Selective Multi-Resolution Motion Estimation Using Half-pixel Accuracy and Characteristics of Motion Vectors

Seong-Geun Kwon, Suk-Hwan Lee, Seung-Jin Lee, Jong-Won Lee, In-Sung Ha, and Kuhn-Il Lee
School of Electronic & Electrical Engineering, Kyungpook National University
1370 Sankyukdong Pukku, Taegu 702-701, Republic of Korea

Abstract

A new multi-resolution motion estimation (MRME) algorithm using half-pel accuracy motion estimation (HPAME) and characteristics of motion vectors in the baseband is proposed in this paper. Generally, MRME method in the WT (wavelet transform) domain needs the exact motion vectors in the baseband because those are used as initial motion vectors in higher frequency subbands. Therefore the proposed method uses the HPAME to estimate the displacement exactly in the baseband. Then the displacement in the higher frequency subbands is selectively estimated based on the characteristics of the motion vector in the baseband. That is, the displacements in the higher frequency subbands are estimated only for the blocks which have motion vectors in the half-pixel accuracy for the corresponding block in the baseband. This is based on the fact that the block having the half-pixel accuracy motion vector in the baseband has a high probability to have a small displacement for the corresponding block in the higher frequency subbands. In the proposed method, by using the HPAME in the baseband and selective motion estimation in the higher frequency subbands, we can obtain motion-compensated image of the similar quality with the conventional method although we reduce the computational amount and the bit rate considerably.

1. Introduction

Video compression is a crucial area of the digital multimedia technologies so that it is used to transmit and store the video sequence through the restricted channel in the Þeld of HDTV, VOD, and video conference. Motion estimation and compensation have an important role in elimination of the temporal redundancy which is present in a video sequence. The video compression standards such as H.261,[1] H.263,[2] and MPEG[3],[4] adopted the BMA (block matching algorithm) as the motion estimation algorithm.

In the BMA, each video sequence is divided into blocks and each block of the current frame is compared to the blocks of the reference frame in the vicinity of its search area, assuming that all the pixels in each block have the same moving activity. The BMA is so simple that its hardware implementation is easy and its execution speed is fast. But it is difficult to decide the proper block size and search area and the blocking artifact occur because of its block-based characteristic. In order to alleviate the blocking artifact, the hierarchical motion estimation algorithm using WT has been developed in recent years.[5]~[8] WT provides multiresolution and multifrequency expression of a signal with localization in both spatial and frequency. Motion activities at different layers of the pyramid structure are different but highly correlated since they actually characterize the same motion structure.

In MRME proposed by Zhang and Zafar,[5] the motion Þeld is first calculated for the lowest frequency subimage, called baseband. Then, motion vectors at higher frequeny subbands are refined using the motion information obtained at the baseband. Due to the nature of the multiresolution, the size of the block at different layers is varied according to the resolution. MRME is free from the blocking artifact due to the nature of global decomposition of WT and represents the smooth motion vector field because of hierarchical motion estimation. But the motion vectors at the baseband should be estimated exactly because those at the higher resolution are refined based on the global motion vector at the baseband and the baseband has most energy of the image. Also, it require a extensive computation and bit overhead due to the hierarchical motion estimation. Thus a new motion estimation algorithm which estimate exactly motion vector in the baseband and reduce the computation and bit overhead in the higher frequency subbands is needed.

We propose an efficient selective multiresolution motion estimation algorithm using the HPAME and characteristics of motion vectors at the baseband. Since the true displacements of the video sequence contents are not in the form of integer-pel accuracy, we adopt the HPAME for the exact motion vectors at the baseband. Based on the characteristics of motion vectors at the baseband, the selective motion estimation is employed at the higher freqyency subbands to reduce the computation and bit overhead. Because blocks which have motion vectors in the half-pixel unit don't move in integer grid unit, the blocks of the corresponding position at the higher frequency subbands have a large probability of having small displacement.

Computer experiment showed that the proposed algorithm represented considerable reduction in the computation overhead and the bit rate at similar image quality. It means that the proposed algorithm results in the significant improvements in coding performance compared with the MRME for video compression.

2. Proposed motion estimation algorithm

We propose a new motion estimation algorithm which reduces the calculation and bit overhead maintaining the image quality through the exact motion estimation in the baseband and the selective motion estimation in the higher frequency subbands. The baseband has most of the energy of the original image and its motion vectors are used as reference motion vectors in the higher frequency subbands. Therefore the accuracy of the motion vectors at the baseband have an important effect on the image quality.

In this paper, we apply the HPAME in the baseband for the exact motion vectors and estimate motion vectors selectively in the higher frequency subbands for the only blocks which need motion estimation. The selective motion estimatin is performed using the characteristics of the motion vectors at the baseband to reduce the calculation and bit overhead. That is, because the blocks which have motion vectors in the half-pixel unit at the baseband don't move in integer sample grid unit, the blocks of the corresponding position in the higher frequency subbands have a large probability of having small displacement. So the motion estimation in the higher frequency subbands is applied for the blocks which have the motion vectors in the half-pixel unit at the corresponding position in the baseband.

2.1. HPAME in the baseband

Since the true frame to frame displacements of the image content are not in the form of integer-pel accuracy, the HPAME[9]∼[11] is used for the improved motion estimation. Girod[9] tested the performance of the HPAME for the videophone and broadcast TV signal, and actually the HPAME is used in the MPEG-2 TM[10],[11] for the exact estimation.

Because the baseband has most of the energy of original image and motion vector at that is used as reference

Table I. The improvement quantity of MSE by HPAME in the baseband.

Sequence	MSE		Improvement quantity [%]
	FPAME	HPAME	
CLAIRE	56.97	43.29	20.41
FOOTBALL	1906.68	1470.17	22.89
TABLE TENNIS	341.43	260.07	23.82

motion vector in the higher frequency subbands, we apply the HPAME in the baseband for the exact motion estimation.

The HPAME is based on the fact that the true frame to frame displacements of the image contents is not always in the integer-pel accuracy. The relationship of the current and reference frame is represented by

$$I_i(x, y) = I_{i-1}(x + d_x, y + d_y) \quad (1)$$

where d_x and d_y is the displacement between two frames, actually

$$d_x = (m_x + v_x)d \quad (2)$$

$$d_y = (m_y + v_y)d \quad (3)$$

m_x and m_y is the integer-pel part and $v_x$and v_y half-pel part.

In HPAME, 8 search points is made by the bilinear interpolation around the motion vector having integer-pel accuracy and the final motion vector is determined as one of the 9 search points which matches best.

The performance of the HPAME in the baseband compared with the FPAME is represented in Table I. From this table, we conÞrmed that the HPAME reduced DFD (displaced frame difference) error compared with the FPAME.

2.2. Selective motion estimation in the high frequency subbands

From the previous section, we saw that prediction error is remarkably reduced by the HPAME in the baseband. While the HPAME reduce the prediction error, it usually introduces the calculation and bit overhead. In this paper, to solve this problem the selective motion estimation is proposed for the higher frequency subbands. In the proposed algorithm, the motion estimation at the higher frequency subbands is selectively performed. That is, only for the blocks having the motion vector in the half-pel accuracy at the corresponding position in the baseband, the motion estimation is performed. The moving characteristic of the blocks having motion vectors in the integer-pixel accuracy in the baseband is the sampling grid unit,

Table II. The improvement quantity of MSE by refinement in the higher frequency subbands.

Sequence	FPAME	HPAME
CLAIRE	11.04	18.32
FOOTBALL	13.79	30.86
TABLE TENNIS	13.92	24.08

Table III. The ratio of blocks having the full pixel motion vectors in the baseband.

Sequence	CLAIRE	FOOTBALL	TABLE TENNIS
Distribution ratio [%]	64.60	54.44	60.04

so the blocks at the corresponding position in the higher frequency subbands have a few probability to have small displacement. In contrary, the blocks having the motion vectors in the half-pixel accuracy in the baseband have a large probability to have small displacement in the higher frequency subbands. To prove this facts, the reduction of MSE by applying the motion estimation at the higher frequency subbands for the both case, the motion vector in the half-pel and integer-pel accuracy at the baseband, is shown in Table II. This table shows while the blocks in the integer-pel accuracy have small reduction of MSE, the blocks in the half-pel accuracy have large reduction of MSE. This result means that the degradation of the image quality is not appeared without motion estimation for the corresponding blocks in the higher frequency subbands because the blocks having the motion vector in integer-pixel accuracy don't represent the small displacement at the higher frequency subbands.

To reduce the calculation and bit overhead caused by the HPAME in the baseband, we adopted the selective motion estimation in the higher frequency subbands based on the characteristics of the motion vector in the baseband.

We can expect that the calculation and bit overhead are reduced at the similar prediction error compared with the MRME by the HPAME in the baseband and selective motion estimation in the higher frequency subbands.

Table IV. The comparison of computational quantity by MRME and the proposed method.

Sequence		CLAIRE	FOOT BALL	TABLE TENNIS
Computational quantity	MRME	3788928	12916800	12916800
	Proposed	2018428	7921573	6938134
Computational quantity reduction [%]		46.73	38.67	46.29

Table V. The bit rate and PSNR by the proposed method and MRME.

Sequence		CLAIRE	FOOT BALL	TABLE TENNIS
Bit rate [bpp]	MRME	0.1172	0.1172	0.1172
	Proposed	0.0591	0.0682	0.0596
Bit rate reduction [%]		49.53	41.74	49.10
PSNR [dB]	MRME	39.86	23.58	26.27
	Proposed	39.89	23.73	25.81

3. Experimental results

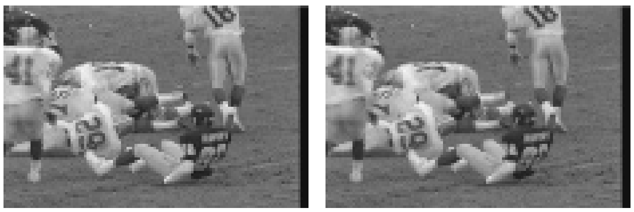
In order to evaluate the proposed motion estimation algorithm, we took 50 frames of the CLAIRE sequence of size 352×288, FOOTBALL and TABLE TENNIS sequence of size 720×480. The biorthogonal discrete WT was used to decompose the original image into 2-levels. The baseband is divided into the blocks of size 4×4 and the search area is -4∼3 in the baseband and -2∼1 in the higher frequency subbands. MAD was used as the metric of the motion estimation and performance measure is the PSNR and the amount of the bit and calculation.

The ratio of blocks which have the motion vector in the integer-pixel accuracy at the baseband is represented in Table III. We can expect that the more the blocks having the motion vector of the integer-pixel accuracy, the more the reduction of the bit and calculation amount. Table IV represent the calculation amount by the MRME and proposed method. This table show that about 38∼46 percent of the calculation amount is reduced by the selective motion estimation. The bit rate and PSNR by the MRME and proposed method is represented in Table V and motion compensated images are represented in Fig. 2, Fig. 3, and Fig. 4. We could see from these results that the proposed method reduced above 50 percent of the bit amount at the similar image quality compared with the MRME. This means that the selective motion estimation at the higher frequency subbands saved the amount of bits and calculation, and the HPAME at the baseband compensated the image degradation by the selective motion estimation at the higher resolution.

In spite of the selective motion estimation, the PSNR is increased by the HPAME for the CLAIRE and FOOTBALL sequence. The background of the TABLE TENNIS sequence shows noise-like characteristic and the decomposition of this image represent large values only in the higher frequency layer. If the blocks having the motion vector in the integer-pixel accuracy at the baseband represent the high frequency components, then motion estimation for the corresponding blocks in the high frequency

(a) (b)

Figure 1: *The motion compensated CLAIRE image by (a) MRME method and (b) the proposed method.*



(a) (b)

Figure 2: *The motion compensated FOOTBALL image by (a) MRME method and (b) the proposed method.*

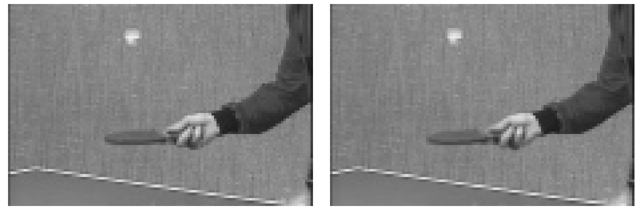
subbands isn't performed and the DFD error is increased. But for the TABLE TENNIS sequence, there are no visual artifacts in the compensated sequence.

The proposed motion estimation method reduce about 40 and 50 of the calculation and bit amount respectively, maintaining the image quality compared with the MRME.

4. Conclusions

We proposed a new hierarchical motion estimation method using the HPAME and selective motion estimation. In the proposed method, the HPAME is used to estimate the exact motion vectors in the baseband and the selective motion estimation in the higher frequency subbands for the reduction of the amount of calculation and bits. That is, the motion estimation is applied only for the corresponding blocks that have motion vectors in the half-pixel accuracy. Because the blocks which have motion vectors in the half-pixel unit don't move in integer grid unit, the blocks of the corresponding position in the higher frequency subbands have a large probability of having small displacement. The degradation of the image quality by the selective motion estimation can be compensated by the HPAME in the baseband.

Experimental results showed that the proposed method reduced the calculation and bit overhead about 40 and 50 percent respectively maintaining the image quality compared with the MRME.



(a) (b)

Figure 3: *The motion compensated TABLE TENNIS image by (a) MRME method and (b) the proposed method.*

References

[1] *ITU-T Recommendation H.261*, "Video codec for audiovisual services at p×64 kbits/s."

[2] *ITU-T Recommendation H.263*, "Video coding for low bit rate communication."

[3] *ISO/IEC 11172-2*, "Coding of Moving Pictures and Associated Audio for Digital Storage Media at up to about 1.5 Mbits/s: Video."

[4] *ISO/IEC 13818-2*, "Information technology - Generic Coding of Moving Pictures and Associated Audio Information: Video."

[5] Y. Zhang and S. Zafar, "Motion Compensated Wavelet Transform Coding for Color Video Compression," *IEEE Trans. on Circuit and System for Video Technology*, vol. 2, no. 3, pp. 285∼296, Sept. 1992.

[6] J. H. Jeon and J. K. Kim, "On the hierarchical edge-based block motion estimation for video subband coding at low bit rates," *SPIE Conf. on Visual Comm. and Image Processing*, vol. 2094, pp. 337∼343, Nov. 1993.

[7] K. H. Goh, J. J. Soraghan, and T. S. Durrani, "MULTIRESOLUTION BASED ALGORITHMS FOR LOW BIT-RATE IMAGE CODING," *Proc. IEEE International Conference on Image Processing*, Austin Texas, pp. 285∼289, Nov. 1994.

[8] S. Panchanathan, E. Chan, and X. Wang, "Fast Multiresolution Motion Estimation Scheme for a Wavelet Transform Video Coder," *SPIE Conf. on Visual Comm. and Image Processing*, vol. 2308, pp. 671∼681, Sept. 1994.

[9] B. Girod, "Motion-compensating prediction with fractional-pel accuracy," *IEEE Trans. on Comn.*, vol. 41, no. 4, pp. 604∼612, Apr. 1993.

[10] Motion Picture Exports Group, "MPEG test model 5 draft revision 2," *ISO-IEC JTC1/SC29/WG11/602*, Nov. 1993.

[11] Motion Picture Exports Group, "MPEG test model 5 draft revision 2," *Doc. ISO-IEC JTC1/SC29/WG11/N0400*, Apr. 1993.