CERIAS Tech Report 2001-118

Preprocessing and Postprocessing Techniques for Encoding Predictive Error Frames in Rate Scalable Video Codecs

by E Asbun, P Salama, E Delp
Center for Education and Research
Information Assurance and Security
Purdue University, West Lafayette, IN 47907-2086

PREPROCESSING AND POSTPROCESSING TECHNIQUES FOR ENCODING PREDICTIVE ERROR FRAMES IN RATE SCALABLE VIDEO CODECS

Eduardo Asbun†, *Paul Salama*‡, *and Edward J. Delp*†

† Video and Image Processing Laboratory (*VIPER*)
School of Electrical and Computer Engineering
Purdue University
West Lafayette, Indiana 47907-1285
U.S.A.

‡ Department of Electrical Engineering
Purdue School of Engineering and Technology
IUPUI
Indianapolis, Indiana 46202-5132
U.S.A.

ABSTRACT

The characteristics of "non natural" images, such as predictive error frames used in video compression, present a challenge for traditional compression techniques. Particularly difficult are small images, such as QCIF, where compression artifacts at low data rates are more noticeable. In this paper, we investigate techniques to improve the performance of a wavelet-based, rate scalable video codec at low data rates. These techniques include preprocessing and postprocessing stages to enhance the quality and reduce the compression artifacts of decoded images.

1. INTRODUCTION

Many techniques have been developed for compression of natural images [1, 2, 3, 4, 5, 6, 7, 8, 9]. When these techniques are used on other types of images, such as synthetic (computer generated) images or predictive error frames (PEF) used in video compression, their performance is poor. Coding artifacts may be introduced, especially at low data rates. An additional challenge is present when coding small images, such as QCIF (176x144 pixels), because artifacts are more noticeable.

Predictive error frames (PEF) are used by video compression algorithms that use motion estimation to reduce the temporal redundancy of video sequences. A PEF, along with a set of motion vectors, is used to reconstruct a frame based on a reference frame. The PEF is usually encoded using transform-based codecs to reduce the spatial redundancy in the image. PEFs typically have low energy content.

Several techniques have been proposed to reduce the coding artifacts of transform-based image compression schemes [10, 11]. These techniques make use of postprocessing to reduce the artifacts introduced by the decoder. In [12], we investigated the use of wavelet shrinkage to improve the performance of a video compression algorithm known as *SAMCoW* [1, 3]. The *SAMCoW*, Scalable Adaptive Motion Compensated Wavelet, video compression technique uses a wavelet decomposition of both intracoded and predictive error frames to remove spatial redundancy, and block-based motion compensation to remove temporal redundancy. *SAMCoW* uses the Color Embedded Zerotree Wavelet (*CEZW*) still image coder [2, 4] on its intracoded and predictive error frames. *CEZW* is an embedded technique that uses a combination of an unique spatial orientation tree and color transform to exploit redundancy across color components. A variation of *SAMCoW*, known as *SAMCoW+*, was described in [12, 13, 14].

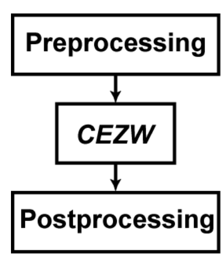


Figure 1: Block diagram of the proposed approach.

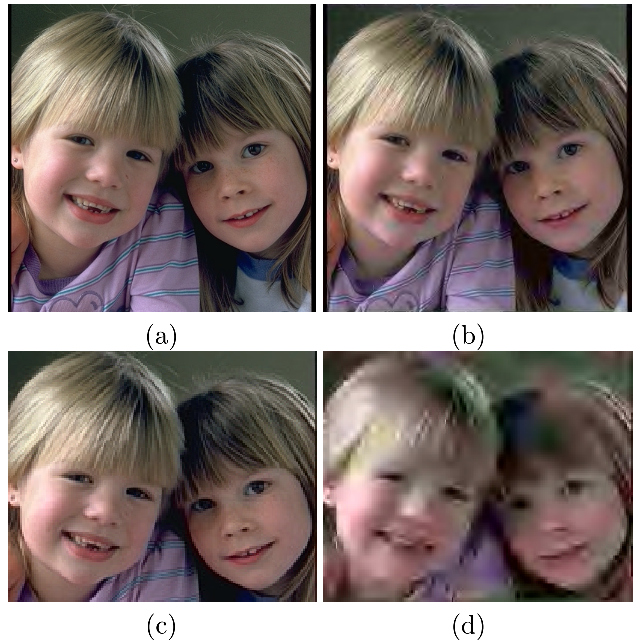
In this paper, we investigate further preprocessing and postprocessing techniques to reduce the coding artifacts of *CEZW* at low data rates. A block diagram of the proposed approach is shown in Figure 1. We are interested in the performance of these techniques when used with QCIF images, because this is the frame size commonly used in low data rate video applications.

This work was supported by a grant from Texas Instruments and an equipment grant from Intel. Address all correspondence to E. J. Delp, ace@ecn.purdue.edu, +1 765 494 1740, or http://www.ece.purdue.edu/∼ace.

The complexity of these techniques is an issue, because *CEZW* is used as part of the *SAMCoW* video compression algorithm, and hence will impact the overall complexity of the codec.

In block-based video compression techniques, such as MPEG-2 and H.263+, PEFs are efficiently encoded using the Discrete Cosine Transform (DCT) on 16x16 blocks of the image. Blocks can be skipped if their energy is below a certain threshold.

Wavelet-based video compression algorithms, such as *SAMCoW*, require that the transform be performed on the entire PEF and, hence, must contend with the global nature of the decomposition and the low pass effect inherent to wavelet filtering.



(a) (b) (c) (d)

Figure 2: Effect of image size of *CEZW*. (a) Original (512x512) (b) Encoded using *CEZW*, decoded at 0.25 bpp. (c) Cropped and resized (176x144) (d) Encoded using *CEZW*, decoded at 0.25 bpp.

2. *CEZW*: EMBEDDED CODING OF COLOR IMAGES

CEZW uses a unique spatial orientation tree (SOT) in the YUV color space [2]. It exploits the interdependence between color components to achieve a higher degree of compression by using the concept that at spatial locations where chrominance components have large transitions, the luminance component also has large transitions [2, 4]. Therefore, each node in the SOT of the luminance component also has descendants in the chrominance components at the same spatial location. The luminance component is scanned first. When a luminance coefficient and all its descendants in both the luminance and chrominance components are insignificant, a zerotree symbol is assigned. Otherwise, a positive significant, negative significant, or isolated zero symbol is assigned. The chrominance components are scanned after the luminance component. *SAMCoW+* uses *CEZW* for coding intracoded (I) frames. A variation of *CEZW*, described in [12], is used for coding PEFs in *SAMCoW+*.

3. CODING ARTIFACTS IN *SAMCOW*

Maintaining acceptable quality in color images coded at rates less than 0.5 bits per pixel (bpp) is a challenge, especially in small images (QCIF or smaller). Ringing artifacts and areas of discoloration are commonly noticeable when using wavelet-based image compression algorithms.

In Figure 2, a 512x512 YUV 4:1:1 image is encoded using *CEZW*, and decoded at 0.25 bpp. The same image, cropped and scaled to 176x144 pixels, is encoded using *CEZW*, and decoded at the same data rate. The subjective quality of the smaller image is lower. The same effect is observed in other wavelet-based video encoders, such as SPIHT [8]. The reason for this effect is that in small images, one pixel represents more "area" than in larger images. Therefore, decoding artifacts are more noticeable.

In Figure 3, two PEFs from the *foreman* sequence, extracted from *SAMCoW+*, are shown. Both are encoded using *CEZW* and decoded at 0.25 bpp. These are 176x144 YUV 4:1:1 images. Blotchiness and ringing artifacts are evident in the decoded frames. In the decoder, the decoded frame is obtained by adding the decoded PEF to the predicted frame produced by motion compensation. The errors propagate to future frames when the decoded frame is used as a reference frame.

In [12], we introduced a variation of *CEZW* that improves its performance when coding PEFs at low data rates. This technique is based on preprocessing the PEFs before encoding, and coding only certain "significant trees" in the wavelet decomposition. We found that the performance of *SAMCoW+* is improved when this technique is used with *CEZW*.

The preprocessing stage used in this paper and in [12], is an adaptive gain (AG) function followed by wavelet shrinkage, as shown in Figure 4. This stage is used to enhance the most important features of a PEF. The parameters of the AG function are dynamically changed, therefore adapting to the varying content of

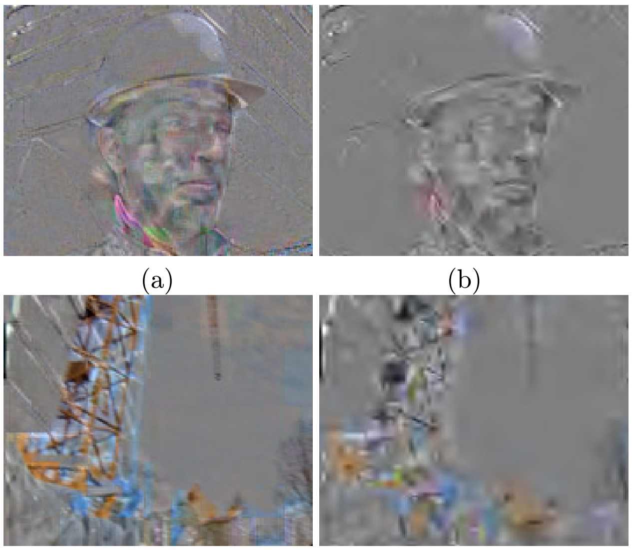


Figure 3: (a) and (c) are original PEFs from frames 35 and 293, respectively, of the *foreman* sequence. (b) and (d) are the same PEFs encoded using *CEZW* and decoded at 0.25 bpp.

PEFs in a sequence. The AG function is defined as

$$H_{AG}(p) = \begin{cases} 0 & , \text{if } 0 \leq |p| < t_1, \\ p & , \text{if } t_1 \leq |p| < t_2, \\ p + K * (t_3 - p) & , \text{if } t_2 \leq |p| < t_3, \\ p & , \text{if } t_3 \leq |p| < max, \end{cases} \quad (1)$$

where t_1, t_2, and t_3 are thresholds that depend on the content of the PEF, K is a constant that controls the feature enhancement, and max is the largest pixel magnitude in the PEF. This function is shown in Figure 5.

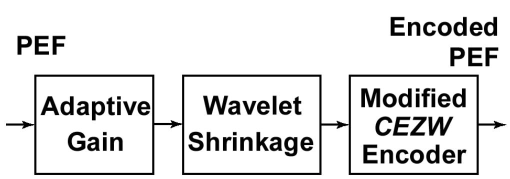


Figure 4: Preprocessing stage in the proposed approach.

In this paper, the postprocessing stage consists of an enhancement stage, as shown in Figure 6. The goal is to sharpen the features of the decoded image. Before obtaining the inverse wavelet transform, the wavelet coefficients are modified to compensate for the preprocessing stage. The neighborhood of those coefficients whose magnitude is larger than a threshold, is examined. If their absolute magnitude is relatively close to zero, the coefficient is multiplied by a scale factor. This

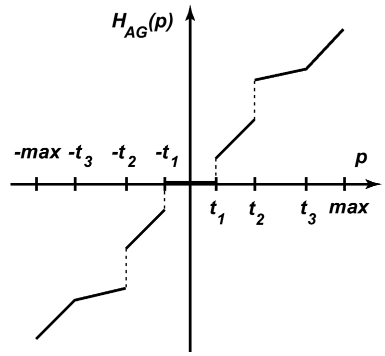


Figure 5: Adaptive gain (AG) function used as part of the preprocessing stage.

is done to avoid having sharp differences of magnitude between adjacent coefficients. These large differences would have been produced during the encoding process, by not allocating enough bits to encode a frame, effectively ignoring nonzero coefficients. Finally, a high pass filter is used on the frame produced by the *CEZW* decoder.

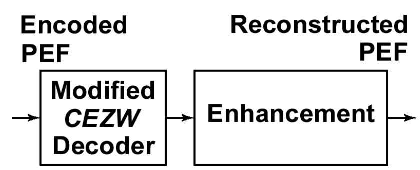


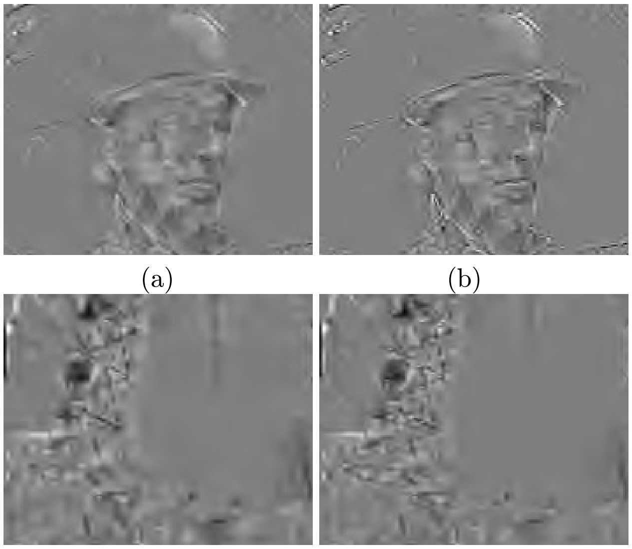
Figure 6: Postprocessing stage in the proposed approach.

In Figure 7 (a) and (c), PEFs from frames 35 and 293, respectively, of the *foreman* sequence are shown after encoding and decoding using *CEZW*. Figure 7 (b) and (d) show the same PEFs after preprocessing and postprocessing, as described in this paper. More detail can be seen in the PEFs obtained using our approach.

4. SUMMARY

In this paper, we presented the use of preprocessing and postprocessing techniques to improve the performance of the *SAMCoW* algorithm by exploiting the characteristics of PEFs (e.g. low energy content). We used image enhancement techniques, including high pass filtering, to postprocess the decoded images. Preprocessing and postprocessing techniques are attractive, because they do not add overhead to the encoded bitstream. The computational requirements of these techniques are low, and do not increase the overall complexity of the video compression algorithm.

Future research includes investigating other filter



(a) (b) (c) (d)

Figure 7: Experimental results. (a) and (c) are PEFs from frames 35 and 293, respectively, of the *foreman* sequence, encoded using *CEZW* and decoded at 0.25 bpp. (b) and (d) are the same PEFs after preprocessing and postprocessing, as described in this paper.

pairs. PostScript and PDF versions of this paper, and the images produced by our algorithm, are available via anonymous FTP to `skynet.ecn.purdue.edu` in the directory `/pub/dist/delp/vlbv99/`.

5. REFERENCES

[1] 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.

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

[3] 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.

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

[5] Z. Xiong, K. Ramchandran, and M. T. Orchard, "Space-frequency quantization for wavelet image coding," *IEEE Transactions on Image Processing*, vol. 6, no. 5, pp. 677–693, May 1997.

[6] C. Chrysafis and A. Ortega, "Efficient context-based entropy coding for lossy wavelet image compression," *Data Compression Conference*, pp. 241–250, Snowbird, UT, March 1998.

[7] J. M. Shapiro, "Embedded image coding using zerotrees of wavelets coefficients," *IEEE Transactions on Signal Processing*, vol. 41, no. 12, pp. 3445–3462, December 1993.

[8] A. Said and W. A. Pearlman, "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.

[9] International Organization for Standardization, *ISO/IEC JTC 1/SC 29/WG 1 (ITU-T SG8), Coding of Still Pictures*, April 1999. (JPEG2000 Verification Model Version 4.0).

[10] T. O'Rourke and R. Stevenson, "Improved image decompression for reduced transform coding artifacts," *IEEE Transactions on Circuits and Systems for Video Technology*, vol. 5, no. 6, pp. 490–499, December 1995.

[11] M.-Y. Shen and C.-C. Kuo, "Artifact reduction in low bit rate wavelet coding with robust nonlinear filtering," *Proceedings of the 1998 IEEE Second Workshop on Multimedia Signal Processing*, pp. 480–485, Redondo Beach, CA, December 1998.

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

[13] E. J. Delp, P. Salama, E. Asbun, M. Saenz, and K. Shen, "Rate scalable image and video compression techniques," *Proceedings of 42nd Midwest Symposium on Circuits and Systems*, Las Cruces, New Mexico, August 1999.

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