SIGNIFICANCE-LINKED WAVELET VIDEO CODER

Jozsef Vass, Bing-Bing Chai[†]*, and Xinhua Zhuang*

Department of Computer Engineering
& Computer Science
University of Missouri-Columbia
Columbia, MO 65211, USA

[†]Sarnoff Corporation
Princeton, NJ 08543, USA

ABSTRACT

Perhaps, Sarnoff Corporation's zerotree entropy (ZTE) coder is the most successful wavelet video coder published so far which exploits the statistical properties of wavelet-transformed images by utilizing novel data representation and organization strategies. In this paper, a high performance hybrid video coding algorithm termed video significance-linked connected component analysis (VSLCCA) is developed. It is quite encouraging that, at least empirically convinced, the wavelet transform with aids of those recently published innovative data representation and organization methods can be an invaluable asset in video coding if motion-compensated error frames are coherent. In VSLCCA, time domain motion estimation followed by *exhaustive* overlapped block motion compensation is utilized to ensure coherency, and then wavelet transform is applied to each error frame with significant wavelet coefficients being encoded by highly efficient SLCCA technique. Experimental results on standard MPEG-4 test sequences show that VSLCCA is superior to H.263 and ZTE by 0.48 dB and 0.77 dB on average, respectively.

1. INTRODUCTION

The recently adopted ITU-T H.263 Recommendation [10] offering solution for very low bit rate videophony applications was the first standard to break the 64k bits-per-second (bps) barrier in audio-visual communications. H.263 represents a modified and enhanced version of previous block-based video coding standards such as H.261, MPEG-1, and MPEG-2 tailored to very low bit rate applications. One of the functionalities of the emerging MPEG-4 standard targeting very low and medium bit rate multimedia communications is still improved coding efficiency [9].

In recent years, we have seen a tremendous success of wavelet image coding which is mainly attributed to innovative data organization and representation strategies [8, 6, 7, 2, 3]. Despite success in still image coding, there have been only a few successful attempts to apply wavelet transform to video coding [5]. It appears that the global wavelet transform by no means tolerates artificial blocking effects introduced by conventional block-based motion estimation and compensation methods.

In the paper, a novel hybrid wavelet video coding algorithm termed video significance-linked connected component analysis (VSLCCA) for very low bit rate applications is developed. It is empirically convinced that wavelet transform with aids of those innovative data organization and representation strategies can provide an invaluable asset with video coding as far as coherency of each motion compensated error frame is guaranteed. In VSLCCA, motion estimation followed by *exhaustive* overlapped block motion compensation is utilized to ensure coherency, and then wavelet transform is applied to each error frame with significant wavelet coefficients being encoded by highly efficient significance-linked connected component analysis technique [2, 3].

Performance evaluation on several MPEG-4 test sequences shows that for intraframe coding, the proposed codec exceeds H.263 and ZTE in PSNR as much as 1.79 dB and 0.92 dB at 14 kbits on average, respectively. For entire sequence coding, VSLCCA is superior to H.263 and ZTE by 0.48 dB and 0.77 dB on average, respectively. The subjective advantage of VSLCCA over H.263 is also distinctive in that the disturbing blocking effects are entirely eliminated.

The rest of the paper is organized as follows. In Section 2 VSLCCA algorithm is described in detail. Performance evaluation is given in Section 3. The last section concludes the paper.

2. VSLCCA VIDEO CODING ALGORITHM

The block diagram of VSLCCA is shown in Fig. 1. The time domain block-based motion estimation and *exhaustive* overlapped block motion compensation (OBMC) follow [10, 4] in spirit. The so-generated coherent blocking-effect-free motion-compensated error frames are each undergone wavelet decomposition and significant wavelet coefficients are encoded by using advanced SLCCA data representation and organization technique.

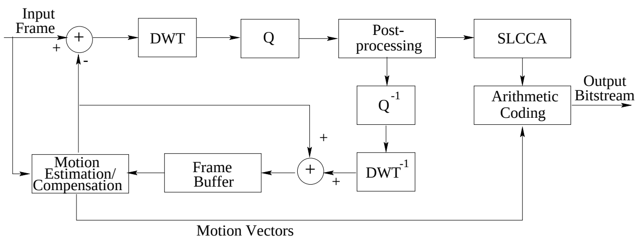

Figure 1: Block diagram of VSLCCA video coding algorithm.

2.1. Motion Estimation

The original frame is divided into non-overlapping 16×16 *macroblocks*. As in H.263 Recommendation, each macroblock can have either zero, one or four motion vectors. Full search block matching algorithm with ± 15 pixels search range and subpixel refinement is used to calculate one motion vector per macroblock. Then each macroblock is split into four 8×8 *blocks* and one motion vector per block is determined by using the same search procedure. Based on two thresholds, zero, one, or four motion vectors are decided for each macroblock. Finally, motion vectors are encoded by adaptive arithmetic coder where a separate context is used for vertical or horizontal vector components.

2.2. Motion Compensation

Blocking-effect-free coherent motion compensation is crucial to success of VSLCCA. In DCT-based hybrid video coding algorithms, the effectiveness of DCT is not significantly degraded by block-based motion estimation and motion compensation since the introduced artificial motion block boundaries are well aligned with DCT block boundaries. In the case of the global wavelet transform, however, many spurious high frequency components can be produced by the artificial blocking discontinuities and effectiveness of wavelet transform in terms of energy compaction is inevitably gone. This argument is also empirically supported. In VSLCCA, *exhaustive* OBMC is utilized to ensure coherency and works as follows. The frame is divided into non-overlapping blocks of 8×8 pixels. Each block is assigned by one motion vector; in case of one motion vector per macroblock being decided during motion estimation, that motion vector is replicated for each constituent block. The predicted frame is composed as follows. First, each block in the predicted frame points to a corresponding block in the previous reconstructed frame by using the assigned motion vector. Then the corresponding block is extended to include its eight neighboring blocks and a weighted sum of these nine reconstructed blocks is used to represent the predicted block. For any two nonoverlapping predicted blocks, their respective nine-reconstructed-block sets might be overlapped. Performance evaluation on several test image sequences shows that *raised cosine* window with 4 pixels overlap is a good choice for weighting.

2.3. Motion-Compensated Error Image Encoding: Significance-Linked Connected Component Analysis

SLCCA represents a high performance wavelet image coding algorithm [2, 3]. As opposed to zerotree-based algorithms such as EZW and SPIHT, it follows MRWD in spirit to directly deal with significant coefficients organized as irregularly shaped clusters. But SLCCA exploits not only the *within subband* clustering property but also the *cross-scale* dependency among clusters. As was implied by the *cross-scale* statistical decaying of magnitude of wavelet coefficients [8, 3], a significant cluster at a course scale likely can be linked to a significant cluster at a finer scale. This significance-linked connected component analysis or clustering renders the highest PSNR with SLCCA in comparison to other three wavelet image coders

3. PERFORMANCE EVALUATION

To conduct a fair performance comparison, the same test sequences used in [5] are used here. The performance of VSLCCA is evaluated on eight standard MPEG-4 test sequences (four Class A sequences “Akiyo,” “Container Ship,” “Hall Monitor,” and “Mother & Daughter;” and four Class B sequences “Coast Guard,” “Foreman,” “News,” and “Silent Voice”) in QCIF resolution. Peak signal-to-noise ratio (PSNR) is used as the objective performance measure.

Motion estimation is carried out only on the luminance component. For motion compensation of the chrominance components the corresponding luminance motion vectors are divided by two and OBMC with 4×4 block size is applied. For wavelet decomposition biorthogonal 9/7 filters [1] are used. The quantizer step size is used to adjust the bit rate.

In interframe coding comparison, for H.263 unrestricted motion vector mode and advanced prediction mode are used. Class A sequences are coded at 5 frame-per-second (fps), 10 kbps, and Class B sequences are coded at 7.5 fps, 48 kbps. First H.263 was run, then the quantizer step size was adjusted in VSLCCA to match the bit rate of H.263. The performance evaluation is tabulated in Tables 1 and 2 for Class A and Class B sequences, respectively. For Class A test sequences for the luminance component VSLCCA is superior to H.263 and and ZTE by 0.73 dB and 1.10 dB on average, respectively. For Class B test sequences VSLCCA slightly outperforms H.263 by 0.23 dB on average and ZTE ranging from 0.24–0.62 dB. For the "Container Ship" sequence, the frame-by-frame luminance PSNR comparison between H.263 and VSLCCA is shown in Fig. 2.

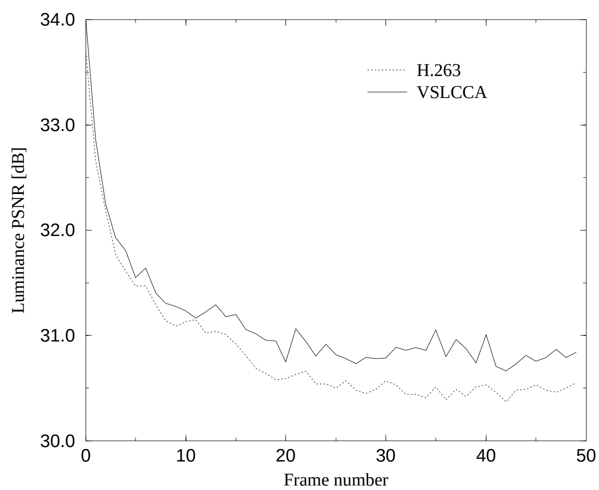

Figure 2: Performance comparison (PSNR, [dB]) between H.263 and VSLCCA for the luminance component of the "Container Ship" sequence at 5 fps, 10 kbps.

Intraframe coding comparison is done at 14k bit (kb) (0.39 bits-per-pixel). The performance comparison among VSLCCA, ZTE and H.263 is tabulated in Table 3. For the luminance component VSLCCA exceeds H.263 by 1.79 dB on average, and ZTE ranging from 0.64–1.39 dB. The original, and decoded first frame of the "Akiyo" sequence by VSLCCA and H.263 are shown in Fig. 3.

4. CONCLUSIONS

The paper provided a strong evidence that wavelet transform could be excellent not only in intraframe coding but also in interframe coding. Therefore, using wavelet techniques in video coding is worth while further research.

5. REFERENCES

[1] M. Antonini, M. Barlaud, P. Mathieu, and I. Daubechies. Image coding using wavelet transform. *IEEE Transactions on Image Processing*, 1(2):205–220, April 1992.

[2] B.-B. Chai, J. Vass, and X. Zhuang. Highly efficient codec based on significance-linked connected component analysis of wavelet coefficients. In *Proceedings of SPIE AeroSense'97*, 1997.

[3] B.-B. Chai, J. Vass, and X. Zhuang. Significance-linked connected component analysis for high performance low bit rate wavelet coding. In *Proceedings of IEEE International Conference of Multimedia Signal Processing*, 1997.

[4] J. Katto, J.-I. Ohki, S. Nogaki, and M. Ohta. A wavelet codec with overlapped motion compensation for very low bit-rate environment. *IEEE Transactions on Circuits and Systems for Video Technology*, 4(3):328–338, June 1994.

[5] S.A. Martucci, I. Sodagar, T. Chiang, and Y.-Q. Zhang. A zerotree wavelet video coder. *IEEE Transactions on Circuits and Systems for Video Technology*, 7(1):109–118, February 1997.

[6] 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*, 6(3):243–250, June 1996.

[7] S. Servetto, K. Ramchandran, and M.T. Orchard. Wavelet based image coding via morphological prediction of significance. In *Proceedings of IEEE International Conference on Image Processing*, pages 530–533, October 1995.

[8] J.M. Shapiro. Embedded image coding using zerotrees of wavelet coefficients. *IEEE Transactions on Signal Processing*, 41(12):3445–3462, December 1993.

[9] T. Sikora. The MPEG-4 video standard verification model. *IEEE Transactions on Circuits and Systems for Video Technology*, 7(1):19–31, February 1997.

[10] Draft ITU-T Recommendation H.263. Video coding for low bitrate communications, December 1995.

Sequence	Bit Rate [kbps]	Luminance, PSNR [dB]			Chrominance, PSNR [dB]				
					U	V	Avg.	U	V
		H.263	ZTE	VSLCCA	H.263		ZTE	VSLCCA	
Akiyo	8.87	34.61	34.61	35.55	38.45	40.52	39.86	39.81	41.43
Container Ship	10.04	30.79		31.13	37.81	36.98		37.94	36.72
Hall Monitor	8.18	30.36	30.25	31.51	36.21	39.35	38.05	37.45	39.95
Mother & Daughter	9.54	33.40		33.87	38.91	39.79		39.96	40.54

Table 1: Performance comparison of coding results for Class A sequences at 5 fps, 10 kbps.

Sequence	Bit Rate [kbps]	Luminance, PSNR [dB]			Chrominance, PSNR [dB]				
					U	V	Avg.	U	V
		H.263	ZTE	VSLCCA	H.263		ZTE	VSLCCA	
Coast Guard	46.03	29.74	29.20	29.82	39.79	41.72	40.88	39.88	41.84
Foreman	46.77	31.91		32.26	37.54	37.71		37.43	37.56
News	49.85	35.10	35.17	35.41	38.76	39.45	40.46	38.77	39.64
Silent Voice	49.94	35.94		36.10	39.36	40.18		39.70	40.52

Table 2: Performance comparison of coding results for Class B sequences at 7.5 fps, 48 kbps.

Sequence	File Size [Bits]	Luminance, PSNR [dB]			Chrominance, PSNR [dB]				
					U	V	Avg.	U	V
		H.263	ZTE	VSLCCA	H.263		ZTE	VSLCCA	
Akiyo	14000	33.06	34.62	35.34	35.23	37.38	36.19	37.12	39.45
Coast Guard	13888	30.12		31.17	41.98	44.94		42.83	45.12
Container Ship	14088	28.81		30.69	36.53	35.15		37.35	36.59
Foreman	13976	30.11	30.86	32.25	38.22	38.31	38.69	39.26	39.03
Hall Monitor	14104	29.68		31.83	36.04	39.99		37.40	40.29
Mother & Daughter	13624	33.78		35.53	39.27	39.67		40.85	41.57
News	14080	28.60	29.38	30.02	33.24	34.40	33.47	34.90	35.78
Silent Voice	13688	30.34		32.02	34.76	36.53		35.74	37.25

Table 3: Performance comparison of intraframe coding results at 14 kb.

(a) (b) (c)

Figure 3: Intraframe coding results of the first frame of the "Akiyo" sequence at 14 kb. (a) The original frame, reconstructed frame by (b) VSLCCA, PSNR = 35.34 dB, and (c) H.263, PSNR = 33.06 dB.