A Low Bit Rate Hybrid Wavelet-DCT Video Codec

Reza Dianat, Mohammed Ghanbari, *Fellow, IEEE*, and Farokh Marvasti

***Abstract*—A hybrid video codec, where the intraframe pictures are wavelet-based coded and the interframe pictures are coded with a standard H.263 codec is proposed. We show that intraframe-wavelet coded pictures not only improve the quality of I-pictures but also result in lower distorted prediction pictures that outperforms a pure H.263 video codec.**

***Index Terms*—Discrete cosine transform (DCT), H.263 codec, set partitioning in hierarchical trees (SPIHT), wavelet.**

I. Introduction

WAVELET decomposition using zero-tree or bit plane entropy coding schemes are the state of the art in image compression algorithms [1]–[3]. For video coding, both three-dimensional (3-D) wavelet decomposition and two-dimensional (2-D) wavelet decomposition with motion estimation have been used [4], [5]. In this letter, we propose a simple hybrid wavelet/discrete cosine transform (DCT) method that can be easily implemented with a slight modification to H.263 codec to improve its performance. The use of wavelets for intraframe pictures has several advantages. First in the case of wavelets by increasing the number of layers, we exploit more intercorrelation among the pixels, while in DCT, the block size is fixed to a length of 8. This is the main reason for use of wavelet in JPEG-2000 [6]. Second in case more frequent I-pictures is required, due to inherent scalability of wavelet, the target bit rate can easily be reached, while that of purely DCT-based has the problem of additional overhead [6]. On the other hand interframe pictures, specially motion-compensated ones are better coded with DCT. This is mainly due to coding tricks used in DCT and also motion-compensated macroblocks create artificial high frequency components that are not suitable for wavelets.

In our proposed method, we exploit the advantages of wavelets for intraframes and those of DCT for interframes. We should emphasize that although in our method only I-pictures are wavelet coded, the impact of this modification can be seen in the interframe pictures as well. With better quality anchor pictures, the remaining interframe pictures are more efficiently coded with DCT. We expect to see considerable improvement, especially at the lower bit rates compared to a pure DCT-based codec such as H.263.

II. Proposed Method

Within an H.263 video codec, for the duration of I-pictures, the intraframe part of the encoder is disabled, instead these pictures are coded with the set partitioning in hierarchical trees (SPIHT) method [2]. At the local loop of the encoder, the reconstructed "I" pictures are also used for the prediction of "B" and "P" pictures similar to the decoder. For ease of implementation we assume all the macroblocks of "B" and "P" frames are DCT coded. This is because, the number of intracoded macroblocks in the "P" and "B" pictures of a very low bit rate codec is very small and there will not be any advantage if they were wavelet coded.

Manuscript received June 20, 2004; revised October 8, 2004. This paper was recommended by Associate Editor R. Lancini.

R. Dianat is with the DSP and Multimedia Laboratory, Iran Telecommunications Research Center, Tehran 14399–55471, Iran (e-mail: dianat@mehr.sharif.edu).

M. Ghanbari is with Department of Electronic Systems Engineering, University of Essex at Wivenhoe Park, Colchester CO4 3SQ, U.K.

F. Marvasti is with Electrical Engineering Department, Sharif University of Technology, Tehran 11365-9363, Iran.

Digital Object Identifier 10.1109/TCSVT.2005.848342

III. Simulation Results and Discussions

We simulated a standard H.263 video codec along with our proposed hybrid method for a variety of raw CIF and QCIF video sequences at different intraframe picture rates. Both subjective and objective results show the superiority of the proposed method over the standard H.263 method in low bit rates. In both methods, arithmetic coding is used for further improvement, with the deblocking filter activated. Fig. 1 shows the average rate-distortion curves (I- and P-pictures) for the "News" and "Coast guard" CIF sequences, at the I-picture rates of 1/10 and 1/5. The simulations were performed on 300 frames at 30 frames/s. As the figures show, there is a 1-dB improvement in PSNR at low bit rates for "News" sequence. This improvement is aboout 0.5 dB for the "Coast guard" sequence.

The curves show that in the high bit rates the standard codec outperforms the proposed codec. The reason is that the wavelet-based image coding like SPIHT performs better than DCT-based image coding in low bit rates. In high bit rates the difference between the DCT-based image coding and their wavelet-based counterparts is negligible. On the other hand the different natures of the incurred distortions in these two class of image coding results in a mismatch. In the pure DCT-based video coding in both intraframes and interframes the distortions have the same nature. In the proposed method the distortion in the intraframes is blurring, but that in the interframe is blockiness. The decoded frames 101 (encoded as intra) and 105 (encoded as inter) of the "Foreman" CIF sequence are shown in Fig. 2. The sequence was coded at 300 kb/s and every fifth frame was coded as "I". The figure shows subjective improvement of the proposed method in comparison to the H.263 standard. The average PSNR for this sequence was 29.6 dB using the proposed method and 28.5 dB using pure H.263 codec. As can be seen not only I-pictures are better coded under the hybrid scheme, the impact of better anchor picture has resulted in a better quality of P-picture despite the fact that P-pictures are coded in a similar way for both methods.

As the intraframe rate decreases, the effect of the intraframes becomes less important. In fact the effect of the coding of an intraframe dies out as the time passes. Therefore it is expected that the superiority of the proposed method can be more pronounced with increasing of the intrarate. Table I compares the compression

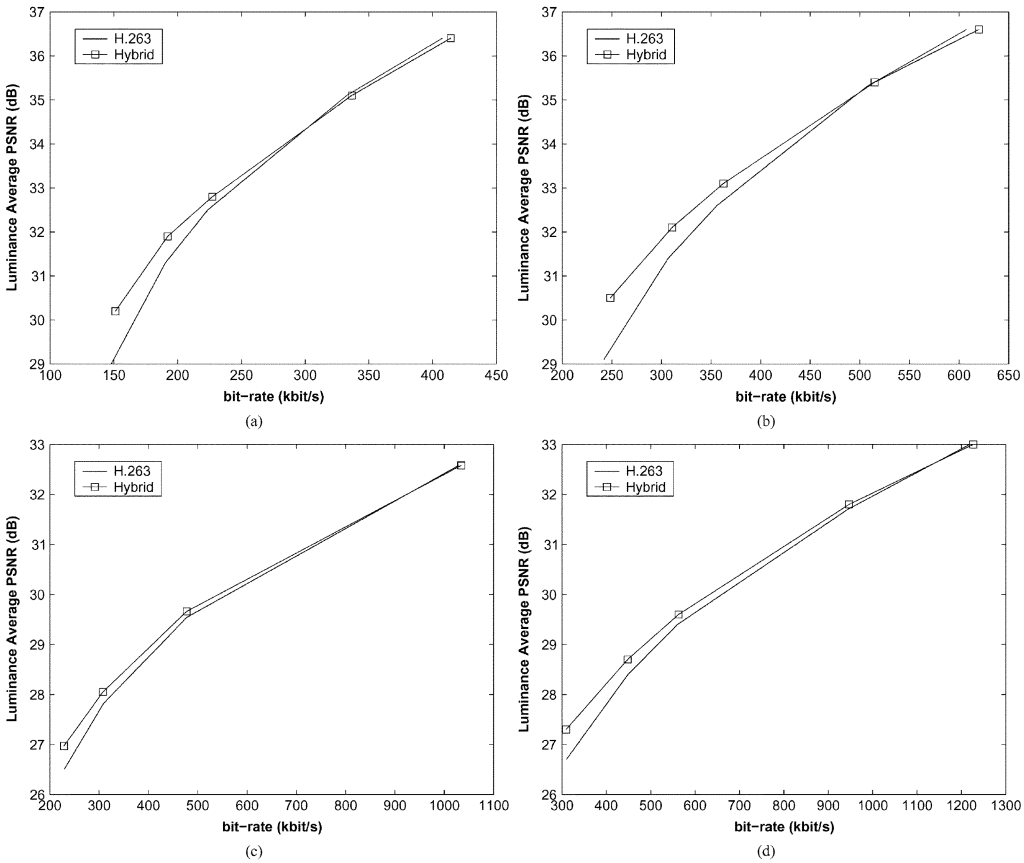

Fig. 1. CIF sequences compression results (a) "News" sequence, intrarate $= 1/10$. (b) "News" sequence, intrarate $= 1/5$. (c) "Coast guard" sequence, intrarate $= 1/10$. (d) "Coast guard" sequence, intrarate $= 1/5$.

Fig. 2. Decoded "Foreman" CIF video at 300 kb/s (One intra, four inter). (a) Decoded frame no. 101 (intra) using H.263. (b) Decoded frame no. 101 (intra) using proposed method. (c) Decoded frame no. 105 (inter) using H.263. (d) Decoded frame no. 105 (inter) using proposed method.

performance for 300-frame CIF "Foreman" sequence for the intrarates of 1/50, 1/10, and 1. As can be seen in the extreme intrarate of 1/50 (one out of 50 frames is intracoded), the proposed method has almost no effect. The largest improvement is at the intrarate

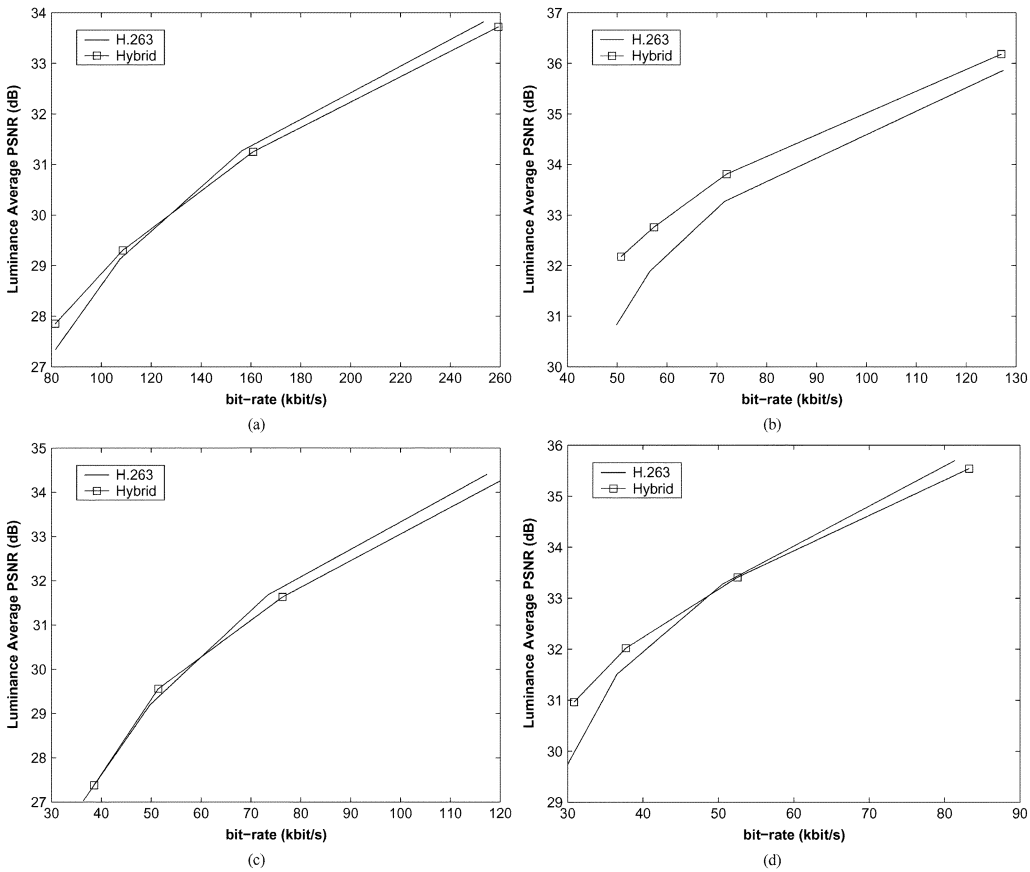

Fig. 3. QCIF sequences compression results. (a) "Foreman," intrarate $= 1/5$, (b) "Suzie," intrarate $= 1/5$, (c) "Container," intrarate $= 1/10$, (d) "Mother & daughter," intrarate $= 1/10$.

TABLE I
COMPRESSION RERSULS FOR 300-FRAME "FOREMAN" CIF VIDEO FOR THREE INTRARATES

intra coding rate	compression method	Bit rate (kbit/s)	Luminance PSNR (dB)
1/50	Hybrid	299.11	31.67
	H.263	299.47	31.61
1/10	Hybrid	396.83	32.06
	H.263	396.25	31.86
1	Hybrid	1600	33.19
	H.263	1600	32.47

TABLE II
COMPRESSION RESULTS FOR 300-FRAME "AKIYO" CIF VIDEO FOR THREE INTRARATES

intra coding rate	compression method	Bit rate (kbit/s)	Luminance PSNR (dB)
1/50	Hybrid	94.80	36.19
	H.263	95.02	36.03
1/10	Hybrid	146.91	35.44
	H.263	147.34	34.63
1	Hybrid	773	34.27
	H.263	773	31.49

of 1, when all the frames are intracoded. Table II shows the results for 300-frame "Akiyo" CIF video. Again the the proposed method has no effect at the intrarate of 1/50 and at the intrarate of 1, the difference is the largest, at about 3 dB.

The other parameter that plays an important role in the wavelet-based compression methods, is the image size. The larger the size of an image, the more decomposition scales is possible and as a result the better performance can be obtained by the wavelet-based methods. The reason is that as the image size decreases, the correlations among the pixels decreases, and the zero-trees vanish. The efficiency of the zero-tree coding of wavelet coefficients relies on the validity of the assumption that most of the trees in the transformed image are zero-trees. Fig. 3 demonstrates the result for the QCIF "Foreman" sequence in different bit rates for the intrarate of 1/5. It can be seen that little can be earned at low bit rates in "Foreman" and "Container" sequences. At high bit rates the standard method is more preferable generally, but in some cases like "Suzie" sequence, the proposed method still performs better. Also it can be seen that as the intrarate decreases, the effect of the proposed method vanishes.

REFERENCES

[1] J. M. Shapiro, "Embedded image coding using zerotrees of wavelet coefficients," *IEEE Trans. Signal Process.*, vol. 41, no. 12, pp. 3445–3462, Dec. 1993.

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

[3] D. Taubman, "High performance scalable image compression with EBCOT," *IEEE Trans. Image Process.*, vol. 9, no. 7, pp. 1158–1170, Jul. 2000.

[4] S. J. Choi and J. W. Woods, "Motion-compensated 3-D subband coding of video," *IEEE Trans. Image Process.*, vol. 8, no. 2, pp. 155–167, Feb. 1999.

[5] D. Mukherjee and S. K. Mitra, "Vector SPIHT for embedded wavelet video and image coding," *IEEE Trans. Circuits Syst. Video Technol.*, vol. 13, no. 3, pp. 231–246, Mar. 2003.

[6] M. Ghanbari, *Standard Codecs: Image Compression to Advanced Video Coding*, ser. IEE Telecommunication 49, London, U.K., 2003.