A Scalable Video Compression Technique based on Wavelet and MPEG Coding

Pao-Chi Chang and Ta-Te Lu

Department of Electrical Engineering, National Central University, Chung-Li, Taiwan 320

Abstract -- **We present a scalable coding scheme based on discrete wavelet transform (DWT) and MPEG coding. It has the advantages of reusing the MPEG hardware and software as well as relieving the hardware limitation on the image size.**

I. INTRODUCTION

With the development of video technology and high bandwidth networks such as ATM, video services are getting more popular in our daily life. With scalable coding, we can reconstruct video sequences with only the partial bit stream so that the coded signal can be stored or transmitted with different resolutions or bitrates. Discrete wavelet transform (DWT) has shown great potential in video coding and its structure is particularly suitable for scalable coding [1]-[3].

We intend to design a scalable coding scheme that only needs DWT as a pre-processor and minor modifications to existing MPEG codecs so that not only the development time can be much shortened but also the cost can be reduced since the existing MPEG hardware and software can be reused. In addition, it also provides the expansion structure for MPEG coding. Namely, the limitation of the image size imposed by MPEG hardware technology can be relieved by using multiple MPEG chips for different bands. Therefore we propose a scalable DWT with MPEG coding scheme (DWT-MPEG) that employs wavelet pyramid structure, hierarchical fixed block-size motion compensation, and modified MPEG coder. In the modified MPEG coder, custom-designed quantization matrices and scanning directions are applied to different bands for achieving better performance. The encoded bit streams are then sent to multiplexer before transmission to the networks. In decoder, depending on different resolution requirements, some or all of these bands can be synthesized to reconstruct video frames.

II. DWT-MPEG SYSTEM OVERVIEW

The block diagram of a three-layer DWT-MPEG encoding system is shown in Fig. 1 as an example. We use a 9-7 taps biorthogonal wavelet filter [3] which has the linear phase to decompose the image into three layers with seven non-uniform bands, labeled as in Fig. 2. Observing that the lowest band, LL2, has similar statistic features with the original image, we apply MPEG coding with default setting to LL2 directly. However, the default parameters of MPEG are not optimal for other bands due to different statistic characteristics. Thus we modify the MPEG scheme for higher bands to improve the coding efficiency.

III. MODIFIED MPEG CODING

The modified MPEG encoder has the same architecture as the ordinary MPEG, including DCT, scalar quantization, scanning and variable length coding (VLC), with the differences in hierarchical motion estimation/compensation, custom-designed quantization matrices, and distinct scanning directions.

A. Hierarchical Motion Estimation and Compensation

The hierarchical fixed size motion estimation has the same block size with MPEG, 16x16 pixels. Because motion activities in different bands are highly correlated, the motion vector in LL2 can be reused in high frequency bands. In LH1, HL1, and HH1 bands, each LL2 motion vector is used for 4 adjacent blocks with the size of 16x16, the MPEG motion compensation block size. The motion compensation in each band is equal to what MPEG performs. By reusing the motion information, this structure is easy to be implemented with low computational complexity.

B. Custom designs of quantization matrices and scanning directions

We observe that the energy of each band is concentrated only on a small but different part of DCT coefficients. As a result, efficient coding can be achieved with custom designed quantization matrices in MPEG coder for each band. The quantization tables in DWT-MPEG are designed based on energy distribution for still image [4]. It suggests to give more bits to the parts with large energy distributed. After analyzing the correlation of each band, we find spatial redundancy can be further removed by utilizing distinct scanning directions in MPEG coder. We apply four scanning methods, shown as in Fig. 3, to different bands according to the energy distribution. The zig-zag scan is

0-7803-5123-1/99/ $10.00 © 1999 IEEE

used in LL2 band. The vertical and horizontal scans are applied to LH1, LH2 bands, and HL1, HL2 bands, respectively. Finally, the inverse zig-zag scan is chosen for HH1 and HH2 bands.

C. Optimal bit allocation

Before encoding, we need to determine the bit rate for each band. We adopt the perceptual weighting bit allocation which gives different weightings to different bands based on the sensitivity of human visual system [5], [6].

IV. SIMULATION RESULTS

The video sequences "MIT" and "Renata" with the size of 1408x960 pixels and 4:2:0 format are used as the test sequences. Other parameters include that the number of frames in a GOP is 12 (M=12), the frame distance between I and P frames is 3 (N=3), and the frame rate is 30 frames/sec. Fig. 4 shows the PSNR of Y component of MPEG and DWT-MPEG at the bit-rate of 0.3 bpp and 0.2 bpp. At a relatively high bit rate 0.3 bpp, DWT-MPEG is slightly better. However our proposed DWT-MPEG yields 1.5~0.3 dB improvement over MPEG for "Renata" and 1.1~0.1 dB improvement for "MIT" at the rate of 0.2 bpp. Thus it is particularly suitable for relatively low bit rates.

V. CONCLUSION

We have proposed DWT-MPEG video coding scheme that supports resolution-scalability. It has the advantages of reusing the existing technology of MPEG codec in software and hardware as well as short development time for implementation. The constraint of the maximum image size imposed by the hardware capability is also relieved. At a relatively low bitrate, our proposed DWT-MPEG coding provides significantly better reconstructed video quality than the ordinary MPEG.

REFERENCES

[1] Y. Q. Zhang and S. Zafar, "Motion-compensated wavelet transform coding for color videocompression," *IEEE Trans. Circ. And Syst. Video Tech.*, vol. 2, pp. 285-296, Sept. 1992.

[2] S. Kim and S. Rhee, "Interframe Coding Using Two-stage Variable Block-Size Multiresolution Motion Estimation and Wavelet Decomposition," *IEEE Trans. Circ. And Syst. Video Tech.*, vol. 8, pp. 399-409, August. 1998.

[3] M. Antonini, M. Barlaud, "Image Coding Using Wavelet Transform", *IEEE Trans. on Image Processing*, vol. 1, no 2, April 1992

[4] H. W. Paik and A. Khubchandani, "Quantization Scheme for JPEG Baseline Sequential Encoding of Still Images", *Proceeding of 35th Midwest Symposium on Circuit and System*, vol. 2, pp. 976-979,1992

[5] G. Strang and T. Nguyen, *Wavelets and Filter Banks*, Wellesley Cambridge, 1996.

[6] K. R. Rao and J. J. Hwang, *Techniques, standards for Image, video, and Audio Coding*, Prentice Hall PTR, 1996.

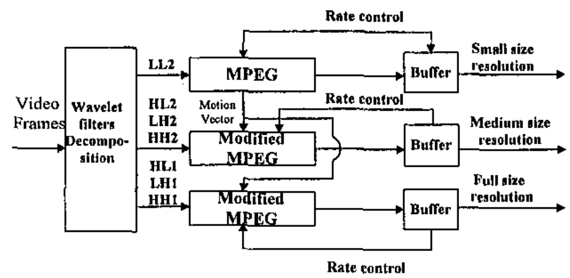

Fig. 1. Encoder of a three-layer DWT-MPEG video coding system

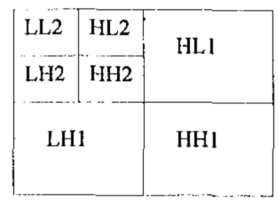

Fig.2. Band decomposition in a three-layer DWT

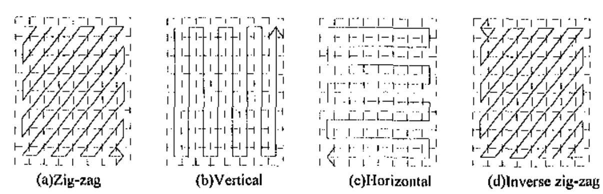

Fig. 3. Custom-designed scanning directions for different bands

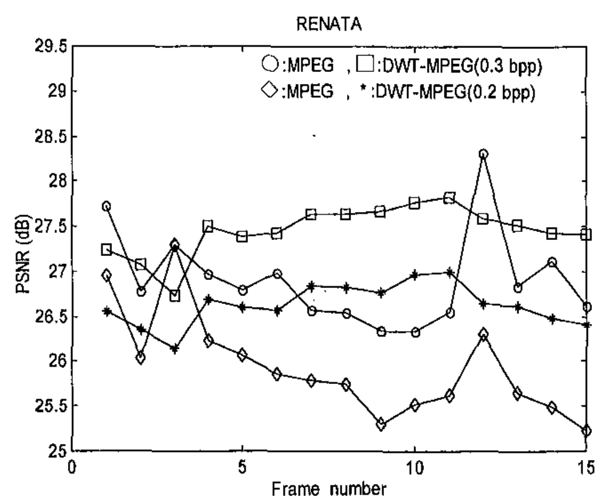

Fig. 4. Comparisons of PSNR between MPEG and DWT-MPEG at 0.3 and 0.2 bits/pixel for Renata