AN EFFICIENT APPROACH OF FAST MOTION ESTIMATION AND COMPENSATION IN WAVELET DOMAIN VIDEO COMPRESSION

Weiting Cai and Malek Adjouadi
Center for Advanced Technology and Education,
Department of Electrical & Computer Engineering, Florida International University
weitingcai@yahoo.com, adjouadi@fiu.edu

ABSTRACT

Motion estimation and compensation combined with wavelet transformation has been beneficially applied to improve video compression efficiency and accuracy. Based on the similar motion structures between the decomposed subimages, wavelet domain video compression can be improved by means of the simplifications in motion field refinement as well as in motion vector representation. Although the utilization of the motion field correlations between the same orientational subbands can achieve better compression performance than that between the approximation subbands, the computational load has to be traded off. Since most of the energies are concentrated in the low resolution subbands while decreased in the high resolution subbands, by considering both the accuracy and complexity, an efficient approach, called level-refined motion estimation and subband compensation (LRSC) method is proposed. It realizes low-entropy coding in the subbands, while keeping low computational load in motion estimation, thus to achieve both temporal compression quality and computational simplicity.

KEYWORDS: Video compression, Motion estimation, Motion compensation, Wavelet transformation, Entropy coding.

I. INTRODUCTION

By the fact that the wavelet transformed coefficients still contain spatial domain information, motion structure within the subbands at the same scale and between different scales are highly correlated. Motion estimation and motion compensation (MEMC) can be implemented globally but not fully, using a multiresolution description of the video frames to obtain initial and refined search inside wavelet domain by building up the frame pyramids. It is expected that wavelet domain video compression can achieve good performance, especially at low bit rates [1], by the experiment results of image compressions using wavelet transform. Wavelet-based video compression has been studied in some papers by applying the search strategies [2-8], complex-valued wavelet transform [9], overcomplete discrete wavelet transform [10] and wavelet packets [11].

The refinement mechanism of wavelet motion estimation method is based on the motion similarities between the low frequency and high frequency subbands. The correlations between the LL subframes of different levels are actually low frequency signals with more details; And the correlations between the approximation subband and the detail subbands are actually adjacent bandpass signals. Therefore, the motion estimation procedure can be simplified so as to save the computational load and motion information bandwidth. Furthermore, appropriate motion compensation in the high frequency subbands may obtain efficient coding when the entropy is getting lower, as the proposed level-refined motion estimation and subband compensation (LRSC) method.

The layout of this paper can be described as follows: section II is the outline of wavelet domain video compression including the introductions of level-refined and subband-refined motion estimation methods. Section III presents the issue of computational complexity by these algorithms. Section IV introduces the LRSC method, an efficient MEMC approach. Section V focuses on the implementations of the entropy-coded wavelet domain video compression. Section VI shows the resulting compression performances in terms of computational load and reconstruction quality. Section VII provides the conclusions by the results attained in this work.

II. WAVELET DOMAIN MOTION ESTIMATION METHODS

Wavelet domain motion estimation is processed by searching the best matching block between the current subframe and the corresponding reference subframe, from coarse to fine resolutions hierarchically. An N-level decomposed multiresolution motion estimation can be refined by:

$$\begin{aligned} &LL_i \to LL_{i-1} \qquad (i = 2 \sim N) \\ &LL_i \to HL_i, LH_i, HH_i \qquad (i = 1 \sim N) \end{aligned} \tag{1}$$

where the motion vectors in the right side subband are taking refinement references from the left side.

In level-refined (LR) motion estimation method, the motion vectors of LL subframe in a higher resolution level are refined by those from the lower resolution LL subframe after doubling, and represent the motion vectors for the three detail subbands at the same level:

$$\begin{aligned} &LL_i \to LL_{i-1} \qquad (i = 2 \sim N) \\ &LL_i = HL_i = LH_i = HH_i \qquad (i = 1 \sim N) \end{aligned} \tag{2}$$

Because of the similar motion activities in the approximation and detailed subbands, motion fields can also be estimated by subband-refined (SR) motion estimation method, in which motion estimation is developed by the correlated motion structures between the successive frequency subbands in horizontal, vertical and diagonal directions:

$$\begin{aligned} &LL_N \to HL_N, LH_N, HH_N \\ &HL_i \to HL_{i-1},\ LH_i \to LH_{i-1},\ HH_i \to HH_{i-1} \quad (i = 2 \sim N) \end{aligned} \tag{3}$$

III. COMPUTATIONAL COMPLEXITY

Although the block size relative to the coverage of the full frame remains the same, the physical block size to compute the matching criterion function is halved. Assume the maximum displacement in the lowest resolution level is d_0-pixel using full search method after N-level decomposition, and the refinements in other subframes and levels are d_r-pixel, the equivalent search displacement d in frame resolution using wavelet domain search method is:

$$d_{WS} = 2^N d_0 + (2^{N+1} - 2) d_r \tag{4}$$

Since the block size is $\frac{N_1}{2^i} \times \frac{N_2}{2^i}$ at level i, the computational load of a current block in all levels for a full frame in size $N_1 \times N_2$ using WS method is:

$$L_{WS} = L_B \{ (\frac{1}{4})^N [(2d_0 + 1)^2 + 3(2d_r + 1)^2] + \frac{4}{3}[1 - (\frac{1}{4})^{N-1}](2d_r + 1)^2 \} \tag{5}$$

assuming that the refinements are performed in each of the subsequent levels and subbands, and

$$L_B \propto N_1 * N_2 \tag{6}$$

is the computational load of the matching criterion function for a full frame current block [12]. Usually, the refinement of motion vector is 1 pixel $(d_r = 1)$, owing to the full search method used in the lowest resolution level and the 2:1 decimation.

The equivalent search displacement in full frame resolution and the corresponding computational load of a current block in all levels for LR and SR methods can be expressed by:

$$d_{LR} = 2^N d_0 + \sum_{i=1}^{N-1} 2^i d_r = 2^N d_0 + (2^N - 2) d_r \tag{7}$$

$$L_{LR} = L_B \{ (\frac{1}{4})^N (2d_0 + 1)^2 + \frac{1}{3}[1 - (\frac{1}{4})^{N-1}](2d_r + 1)^2 \} \tag{8}$$

$$d_{SR} = 2^N d_0 + \sum_{i=1}^{N-1} 2^i d_r = 2^N d_0 + (2^N - 2) d_r \tag{9}$$

$$L_{SR} = L_B \{ (\frac{1}{4})^N (2d_0 + 1)^2 + [1 - (\frac{1}{4})^N](2d_r + 1)^2 \} \tag{10}$$

assuming that the motion vectors are available in each of the reference subbands.

IV. LEVEL-REFINED MOTION ESTIMATION AND SUBBAND COMPENSATION METHOD

From the level-refined motion estimation method, motion compensations in the detail HL, LH and HH subbands use the same positioned compensation blocks in order to save the motion information bandwidth and computational load of the motion field refinements. Would the intermode compression be more efficient than the intramode compression in the wavelet domain decomposed blocks? The answers are quite different for approximation and detail subband signals. Based on the compensated blocks in the subband-refined motion estimation method, the increased intrablocks after motion compensation indicates the energy distributions in the decomposed subbands. Most of the energies are concentrated in the low resolution approximation subband, while decreased much in the high resolution detail subbands, and even possible to be less than the differential blocks after motion estimation. If the energy comparison between the current block and the differential block is computed before motion compensation by the same unrefined motion vectors, lower entropy coding can be achieved. This is the level-refined motion estimation and subband compensation (LRSC) method. If the energy of the differential block is larger than the current block, the compression mode bit is set as 1 and intermode compression is coded with the corresponding motion vectors and differential data. Otherwise, the compression mode bit is set as 0 to represent intramode compression on the current block. If a current block in the LL subframe is in intramode compression without motion vectors found, the assumption is that all the detail subbands at the same level are using intramode compression, so the compression mode bit is set as 0 without motion compensation in the decoder. The intramode compression of the current block in the detail subbands does not change the motion field reference for the next level refinements, but it will in the subband-refined method and cost the computational load of motion estimation by reinitializing full search around the stationary center of the current block.

Based on the fact that only one motion vector array is shared by all the subbands in one level, computational loads of LRSC and LR methods are identical in terms of motion estimation, the most exhausted part in the encoder. The additional computation from LRSC method is the comparison of the energies contained in the current block and the differential block after motion estimation, which is very insignificant. Furthermore, the motion vector refinement in each level as well as the differential motion vector coding remain the same as in LR method.

Besides the motion vector array to specify the locations of the best matching blocks for the intermode compressed blocks, LRSC method contains an extra compression mode bit array to distinguish intramode or intermode compression for each detail subband. Only one bit for each current block in each level is needed, this bandwidth is quite small comparing to the block of frame data. On one hand, LRSC method saves the refinement and description of the motion vector array for each detail subband, by sharing the same motion vector array at one scale level as in LR method; On the other hand, since each current block from LRSC method has lower energy than that from the LR method, data distribution is more centralized to represent lower entropy coding; Therefore, improved compression efficiency is obtained from the LRSC method.

V. IMPLEMENTATION

Discrete wavelet transform (DWT) is applied on the current frame and the reference frame before motion estimations between the subframe pyramids [2, 13]. Minimum absolute difference (MAD) is used as the matching criterion at the lowest resolution level as well as the reinitialized current blocks when the best matching block from the last level is not found. Dynamic Huffman coding is performed, to deal with the uncertain distribution of the wavelet differential frame. The bit stream is ordered from motions to coefficients, from luminance to chrominances, from approximation to details and from low to high resolution levels.

If the best matching blocks are found in both the current and its reference subband, the differential motion vectors, which include the horizontal and vertical refinements are coded; If the best matching block is found in the current level but not in its reference subband, real motion vectors are coded. In both cases, compression mode bit of 1 is assigned in front of the motion vector and the quantized differential data to represent intermode compression. Define level 2 be in half resolution of level 1, the first decomposition level of the full frame. Table 1 summarizes the motion vector refinements as well as the differential coding

Table 1. Motion vector refinement (left) and differential coding (right) for the subbands in wavelet domain motion estimation methods

	LL_2		HL_2, LH_2, HH_2		LL_1		HL_1		LH_1		HH_1	
WS	N/A	0	LL_2	LL_2	LL_2	$2LL_2$	LL_1	LL_1	LL_1	LL_1	LL_1	LL_1
LRSC, LR	N/A	0	LL_2	N/A	LL_2	$2LL_2$	LL_1	N/A	LL_1	N/A	LL_1	N/A
SR	N/A	0	LL_2	LL_2	N/A	N/A	HL_2	$2HL_1$	LH_2	$2LH_2$	HH_2	$2HH_2$

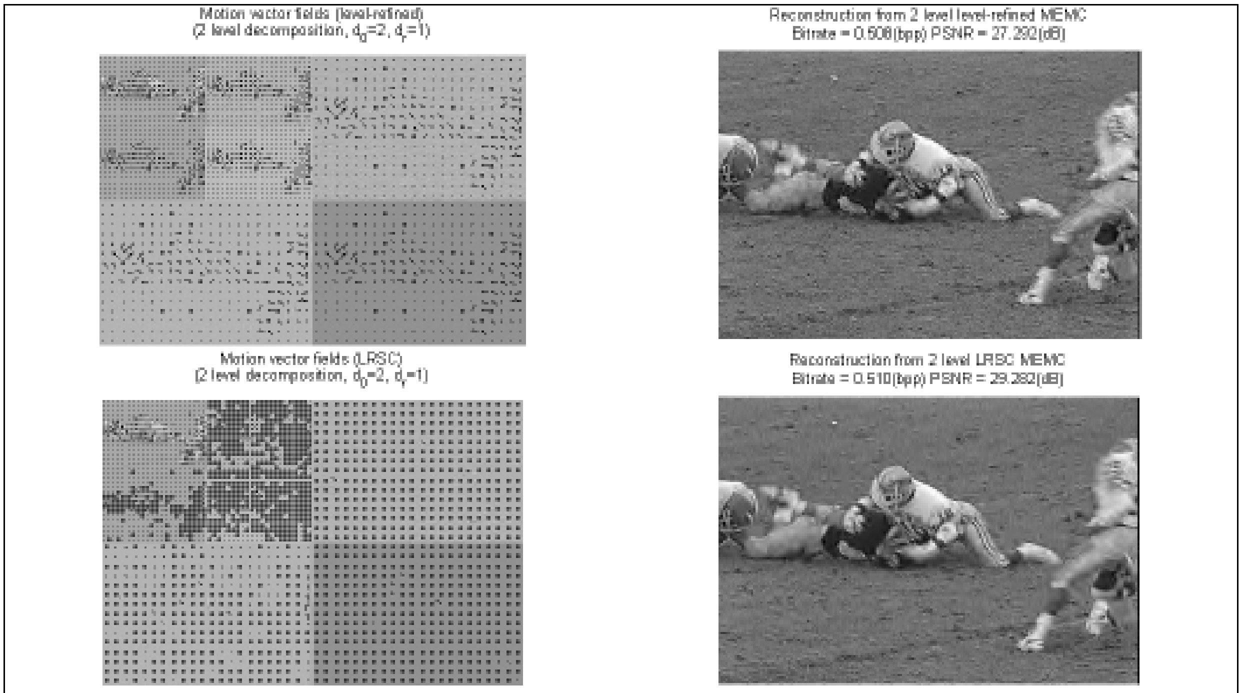

Figure 1. Wavelet domain compression using level-refined and LRSC MEMC methods for "football"

relationships between two adjacent decomposed subbands, assuming that the best matching block is found from the reference subband, otherwise full search algorithm is used instead.

VI. RESULTS

The compression performances of the wavelet domain motion estimation and compensation methods are evaluated by the reference-current frame pairs of the test video sequences. The experiment takes the 2nd frame from "football" as the current frame and uses the 1st frame as the reference. Multiresolution motion estimation is performed on the subframe pairs using 2-pixel full search in the lowest resolution level and 1-pixel refinements after 2-level decomposition. The results from the LR method as well as the LRSC method are compared in Figure 1 by illustrations in each row respectively. The motion vectors of the intermode compressed blocks are drawn after the best matching block is found, otherwise a square is marked for an intramode compressed block. The motion fields of the subbands at the same level are identical in LR method since one motion vector array from the approximation subband are shared for all the detail subbands, but modified by the LRSC method to realize most of the possible intramode compressed blocks. The reconstructed frame from LRSC method achieves 29.3dB in PSNR at bit rate of 0.5bpp after 50 times compression, which is 2dB above that from the LR method. The low bit rate compressions using the WS, LR, SR and LRSC methods are compared in Figure 2 for sequence "carphone", in which the 5th frame is the current processing frame and the 1st frame is the reference frame, as the interframe compression pair.

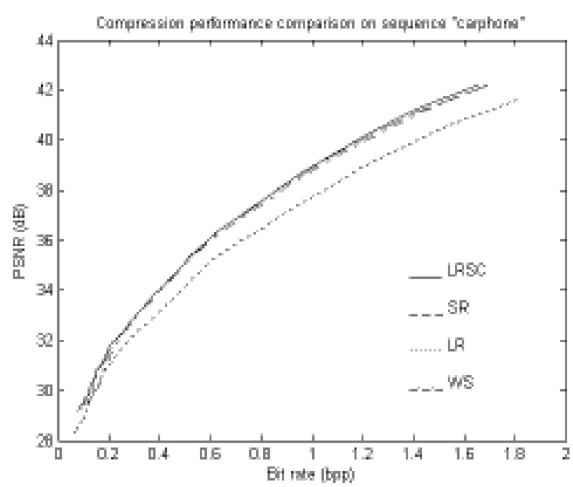

Figure 2. Wavelet domain MEMC methods for "carphone"

Besides compression quality, another important issue as mentioned in section III is the computational complexity of motion estimation. Figure 3 illustrates the computational load L including all the decomposed levels for a current block with respect to the search displacement d using WS, LR, SR and LRSC methods. A 2-level decomposition with $d_r = 1$ pixel refinement is applied and the evaluations are in unit of L_B. It shows that the LR and LRSC methods contain only $1/3$ of the computational complexity in WS and SR methods.

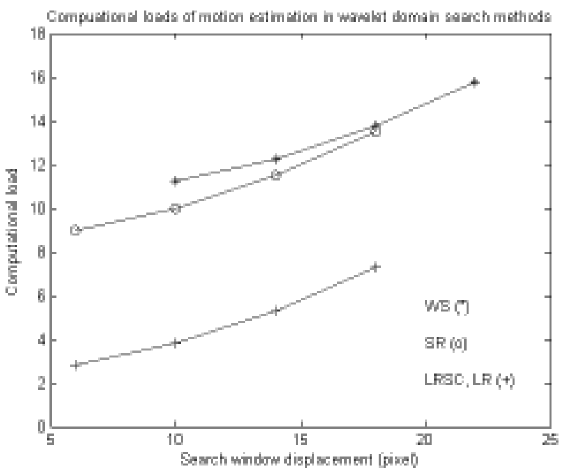

Figure 3. Computational load of wavelet domain MEMC methods

The results illustrate the improved performance of the LRSC method than the other wavelet domain MEMC methods by the low computational load and motion bandwidth as the level-refined method, as well as the good reconstruction quality as the subband-refined method.

VII. CONCLUSIONS

The efficiency obtained by LRSC method is based on the similar motion structures and compact energy distributions in the wavelet decomposed subbands. Because of the strong motion correlations between the adjacent frequency passbands, motion vector fields from the level-refined and subband-refined methods simplify the wavelet domain motion estimation and compensation procedure, as compared by the computational load formulations. The energy distributions and the increased intramode compressed blocks in the decomposed subbands indicate the energy contained in the high frequency subbands, which is even lower than the differential error blocks after motion estimation. The proposed level-refined motion estimation and subband compensation method applies selective motion compensation in the detail subbands using the estimated motion vectors from the approximation subband. By comparing the compression performances, LRSC method obtains close or better signal to noise ratio than the WS and SR methods, which is much better than the LR method, while the motion estimation process is as fast as the LR method. Therefore, referring to reconstruction quality and computational load as the main factors of performance measurements, LRSC method is the most efficient technique in the wavelet domain video compressions discussed in this research work.

REFERENCES

[1] Detlev Marpe and Hans L. Cycon, "Very Low Bit-Rate Video Coding Using Wavelet-Based Techniques", *IEEE Transactions on Circuits and Systems for Video Technology*, Vol. 9, No. 1, Feb. 1999.

[2] Sohail Zafar, Ya-Qin Zhang and Bihan Jabbari, "Multiscale Video Representation Using Multiresolution Motion Compensation and Wavelet Decomposition", *IEEE Journal on Selected Areas in Communications*, Vol. 11, No. 1, Jan. 1993.

[3] Pankaj N. Topiwala, *Wavelet Image and Video Compression,* Kluwer Academic Publishers, 1998.

[4] Mi-Sook Chung, Suk Kang, Jae-Ho Choi and Hoon-Sung Kwak, "New Multiresolution Motion Estimation for Wavelet Transform Video Coding", *Proceedings of IEEE Asia Pacific Conference on Circuits and System '96*, Nov. 18~21, 1996, Seoul, Korea, pp. 10-13.

[5] Suhuai Luo, "On the Improvement of Motion Compensation in Wavelet Transform Video Coding", *1997 IEEE TENCON – Speech and Image Technologies for Computing and Telecommunications*, pp. 735-738.

[6] El-Sayed Youssef, Ehab Sabry and Mohamed Khedr, "Hybrid Adaptive Motion Search with Wavelet/Wavelet Packet Based Motion Residual Video Coding", *Seventeenth National Radio Science Conference*, Feb. 22-24, 2000, Minufiya University, Egypt, pp. C2 1-11.

[7] G. Van der Auwera, A. Munteanu, G. Lafruit and J. Cornelis, "Video Coding Based on Motion Estimation in the Wavelet Detail Images", *ICASSP '98*, pp. 2801-2804.

[8] Seongman Kim and Seunghyeon Rhee, "Interframe Coding Using Two-Stage Variable Block-Size Multiresolution Motion Estimation and Wavelet Decomposition", *IEEE Transactions on Circuits and Systems for Video Technology*, Vol. 8, No. 4, Aug. 1998.

[9] Julian Magarey and Nick Kingsbury, "Motion Estimation Using Complex Wavelets", *IEEE International Conference on Acoustics, Speech and Signal Processing*, Vol. 4, pp.2371-2374, 1996.

[10] Radu Zaciu, Claudiu Lamba, Constantin Burlacu and Gabriel Nicula, "Motion Estimation and Motion Compensation Using an Overcomplete Discrete Wavelet Transform", *ICIP96*, pp. 973-976.

[11] Michael B. Martin and Amy E. Bell, "New Image Compression Techniques Using Multiwavelets and Multiwavelet packets", *IEEE Transactions on Image Processing*, Vol. 10, No. 4, Apr. 2001.

[12] Luis Miguel Lopes Teixeira and Artur P. Alves, "Block Matching Algorithms in MPEG Video Coding", *Proceedings of ICPR '96*, pp. 934-938.

[13] Ya-Qin Zhang and Sohail Zafar, "Motion-Compensated Wavelet Transform Coding for Color Video Compression", *IEEE Transactions on Circuits and Systems for Video Technology*, Vol. 2, No. 3, Sep. 1992.

ACKNOWLEDGEMENTS

The authors gratefully acknowledge the support from the National Science Foundation under grant EIA-9906600 and the Office of Naval Research under grant N00014-99-1-0952.