Review on The Latest H.264 Motion Estimation Techniques

Arief Affendi Juri
School of Microelectronic Engineering
Universiti Malaysia Perlis
s0930110441@studentmail.unimap.edu.my

Asral Bahari Jambek
School of Microelectronic Engineering
Universiti Malaysia Perlis
asral@unimap.edu.my

***Abstract*— Motion estimation is the most time consuming part in the H.264 video coding standard. However, this part is vital where it determines the quality and speed of an encoding process. The UMHexagonS algorithm has been adopted for H.264 fast search algorithm to overcome the problem on encoding speed. In thorough study on the UMHexagonS algorithm, there are numerous methods to improve the UMHexagonS speed whilst maintaining the quality of the encoding process. This paper reviews several techniques that can be implemented in each step in the UMHexagonS algorithm. The results show that the UMHexagonS performance can be improved further by 1.06% to 17.31% whilst maintaining good image quality.**

Keywords-UMHexagonS; H.264; motion estimation; octagon pattern; irregularity-cross; hexagon pattern

I. INTRODUCTION

Motion estimation (ME) is a vital part in video compression since it takes most of the encoding time and it determines the quality of the compressed video. Due to its high computational complexity, ME consumes 70% to 90% of the total encoding time [1]. Due to this problem, fast search algorithm plays a major role in improving the encoding speed.

Even though full search algorithm (FS) will give the most accurate result, it's overly time consuming process makes it unsuitable for a real time video application. Fast search algorithm significantly improves the encoding speed with negligible loss in picture quality. There are many fast search technique that have been proposed such as three step search (TSS)[3], new three step search (NTSS)[4], four step search (FSS)[5], diamond search (DS)[6] and hexagon-based search (HEXBS)[7]. These techniques reduce the number of candidates by choosing only the best possible candidate location. These fixed pattern search algorithm will effectively reduce the number of candidates, and thus reducing the computational load. However, these algorithms tend to be trapped in local minimum point when the motion does not match well with the predefined pattern. This will greatly significantly degrade the video quality.

To overcome this problem, fast hybrid algorithms were introduced such as Prediction Motion Vector Field Adaptive Search technique (PMVFAST) and Unsymmetrical-Cross Multi-hexagon-grid Search (UMHexagonS)[8]. These algorithms combine several fixed technique algorithms to balance the encoding speed and the quality of the video compression.

II. UMHEXAGONS

UMHexagonS has been adopted as fast search algorithm for H264 video compression standard since it performs well in video with small motion and also with large motion situation. UMHexagonS accuracy and rate distortion performance is very close to FS while maintaining the computational complexity low, up to one tenth of the FS [2]. The hybrid algorithm UMHexagonS consists of five steps. These are initial search point decision, unsymmetrical-cross search, small rectangular full search, uneven multi-hexagonal-grid search and extended hexagon-based search [2].

In the first step, initial search point prediction is performed using five different types of prediction. These are median prediction, (0,0) prediction, uplayer prediction, corresponding block prediction and neighboring reference picture prediction. The best match amongst the prediction is chosen as the initial search point. The second step performs unsymmetrical-cross search where the horizontal is twice of vertical search. For the third step a small rectangular full search is done within a 5x5 search area of the search center. Uneven multi-hexagon-grid search is performed in the fourth step. This step has the highest computational complexity as it contains 16 points per hexagon as shown in Figure 4. With a total of four hexagon rings, the total search point in this step is 64 points. The last step performs extended hexagon-based search where small hexagon search is performed repeatedly until minimum rate-distortion lies in the center of hexagon. This will be followed by a small diamond search pattern which is done for final tuning.

III. UMHEXAGONS' IMPROVEMENTS

A number of improvements have been proposed for UMHexagonS algorithm, each of these focusing on certain step in the algorithm. Most of the proposed algorithm focuses on increasing the speed by reducing the number of search points. In this section, these proposed techniques and the impact to the overall UMHexagonS performance will be discussed.

10.1109/ICEDSA.2011.5959075

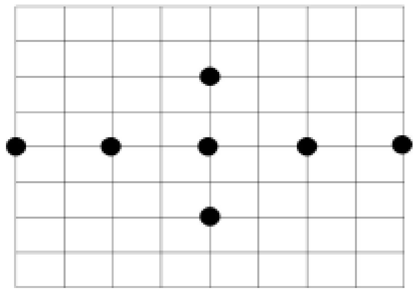

Figure 3. Unsymmetrical-cross search pattern

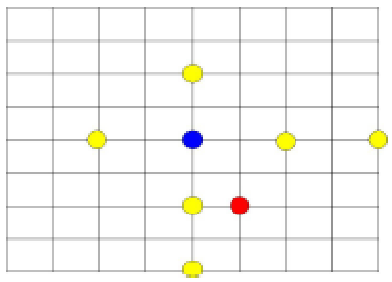

Figure 4. Irregularity-cross search pattern: Case 1

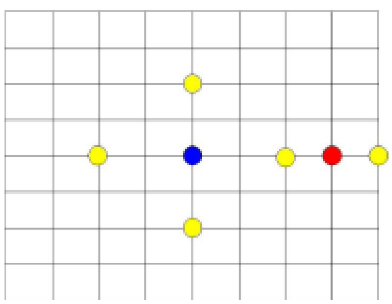

Figure 5. Irregularity-cross search pattern: Case 2

A. Irregularity-Cross Search Pattern

As discussed in the previous section, , the unsymmetrical-cross search is performed within the UMHexagonS algorithm with the horizontal search range as twice of the vertical search range as shown in Figure 1.

The horizontal range is set longer based on the assumption that most of the motion vector (MV) in aggressive video sequence are on the horizontal plane. But with emphasis only on the horizontal plane, the aggressive motion on the vertical plane will be less emphasized. An improvement of this problem has been proposed in [2] where it considers the motion trend to predict the possible area of the motion vector. In other words, the improvement analyses the initial search point and the current location of the block.

In this irregularity-cross step search, the authors divide the cross into four quadrants as shown in Figure 2 where the red point represents the current block, the blue point represents the initial search location and the yellow points represent the search points for this step. As the current block is located in

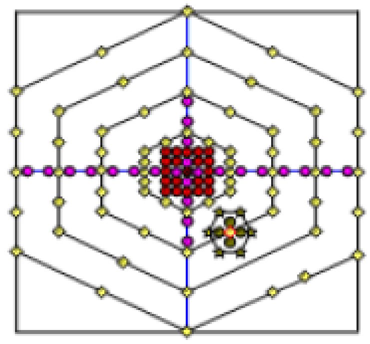

Figure 1. Hexagon search pattern

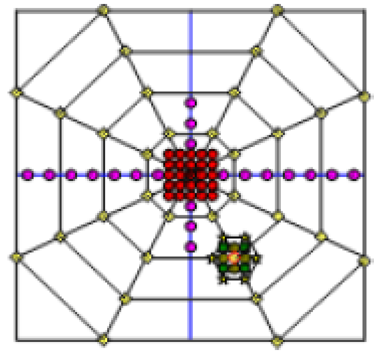

Figure 2. Octagon search pattern

the fourth quadrant, the search range will be longer towards this quadrant as shown in Figure 2. Thus, the match location will be located more likely between the current block and the initial search location. Another situation that might occur using this method is that the current block is located exactly between the two quadrants. In this case, only the range of the related arm is set longer as shown in Figure 3, which will results in less search points since only one arm is longer compared to other cases.

B. New Square Search Pattern

As mentioned in the previous section, small full search will perform the search within 5x5 of the search center which makes a total of 25 search points. More than 80% of the motion vector is located within the 5x5 region around the search center and more than 70% is located in the 3x3 region [10]. With only approximately 10% difference, the 3x3 search area is a better choice as this choice reduces more than 60% of the total search points in this step.

C. Multi-Octagon-Grid Search

Uneven multi-hexagon-grid search consists of 16 search points per hexagon. This results in 64 search points with four hexagon rings as shown in Figure 4. To reduce the search point, an octagon-grid search is proposed in [9]. Using this method, the search points per ring are eight with a total of 32 search points. This is a 50% reduction compared to the original algorithm. The points are hereby distributed evenly as

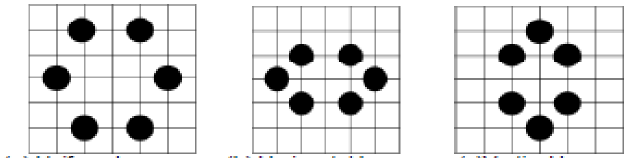

Figure 6. Three hexagonal pattern

shown in Figure 5 and this search method is particularly suitable for high or aggressive video motions.

D. Horizontal and Vertical HexagonSearch

For the last step of the UMHexagonS algorithm, an improvement proposed in [10] is added to increase sensitivity towards non-uniform distribution of gradient descent direction. This proposed algorithm can adaptively adopt different search pattern as shown in Figure 6.

When the current block size is 16x16 or 8x8, the uniform hexagon pattern as shown in Figure 6(a) is used. For 16x8 or 8x4 macroblocks, the horizontal hexagon pattern as shown in Figure 6(b) is adopted. On the other hand, vertical hexagon as shown in Figure 6(c) will be used for 8x16 or 4x8 block size. If a 4x4 macroblock is used, the block will be finalized with small diamond search as the last step for UMHexagonS algorithm.

IV. Result And Analysis

In this paper, the simulation is performed on Intel (R) Core (TM) i5 processor 430M (2.26GHz, 3MB L3 Cache), 2GB RAM and Windows 7 Home Premium 64-bit using Cygwin compiler. The parameters for the simulation are listed in Table 1. The simulation is based on four test video sequences: Foreman, News, Bus and Football with each of them using CIF format (352x288 pixels).

Most of the proposed techniques focus on reducing the ME time while maintaining the picture quality and the bitrate. The test videos are divided into two categories: videos with low video motion and with aggressive video motion. Foreman and News sequence belong to low video motion, while Bus and Football belong to aggressive video motion.

As shown in Table 1, for low motion videos, the improvement in ME computation times vary from 0.64% to 3.03% for individual techniques and up to 6.05% when these techniques are applied together. The highest increase in PSNR is from Multi-Octagon Grid Search for about 0.022dB while the highest decrease in PSNR is from New Square Search Pattern for only -0.01dB. For the bitrate, the highest increase and decrease are only about 1.69kb/s (0.43%) from Multi-Octagon Grid Search and -0.61kb/s (-0.28%) from New Square Search Pattern respectively, which will create only a little or no noticeable effect.

For the aggressive motion, the improvement in ME computational time vary from 0.04% to 11.77% for individual

TABLE I. H.264 Parameter For The Simulation

Profile	Baseline
Level	4.0
Codec	JM17.2
Image Format	CIF (352x288 pixels)
MV Search Range	32
Frame Rate	30 fps
RD optimization	On
Total Number Of Reference	5
Sequence Type	IPPP
Entropy Coding	CAVLC
Encoded Frames	100

techniques applied separately and up to 17.31% when these techniques are combined . In most simulation, the PSNR will have an increase up to 0.015dB from Multi-Octagon Grid Search and although in some simulation the PSNR will decrease about -0.013dB from Multi-Octagon Grid Search. This is acceptable as these values will not have much noticeable difference for the human eye. For the bitrate, the highest increase and decrease are about 9.36kb/s (0.78%) from Multi-Octagon Grid Search and -0.47kb/s (-0.04%) from Irregular-Cross Search Pattern respectively.

Most of the proposed techniques give a better result when applied to aggressive motion especially the Multi-Octagon Grid Search. The highest improvements are obtained from the Multi-Octagon Grid Search technique followed by New Square Search Pattern. The least improvement results are from Irregular-Cross Search Pattern followed by Horizontal and Vertical Hexagon. This is because Multi-Octagon Grid Search reduces 50% of the original step search points while New Square Search Pattern reduces more than 60%, compared to Irregular-Cross Template and Horizontal and Vertical Hexagon which focus on relocating the search points.

V. Conclusion

In this paper, the improvement of each step of the UMHexagonS was discussed. By modifying each step in the UMHexagonS algorithm, the existing standard UMHexagonS algorithm can be improved further. The simulation results show that the UMHexagonS algorithm can be improved by 1.06% to 17.31%. Even though each individual technique provides only a small improvements, the combination of all techniques results in a significant improvement on ME encoding time while still maintaining high-quality picture.

TABLE II. COMPARISON BETWEEN UMHEXAGONS AND PROPOSED IMPROVEMENT

		PSNR(db)	Bit Rate (kb/s)	ME Time(ms)
UMHexagonS	Foreman	36.855	394.95	327.91
	Football	36.696	1566.99	628.63
	News	38.35	214.86	223.23
	Bus	34.931	1202.44	463.58
Irregular-cross search pattern	Foreman	36.858	394.93	328.68
	Football	36.705	1566.82	628.34
	News	38.345	214.73	221.81
	Bus	34.934	1201.97	462.28
New square search pattern	Foreman	36.865	396.30	322.54
	Football	36.70	1568.93	602.18
	News	38.356	214.50	220.85
	Bus	34.921	1202.72	457.08
Multi-octagon-grid search	Foreman	36.877	396.64	317.99
	Football	36.711	1576.38	554.66
	News	38.342	215.34	217.23
	Bus	34.918	1211.80	426.23
Horizontal and Vertical Hexagon	Foreman	36.884	396.42	324.09
	Football	36.706	1569.57	614.38
	News	38.34	214.25	219.52
	Bus	34.934	1204.75	457.15
Combined Improvements	Foreman	36.865	397.49	308.07
	Football	36.715	1570.03	519.81
	News	38.342	214.40	214.42
	Bus	34.921	1210.48	408.81

REFERENCES

[1] Z.Chen, P.Zhou, and Y.He, ”Fast Integer and Fractional Pel Motion Estimation for JVT,” JVT-F017, Awaji, Japan, December, 2002,pp.5-13.

[2] Peng Huang, and Cui-hua Li, “Irregularity-cross multi-hexagon-grid Search Algorithm for Fast Motion Estimation on H.264” in 2nd International Conference on Computer Engineering and Technology 2010 (ICCET), vol. 3, pp. 587-592.

[3] T. Koga, K. Linima, A. Hirano, and T. Ishi-guro, “Motion-compensated interframe coding for video conferring,” in Proceedings of – National Telecommunication Conference, pp. C9.6.1-C9.6.5, New Orleans, USA, Nov. 1981.

[4] C. Karthikeyan. "Performance Analysis of Block Matching Algorithms for Highly Scalable Video Compression", International Symposium on Ad Hoc and Ubiquitous Computing 2006 (ISAUHC ‘06), pp. 179-182, Dec 2006

[5] Lai-Man Po and Wing-Chung Ma, “A novel four-step search algorithm for fast block motion estimation” IEEE Transactions on Circuits And Systems For Video Technology, vol. 6, pp.313-317, June 1996.

[6] J. Y. Tham, S. Ranganath and A. Al. Kassim, “A novel unrestricted center-biased diamond search algorithm for block motion estimation,” IEEE Transactions on Circuits and Systems for Video Technology, vol.8, pp.369-377, Aug 1998

[7] C. Zhu, X. Lin and L. Chau, “Hexagon-Based Search Pattern for Fast Block Motion estimation,” IEEE Transaction on Circuits and Systems for Video Technology, May 2002, pp.349-355.

[8] Zhibo Chen, Jianfeng Xu, Yun He and Junli Zheng, “Fast integer-pel and fractional-pel motion estimation for H.246/AVC” Journal of Visual Communication & Image Representation, April 2006, pp. 264-290.

[9] Zhu Weixing, Chen Xianyong and Li Xincheng, “A New Search Algorithm Based On Multi-Octagon-Grid” in CISP '09. 2nd International Congress Image and Signal Processing, 2009,pp.1-5.

[10] Li Hong-ye, Liu Ming-jun and Zhang Zhi-qiang, “ A New Fast Motion Estimation Algorithm Based on H.264” in MINES '09. International Conference on Multimedia Information Networking and Security, 2009, vol. 1,pp.287-290.