An Adaptive Mutiresolution Estimation Considering the Spatial and Spectral Characteristic

Kwang-Yong KIM and Kyung-Ok KIM
Spatial Information Technology Center,
Electronics and Telecommunications Research Institute,
161, Kajong-Dong, Yusong-Gu, DaeJeon, 305-600, Korea
Tel : +82-42-860-1628, Fax : +82-42-860-4844
e-mail: kimky@etri,re,kr

Abstract: In this paper, we proposes an adaptive method for reducing the computational overhead of fine-to-coarse MRME at the finest resolution level by considering for the spatial and spectral characteristics between wavelet decomposition levels simultaneously.

As we know, there is high correlation between the adjacent blocks and it can give the very important clue to estimate motion at finest level.

So, in this paper, using the initial motion vector and the adjacent motion vector in the coarsest level, we determine the optimal direction that will be minimized the estimation error in the finest level. In that direction, we define the potential searching region within the full searching region that is caused to increase much computational overhead in the FtC method. Last, in that region, we process the efficient 2-step motion estimation. and estimate the motion vector at finest resolution level. And then, this determined motion vector is scaled to coarser resolutions. As simulation result, this method is similar to computational complexity of the CtF MRME method and very significantly reduces that of the FtC MRME method. In addition, they provide higher quality than CtF MRME, both visually and quantitatively

1. Introduction

Wavelet transform can express the signal with the locality about time and frequency, so profit to analyze the image data with the non-stationary process. Moreover, in the case of low compression, because there is no blocking effect like DCT-based compression, it can enhance subjective image quality in the restored image. For these advantages, also in the field of moving-image, wavelet analysis rises as alternative method of MPEG-4. In particular, the key point of moving image compression is the efficient motion estimation technique, due to it's much computational complexity.

Several motion estimation methods have been proposed in the wavelet domain. Among them, the coarse-to-fine (CtF) motion estimation methods generally have a lower complexity at the expense of inaccurate estimation. Several reason of inaccurate estimation of the CtF method is following. Firstly, it is the potential for inaccurate motion estimation at the coarsest resolution, due to lack to detail and aliasing effects. These inaccuracies result in suboptimal motion estimation at finer resolutions. Secondly, the small block-size at a coarser resolution level does not provide robust motion estimation. On the other hand, the fine-to-coarse (FtC) motion estimation methods provide a superior estimation, but at a higher complexity. In this method, because accurate motion estimation are formed at the finest resolution and then scaled to coarser resolutions in the encoded process, these motion estimates better track the true motion and exhibit lower entropy than coarse-to-fine estimations, providing higher quality, both visually and quantitatively. But, because FtC MRME is executed in the full search region at the finest resolution level, it is caused to increase much computational complexity in relatively coarser energy level.

In this paper, we proposes an adaptive method for reducing the computational overhead of fine-to-coarse MRME at the finest resolution level by considering for the spatial and spectral characteristics between wavelet decomposition levels simultaneously. Using the initial motion vector (MV) and the adjacent motion vector in the coarsest level, we determine the optimal direction that will be minimized the estimation error in the finest level. In that direction, we define the potential motion region within the full search region that is caused to increase much computational overhead in the FtC method. Last, in that region, we process the efficient 2-step motion estimation. and estimate the motion vector at finest resolution level. And then, this determined motion vector is scaled to coarser resolutions. Therefore, this method is similar to computational complexity of the CtF MRME method and very significantly reduces that of the FtC MRME method. In addition, they provide higher quality than CtF MRME, both visually and quantitatively

This paper is organized as follows. Section 2 provides several brief MRME techniques; like CtF method, FtC method. In Section 3, the proposed algorithm is detailed. The simulation results are provided in Sec.4, which is followed by conclusions in Sec.5.

2. CtF and FtC method

2.1. Coarse-to-Fine MRME method.

This approach exploits the multiresolution property of the wavelet pyramid in order to reduce the computational complexity of the motion estimation process. We note that Zhang et al.[1] have considered several techniques for motion estimation. Here, we choose the "S_8, W_8+refine"

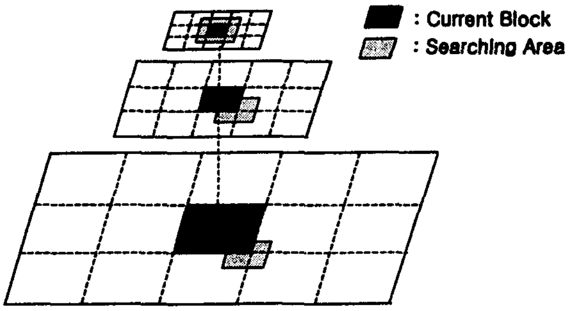


Figure. 1. Coarse-to-Fine MRME method

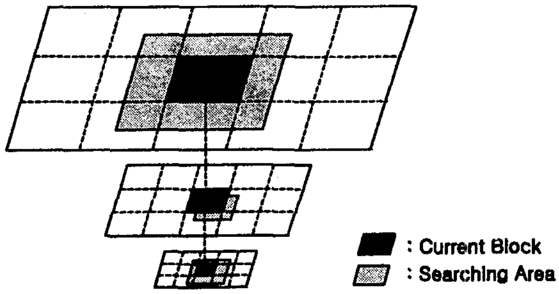


Figure. 2. Fine-to-Coarse MRME method

techniques for motion estimation since it provided superior motion estimation. Fig.1 is shown the typical CtF pyramid construction. In this MRME scheme, the motion vectors at the coarsest level of the wavelet pyramid are first estimated using the conventional block-matching-based motion estimation algorithm. Then the motion vectors at the next level of the wavelet pyramid are predicted from the motion vectors of the preceding level, which are refined at each step. This procedure is called "refinement".

So, this "refinement" has lower complexity needed block-matching. But, if there exits mis-estimation motion vector in the coarsest level, these estimation error is propagated to other levels, so there is every probability of having local minima problem in the finest level. And these methods have the weak point to discard the detail information

For the rest, M.K. Mandal[2] has proposed the method that the motion vectors are dropped adaptively depending on the motion compensation performance, and Karlekar Jayashree, Desai U[3] has proposed the MSAD method to estimate the motion.

2.2. Fine – to – Coarse MRME method

Conklin et al.[5] have proposed a MRME scheme based on a fine-to-coarse approach. Here, comparing with CtF method, the initial motion estimation is executed in the pixel domain. In the other word, the motion vectors at the finest level of the wavelet pyramid are first estimated using the conventional block-matching-based motion estimation *algorithm. Then, scale and refine that at coarser resolutions.* Here, for the ordering transmission of the level, FtC method is needed to code the refinements of the MVs and compensated prediction image after the total motion estimation process.

In this technique, because accurate motion estimation are formed at the finest resolution and then scaled to coarser resolutions in the encoded process, these motion estimates better track the true motion and exhibit lower entropy than coarse-to-fine estimations, providing higher quality, both visually and quantitatively, but because fine-to-coarse MRME is executed in the full search region at the finest resolution subband, it is caused to increase much computational complexity in relatively coarser energy level.

3. Proposed Method

Generally, the basic reason that can propagate the motion vector of coarsest level to the next level in the CtF method is to give the very important clue to estimate motion at other level.

So, if we reduce the computational overhead of fine-to-coarse MRME at the finest resolution level by considering for the spatial correlation between the coarsest motion vector and the adjacent block motion vector, we can have a relatively little complexity and have similar performance.

3.1 Potential Searching Area

So, this paper sets up two candidate searching points (SP_1, SP_2) in consideration of spatial and spectral characteristics in the coarsest level. (See also Fig. 3)

- SP_1: the point scaled the motion vector in the coarsest level.
- SP_2: the point scaled the motion vector which has minimum MAD among the adjacent blocks and current block.

About 2 candidate searching points, we process the mean absolute difference (MAD) method, and determine the center point of potential searching area which minimized estimation error between 2 points. Here, for searching efficiency, we put first order on SP_1, and, if first potential searching area is overlaid with the other region which sets up 1/4 area compared with full searching area, we skip the SP_2 *searching.*

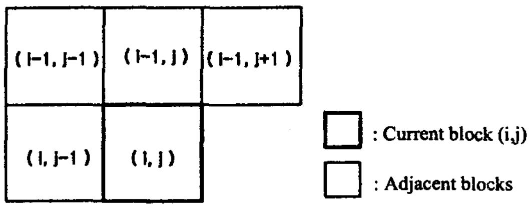


Figure. 3. Adjacent Candidate Blocks (Coarsest Level).

Using these SPs, we then define the potential searching area reducing full searching area by a factor of 2, comparing the original searching area ($N \times N$) and in order to increase the searching performance, we shift the limited searching area ($N/2 \times N/2$) to inner point.

As the result, considering the spatial and spectral characteristics in the wavelet transformed multi-resolution image layers, we intensively estimate the potential searching area where is every probability laying the true motion vector.

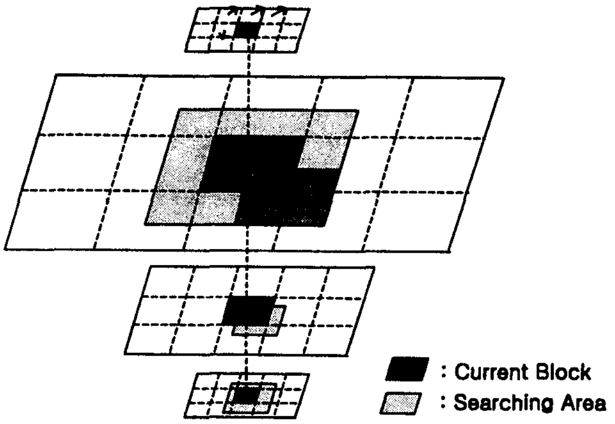


Figure.4. The Proposed MRME method.

3.2 2-step motion estimation.

In the potential searching area, we do 2-step motion estimation.

Firstly, we choose the alternatively subsampling point to estimate the motion in a limited (potential) searching area, using block-matching algorithm, and choose the vector for which the MAD is the smallest, then, we again estimate the motion around the chosen vector.

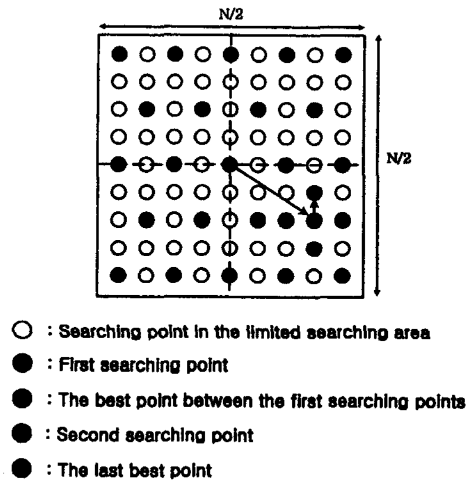


Figure. 5. 2-step searching pattern in the potential searching area

3.3 Procedure

The Process order is following.

Step1.Construct Discrete Wavelet Transformed image with N-level

Step2. Estimate the motion at the coarsest level(S_8).

Step3. Determine the Potential Searching area using the SP_1, SP_2.

Step4. Using 2-step motion Estimation,

estimate adaptively the motion at finest level (W_2^o) and as it decrease the level, scale and refine the MVs to coarsest level.

$$V_2^o(x,y) = V_8'(x,y) \times p(x',y') + \Delta'(\delta x + \delta y)$$

$$\bullet \ where \ p(x',y') \begin{bmatrix} x' = x \times 2 & if \ x = 2 \\ x' = x \times 4 & else \\ y' = y \times 2 & if \ y = 2 \\ y' = y \times 4 & else \end{bmatrix} \qquad \text{Eq.(1)}$$

- $\Delta'(\delta x + \delta y)$: *2-step moti on searhin g po* int

$$V_j^o(x,y) = V_2^o(x,y) + 2^{N-j} + \Delta(\delta x + \delta y)$$

$$for \quad j = 1,2,,,N \qquad \text{Eq.(2)}$$

Step5. Increase the level.

Using the best set of MVs, form a motion compensated prediction image ($\hat{W}_{2^{N-j}}^o$), and prediction residual ($\overline{W}_{2^{N-j}}^o$),

Code the MVs and the prediction residual information.

Step6. If the proceeding is the finest level, go the next image.

4. Simulation and Results

To test the performance of this approach, simulation were run on 720×480 three sequences; 50 frames of Football, Susie, and 40 frames of Table tennis. Football and table tennis sequences offer an interesting combination of still, slow- and fast-moving objects, camera zoom and panning, and objects with relatively difficult sizes. On the other hand, Susie sequences offer slow motion and low spatial detail.

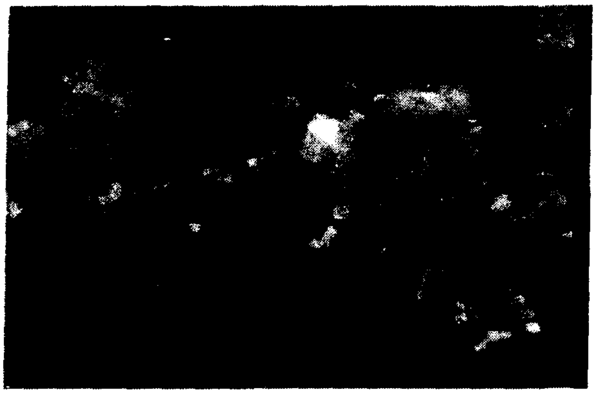
We have used the Daubechies-9/7 Tab wavelet, which provides good coding performance. The peak signal-to-ratio (PSNR), which defined as

$$PSNR = 10\log_{10}\left(\frac{255^2}{MSE}\right) \qquad \text{Eq. (3)}$$

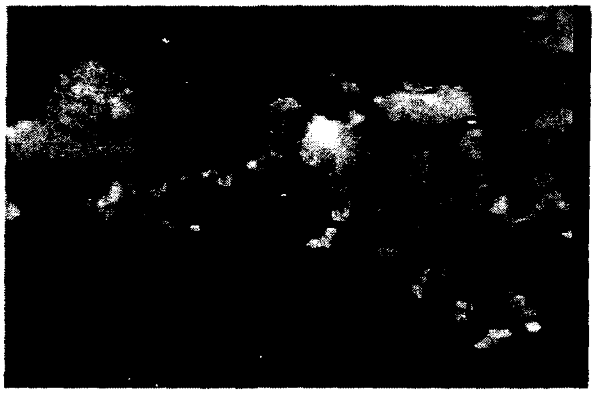
(Where MSE is the mean square error between the signal/images)

has be employed as a measure of the quality of the reconstructed images. For the relative comparison, we have run the CtF method, FtC method, and proposed method. In CtF method, the searching area is –2 ~ +2. And, in FtC method, the searching area at finest level is –8~+8 and that of other levels is –2~+2.

That result is shown the table 1.As shown the table 1, the computational complexity of the proposed method is about 10% than that of FtC method, similar to that of CtF method. And the increase of PSNR is 0.5~1.0dB. Comparing the PSNR's increase of FtC method with that of CtF method, the increase of the proposed method is about 70%.



(a)



(b)

(c)

Figure.6. the Motion Compensated predicted image of the 11th football frame:
(a) CtF method, (b) FtC method, (c) Proposed method

Table 1. PSNR and Computational complexity

[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]
Fb	23.15	24.45	24.11	25	289	27 (9.8%)*
Tt	25.54	26.52	26.22			
Ss	33.59	34.32	34.07			

()*: the ratio of matching point of paper to that of FtC

Fig.6 is shown the motion compensated predicted image of the football 11th frame. As shown the Fig.6, we can see that our proposed method track the motion at the fine level very well.

5. Conclusion

In this paper, we proposes an adaptive method for reducing the computational overhead of fine-to-coarse MRME at the finest resolution level by considering for the spatial and spectral characteristics between wavelet decomposition levels simultaneously.

As shown the result, our method of the computational complexity is reduced about 10%, that value is similar to computational complexity of the CtF MRME technique and very significantly reduces that of the FtC MRME technique.

In addition, flicking in textured regions, such as the textured background in Table Tennis or Football, is significantly reduced using our method.

As a result, this method ME provides both higher PSNRs and better visual quality.

References

[1] Y. Q. Zhang and S. Zafar, "Motion-Compensated transform coding for color video compression." IEEE Trans. Circuits Syst. Video Technol., Vol. 2, No. 3, pp 285-296, 1992.

[2] M. K. Mandal, E. Chan, X. Wang, S. Panchanathan, "Multiresolution motion estimation techniques for video compression", Optical Engineering, Vol 35, No 1, pp 128-136, Jan, 1996.

[3] K. Jayashree, Desai U., "A New Multiresolution Motion Estimation and Compensation Scheme", Proc. IEEE ISCS'99 , pp459-462, 1999

[4] K. M. Uz, M. Vetterli, and D. J. Legall, "Interpolative multiresolution coding of advanced television with compatable subchannels." IEEE Trans, on CSVT, Vol 1, No 1, March, 1991.

[5] G. J. Conklin, S. S. Hemami, "Multiresolution Motion Estimation", Proc. of ICASSP, pp 2873-2876, 1997.