Fast Lossless Multi-Resolution Motion Estimation for Scalable Wavelet Video Coding

Yu Liu and King Ngi Ngan
Department of Electronic Engineering,
The Chinese University of Hong Kong, Shatin, N.T., Hong Kong, China
Email: {yuliu, knngan}@ee.cuhk.edu.hk

***Abstract*— In this paper, we propose a fast lossless multi-resolution motion estimation algorithm in shift-invariant wavelet domain by using wavelet matching error characteristic based partial distortion elimination (MR-WMEC-PDE). Due to its multi-resolution nature, the proposed approach can be applied to scalable video coding. The proposed algorithm can achieve the same estimate accuracy as full search algorithm (FSA), while needing much less computation requirement than FSA.**

I. Introduction

Motion estimation and motion compensation (ME/MC) combined with discrete wavelet transform (DWT) has received much attention recently due to its superior performance compared to the conventional spatial domain ME/MC using the DCT, such as blocking artifacts. However, ME/MC in critically-sampled wavelet domain is inefficient in high bands and suffers significant performance loss when compared to ME/MC in spatial domain because of the shift-variant property of wavelet decomposition. To overcome the shift-variant property of the wavelet transform,the low-band-shift (LBS) method [1] and the complete-to-overcomplete discrete wavelet transform (CODWT) [2] method are proposed for ME/MC in the shift-invariant wavelet domain. These methods avoid the shift-variant property of the wavelet transform and perform motion compensation more precisely and efficiently. However, a major disadvantage of these methods is the computational complexity. Although the CODWT method has reduced a part of computational complexity by skipping the IDWT and making the direct link between the critically-sampled subbands and the shift-invariant subbands, another part of computational complexity comes from full search ME algorithm.

Full search ME algorithm can obtain the optimal result by searching exhaustively for the best matching block within a search window, but its high computational cost limits its practical applications. Due to the multi-resolution nature in the wavelet pyramid, many fast lossy multi-resolution motion estimation (MRME) algorithms, such as MRME [3], adaptive threshold MRME (AMRME), bi-directional MRME (BM-RME) and fast MRME (FMRME) [4], are proposed to reduce the computational complexity. However, these fast algorithms reduce the computations at the expense of the accuracy of motion estimation. In addition, these fast lossy algorithms did not overcome the shift-variant property of wavelet transform because all these algorithms use the critically-sampled subbands of the wavelet reference frame for ME/MC.

In this paper, we take advantage of the new discovery about wavelet matching error characteristic to develop three new strategies for PDE and propose a fast lossless multi-resolution motion estimation algorithm, MR-WMEC-PDE. The proposed MR-WMEC-PDE algorithm can achieve the same estimate accuracy as full search algorithm, while possessing multi-resolution motion vector information which is suitable for scalable video coding and needing much less computation requirement than full search algorithm.

II. Background

A. Motion Estimation in Shift-Invariant Wavelet Domain

In LBS or CODWT, the shift-invariant wavelet coefficient of v-pixel-shifted reference frame t' can be represented by

$$S_{t',l,k,v}\left(i,j\right) = f_{t',l,k}\left(dx\%2^l, dy\%2^l, i + \left\lfloor dx/2^l \right\rfloor, j + \left\lfloor dy/2^l \right\rfloor\right)$$

where l denotes the wavelet decomposition level, k denotes the LL/LH/HL/HH subband type, $v = \{dx, dy\}$ denotes the shifting vector or displacement vector, $\{dx\%2^l, dy\%2^l\}$ indicates the number of shifts in the level l, and $\{i + \lfloor dx/2^l \rfloor, j + \lfloor dy/2^l \rfloor\}$ is the location in the subband k. And the wavelet coefficient of the current frame t can be represented by $S_{t,l,k}\left(i,j\right) = f_{t,l,k}\left(i,j\right)$.

For ME/MC in shift-invariant wavelet domain, the coefficients of each wavelet tree rooted in the lowest subband are rearranged to form a wavelet block. The purpose of the wavelet block is to provide a direct association between the wavelet coefficients and what they represent spatially in the image. Related coefficients at all scales and orientations are included in each wavelet block. The wavelet blocks in search window W in the reference frame are compared to the current wavelet block, and a reference wavelet block that leads to the best match is selected. The SAD of the pth wavelet block for the displacement vector v is computed as follows:

$$\begin{aligned} SAD\left(p,v\right) = &\sum_{i=1}^{N/2^L}\sum_{j=1}^{N/2^L}\left|S_{t,L,0,p}\left(i,j\right) - S_{t',L,0,p+v}\left(i,j\right)\right| \\ &+ \sum_{l=1}^{L}\sum_{k=1}^{3}\sum_{i=1}^{N/2^l}\sum_{j=1}^{N/2^l}\left|S_{t,l,k,p}\left(i,j\right) - S_{t',l,k,p+v}\left(i,j\right)\right| \end{aligned} \quad (1)$$

The optimum motion vector v^* of the pth wavelet block, which has minimum displacement error, is given by:

$$SAD\left(p,v^*\right) = \min_{v\in W} SAD\left(p,v\right) \quad (2)$$

B. Partial Distortion Elimination

The PDE [5] approach uses the partial sum of differences (PSAD) to eliminate impossible candidates before the complete calculation of the SAD, as follows:

$$\begin{aligned} PSAD_m(p,v) &= \sum_{n=1}^{m\leq N\times N} \left| S_{t,p}(O(n)) - S_{t',p+v}(O(n)) \right| \\ &\geq SAD_{min}(p) \end{aligned} \quad (3)$$

where m is the number of differences needed to reach the SAD_{min} which has already been found so far, and $O(n)$ represents a generic matching order.

III. Proposed Algorithm

A. Wavelet Matching Error Characteristic

To improve the computational saving of PDE, if the expected values of matching error $d_{p+v}(i,j)$ in the search window W fulfills the following criterion:

$$E\left[d_{p+v}(i',j')\right] \geq E\left[d_{p+v}(i'',j'')\right] \quad (4)$$

where $d_{p+v}(i,j) = |S_{t,p}(i,j) - S_{t',p+v}(i,j)|$, then we assume that this matching order will be generally effective for all of the candidate vectors. Therefore, to fulfill the above objective, we have to predict the matching error with $e_p(i,j)$. In order to obtain the predicted matching error $e_p(i,j)$, we consider

$$e_p(i,j) = \arg\min_{e_p(i,j)} E\left[\left((e_p(i,j))^2 - (d_{p+v}(i,j))^2\right)^2\right] \quad (5)$$

To solve this equation, we have

$$\frac{\partial}{\partial(e_p(i,j))} E\left[\left((e_p(i,j))^2 - (d_{p+v}(i,j))^2\right)^2\right] = 0 \quad (6)$$

Since it is well-known that the wavelet coefficients follow the Laplacian distribution, let's assume that $E[S_{t',p+v}(i,j)] \approx 0$. Thus, we finally can obtain an approximate solution of Eq. (6):

$$\begin{aligned} e_p(i,j) &= \sqrt{E\left[(S_{t,p}(i,j) - S_{t',p+v}(i,j))^2\right]} \\ &\approx \sqrt{(S_{t,p}(i,j))^2 + E\left[(S_{t',p+v}(i,j))^2\right]} \\ &\approx |S_{t,p}(i,j)| + E\left[|S_{t',p+v}(i,j)|\right] \end{aligned} \quad (7)$$

The solution of the predicted matching error $e_p(i,j)$ consists of two parts: one is the wavelet coefficient magnitude in the current wavelet block, $|S_{t,p}(i,j)|$; another is the expected value of the wavelet coefficient magnitude in the reference wavelet block, $E[|S_{t',p+v}(i,j)|]$. Since the expected value of the wavelet coefficient magnitudes in a search window varies smoothly, the second part can be considered as a constant for different $e_p(i,j)$. Therefore, *the predicted matching error $e_p(i,j)$ is proportional to the wavelet coefficient magnitude in the current wavelet block.* That is to say, larger wavelet magnitude in the current wavelet block tends to produce larger matching error. This conclusion is also in accordance with the fact that the edges and texture contribute most to the SAD matching error, and high magnitude wavelet coefficients usually correspond to edges and texture.

B. MR-WMEC-PDE

Observed from Eq. (3) in Section II-B, we can sum up three key factors which affect the performance of PDE:

- *the searching order*, in which the wavelet blocks are tested during the searching phase. In fact, if a good SAD_{min} is found early, then many more successive tests have a tighter distortion bound and may be skipped.
- *the matching order*, in which the coefficients within a wavelet block are picked up to compute the SAD. In this case, if the highest contributions to SAD are found early, then the distortion bound may be reached after a small number of differences and the PSAD can be stopped.
- *the comparison interval*, in which comparison between PSAD and SAD_{min} is performed. Here, if a good tradeoff between the cost of comparisons and the number of useless differences (when the PSAD is already greater than SAD_{min}) is achieved, then the number of comparisons and differences can be reduced.

In order to make full use of these three factors, the wavelet matching error characteristic (WMEC) is employed to exploit the new strategies for PDE: Searching Order Strategy based on Wavelet Multi-Resolution Property, Matching Order Strategy based on Wavelet-tree Grouping Scheme, Comparison Interval Strategy based on Adaptive Sub-blocks Checking Unit.

1) Searching Order Strategy based on Wavelet Multi-Resolution Property: Due to the fact that wavelet coefficients still contain spatial information, the LL subband actually corresponds to a lowest resolution version of an original image frame. Therefore, the motion information of the LL subband in a wavelet block is highly correlated with that of the whole wavelet block. This property can be used for a new searching order strategy to get a good match sooner. Different from the spiral searching order [5] which uses a spiral outward trajectory starting from the center of the search window according to the statistical distribution of the predicted motion vector, the new proposed searching order strategy uses the normalized partial SAD in LL subband level as the estimated SAD (ESAD):

$$SAD_{LL}(p,v) = \sum_{i=1}^{\frac{N}{2^L}} \sum_{j=1}^{\frac{N}{2^L}} \left| S_{t,L,0,p}(i,j) - S_{t',L,0,p+v}(i,j) \right| \quad (8)$$

$$ESAD(p,v) = \frac{1}{2^L} \cdot \frac{1}{(\frac{N}{2^L})^2} SAD_{LL}(p,v) = \frac{2^L}{N \times N} SAD_{LL}(p,v) \quad (9)$$

The value of ESAD is between 0 and 255. This fixed range allows us to use a fast sorting technique, the counting sort algorithm [7] which has linear complexity, to obtain the searching order $SO = \{v_n | n = 0, ..., W-1\}$, as follows:

$$ESAD(p,v_0) \leq ... \leq ESAD(p,v_n) \leq ... \leq ESAD(p,v_{W-1}) \quad (10)$$

2) Matching Order Strategy based on Wavelet-tree Grouping Scheme: Since the wavelet coefficient magnitude tends to be larger in the higher decomposition level, the matching error also tends to be larger in the higher decomposition level according to above reasoning. Therefore, the trend of matching order is from the higher decomposition level to the lower decomposition level. In addition, it is well-known

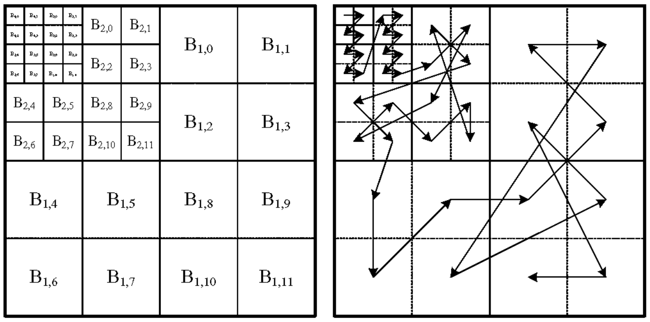

Fig. 1. (a)Wavelet-tree grouping scheme in a wavelet block, (b)an example of adaptive matching order based on wavelet-tree grouping scheme

that significant wavelet coefficients with similar magnitude tend to appear in clusters. According to our observation, this clustering property is also found in the matching errors. That is, the matching errors with similar magnitude in the same level tend to appear in clusters. Thus, we propose a wavelet-tree grouping scheme which groups the coefficients into sub-blocks according to spatial self-similarity property and matching error clustering property of wavelet, as shown in Fig.1(a). To determine the matching order of sub-blocks, the estimated matching error of each sub-block $B_{l,b_{l,m}}$ is calculated by

$$E_p(B_{l,b_{l,m}}) = \sum_{(i,j)\in B_{l,b_{l,m}}} |S_{t,l,b_{l,m}/4,p}(i,j)| \qquad (11)$$

where $b_{l,m}$ is the index of each sub-block of level l, and $m = 0, ..., M-1$. Then, sorting the estimated matching error $E_p(B_{l,b_{l,m}})$ in descending order,

$$E_p(B_{l,b_{l,0}}) \geq ... \geq E_p(B_{l,b_{l,m}}) \geq ... \geq E_p(B_{l,b_{l,M-1}}) \qquad (12)$$

to obtain the matching order of level l: $MO_l = \{b_{l,m} | m = 0, ..., M-1\}$, as shown in Fig.1(b). Since the value of M is small and equal to 12, the quick sort algorithm [7] which has $O(n \log n)$ complexity is used to obtain the matching order.

3) Comparison Interval Strategy based on Adaptive Sub-blocks Checking Unit: In conventional PDE methods, fixed comparison interval, such as eight-pixels or sixteen-pixels checking unit, is usually adopted. Since the matching error tends to be larger in the higher decomposition level in the wavelet block, it is preferred to choose a smaller comparison interval for the higher decomposition level and a larger comparison interval for the lower decomposition level. Combined with the wavelet-tree grouping scheme, to achieve a good tradeoff between the cost of comparisons and the number of useless differences, an adaptive comparison interval strategy is proposed. In the proposed scheme, every 2^{l-1} sub-blocks in the decomposition level l are used as the checking unit, that is, the comparison between PSAD and SAD_{min} is performed every 2^{l-1} sub-blocks for the decomposition level l. The advantage of the adaptive comparison interval scheme lies in that the distortion bound may be reached after a small number of differences and the partial sum of differences can be stopped early thereby maintaining the cost of comparisons remains at a reasonable level.

IV. Experimental Results

In order to compare the performance of the proposed algorithms with other methods, we use the first 150 frames of a large variety of video sequences for evaluation, including ten QCIF sequences, ten CIF sequences, and four 4CIF sequences. In the experiments, Daubechies 9/7 biorthogonal wavelet filter is employed for the wavelet transform, and the maximum wavelet decomposition level $L = 3$ is applied. The wavelet block size is 16x16 pixels, while the search windows are [-15, 15] for QCIF, [-31, 31] for CIF, and [-63, 63] for 4CIF, respectively.

To demonstrate the performance of the proposed algorithms, we compare the proposed lossless MR-WMEC-PDE algorithm with Spiral-PDE [5] and modified CPME-PDS [6]. All these algorithms were implemented in shift-invariant wavelet domain. Although the original CPME-PDS are implemented in critically-sampled wavelet domain, we modified the algorithm to work in shift-invariant wavelet domain for *fair* comparison. Simulation results are reported in the following ways:

- *operation number per block* used to compute the partial distortion;
- *execution time per frame* for motion estimation including the required overheads for comparison.

Table I and II contain the comparisons between the computational efficiency of the tested algorithms in terms of average operation number per block and average execution time per frame. On average speed-up ratio in terms of operation number per block, the performance of MR-WMEC-PDE is better than Spiral-PDE and CPME-PDS by about 92% and 34%, respectively. On average speed-up ratio in terms of execution time per frame, the performance of MR-WMEC-PDE is better than Spiral-PDE and CPME-PDS by about 84% and 63%, respectively. Experimental results show that the proposed MR-WMEC-PDE algorithm can successfully improve computational efficiency of the conventional PDE algorithms.

It is also worth noting that there are additional computational overheads in CPME-PDS and MR-WMEC-PDE algorithms, except for Spiral-PDE algorithm. The performance influences from the additional computational overheads can be observed from the speed-up ratio of execution time per frame when compared to that of operation number per block. However, the overhead of the proposed MR-WMEC-PDE algorithm can still be tolerable since the loss of performance that comes from the overhead is only 10% on average.

V. Conclusion

In this paper, we took advantage of the new discovery about wavelet matching error characteristic to develop three new strategies for PDE and proposed a fast lossless multi-resolution motion estimation algorithm, MR-WMEC-PDE. Due to its multi-resolution nature, the proposed approach can be applied to scalable wavelet video coding. Experimental results show that MR-WMEC-PDE can reduce significantly computational cost of full search algorithm and also improve computational efficiency of conventional PDE algorithms while keeping the same estimate accuracy.

REFERENCES

[1] H.W. Park and H.S. Kim, "Motion estimation using low-band-shift method for wavelet-based moving-picture coding," IEEE Trans. Image Process., Vol.9(4), pp.577-587, Apr. 2000

[2] Y. Andreopoulos, A. Munteanu, G. Van der Auwera, J.P.H. Cornelis, P. Schelkens, "Complete-to-overcomplete discrete wavelet transforms: theory and applications," IEEE Trans. Signal Process., Vol.53(4), pp.1398-1412, Apr. 2005

[3] S. Zafar, Y.Q. Zhang, B. Jabbari, "Multiscale video representation using multiresolution motion compensation and wavelet decomposition," IEEE Journal on Selected Areas in Commun., Vol.11(1), pp.24-35, Jan. 1993

[4] M.K. Mandal, E. Chan, X. Wang and S. Panchanathan, "Multiresolution motion estimation techniques for video compression," Optical Engineering, Vol.35, pp.128-136, 1996

[5] ITU-T Recommendation H.263 software implementation, Digital Video Coding Group, Telenor R&D, 1995

[6] K.Y. Wong, W.C. Siu, K.C. Hui, "Fast motion estimation for wavelet-based video coding," *Proc. of International Symposium on Intelligent Multimedia, Video and Speech Processing*, 2004, pp.398-401, Oct. 2004

[7] H.M Mahmoud, "Sorting: a distribution theory," New York, Wiley, 2000

TABLE I

AVERAGE OPERATION NUMBER PER BLOCK FOR THE TESTED ALGORITHMS

Format	Video	FSA		Spiral-PDE		CPME-PDE		MR-WMEC-PDE	
		Operation Number	*Speed-up ratio*	Operation Number	*Speed-up ratio*	Operation Number	*Speed-up ratio*	Operation Number	*Speed-up ratio*
QCIF (176x144)	Carphone	246016	*1.00*	19688	*10.17*	13781	*14.53*	10433	*19.19*
	Claire	246016	*1.00*	19807	*10.11*	14160	*14.14*	11065	*18.10*
	Foreman	246016	*1.00*	17305	*11.57*	11998	*16.69*	8490	*23.59*
	Hall	246016	*1.00*	19953	*10.04*	13985	*14.32*	11048	*18.13*
	Mobile	246016	*1.00*	24253	*8.26*	19882	*10.07*	16895	*11.85*
	Mother	246016	*1.00*	17734	*11.29*	12113	*16.53*	8865	*22.59*
	Salesman	246016	*1.00*	13729	*14.59*	8264	*24.23*	4899	*40.87*
	Stefan	246016	*1.00*	26736	*7.49*	21010	*9.53*	18241	*10.98*
	Suzie	246016	*1.00*	25340	*7.90*	19314	*10.37*	16413	*12.20*
	Tempete	246016	*1.00*	19455	*10.29*	14388	*13.92*	11200	*17.88*
CIF (352x288)	Akiyo	871659	*1.00*	57254	*15.22*	30989	*28.13*	17245	*50.55*
	Bus	871659	*1.00*	119625	*7.29*	91498	*9.53*	78873	*11.05*
	Coastguard	871659	*1.00*	122215	*7.13*	90471	*9.63*	79983	*10.90*
	Container	871659	*1.00*	96045	*9.08*	70695	*12.33*	57952	*15.04*
	Flower	871659	*1.00*	120872	*7.21*	93563	*9.32*	81970	*10.63*
	Foreman	871659	*1.00*	102687	*8.49*	75176	*11.59*	63432	*13.74*
	Mobile	871659	*1.00*	113152	*7.70*	88694	*9.83*	76253	*11.43*
	Paris	871659	*1.00*	71213	*12.24*	44887	*19.42*	31401	*27.76*
	Tempete	871659	*1.00*	100976	*8.63*	75234	*11.59*	62700	*13.90*
	Waterfall	871659	*1.00*	72515	*12.02*	53450	*16.31*	39918	*21.84*
4CIF (704x576)	City	3632446	*1.00*	395276	*9.19*	293051	*12.40*	243722	*14.90*
	Crew	3632446	*1.00*	551500	*6.59*	414204	*8.77*	362152	*10.03*
	Harbour	3632446	*1.00*	283106	*12.83*	182161	*19.94*	130904	*27.75*
	Soccer	3632446	*1.00*	441718	*8.22*	323431	*11.23*	277418	*13.09*

TABLE II

AVERAGE EXECUTION TIME PER FRAME FOR THE TESTED ALGORITHMS

Format	Video	FSA		Spiral-PDE		CPME-PDE		MR-WMEC-PDE	
		Execution Time	*Speed-up ratio*	Execution Time	*Speed-up ratio*	Execution Time	*Speed-up ratio*	Execution Time	*Speed-up ratio*
QCIF (176x144)	Carphone	802	*1.00*	82	*9.78*	74	*10.84*	46	*17.43*
	Claire	782	*1.00*	81	*9.65*	74	*10.57*	47	*16.64*
	Foreman	796	*1.00*	73	*10.90*	66	*12.06*	39	*20.41*
	Hall	802	*1.00*	83	*9.66*	75	*10.69*	48	*16.71*
	Mobile	814	*1.00*	110	*7.40*	101	*8.06*	82	*9.93*
	Mother	800	*1.00*	72	*11.11*	64	*12.50*	38	*21.05*
	Salesman	809	*1.00*	55	*14.71*	44	*18.39*	20	*40.45*
	Stefan	798	*1.00*	114	*7.00*	112	*7.13*	85	*9.39*
	Suzie	813	*1.00*	106	*7.67*	103	*7.89*	74	*10.99*
	Tempete	810	*1.00*	81	*10.00*	79	*10.25*	53	*15.28*
CIF (352x288)	Akiyo	13921	*1.00*	907	*15.35*	645	*21.58*	285	*48.85*
	Bus	13843	*1.00*	1991	*6.95*	1961	*7.06*	1475	*9.39*
	Coastguard	13854	*1.00*	2052	*6.75*	1960	*7.07*	1509	*9.18*
	Container	13903	*1.00*	1650	*8.43*	1574	*8.83*	1084	*12.83*
	Flower	13856	*1.00*	2040	*6.79*	2038	*6.80*	1532	*9.04*
	Foreman	14003	*1.00*	1711	*8.18*	1601	*8.75*	1120	*12.50*
	Mobile	14000	*1.00*	1915	*7.31*	1882	*7.44*	1441	*9.72*
	Paris	13985	*1.00*	1265	*11.06*	1040	*13.45*	580	*24.11*
	Tempete	14085	*1.00*	1678	*8.39*	1620	*8.69*	1160	*12.14*
	Waterfall	14184	*1.00*	1194	*11.88*	1185	*11.97*	765	*18.54*
4CIF (704x576)	City	236156	*1.00*	32164	*7.34*	27732	*8.52*	19873	*11.88*
	Crew	238409	*1.00*	45796	*5.21*	38275	*6.23*	28377	*8.40*
	Harbour	238562	*1.00*	24583	*9.70*	18307	*13.03*	9706	*24.58*
	Soccer	233424	*1.00*	36268	*6.44*	30512	*7.65*	22199	*10.52*