LOW COMPLEXITY VIDEO CODING FOR SENSOR NETWORK

Athar Ali Moinuddin[1]

[1]*Department of Electronics Engineering, AMU, Aligarh (UP) 202002, India*

Abstract

Modern video codecs such as H.264/AVC give state-of-the-art compression performance. However, extensive use of optimization tools makes them highly complex and hence not suitable for wireless video sensor network. In this paper an efficient video codec with substantially reduced complexity is proposed. Simulation result shows that the proposed video codec gives comparable compression performance compared to H.264/AVC but at substantially reduced computational complexity.

Keywords—*Low complexity coding, Sensor network, Video coding, Wavelet transform.*

--***--

1. INTRODUCTION

Modern video coding techniques today offer the possibility to store or transmit the vast amount of data necessary to represent digital videos in an efficient way. Meanwhile, the growing popularity of wireless network and the availability of low cost portable devices have fuelled a host of new sets of applications and services [1]. An example of such an application is a wireless video sensor network for surveillance. A wireless video sensor network comprises a dense field of battery-powered miniature video cameras, each packaged with a low power wireless transceiver that is capable of processing, sending, and receiving data. Such wireless systems are usually not expensive to install, but requirement of low power consumption and thus low computational demand at the video sensors is of prime concern. Since every sensor also acts as a network node with limited throughput, video sensor data needs to be compressed efficiently-yet with very low encoder complexity. There is usually limited constraint at the decoder [2]. Similarly, efficient compression with low complexity is also required for wireless PC cameras, mobile camera phones, disposable video cameras, and networked camcorders.

The latest standard video codec H.264/AVC [3] give state-of-the-art compression, albeit at the cost of increased computational complexity. As a consequence, their software or even hardware implementations require powerful processors with enough computational resources. All the existing video coding standards such as MPEG-2, MPEG-4, H.263 and H.264/AVC are based on the hybrid scheme of temporal motion compensation and the discrete cosine transform (DCT) [4]. Wavelet with multi-resolution decomposition coupled with a progressive encoding has recently received much attention and has emerged as a powerful competitor against the traditional hybrid video coding scheme. The key advantage of the wavelet-based approaches is their scalability functionality [5].

Essentially the design of wavelet video codecs can be based on either a 3-D wavelet transform [6]-[9] or an adaptation of the traditional feedback architecture [10]-[16] similar to the standard video codecs such as H.264/AVC. Although in the past both approaches have been used but lately there is more emphasis on wavelet video coding using 3-D wavelet transform. However, recent investigations have shown that open-loop architecture of 3-D wavelet transform has its own drawback which can be overcome by using closed-loop architecture only [8]. Further, storage requirement and the large delay involved limits its use for low memory, delay sensitive video services.

In this paper, a wavelet hybrid video coder with low complexity is presented. The video coder is based on temporal motion compensation and the discrete wavelet transforms (DWT). The key element of the video coder is the use of an efficient low complexity algorithm to quantize and code the motion compensated residual frames. The rest of the paper is organized as follows. An overview of the video coding system is given in Section 2. Section 3 gives the implementation detail. Simulation results and discussions are presented in Section 4 and finally the paper is concluded in Section 5.

2. SYSTEM OVERVIEW

A simplified block diagram of the video coder is shown in Fig. 1. Essentially, it consists of: a temporal predictive feedback loop for motion estimation and compensation, DWT, WBTC encoder, arithmetic coder and a local decoder [17]. In the temporal prediction loop, prediction is performed using motion estimation and compensation. Conventional block-based algorithm is used for motion estimation. The differential motion vectors are lossless coded with adaptive arithmetic coding and are multiplexed with the bits generated by the WBTC encoder (with or without arithmetic coding). Overlapped block motion compensation (OBMC) [18] is employed to remove temporal redundancy. Overlapping is

necessary in order to remove the blocking artifacts in residual frames resulting from block matching.
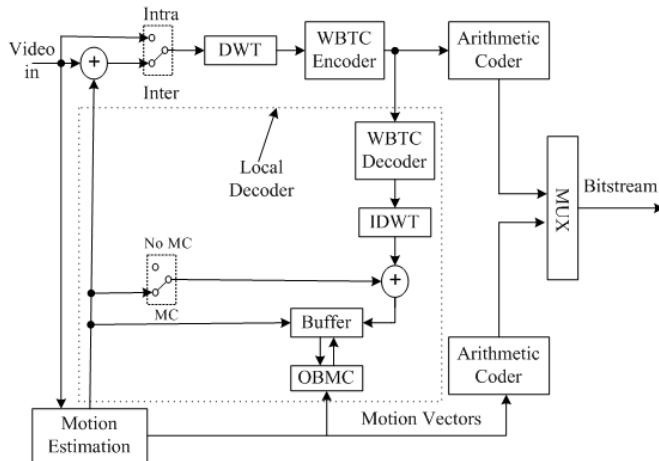

**Fig 1** Block diagram of the proposed video encoder

Frames are encoded in either intra (I) or inter (P) modes only. For I-mode, the predictive feedback loop is bypassed. In the present work, first frame is always coded using I-mode and P-mode is used for the remaining frames of the sequence. Each color plane of I- or P-frames is wavelet transformed to suitable number of decomposition levels. The resulting wavelet coefficients are then encoded using WBTC algorithm [17]. Since the proposed algorithm performs both quantization and encoding, it generates binary symbols that may optionally be entropy coded. For entropy coding of the coded bitstream, we have used a bitplane-based arithmetic coding with non-uniform initialization model as given in.

The precise rate control is achieved by first encoding the I-frame at a fixed bitrate and then allocating the remaining bit budget equally among all the P-frames. Due to use of WBTC algorithm, which generates fully embedded bitstream for each frame, the bit budget can precisely be matched. This is in contrast to standard coders, where it is very difficult to match the exact bitrate.

After dyadic wavelet decomposition, the wavelet coefficients in each colour plane are divided into blocks of $n \times n$ coefficients. Inter-dependency among the three color planes is exploited through the use of a composite spatial orientation block-tree as shown in Fig. 2. Since the number of blocks in the LL-subbands of U and V color plane is four times the number of descendent free blocks in the LL-subband of Y color plane, therefore each descendent free block in the LL-subband of Y color plane has four children in the LL-subband of each of U and V color planes as shown in Fig. 2 (a). This will ensure that all the root blocks of each of the chrominance plane are linked with the root from luminance plane only. The detailed composite spatial orientation block-tree structure is shown in Fig. 2 (b).
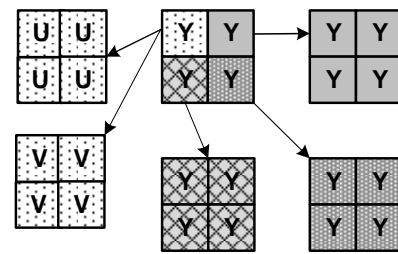
(a)
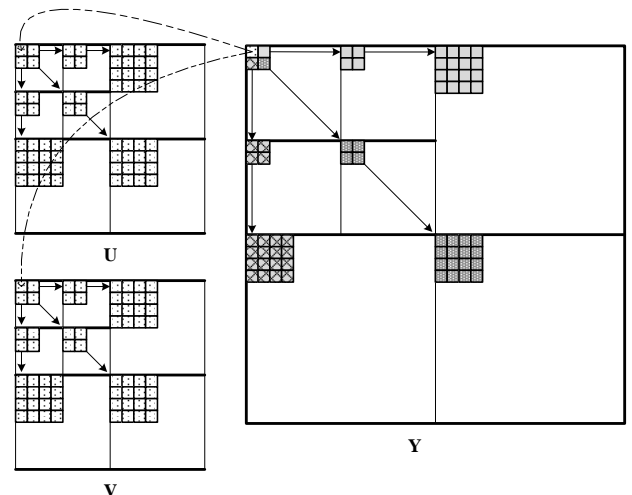
(b)

Fig 2 Parent-child relationships in the composite block-tree structure (a) 2×2 root blocks in Y color plane and their off springs and (b) the detailed composite block-tree.

3. IMPLEMENTATION DETAIL

Performance of the video coding system with the same basic algorithm can be quite different according to the actual implementation. Thus it is necessary to specify the main features of the implementation. In this section, issues that are important from the point of view of practical implementation are described.

The video sequences are coded as IPPPP....., that is the first frame is intra-coded (I-frame) and all other frames are inter-coded (P-frame). The I-frame is coded at a fixed rate of 1.0 bits/pixel (bpp), which is taken into account to allocate bits for the remaining P-frames, and the overall target bitrate is precisely achieved. The 9/7 biorthogonal wavelet [19] is used for I-frame because of its overall good performance for still images [20]. For P-frames, we use Haar wavelet because it is simpler and performs better [21]. For QCIF sequences 3-level decomposition is used for Y plane and 2-level for each of the chrominance planes. And for CIF sequences, Y plane is decomposed to 4-levels while U and V planes are decomposed to 3-levels. The number of decomposition levels is included in

the header of the bitstream. For I-frame, image mean of the three color components are extracted and transmitted as the header of that frame. However, since the P-frames have nearly zero mean values individually for each color components, there is no need to extract the image mean before they are wavelet transformed.

In order to remove the temporal redundancies, simple full search block motion estimation with half-pixel accuracy is used. Frames are divided into blocks of 16×16 pixels and one motion vector per block is estimated with in the search range of ±7 and ±15 pixels (for QCIF and CIF sequences respectively) in both directions relative to the block. The motion vectors are restricted such that all pixels referenced by them are within the coded frame area. The horizontal and vertical components of the motion vectors are assumed to be independent of each other and thus they are coded separately. The differential motion vectors (difference of actual and its predicted value) are losslessly coded with adaptive arithmetic coding. For each component, the predicted value is the median of the corresponding components from the three surrounding blocks. The motion compensation is performed using OBMC. For the chrominance plane (in 4:2:0 format), the motion vectors are derived by dividing each component of the luminance motion vector by a factor of two due to the lower chrominance resolution. The component values of the resulting quarter pixel resolution vectors are modified towards the nearest half pixel position.

4. SIMULATION RESULTS

The video coding system with WBTC has been implemented in software and the performance is evaluated on four color video sequences (YUV 4:2:0 format), namely; Mother-Daughter (QCIF), Carphone (QCIF), Salesman (CIF) and Akiyo (CIF). Tests were performed using 96 frames at a rate of 30 frames per second (fps) for CIF sequences while the QCIF sequences are coded at a rate of 10 fps with a total of 286 frames. The results are compared with video coders based on SPIHT [22], VSPIHT [14] and H.264/AVC [3]. The objective quality of the decoded frames is expressed in terms of the peak signal-to-noise ratio (PSNR) of the three colour components. The initial results of the proposed video coder are found to be superior at low bitrates compared to H.263+ [4] and 3-D SPIHT [6] as reported in [17].

Figures 3 show the frame-by-frame luminance PSNR for the QCIF sequences at 30 kbps. A comparison reveals that the WBTC is quite efficient in coding the residual frames as compared to SPIHT and VSPIHT. In particular, the average luminance PSNR of WBTC-based video coder is superior to VSPIT and SPIHT-based video coders by 1.1-1.4 dB and 1.9-2.2 dB, respectively.

The coding performances of various coders are also compared with arithmetic coding. The wavelet coefficients bits

generated by SPIHT, VSPIHT and WBTC are passed through a bitplane-based arithmetic coder. Also included for comparisons are the performance figures for H.264/AVC [3], representative of the state-of-the-art in hybrid video coding. The H.264/AVC encoding is configured so as to operate in the low complexity mode. Specifically, H.264/AVC JM 14.0 reference software [23] operating in the Baseline Profile and without rate-distortion optimization is used. Motion estimation is performed using Fast Full Search algorithm having quarter pixels accuracy.

Table 1 compares the coding performance in terms of average PSNR (dB) of the three color components at three different bitrates for Mother-Daughter (QCIF) and Akio (CIF) sequences. As can be seen, coding performance of the video coder based on WBTC is superior to that of SPIHT and VSPIHT. Specifically, its luminance PSNR is superior to SPIHT and VSPIHT by up to 0.6 dB for both the sequences. Also, its chrominance planes PSNR is superior (with few exceptions) to SPIHT and VSPIHT by 0.3-0.7 dB. When compared to H.264/AVC, the luminance PSNR of the proposed coder is inferior by 0.3-0.6 dB. The chrominance PSNR of H.264/AVC is always better than WBTC for Mother-Daughter sequence, whereas for Akio sequence its chrominance planes performance is comparable (sometime better and sometime inferior). We believe that this performance gap is mainly due to use of motion estimation with half pixel in WBTC, whereas H.264/AVC uses quarter pixel accuracy. Further, it should be noted that H.264/AVC is capable of substantially superior rate-distortion performance when its advanced coding modes are employed, however, at the prohibitively increased cost of the computational complexity. It is anticipated that WBTC may also benefit from many of the same advanced coding techniques also.
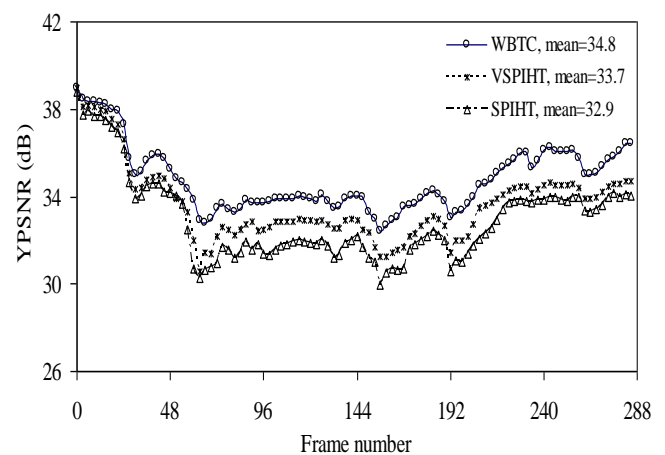
(a)
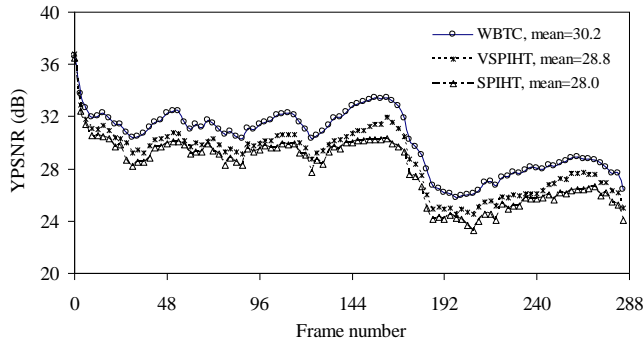
(b)

Fig 3 Frame-by-frame luminance PSNR (dB) comparisons at 30 kbps for QCIF sequences (a) Mother-Daughter and (b) Carphone.

Table 1: Average PSNR (in db Comparison) of Different Video Codecs

Bitrate (kbps)	SPIHT			H.264			WBTC		
	Y	Y	U	Y	U	V	Y	U	V
Mother-Daughter (QCIF)									
30	34.6	40.7	41.2	35.5	42.0	42.4	35.2	41.0	41.0
60	37.3	41.9	42.4	38.2	43.0	43.5	37.8	42.4	42.8
180	43.1	45.1	45.7	44.0	46.6	47.0	43.5	45.8	46.3
Akiyo (CIF)									
300	40.4	45.5	46.3	41.6	45.6	46.3	41.0	46.0	46.8
600	43.2	46.9	47.8	44.3	47.3	48.2	43.7	47.5	48.2
900	44.8	47.8	48.6	45.8	48.6	49.2	45.2	48.2	49.2

The reason for the superior performance of WBTC-based video coder is due to the better exploitation of intra-band correlations in the form of zeroblock. Since residual frames have low variance of pixel intensities as compared to the natural images, which will also be reflected among the wavelet coefficients. As a result, in wavelet transformed residual frames, the intra-subband correlation is more dominant than the inter-subband correlation. Also, the proposed WBTC algorithm results in better aggregation of zero-trees as well.

Finally, the computational complexity of different video coders is assessed by measuring the run times of the entire coding process. Table 2 summarizes these timings (in seconds)

for encoding Mother-Daughter (QCIF) and Akio (CIF) sequences at three different bitrates. Only encoding time is compared as encoder includes local decoder. It is observed that in general WBTC based video coder requires the least execution time, whereas H.264/AVC requires the maximum execution time. In particular, WBTC code a video sequence 11-15 times faster than the H.264/AVC. Similar trends were also observed for other sequences. The low complexity and better coding performance features of WBTC is quite attractive for real-time video communication using handheld portable devices having limited processing power.

Table 2: Execution Time (in seconds) Comparisons of the Coding Process

Bitrate (kbps)	Coding time (sec.)			
	SPIHT	VSPIHT	H.264	WBTC
Mother-Daughter (QCIF)				
30	1.7	1.9	20.3	1.5
60	1.7	2.0	21.1	1.6
180	2.3	2.6	21.8	2.0
Akiyo (CIF)				
300	6.2	7.2	80.0	5.5
600	6.8	8.1	81.4	6.2
900	7.4	9.1	84.7	6.9

5. CONCLUSIONS

This paper presents a hybrid wavelet-based video coder. The heart of the video coder is a new technique to efficiently encode the wavelet coefficients of I- and P-frames. The video coder presented in this paper outperforms state-of-the-art wavelet-based video coders. Though it has slightly inferior/comparable coding performance compared to H.264/AVC, its execution time is 11-15 times lesser than H.264/AVC. Therefore, the proposed video coder is very attractive for wireless video sensor network applications, where processing power and battery life are scares resources.

REFERENCES

[1] M. E. Al-Mualla, C. N. Canagarajah, and D. R. Bull, Video Coding for Mobile Communications: Efficiency, Complexity, and Resilience. San Diego: Academic Press, 2002.

[2] T. Sikora, "Trends and perspectives in image and video coding," Proc. of the IEEE, vol. 93, pp. 06-17, Jan. 2005.

[3] Advanced Video Coding for Generic Audiovisual Services, ITU-T Recommendation H.264, May 2003.

[4] M. Ghanbari, Standard Codecs: Image Compression to Advanced Video Coding. London, UK: IEE, 2003.

[5] J. E. Fowler and B. Pesquet-Popescu, An overview on wavelets in source coding, communications, and

networks, EURASIP Journal on Image and Video Processing, 2007, (Oct. 2007) 1-27.

[6] B. J. Kim, Z. Xiang, and W. A. Pearlman, Low bit-rate scalable video coding with 3-D set partitioning in hierarchical trees (3-D SPIHT), IEEE Trans. on Circuits and Systems for Video Technology, 10, (December 2000) 1374-1387.

[7] S.-T. Hsiang and J. Woods, Embedded video coding using invertible motion compensated 3-D subband/wavelet filter bank, Signal Processing: Image Communication, 16, (May 2001) 705-724.

[8] N. Adami, A. Signoroni, and R. Leonardi, State-of-the-art and trends in scalable video compression with wavelet-based approaches, IEEE Trans. Circuits and Systems for Video Technology, 17, (Sept. 2007) 1238-1255.

[9] L. Zhang, D. Wang, and André Vincent, Decoupled 3-D zerotree structure for wavelet-based video coding, IEEE Trans. on Broadcasting, 54, (Sept. 2008), 430-436.

[10] S. A. Martucci, I. Sodagar, T. Chiang, and Y. Zhang, "A zerotree wavelet video coder," IEEE Trans. on Circuits and Systems for Video Technology, vol. 7, pp. 109–118, Feb. 1997.

[11] Q. Wang and M. Ghanbari, Scalable coding of very high resolution video using virtual zero tree, IEEE Trans. on Circuits and Systems for Video Technology, 7, (October 1997) 719-727.

[12] D. Marpe and H. L. Cycon, "Very low bitrate video coding using wavelet-based techniques," IEEE Trans. on Circuits and Systems for Video Technology, 9, (Feb. 1999) 85-94.

[13] K. Shen and E. J. Delp, Wavelet based rate scalable video compression, IEEE Trans. on Circuits and Systems for Video Technology, 9, (Feb. 1999) 109-122.

[14] E. Khan and M. Ghanbari, An efficient and scalable low bit-rate video coding with virtual SPIHT, Signal Processing: Image Communication, 19, (April 2004) 267-283.

[15] J. Vass, B.-B. Chai, K. Palaniappan, and X. Zhuang, Significance-linked connected component analysis for very low bit rate wavelet video coding, IEEE Trans. on Circuits and Systems for Video Technology, 9, (June 1999) 630-647.

[16] D. Marpe, T. Wiegand, and H. L. Cycon, Design of a highly efficient wavelet-based video coding scheme, Visual Communications and Image Processing (VICIP 2002): Proc. of SPIE, 4671, (Jan. 2002)1133-1141.

[17] A. A. Moinuddin, E. Khan and M. Ghanbari, "Wavelet-based low bitrate video coding using block-tree approach," IET Int. Conf. on Visual Information Engineering (VIE 2006), Bangalore, India, Sept. 26-28, 2006, pp. 287-291.

[18] M. T. Orchad, G. J. Sullivan. "Overlapped block motion compensation: an estimation-theoretic approach", IEEE Trans. on Image Proc., 3, pp. 693–699, (1994).

[19] M. Antonini, M. M. Barlaud, P. Mathieu, and I. Daubechies, "Image coding using wavelet transform," IEEE Trans. on Image Processing, vol. 1, pp. 205-220, April 1992.

[20] J. D. Villasenor, B. Belzer, and J. Liao, Wavelet filter evaluation for image compression, IEEE Trans. on Image Processing, vol. 4, pp. 1053-1060, August 1995.

[21] K. K. Lin, and R. M. Gray, "Wavelet video coding with dependent optimization," IEEE Trans. on Circuits and Systems for Video Technology, 14, (April 2004) 542-554.

[22] E. Khan, Efficient and robust wavelet based image/video coding techniques, PhD Thesis, Dept. of Electronic Systems Engineering, University of Essex, UK, 2003.

[23] Available at http://iphome.hhi.de/suehring/tml/