation codes for unequal error protection," *IEEE Transactions on Information Theory*, vol. 41, no. 2, pp. 576-581, March 1995.

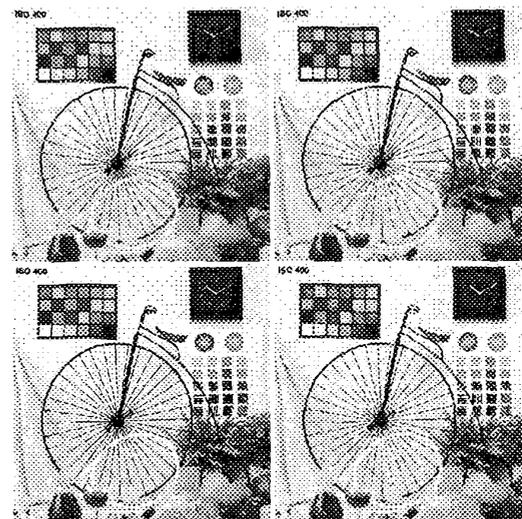
- [15] R. H. Morelos-Zaragoza and S. Lin, "On primitive BCH codes with unequal error protection capabilities," *IEEE Transactions on Information Theory*, vol. 41, no. 3, pp. 788–790, May 1995.
- [16] M. Barazande-Pour, J. W. Mark, and A. K. Khandani, "Multi-level transmission of images over a turbo-coded channel," in *Proceedings of the IEEE Pacific-Rim Conference*, Victoria, BC, Canada Utah, August 20–22 1997, pp. 904–907.
- [17] V. Parthasarathy, J. Modestino, and K. S. Vastola, "Design of a transport coding scheme for high quality video over ATM networks," *IEEE Transactions on Circuits and Systems for Video Technology*, vol. 7, no. 2, pp. 358–376, April 1997.
- [18] H. Ohta and T. Kitami, "A cell loss recovery method using FEC in ATM networks," *IEEE Journal on Selected Areas in Communications*, vol. 9, no. 9, pp. 1471–1482, December 1991.



**Fig. 2.** (a) Image decoded at 1 bpps after 10% of the ATM cells were lost, (b) image decoded at 0.25 bpps after 10% of the ATM cells were lost, (c) uncorrupted image decoded at 1 bpps, (d) uncorrupted image decoded at 0.25 bpps.



**Fig. 1.** (a) Original image, (b) damaged image due to the misinterpretation of a ZTR symbol as a POS symbol.



**Fig. 3.** (a) Image decoded at 1 bpps after 10% of the ATM cells were lost, (b) image decoded at 0.25 bpps after 10% of the ATM cells were lost, (c) uncorrupted image decoded at 1 bpps, (d) uncorrupted image decoded at 0.25 bpps.