Journal of Information Hiding and Multimedia Signal Processing ©2011 ISSN 2073-4212
Ubiquitous International Volume **2**, Number **3**, July **2011**

Research and Optimization of Low-Complexity Motion Estimation Method Based on Visual Perception

Pengyu Liu[1] and Kebin Jia[2]
School of Electronic Information & Control Engineering
Beijing University of Technology
Beijing, China
Kebin Jia
School of Electronic Information & Control Engineering
Beijing University of Technology
Beijing, China
[1]liupengyu@bjut.edu.cn, [2]kebinj@bjut.edu.cn

Received July 2010; revised October 2010

Abstract. *H.264/AVC was the outstanding and significant video compression standard developed by ITU-T/ISO/IEC Joint Video Team. It achieved excellent encoding performance of at the cost of increased computational complexity and falling encoding speed. In order to overcome poor real-time encoding performance of the important and time consuming motion estimation technology in H.264/AVC standard, aiming at computing redundancy of UMHexagonS algorithm which was selected by H.264/AVC latterly, with special considerations on visual perception, using macro-block's motion characteristics, formulated the adaptive motion estimation strategies and search points distribution rule. A low-complexity motion estimation scheme is proposed in this paper to improve the video coding performance significantly. The simulation results show that the proposed motion estimation scheme achieve reduction of 23.11on average compared with UMHexagonS algorithm, in the premise of effectively improving the encoding speed, keep the low bit rate and high quality of encoding advantage for H.264/AVC compared with full search algorithm.*
Keywords: H.264/AVC; Motion estimation; macro-block; motion characteristics; visual perception

1. **Introduction.** With the development of multimedia information processing and communication technologies, video encoding has broad application prospects in information processing, national defense, research & education, medical treatment, aerospace engineering and other fields, and become one of hot spots for multimedia information processing technology.

H.264/AVC[1] is the most recent and promising international video coding standard developed by the ITU-T Video Coding Experts Group in conjunction with the ISO/IEC Moving Picture Experts Group. The new standard H.264/AVC aims at high-quality coding of video content at very low bit-rates and is designed for application in the areas such as broadcast, video-on-demand or multimedia streaming, multimedia messaging etc. over ISDN, DSL, Ethernet, LAN, wireless and mobile networks. Some new features and capabilities of the H.264/AVC standard such as variable block size (7 types), quarter-sample-accuracy and multiple reference frames enable enhanced coding efficiency,compared with advanced standards, up to 50% of bit-rate reduction can be achieved[2], but at the same time increase the complexity and computation load of motion estimation greatly in H.264/AVC encoder.

Motion estimation (ME) is the basic bandwidth compression method adopted in the video coding standards, while it is the most time consuming module. The block matching algorithm (BMA) is the most implemented one in real time for ME. BMA for ME is the mainstream algorithm for video compression, which has been adopted by many standards such as MPEG-1/2/4, H.263, H.264/AVC, etc. The key problem for BMA is to find the best matched motion vector in their reference frames for every macro-block.

Full search (FS) algorithm is well-known and widely used because it is simple and accurate. It exhaustively tests all the candidate points within a pre-defined searching area in a reference frame to get the best matched motion vector, the calculation is huge. Especially for H.264/AVC, it adopts some effective features mentioned above to improve the compression quality. So FS have to calculate all the search points in the search window for all the variable 7 types block sizes (1616,168,816,88,84, 48,44) and reference frames.

Studies show that about 60% 80% of the computational complexity come from ME, and the proportion of the computational complexity improved with the increase of the number of the reference frames The encoding complexity of H.264/AVC depends on the complexity of ME scheme structure largely[3]. To reduce the calculation burden of ME, in the past years, many fast ME algorithms were presented. They can be classified in two categories. One type is to reduce the search points, such as three step search algorithm, four step search algorithm, diamond search algorithm, hexagon-base algorithm etc. These fixed pattern search algorithms can reduce the search points largely and get a good image quality. But when the actual motion does not match the pattern well, the image quality will decrease. The other type of algorithm is to reduce the calculation for every search candidate. It uses sub-sampling method and partial calculation. Sub-sampling method is efficient but not accurate, while the partial calculation is accurate but not enough efficient.

Recently several fast ME algorithms were proposed for H.264/AVC. They combine many methods together and achieve both fast speed and good image quality. Especially, UMHexagonS[4] (Unsymmetrical-cross Multi-Hexagon-gird Search) algorithm has been adapted by H.264/AVC reference software, claiming that nearly 90% computations can be saved on average compared with the fast FS algorithm in JVT reference software with a fairly good PSNR performance. However, based on many experiments results analysis, find that the UMHexagonS algorithm is also complicated, and can be speed-up further.

In this paper, bases on the analysis of UMHexagonS algorithm, aiming to reducing the number of ME search points and improving ME speed effectively, maintaining or even improving the accuracy of ME, defined the associated characteristics between the current macro-block motion intensity and it's ME search range, and then designed a new ME template, established ME search strategies, striving to reduce the contradiction between computational complexity and accuracy in ME effectively and enhance the coding efficiency of H.264/AVC comprehensively .

The second part of the paper analyzed UMHexagonS algorithm and its optimization space. A new ME template and a novel fast integer pixel ME search strategy with adaptive search strategies for H.264/AVC according to the motion activity of the current macroblock were elaborated in the third part. The simulation results were given in the fourth part. Final part is the summary of this paper.

2. **Principle of UMHexagonS Algorithm.** UMHexagonS algorithm has been accepted for integer-pixel block motion estimation by H.264/AVC reference software. It can

actually reduce the computational load for ME by reducing the number of candidate blocks within a search window. To achieve the high coding efficiency and avoid the local-minimum problem, it is widely conducted into two parts: the first is initial search center prediction and the second is use the hybrid of integer pixel search.

2.1. **Initial search center prediction.** Generally speaking, spatial and temporal predictions are the main mechanisms for ME to find the motion vectors (MV) of the current block. These mechanisms generate four types of prediction means[5]: Median Prediction, Up-Layer Prediction, Corresponding block Prediction, Neighboring Ref-frame Prediction. UMHexagonS algorithm uses these prediction means to predict the initial search center with high veracity.

2.2. **Search strategies of UMHexagonS algorithm.** UMHexagonS algorithm can predict the motion vector accurately. There are three main steps in its search algorithm.

Step1. The search begins with the unsymmetrical cross search. Taking a search range of 16 the defined search window is shown with the search points in Fig.1 (step1).

Step2. The best match of step1 gives the center point for the step2 search which are 55 full search with the search points are shown in Fig.1 (step2-1) and the 4 layers uneven multi-hexagon-grid search. The search points for this search are shown in Fig.1 (step2-2). From inside to outside is layer 1 to layer4, and 16 search points are distributed on each layer.

Step3. The last search process uses extended hexagon-based search, composed of symmetrical- hexagon-grid search shown in Fig.1 (step3-1) and a small diamond search shown in Fig.1 (step3-2) until the center of the search pattern is the best candidate point.

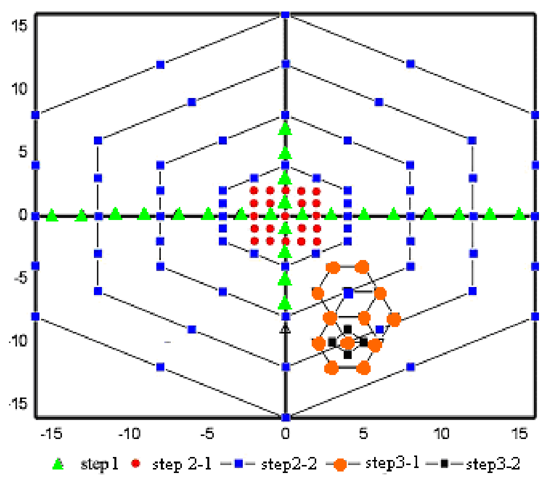


FIGURE 1. Flow chart and the search pattern of UMHexagonS algorithm

To find the optimum MV in these steps, UMHexagonS algorithm uses the hierarchical and hybrid motion search strategies. Obviously, the hybrid strategies exploit the irregularity of search patterns to find the best MV. The irregularity pattern search still causes a heavy computation overhead and limits the performance of ME speed.

2.3. **redundancy analysis of UMHexagonS algorithm.** Take JM12.2 test model of H.264/AVC as the experimental platform. Table1 shows the appearance probability of optimal matching point on different search layer in each step for UMHexagonS algorithm.

Statistical data in table shows that in UMHexagonS algorithm, for Akiyo, Silent and other moderate motion video sequences, the 55 integer pixel full search in step2-1 and the

Table 1. THE APPEARANCE PROBABILITY OF OPTIMAL MATCHING POINT ON DIFFERENT SEARCH LAYERS IN UMHEXAGONS ALGORITHM (%)

Sequences	Step2-1 and step2-2 Layer 1	Step2-2 Layer 2	Step2-2 Layer 3	Step2-2 Layer 4	Step1 and Step3
Akiyo	51.12%	27.43%	10.65%	5.92%	4.88%
Silent	57.85%	21.97%	9.96%	7.02%	3.20%
Container	50.82%	25.36%	7.88%	6.94%	9.00%
Suzi	40.70%	26.85%	15.22%	8.26%	8.97%
Salesman	31.85%	22.15%	21.28%	17.76%	6.96%
Coastguard	22.52%	16.43%	27.89%	27.67%	5.49%
Foreman	21.15%	19.89%	28.02%	22.82%	8.12%
Football	23.67%	21.75%	25.29%	24.12%	5.17%
Computer configuration: 1G memory, 1.6 GHz frequency.					

inner layer in step2-2 (layer 1 or layer 2), has higher probability to appear the optimal matching point. It shows that most encoding macro-block motion vectors are small. For Foreman, Football and other high-motion video sequences, the outer in step-2 (layer 3 or layer 4), has increased probability to appear the optimal matching point. It shows that most encoding macro-block motion vectors are big. There are only a small number of optimal matching points appeared in the step1 and step3.

Because a majority of video sequence are the background with moderate motion, the appearance probability of optimal matching points in inner layer of step 2-1 and step 2-2 is higher than that in outer layer of step 2-2.The statistics in table also shows that there is a necessary relevance between the motion intensity of macro block and it's ME search layers with optimal matching points.

Based on the above analysis, the optimal matching points are non-uniformly distributed in each search steps for UMHexagonS algorithm. Although UMHexagonS algorithm improved ME encoding speed significantly, it has to calculate all the templates and search points to determine the optimal matching point. In this process, UMHexagonS algorithm ignored the motion characteristics of macro block, the number of search points is still large, exist computational redundancy. So the further optimization of search patterns and search strategies for UMHexagonS algorithm is necessary and feasible.

3. **Principle of The Proposed Fast Me Algorithm.** In order to overcome the time-consuming ME, reduce the search points and the computational complexity in H.264/AVC, a fast ME algorithm for variable block sizes by classifying motion activity of macro-block based on the UMHexagonS algorithm is proposed in this paper. The proposed method is composed of two parts.

3.1. **Prediction of motion vector.** Motion prediction is an important part in the ME. If get a good MV predictor, it means the search center is much nearer the best MV. It will need to calculate and compare less search points and have higher possibility to get the optimal MV. So we utilize four kinds of MV predictors, which are median MV predictor, up-layer MV predictor, corresponding block MV predictor and neighboring ref-frame MV predictor.

Median MV predictor is exploited the spatial relationship of neighbor macro-blocks. It is easy to find that neighbor macro-blocks have similar MV. So use the median value of the adjacent blocks on the left block (block A), top block (block B) and top-right (block C) of the current block (block E) shown in Fig.2 to predict the MV of the current block. The equation of the median predictor is described in equation (1).

$$Median\_preditor = median[MV_A, MV_B, MV_C] \tag{1}$$

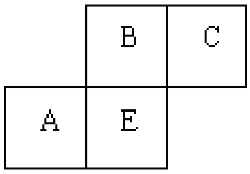


FIGURE 2. Median MV predictor

Up-layer MV predictor shown in Fig.3 is to utilize in the variable block sized ME. In H.264/AVC, test all the 7 types of the current macro-block, choose the partition with lowest cost. Test big partition first, and then turn to smaller ones gradually. The MV search of big partition is a guide for the small partition. It shows the trend of the movement of the macro-block. So the 1616 macro-block's MV can be referenced by 168 or 816 macro-block, etc. The equation of up-layer motion vector predictor is described in equation (2).

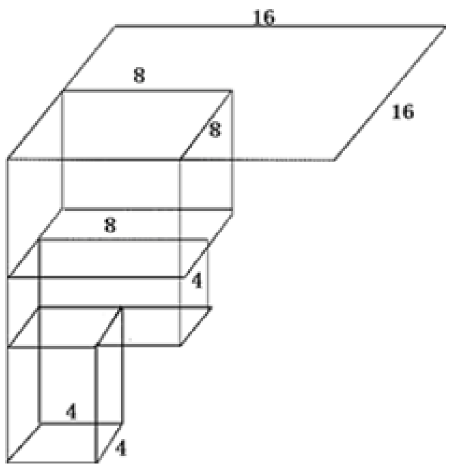


FIGURE 3. Median MV predictor

$$Up\_layer\_preditor = up\_layer[MV] \tag{2}$$

The moving track of a moving object is continuous in the major portion of the video sequence. Corresponding block MV predictor utilizes this characteristic to calculate the MV of the corresponding block in the last frame which is used as one MV candidate. The equation of corresponding block motion vector predictor is described in equation (3).

$$CP\_predictor = lastframe\_MV \tag{3}$$

Reference frame MV prediction is to utilize the temporal relation ship of the same macro-blocks in neighbor frames. The temporal neighbor of reference frame has the

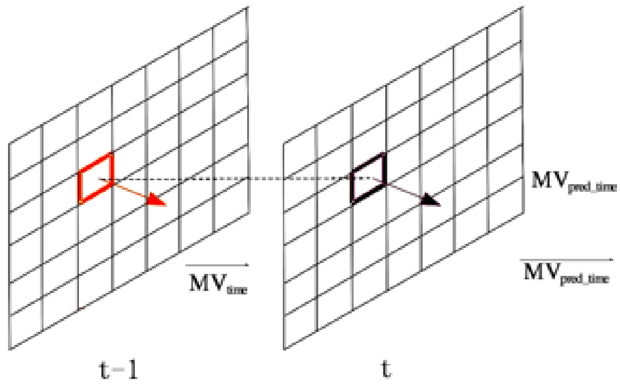


FIGURE 4. Median MV predictor

similar MV. So we can use this similarity to do MV prediction. The current block's MV in reference frame t' can be predicted by scaling the current block's MV in the reference frame t'+1, equation (4) and Fig.5 shows the approach.

$$NR\_predictor = MV_{NR} \times \frac{t-t'}{t-t'-1} \tag{4}$$

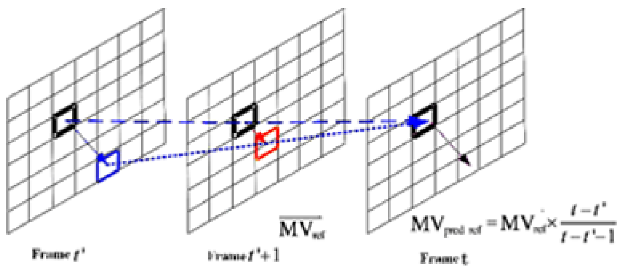


FIGURE 5. Median MV predictor

3.2. **Flexible search criterion.** After the initial search center having been predicted high accurately, adaptive search pattern will be selected according to the motion activity of the macro-block based on the original search pattern of UMHexagonS algorithm. This part includes three main techniques.

1) Modify the search pattern: For original search pattern of UMHexagonS algorithm in the step2-2, it can be find that the uneven multi-hexagon-grid search adopts hexagon-grid search pattern with 16 points all the time. Assume that 16 points of the outmost layer can satisfy the search demand, it will be redundant for the benmost layer still adopts 16 points in the search criterion. So in the proposed algorithm, we modified the original uneven multi layers-hexagon-grid search pattern shown in Fig.6. In step2-2, the search points of each layer will increase with the extent of the search radius. From inside to outside layer, the search points are 8, 8, 12 and 16 respectively. The modified search pattern can maintain the search precision and reduce the unnecessary search points to enhance the efficiency of the ME.

2) Search layer changed according to motion activity: According to the RD (Rate-Distortion) cost[6] of MVs in adjacent blocks, the motion search pattern is classified to three categories: low, medium, and high motion activity. The categories determine the search pattern respectively shown in Fig.6. The detail search strategy is as following.

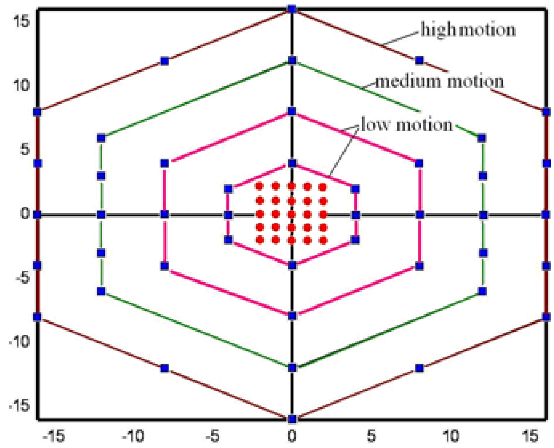


FIGURE 6. Median MV predictor

a) In case of low motion activity: The uneven 2 layers-hexagon-grid search pattern is used because it is expected that the optimal current MV would be near the origin.

b) In case of medium motion activity: The uneven 3 layers-hexagon-grid search pattern is performed.

c) In the case of high motion activity: It is easily expected that the optimal motion vector would be far from the initial search center, so the uneven 4 layers-hexagon-grid search pattern is selected for search pattern.

The motion activity of the current macro-block is defined as follows.

$$\begin{aligned} motion\ activity &= low; & minRDcost &< (1+\varepsilon)\times predminRDcost \\ motion\ activity &= medium; & (1+\varepsilon)\times predminRDc &< minRDcost \\ & & &< (1+\delta)\times predminRDcost \\ motion\ activity &= high; & minRDcost &> (1+\delta)\times predminRDcost \end{aligned} \tag{5}$$

Here,minRDcost is the minimum RD cost of the current MV,predminRDcost expresses the minimum RD cost of the prediction MV.
ε and$\delta$are the adjustable coefficient[7].

$$\varepsilon = \frac{Bsize\ [blocktype]}{pred\,min\,RD\,cos\,t^2} - \alpha_{Radii\,1}[blocktype]$$

$$\delta = \frac{Bsize\ [blocktype]}{pred\,min\,RD\,cos\,t^2} - \alpha_{Radii\,2}[blocktype]$$

$$\alpha_{Radii\,1}[blocktype] = [-0.23, -0.23, -0.23, -0.25, -0.27, -0.27, -0.28]$$

$$\alpha_{Radii\,2}[blocktype] = [-2.39, -2.40, -2.40, -2.41, -2.45, -2.45, -2.48]$$

3) Selected 55 full search technique: In the original UMHexagonS, after doing initial search center prediction, it is expected that the optimal MV would be close to the origin search center, so it do 55 full search primarily in step2-1. But according to the analysis above, if the motion activity is higher, the optimal motion vector is not near the origin search center, the 25 search points are unnecessary and time-consuming. So in the proposed algorithm, only when the motion activity belongs to the low motion, 55 full search (shown in Fig.6) will selected.

4. **Experimental Results and Comparison.** In order to test the effectiveness of the proposed fast ME algorithm in this paper, choose 6 video sequences consisted of different degrees and types of motion content in QCIF format[8], those are fast motion sequence Coastguard, middle-speed motion sequence Forman, slow motion sequence Miss America and Akiyo, with a lot of detail and scene horizontal motion sequence Mobile. The proposed algorithm is integrated with JM 12.2 of the H.264/AVC software for verification with the simulation experiment parameter setting is as follows. Each test sequence contains 100 frames, IPPP structure. The quantization parameter QP is fixed at 28, B frame option is turn off, 5 reference frames, search range is 16 pixels, use Hadamard transform and CAVLC entropy coding. Compare the performance of the proposed algorithm with that of FS and UMHexagonS algorithms in ME encoding time, Bit-rate and PSNR-Y. The simulation results show that the proposed algorithm consistently produces good performance of motion estimation time. Compared with UMHexagonS, it has saved 23.11% ME time, PSNR-Y is compatible to that of FS, while maintaining the same coding efficiency level. Table shows the compared simulation results of above three algorithms. The calculation of ME-time%, Bit-rate% and PSNR-Y is shown in equation (6):

$$\Delta Time = \frac{(ME\, time\, of\ FSor\, UMHexagonS\, algorithm) - (ME\ time\, of\, proposed\, algorithm)}{ME\, time\, of\, FSor\, UMHexagonS\, algorithm} \times 100\%$$

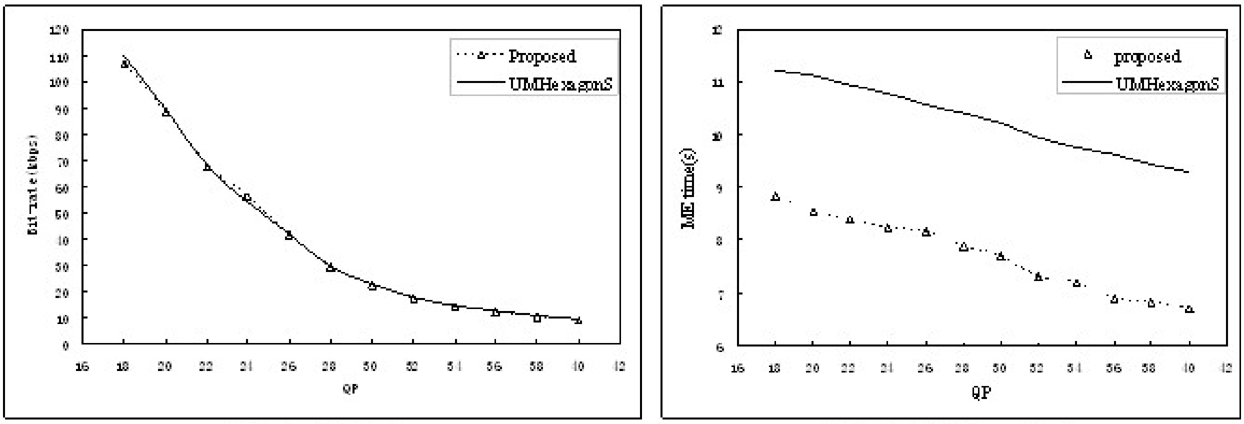
$$\Delta Bit - rate = \frac{(Bit - rate\, of\, FSor\, UMHexagonS\, algorithm) - (Bit - rate\, of\, proposed\, algorithm)}{Bit - rate\, of\, FSor\, UMHexagonS\, algorithm} \times 100\%$$

$$\Delta PSNR - Y = (PSNR - Y\ of\ FSor\, UMHexagonSalgorithm) - (PSNR - Y\ of\, proposed\, dalgorithm) \tag{6}$$

Table 2. THE COMPARED SIMULATION RESULTS OF THE PROPOSED ALGORITHM WITH FS AND UMHEXAGONS ALGORITHM

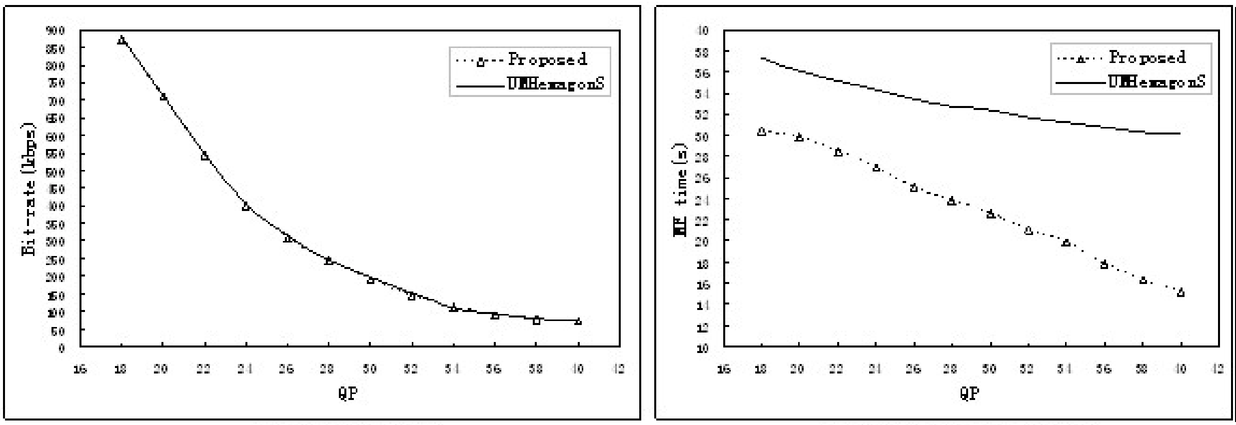
Compared Algorithm / Simulation Results	Full Search algorithm			UMHexagonS algorithm			Proposed algorithm		
	ME-time (sec)	Bite-Rate (kbps)	PSNR-Y (dB)	ME-time (sec)	Bite-Rate (kbps)	PSNR-Y (dB)	ME-time (sec)	Bite-Rate (kbps)	PSNR-Y (dB)
1.Coastguard	131.989	244.85	34.01	32.678	245.66	34.01	23.846	245.16	34.02
2.Forman	128.340	132.62	36.44	24.251	131.50	36.43	18.981	131.87	36.43
3.Akiyo	128.581	29.47	38.25	10.397	29.53	38.28	7.986	29.45	38.28
4.Miss America	132.322	32.39	40.15	11.099	32.16	40.13	8.765	32.48	40.13
5.Mother & Daughter	131.103	46.99	37.44	14.472	46.88	37.40	11.154	47.11	37.43
6.Mobile	128.202	422.85	33.34	27.861	420.01	33.34	21.521	421.83	33.34
Performance comparison with FS, UMHexagonS algorithms on average	ΔME-time (%)	ΔBite-Rate (%)	ΔPSNR-Y (dB)	ΔME-time (%)	ΔBite-Rate (%)	ΔPSNR-Y (dB)	**Annotations:** ME-time: Motion Estimation time —: almost the same level "-" means reduce "+" means increase		
	-88%	—	—	-23.11%	+0.29%	+0.007			
Sequences	 1	2	3	4	5	6			

Under different QP conditions, the contrasts between the proposed algorithm and UMHexagonS algorithm in Bit-rate and ME encoding time for Akiyo sequence with low motion activity and Coastguard sequence with high motion activity are shown in Figure 7 and Figure 8. The graphics show that the proposed algorithm can not only keep the encoding advantage of low compression bit-rate and high reconstructed quality of H.264/AVC, but also improve ME real-time encoding performance. The graphics are smooth, which indicate that the optimization results are obvious and stable, the proposed algorithm can reduce the complexity of ME architecture effectively while maintaining an unnoticeable quality loss in terms of PSNR and bit-rate compared with UMHexagonS algorithm.



(a) Contrast in Bit-rate (b) Contrast in ME encoding time

Figure 7. Contrast of the two algorithms in Bit-rate and ME encoding time (encoding Akiyo sequence)



(a) Contrast in Bit-rate (b) Contrast in ME encoding time

Figure 8. Contrast of the two algorithms in Bit-rate and ME encoding time (encoding Coastguard sequence)

5. **Conclusions.** A fast ME scheme and a concise ME template based on the characteristics of macro-block and visual perception for H.264/AVC encoder was proposed in this paper, which matches different motion contents of video sequence. It can not only find the initial search center accurately, but also modified the search pattern and reduce the search points further to enhance the ME efficiency. Simulation experimental results indicate that, the proposed algorithm achieves the significant calculation burden reduction with almost the same level in PSNR performance compared with that of FS and

UMHexagonS algorithms. The proposed method is a very efficient and robust ME algorithm with low-complexity for real-time video coding applications. The fast speed-up performance and unnoticeable quality losses make the proposed search criterion outperform the famous UMHexagonS ME algorithm.

Acknowledgment. The research work is supported by the National Key Specialized Project Subject of China under Grant No. 2009ZX07212-003-02 and the National Natural Science Foundation of China under Grant No. 30970780.

REFERENCES

[1] Joint Video Specification (ITU-T Rec. H.264/AVC/ ISO/IEC 14496-10 AVC) Joint Committee Draft.Joint Video Ream (JVT) of ISO/IEC MPEG and ITU-T, *Proc. of 7th Meeting: Pattaya*, pp. 7-14, March. 2003.

[2] Gary J. Sullivan, Thomas Wiegand, Thomas Stochammer, Using the Draft H.26L Video Coding Standard for Mobile Applications, *Proc. of IEEE International Conference on Image processing*, Thessaloniki, Greece, Sep. 2001.

[3] Lee Szu-wei, C. C. J. Kuo , Low-complexity H.264/AVC motion compensated prediction for mobile video applications, *Proc. of The International Society for Optical Engineering*, pp. 2-4, Aug. 2010.

[4] Zhibo Chen, Peng Zhou, Yun He, Fast Integer Pel and Fractional Pel Motion Estimation for JVT, *Joint Video Team (JVT) of ISO/IEC MPEG and ITU-T VCEG*, pp. 5-13 , December. 2002.

[5] C. C. Lin, Yinyi Lin, and H. J. Hsieh, Multi-direction search algorithm for block motion estimation in H.264/AVC, *Proc. of IET Image* , vol. 3, no. 2, pp. 88-99, April 2009.

[6] Qiang Tang, Efficient motion re-estimation with rate-distortion optimization for MPEG-2 to H.264/AVC transcoding, *IEEE Trans. Circuits and Systems for Video Technology*, vol. 20, no. 2, pp. 262-274, Feb. 2010

[7] G. J. Sullivan, T. Wiegand, Rate-Distortion Optimization for video compression, *IEEE Trans. Signal Processing Magazine*, vol. 15, no. 6, pp. 74-90, Nov. 1998.

[8] JM Reference Software[EB/OL]. 2006, *http://bs.hhi.de/ suehring/tml*