An Efficient Spatial Scalable Video Encoder Based on DWT

Saoussen Cheikhrouhou[1], Yousra Ben Jemaa[2], Mohamed Ali Ben Ayed[1], and Nouri Masmoudi[1]

[1] Laboratory of Electronics and Information Technology
Sfax National School of Engineering, BP W 3038, Sfax, Tunisia
[2] Signal and System Unit, ENIT, Tunisia
nouri.masmoudi@enis.rnu.tn

Abstract. Multimedia applications reach a dramatic increase in nowadays life. Therefore, video coding scalability is more and more desirable. Spatial scalability allows the adaptation of the bit-stream to end users as well as varying terminal capabilities. This paper presents a new Discrete Wavelet Transform (DWT) based coding scheme, which offers spatial scalability from one hand and ensures a simple hardware implementation compared to the other predictive scalable compression techniques from the other. Experimental results have shown that our new compression scheme offers also good compression efficiency. Indeed, it achieves up to 61\% of bit-rate saving for Common Intermediate Format (CIF) sequences using lossy compression and up to 57\% using lossless compression compared with the Motion JPEG2000 standard and 64\% compared with a DWT-coder based on Embedded Zero tree Wavelet (EZW) entropy encoder.

Keywords: video conference, temporal redundancy, spatial scalability, discrete wavelet transform dwt, tier1 and tier2 entropy coding.

1 Introduction

Advances in video coding technology are enabling an increasing number of video applications ranging from multimedia messaging, and video conferencing to standard and high-definition TV broadcasting. Furthermore, video content is delivered to a variety of decoding devices with heterogeneous display and computational capabilities [1]. In these heterogeneous environments, flexible adaptation of once-encoded content is desirable. This heterogeneity warranted the development of many scalable video codecs [2]. Our main interest is to study naturally scalable video codecs based mainly on DWT compression since the DWT-based video coding has been a fast-expanding research field in the last years[3] [4]. This paper is organized as follows. Section 2 enumerates some limitations of the existing DWT-based scalable video codecs. In section 3, we present the details of our new DWT-based video codec. The experimental results and analysis of the results are presented in section 4. Finally, concluding remarks and future work are presented in section 5.

V.V. Das, E. Ariwa, and S.B. Rahayu (Eds.): SPIT 2011, LNICST 62, pp. 70–75, 2012.

2 Limitation of the Existing Scalable Video Compression Schemes

There are multitudes of related research in the field of scalable video coding. The Joint Video Team of the ITU-T VCEG and the ISO/IEC MPEG has recently standardized the Scalable Video Coding extension (SVC) of the H.264/MPEG-4 Advanced Video Coding (AVC) standard (H.264/AVC) which is the latest amendment for this successful specification.SVC enables the transmission and decoding of partial bit streams to provide video services with lower temporal or spatial resolutions or reduced fidelity. SVC is highly efficient but this efficiency is on the expense of the coding/decoding compression complexity scheme since the scalability function in the SVC is added to another compression scheme, the AVC, which is initially very complex [2].

Thus, we have focused our studies on DWT-based video codecs which offer naturally scalable compression schemes.The DWT is used to ensure scalability in the JPEG 2000 image compression standard. Embedded ZeroTree Wavelet "EZW" [5] and Set Partitioning in Hierarchical Trees "SPIHT" [6] are DWT-based entropy coders. In this context, many video compression schemes are proposed in the literature. These methods, however, suffer from a low performance problem since they do not include any motion compensation feedback loop [7].A commonly used alternative in the literature is to extend the 2D DWT-based algorithms to the time component [8].In [9], three dimensional DWT are used with the "SPIHT" coding, extended into the temporal dimension, and used without motion compensation.

We should notice that, many other approaches have combined the three dimensional DWT and motion compensation [10] [11] [12].However, temporal filtering always produces very clear disturbing ghosting artifacts, also called the drift error, especially in the low-pass temporal subband.

It is clear that the challenge now is how to exploit the motion within the spatio-temporal transform.It is interesting to mention that none of the related work cited above is based on a pre existing standardized compression scheme. On the other side, Shih-Ta Hsiang [13], has proposed a new intra-frame dyadic spatial scalable coding framework based on a subband/wavelet coding approach for MPEG-4 AVC/H.264 scalable video coding (SVC). It is the first attempt in the literature to join the subband filter banks with the traditional macroblock and DCT based video coding system.

Although this approach has achieved a bit rate saving of 13.75\%, compared with the standard H.264/MPEG-4 SVC, it is considered to be very complex since it is based originally on a complex compression scheme.As mentioned before, the two most annoying problems are even the drift error due to three dimensional wavelet transform or the high complexity of the compression scheme. Thus, we propose to ameliorate the Motion JPEG2000 which is a relatively simple DWT-based video compression standard. Indeed, it compresses frames independently each other without utilizing any interframe redundancy. Consequently, in this paper, we propose to code not the original frame but the difference between the source (I) frame and the reconstructed (I-1) frame along with the heritage of many features from Motion JPEG2000 compression schemes especially its entropy coder Tier1 and Tier2 [14] [15].

3 Structure of the Proposed DWT-Based Video Codec

This video compression scheme (Fig. 1) aims to offer spatial scalable coding. Its main idea is to eliminate the temporal redundancy by coding only the first frame of the video followed by the residuals resulting from the differential frames. In fact, these residual frames are obtained by subtracting the source image (I) from the reconstructed (I-1) which is generated by the encoder similar to the one in the decoder side by undergoing the inverse quantization and inverse DWT on the close loop path. The DWT will ensure the spatial scalability and thus, adapt the encoded stream to the corresponding terminal display. Indeed, the same transform as that used in Motion JPEG2000 is used to process the first image of the video. For the residual frames, a new transform and a new disposition of the different sub-bands for the multi-resolution analysis are used and are detailed in [16].The wavelet coefficients obtained will be quantized based on a perceptual quantization matrix "DWTune"[17] which differs from the quantization method of the Motion JPEG2000 standard. Finally, the quantized DWT coefficients will be entropy coded similar to the Motion JPEG 2000's entropy coder based on Tier1 and Tier2 [14] [15].The same process will be conducted for the luminance (Y) and the two chrominance components (Cb, Cr) for each frame.

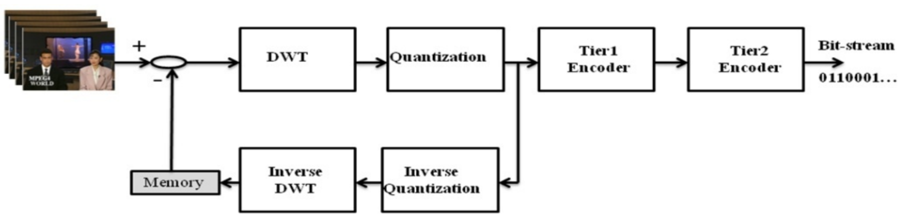

Fig. 1. Structure of the proposed DWT-based video codec

4 Experimental Results and Discussion

4.1 Performances Comparison with the Motion JPEG2000 Standard

During these experiments, the rate controller block for the Motion JPEG2000 standard is disabled in order to ensure more objective comparison and inhibit any other interference.Experimental results show that the proposed approach presents a very clear improvement in terms of PSNR for the sequences akiyo and silent compared to the Motion JPEG2000 standard. Nevertheless, we note that the PSNR of the foreman sequence obtained by the new codec is lower than that obtained by the standard Motion JPEG2000. This proves that our proposed encoder goes along with video conference scenes more than any other content. In fact, Akiyo and Silent sequences present less motion than Foreman sequence.Figures 2, 3 and 4 depict the comparative results for the different video sequences in a lossy mode. Table 1 shows the comparative results for the different video sequences in lossless mode.

Table 1. Bit Rate Comparaison Between Motion Jpeg2000 And The Proposed Codec For The Lossless Compression For The Sequences Foreman.Cif And Akiyo.Cif And News.Cif

Sequence	Bit rate (Kbits/s)	
	Motion JPEG2000	Proposed codec
Akiyo.cif	10479,33	4488,66
News.cif	12038,66	6534,66
Foreman.cif	13626,66	14043,33

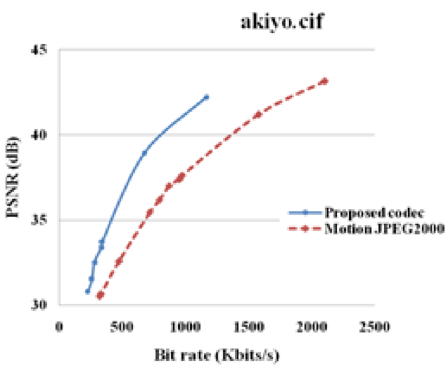

Fig. 2. Rate distortion curve comparison betweenMotion JPEG2000 and the proposed codec for thesequence akiyo.cif

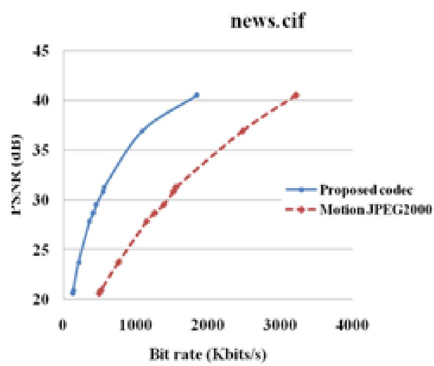

Fig. 3. Rate distortion curve comparison betweenMotion JPEG2000 and the proposed codec for the sequence news.cif

4.2 Performances Comparison with a DWT-Coder Based EZW

On the other hand, our proposed coding scheme is compared with a DWT-codec based on EZW [16]. It is very similar to our new coder with one major difference: it uses the EZW as an entropy coder instead of the Tier1 and Tier2.

As shown by figure 5, the proposed video codec outperforms the EZW-based coder. In fact, more than 50\% reduction in bit rate has been obtained for the same PSNR.

Table 2. Bit rate comparaison between the proposed codec and a scalable dwt-codec based on ezw for the lossless compression for the sequence foreman.cif

Codec	Bit rate (Kbits/s)
EZW-based codec	23203,30
Proposed codec	14043,33

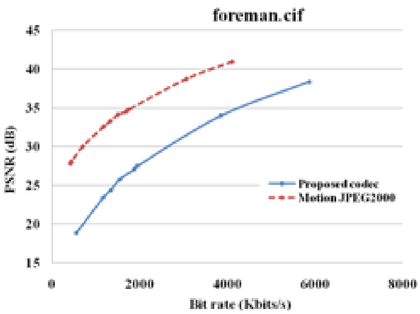

Fig. 4. Rate distortion curve comparison between Motion JPEG2000 and the proposed codec for the sequence foreman.cif

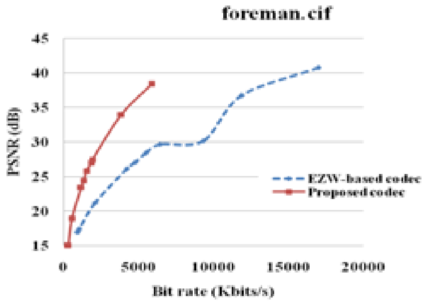

Fig. 5. Rate distortion curve comparison between theproposed codec and a scalable DWT-codec based EZWfor the sequence foreman.cif

5 Conclusion and Future Work

In this paper, a new DWT-based video codec which aims to offer spatial scalable coding is proposed. In fact, this video codec improves the existing predictive coding standards by reducing dramatically the complexity of implementation.The new compression scheme induces a significant improvement for objective quality tests compared to the Motion JPEG2000 standard. Indeed, it achieves up to 61\% of bit-rate saving for Common Intermediate Format (CIF) sequences using lossy compression and up to 57\% using lossless compression compared with the Motion JPEG2000 standard. Furthermore, the proposed video codec outperforms the EZW-based coder by more than 50\% in terms of bit rate reduction.As perspectives, we propose to add an inter-frequency prediction block exploiting the inter-band redundancies. Likewise, better results are promised by implementing an Adaptive Loop Filter (ALF) in order to reduce the gap between the source frame (I) and the reconstructed frame (I-1) from the statistical point of view prior to the determination of the residual frame.

References

1. Man, H., Kossentini, F., Smith, M.J.T.: A family of efficient and channel error resilient wavelet/subband image coders. IEEE Transactions on Circuits and Systems for Video Technology 9(1), 95–108 (1999)
2. Schwarz, H., Marpe, D., Wiegand, T.: Overview of the scalable video coding extension of the h.264/avc standard. IEEE Transactions on Circuits and Systems for Video Technology 17(9), 1103–1120 (2007)
3. Marpe, D., Cycon, H.L.: Very low bit-rate video coding using wavelet-based techniques. IEEE Transactions on Circuits and Systems for Video Technology 9(1), 85–94 (1999)

4. Shen, K., Delp, E.J.: Wavelet based rate scalable video compression. IEEE Transactions on Circuits and Systems for Video Technology 9(1), 109–122 (1999)
5. Shapiro, J.M.: Embedded image coding using zerotrees of wavelets coefficients. IEEE Transactions on Signal Processing 41(12), 3445–3462 (1993)
6. Said, A., Pearlman, W.A.: 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 (1996)
7. Choupani, R., Wong, S., Tolun, M.R.: A drift-reduced hierarchical wavelet coding scheme for scalable video transmissions. In: First International Conference on Advances in Multimedia, pp. 68–73 (2009)
8. Karlsson, G., Vetterli, M.: Three-dimensional subband coding of video. In: Internatioanl Conference on Acoustics Speech and Signal Processing, vol. 2, pp. 1100–1103 (April 1988)
9. Kim, B.J., Pearlman, W.A.: An embedded wavelet video coder using three-dimensional set partitioning in hierarchical trees (spiht). In: IEEE Data Compression Conference, pp. 251–260 (March 1997)
10. Asbun, E., Salama, P., Shen, K., Delp, E.J.: Very low bit rate wavelet-based scalable video compression. In: International Conference on Image Processing, vol. 3, pp. 948–952 (1998)
11. Marpe, D., Cycon, H.L.: Very low bit-rate video coding using wavelet-based techniques. IEEE Transactions on Circuits and Systems for Video Technology 9(1), 85–94 (1999)
12. Mehrseresht, N., Taubman, D.: An efficient content-adaptive motion-compensated 3-d dwt with enhanced spatial and temporal scalability. IEEE Transactions on Image Processing 15(6), 1397–1412 (2006)
13. Hsiang, S.: Intra-frame dyadic spatial scalable coding based on a subband/wavelet framework for mpeg-4 avc/h.264 scalable video coding. In: IEEE International Conference on Image Processing, pp. 73–76 (November 2007)
14. Iso/iec, jtc1/sc29 and wg1, jpeg 2000 final committee draft version 1.0 fcd15444-1, jpeg 2000 image coding system (March 2000)
15. Taubman, D., Ordentlich, E., Weinberger, M., Seroussi, G., Ueno, I., Ono, F.: Embedded bloc coding in jpeg 2000. In: IEEE International Conference on Image Processing, vol. 2, pp. 33–36 (September 2000)
16. Cheikhrouhou, S., Ben Jemaa, Y., Samet, A., Ben Ayed, M.A., Masmoudi, N.: Toward an optimal residual frame coding for dwt based video codec. In: IEEE International Conference on Electronics and Systems, pp. 876–879 (December 2009)
17. Ben Ayed, M.A., Samet, A., Loulou, M., Masmoudi, N.: Dwtune: a technique for visual optimization of dwt quantization matrix for continuous still images. In: ISSPIT 2002, marroco (2002)