Motion Estimation Algorithm for HEVC Suitable for Hardware Implementation

SHIKAI ZUO, MINGJIANG WANG, LIYI XIAO
Microelectronics Center, Key Laboratory of the Technology of the Internet of Things
Harbin Institute of Technology
Harbin, Heilongjiang
China
zuoshikai@aliyun.com, http://www.hitsz.edu.cn

Abstract: - HEVC being developed in the video coding standard(The video coding standard of HEVC being developped) is a new generation of video coding standards for HD resolution video encoding. It is difficult to implement HEVC real time coding in current hardware for the algorithm's complextivity, so the algorithm optimization and hardware acceleration of HEVC become research hot spots. The motion estimation unit of HEVC occupy more than 50% of the computation time in the ITU-T standard coding software, so the algorithm optimization can reduce the encoding time largely. Therefore，this paper proposes a new parallel processing method that the LCU and the PU block are divided for computing motion estimation parameters at the same time. The test results shows that, PSNR is decreased 0.005db and compression rate is decreased 0.026 percentage , the algorithm can increase the processing speed of motion estimation by 77.2%, compare with full search algorithm. The paper algorithm is suitable for hardware implementation, for the parameters of various motion estimation block in the same LCU can be computed parallelly and memery access is regular.

Key-Words: - Video Coding, HEVC/H.264 encoder, Fast Motion Estimation, 1080P, Parallel Computing

1 Introduction

In order to meet the higher video quality and video coding requirements of higher resolution , ISO / IEC Moving Picture Experts Group (MPEG) and ITU-T Video Coding Experts Group (VCEG) have proposed the latest video coding standard HEVC (H.265)[1-2]. The motion estimation is used to eliminate redundant time of the video sequence within the video coding standard HEVC [3]. While studying the standard software HEVC, owing to the complexity of the encoding process, large amount of computation, during the encoding of the video image process, the time of the motion estimation counted over 50% calculation of the whole encoding system [4], which makes it difficult to realize high-quality real-time video communication. Thus, process optimization of motion estimation for the HEVC standard encoding algorithm becomes the research focus nowadays. Meanwhile, motion estimation needs large amounts of computing absolute value of the corresponding pixel, and the existing computing platforms are time-consuming; therefore, in order to realize real-time communication of high-quality, it is of great necessity to develop IP-related motion estimation.

Currently, researchers focus on algorithm optimization of HEVC motion estimation. One is to reduce the predicted block and to lower the complexities of encoding process which are based on inter-frame or intra-frame temporal image correlation, as described in the paper [5-8]. In paper [5], the CU block correlation of the same LCU block can filter part of the CU depth segmentation, so as to lower the complexity of the motion estimation algorithm. Although the algorithm has reduced the computational complexity of motion estimation, but part of the pixel matching operation is still repeated while calculating CU block motion matching of LCU block and only a portion of redundancy was reduced. The other motion estimation process is serial calculation that tree extract the LCU block first and then successively calculate the best match vector under each depth. As for the LCU(64×64)which has not eliminate the part depth Cu block, while the maximum depth is 3, rate -distortion cost of the prediction block requires 1106 times. In this paper, although part of the recycle has been reduced, it is not conducive to realize hardware of motion estimation algorithm. Similarly the paper [6] adopts the same method to end the TZ search searching process in advance. The drawback is the same to paper [5]. In addition, since the powerful searching method of TZ search makes it irregular to read the pixel of the searching area, which makes is more difficult to realize the hardware implementation. In paper [7], the redundancy is reduced on the basis of temporal and spatial correlation of each frame, in particular, to judge the intensity of the movement on the basis of trajectory of several frames. If the prediction block moves vigorously, the corresponding matched area need to be increased. If moves slowly, then the area need to be reduced. Although the algorithm has reduced the amount of calculation during the motion estimation process, still need calculate 1106 times. Method in this paper can reduce the predicted time, but owing to large amounts of serial calculation, it is not suitable to realize the hardware. In paper [8], temporal and spatial correlation of intra-frame Cu block is used to predict on the basis of minimum 8×8 CU block, then to merge the minimum CU block by the motion vector to obtain the best motion prediction mode parameters. Although this method is conductive to hardware implementation of motion estimation, the video bit which is compressed by this method averagely increase 3% and have greater transmission bandwidth requirements. Method of the second category is around the best matching points. Singlet phenomenon of motion estimation can reduce the searching points during the moving process in order to reduce the calculating amounts when the compression ratio stays still and to improve the motion estimation speed. In paper [9-11], the searching path of cross search, diamond search, and sub-pixel search have been studied. The method also goes to HEVC encoding standard reference software HM which is applied by the official, which becomes the TZ searching method for the combination of power search and diamond search [12]. These methods have increased the search speed, and reduced the amount of computation, but still cannot being realized since the serial search method and the search area data are irregularly read.

Obviously motion estimation method hereinabove is not suitable for HEVC motion estimation algorithm in hardware. It is of great necessity to develop new ways to resolve these conflicts. Researches show that parallel motion estimation search method is more efficient than serial search. Therefore, the motion estimation calculation can raise 10 times computing efficiency while changing the serial to parallel computing.

Based on HEVC agreement, a series of coding parameters of every LCU must be determined while the motion estimation unit process rate distortion mode. These parameters include the CU tree segmentation mode; PU segmentation mode, and the best post- match vector motion segmentation. When encoding the video, the encoder needs to traverse all the depth of CU blocks, each block CU or PU need to traverse all possible segmentation which made the motion estimation high complexity in coding complexity. To reduce complexity, this paper will calculate LCU tree segmentation and motion prediction coding parameters PU block parallel execution, also using dynamic motion estimation motion vector search area in order to improve prediction accuracy.

2 Parallel analysis of HEVC inter-frame motion estimation algorithm

In video encoding systems, the motion estimation (ME) can remove most inter-frame redundancy, so a high compression ratio can be realized [13]. Among various motion estimation algorithms, fast full-search algorithm is usually used because of its perfect effect and regular computation.

2.1 HEVC basic computational unit

Compared with the video coding standard H.264, HEVC has adopted more flexible coding structure, including coding unit (CU), predict unit (PU) and transform unit (TU), among which CU is the basic unit of video coding and CU can be further divided into the size of 64×64 ; PU is the basic unit for estimation of intra- and inter-frame [14]. In HEVC, Four fork tree segmentation technology of motion estimation unit and asymmetric division technology of PU unit are also greatly enhancing the complexity of calculation, which counts the major causes for the rising of HEVC calculation. Those complicated calculation is no good for the application of video chatting.

2.2 Quadtree Partitions Of LCU

Compared with the H.264 encoding algorithm, HEVC had adopted larger encoding unit size, of which the biggest size is 64×64, meanwhile, largest coding unit (LCU) does recursive process to CU by quad tree. The largest coding depth of CU is 3. Take CU for example, the depth of 0 stands for undivided one; while depth is 1, LCU is divided into 4 CU of size 32×32, the dividing point is called leaf node; while depth is 2, every 32×32 CU is divided into 4 CU of size 16×16; while depth is 3, every 16×16 CU is divided into 4 CU of size 8×8[15]. The motion estimation unit takes every divided part after every depth as processing unit to process the optimized PU and TU dividing and parameter calculation.

Various computing unit and dividing method make HEVC motion estimation more complex. Based on HEVC agreement, while the depth is 3, the LCU needed to be divided into 85 CU blocks of different sizes by tree dividing method[16]. Take the divided CU block to process calculation of motion vector, and then carry rate distortion calculation. From the analysis hereinabove, we know that during the tree dividing process, every smaller CU block is the sub-block of the larger CU; its matching price has high and strong intensity. Thus, it is possible to calculate the motion estimation parameter of every CU block by the parallel method. Only in this way, the 85 times CU block parameter calculation can be turned into 1, meanwhile, the motion estimation time is greatly saved and the energy-consuming of the processor is reduced.

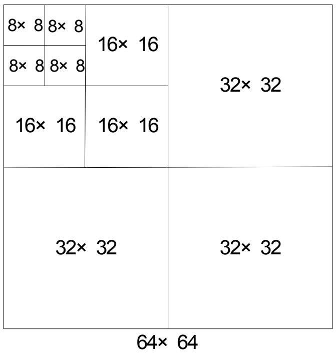

Fig.1. The Quadtree Partitions of LCU

2.3 CU motion estimation mode

As showed in figure 2, motion estimation unit adopts skip of each depth , four symmetric split mode 2N × 2N, 2N × N, N × 2N, N × N, and four kinds of asymmetric division model 2N×nD, 2N×nU, nL×2N, nR×2N to predict coding[17]. General speaking, CU in every layer need to calculate 14 predicting blocks, of which including one skip, 1(2N×2N), 2 (N×2N), 2 (N×2N), 2 (2N×N), 2 (2N×nD), 2(2N×nU), 2(nL×2N), 2(nR×2N). CU of the maximum depth should calculate another 4 (N×N) block, i.e., CU at maximum depth needs 18 predicting block. As to the depth of 3 for LCU (64×64) , totally 1106 predicting block is needed.

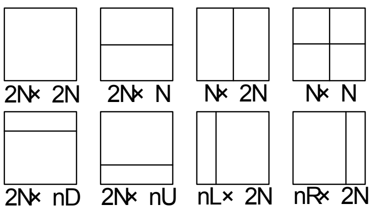

Fig.2. The segmentation of PU

If HEVC standard reference software algorithms for solving an optimal motion estimation parameters LCU, after 1106 cycles needed to get the best split mode and CU PU best split mode, a huge amount of computation. However, the encoder unit CU calculate various split mode, the mode can be based on various split between the matching reference pixel with a pixel array to calculate the difference between the combination, it can be processed in parallel for each division pattern, i.e., the block unit CU of each parallel processing mode is divided Distortion consideration, thereby reducing the redundancy calculation of the matching difference between the pixel and improve the processing speed.

3 Fast Algorithms Suitable for Hardware Implementation

3.1 Basic principles of algorithms

The analyses in chapter 2 show that when we encode LUC block of 64×64, we can merge 1106 circulation calculation into one through parallel processing with the maximum depth being 3. Compared with full search algorithm, it saves much of the calculation time. The using of parallel processing method makes motion estimation algorithm more suitable for hardware implementation.

3.2 Specific Implementation Procedure of Fast Algorithm

This paper decreases the complexity of motion estimation by performing the parallel calculation of CU block of every LCU in HEVC standard algorithm and corresponding PU block parameters. The specific procedure is seen in figure 3. At first, motion estimation model reads the pixel value of the whole LCU (64×64) from memory unit and put it into calculation unit, and then forecast the motion vector of present block according to the motion vector of existing neighboring blocks. Then obtain the pixel data of search region according to this forecasting vector and search region. Get the differentials between total pixel value of LCU and that corresponding to research region, then get the SAD value of every PU model under every CU block by adding the pixel values of all kinds of forecasting model PU blocks under every CU level. Then get the best matching motion vector according to SAD value so as to get the best PU segmentation model under every CU level. Then encode the algorithm and calculate the rate distortion cost of every level of CU block according to the parameters of CU and PU, the PU parameter in the frame we get and the change of corresponding TU block. And get the series of change corresponding to the minimum distortion rate at last.

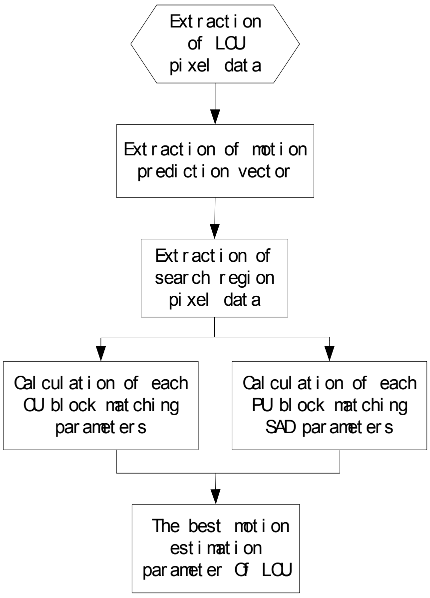

Figure3.The procedure of fast algorithm

Motion estimation needs a large amount of calculation of the differences between present PU block and the pixels in research region. The calculation is seen in formula (1)[18]. $I_k(m, n)$ refers to the pixel value of present PU block, and $I_k(m+dx, n+dy)$ refers to the pixel value in search region, and SAD refers to the sum of difference between the total pixel value of this PU block and that of corresponding search block. We can ascertain the best matching block of PU block by comparing the SAD value and then we will get motion vector, whose calculation is seen in formula (2) , and MV is the best matching vector of PU block[19].

$$SAD(dx,dy) = \sum_{m=x}^{x+N-1} \sum_{n=y}^{y+N-1} | I_k(m,n) - I_{k-1}(m+dx, n+dy) |$$

-- (1)

$$(MV_x, MV_y) = (dx,dy)|_{\min SAD(dx,dy)} \quad (2)$$

Fast algorithm divides LCU into 1024 2×2 pixel unit from PE0-PE1023, which can be seen in figure 4. Every PE is the calculation unit of the calculation of 2×2 pixel matrix SAD. The SAD value under all kinds of segmentation models of PU can be realized by different combination of PE[20]. Take LCU as example, when depth is 3, LCU has the biggest tree segmentation—the divided block of CU is 8×8.

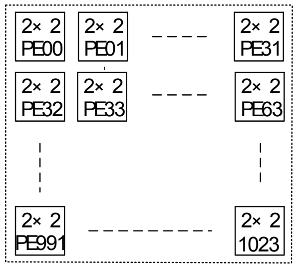

Figure4 LCU calculation unit

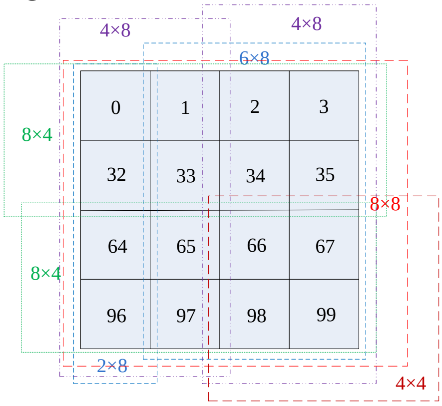

Figure 5 the calculating method of SAD of motion estimation PU

Take tree segmentation depth=3 and the 8×8 CU in top left corner as example, this CU includes PE0-PE3, PE32-PE35, PE64-PE67 and PE96-PE99. The inter frame motion estimation calculation of CU block is seen in figure 5: when PU is one 8×8 block, put all of the sum of the PE unit SAD contained in this block into formula 2, and we can get the best motion estimation value; when PU is divided into 2 8×4 blocks, taking the SAD value of upper part of CU block as example, the value is the sum of PE0-PE3 and PE32-PE35. Similarly we can get the SAD of lower part; when PU is divided into 2 4×8 blocks, taking the SAD value of the left part of CU block as example, the values is the sum of PE0-PE1, PE32-PE33, PE64-PE65 and PE96-PE97. Similarly we can get the SAD of right part; when PU is divided into 4 4×4 blocks, the SAD in the top left corner is the sum of PE0, PE1, PE32 and PE33. Similarly we get the SAD of other PU blocks; when PU is divided asymmetrically—the 8×2 and 8×6 forecasting blocks, SAD(8×2) is the sum of PE0-PE3 and SAD (8×6) is the sum of PE32-PE35, PE64-PE67 and PE96-PE99. Similarly we can get the SAD value of PU blocks under other asymmetric division mode. This algorithm takes the parallelism of hardware into consideration, each PE just needs calculating once, and SAD value under all kinds of PU division mode and depths by the various combinations with other PE, and then we can get the motion vector of PU block in search region.

3.3 Dynamic Search Procedure

To When we use the method above to forecast the search region of present block, it will cause the distortion of motion estimation search region that every sub-CU block and LCU uses the same forecast vector. To overcome this disadvantage, this paper improves the accuracy of motion estimation search region by two methods. One is using step-by-step forecast search region—use 32×32 CU block as the basic unit to estimate motion vector. Then we should test that whether the best matching point is at the edge of search region. If it is, we should expand the search region 2 pixels outward, but if it is not, the search is over. The procedure of dynamic search is seen in figure 6. Divide the LCU into 4 32×32 CUs with depth=1 and mark them as CU10, CU11, CU12 and CU13 respectively. Estimate the motion vector of CU10 according to finished motion estimation on the left, on the top and in the top left corner, which is seen in figure 7: when the motion vector of A, B and C co-exist, its motion vector is the mid-value; if motion vector A and C do not exist, we choose motion vector B; when motion vector motion vector A and B do not exist, we choose motion vector C. In next step, we take the estimate vector of CU10 as center and rectangular search box as the search region to search motion matching parameter of CU of present LCU（64×64）with all depths and record the best matching vector of PU under all kinds of models in next step. To increase the accuracy of motion estimation, we should judge that whether the best matching point of each PU block is at the edge of search region. If it is, we should expand the search region 2 pixels outward. Especially if the best matching point of any PU block is in the 4 corners of search region, we should expand it the two corresponding directions. After that we should calculate every matching parameter in the new search region and refresh the best matching point. If the best matching point is still on the edge, we

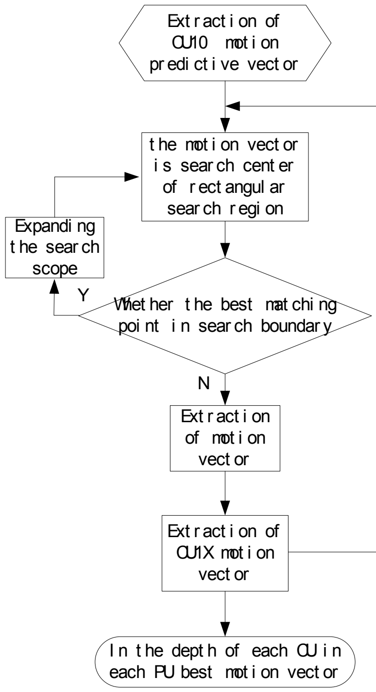

Figure 6 Algorithm procedure of dynamic search region

should continue to expand the search region until all of the best matching points is inside the search region. Save every parameter of LCU which takes the estimation vector of CU10 as searching center, and then take the estimation vector of CU11, CU12 and CU13 as center to search for new search region so as to ascertain every motion parameter of LCU. During the process of calculation, new search region may cover the old ones. This paper reduces the repeated calculations by overlap region elimination method. To sum up, we can get the motion estimation matching parameter of LCU at all depths through this process.

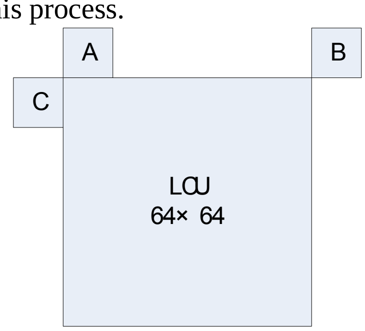

Figure 7 the estimation of the motion vector of CU

4 Simulation Result and Analysis

The Use PC platform to test the algorithm in this paper, and we test three groups of video sequences, which includes three 1920×1080 sequences and three 3840×2160 sequences. We use HM6.0 as test software, 3 as the largest depth of tree segmentation, and 128×96 rectangular region as search range. The result is seen in table 1, where QP is quantization parameter with the quantization value of 28, 32 and 36 respectively. PSNR（peak signal to noise ratio）is peak signal-to-noise ratio and △PSNR refers to the difference between PSNRs under the algorithm in this paper and full search algorithm. △BR(bit ratio) refers to the increase percentage rate after compression. △METM（Motion Estimation Time）refers to the percentage of reduced time under the motion estimation algorithm in this paper and full search algorithm.

Table 1 The Result of Fast Motion Estimation Algorithm

Test Sequence		△BR（%）			△PSNR（db）			△METM（%）		
		QP=28	QP=32	QP=36	QP=28	QP=32	QP=36	QP=28	QP=32	QP=36
1280×720	Vidyo1	0.011	0.008	0.012	-0.002	-0.001	0.000	-85.4	-85.5	-88.2
	Vidyo3	0.015	0.023	0.030	-0.012	-0.002	-0.004	-87.3	-84.0	-86.9
	Vidyo4	0.006	0.009	0.002	-0.001	-0.002	-0.001	-85.1	-88.1	-84.8
1920×1080	Controlled_burn	0.038	0.020	0.026	-0.018	-0.020	-0.005	-88.1	-84.0	-89.3
	Touchdown_pass	0.028	0.017	0.043	-0.001	-0.000	-0.004	-87.9	-86.5	-86.8
	Ducks_take_off	0.082	0.142	0.131	-0.011	-0.006	-0.016	-85.3	-85.6	-85.6
3840×2160	Park_joy	0.002	0.001	0.001	-0.001	0.000	0.000	-90.2	-92.6	-86.7
	Old_town_cross	0.007	0.003	0.005	-0.002	-0.003	-0.003	-90.0	-89.9	-87.9
	Crowd_run	0.010	0.009	0.014	-0.003	-0.002	-0.005	-87.5	-88.2	-88.1
Mean		0.026			-0.005			-87.2		

Table 1 shows that there is an average 0.005db decreasing of , an average 0.085% increasing of output bit rate and an average 87.2% decreasing of time motion estimation algorithm takes when we compare the video sequences of the algorithm in this paper with that of full search algorithm. The result shows that this algorithm achieves the collateral execution and reduces much of the time motion estimation unit takes with constant monomeric efficiency. The test of this motion estimation algorithm to all kinds of videos with various resolution ratio shows that the compression time of full search algorithm is relatively more stable-the leg of video transmitter is stable by comparison, which shows that it is more suitable for real-time video communication.

Table 2 compares the algorithm in this paper with other research results, and the result shows that the algorithm in this paper is better than the fast motion estimation algorithm mentioned in paper [19] but a little worse than that in paper[20] and [21]. It is obvious that this paper develops parallel system algorithm for the hardware achievement of HEVC motion estimation algorithm. But it is hard to achieve the performance of fast parallel calculation because of the limit of experiment software platform, which decreases other algorithm performance.

Table 2 Algorithm Comparison

Sequences	BD-PSNR(dB)	BD-BR(%)	△ME Time(%)
[21]	-0.0013	0.290	-67.2
[22]	-0.080	0.280	-75.9
[23]	-0.060	2.210	-87.0
Proposed	-0.005	0.026	-77.2

This paper, aiming to the hardware achievement latest video encoding standard HEVC motion estimation unit, comes up with motion estimation unit parallel processing method-the process and method to divide the LCU into 4 equal CU blocks and divide every CU block into motion estimation PU block. The result shows that compared with motion estimation full search algorithm, this encoding method can increase the processing speed of motion estimation unit by 77.2% on the base of decreasing PSNR by 0.005db and decreasing compressing rate by 0.0265%. This algorithm structures the data of search region, and neighboring pixels can be reutilized to reduce the number of memory access and the calculation and processing of the best matching point. To sum up, the algorithm in this paper can improve hardware utilization and save hardware cost during the process of hardware achieving, so it is good for the real-time video communication which takes HEVC as standard.

V. Conclusion

Project Name: Research on the Key 3D Technology of NoC Video Encoder Architecture
Project Number: JCYJ20120829102244551
ShenFaGai[2012]1583

References:

[1] ITU. Parameter values for ultra-high definition television systems for production and international programme exchange [S], ITU-T,2012.

[2] Sulivan G J, Ohm J R. Video Coding of ITU-T SG16 WP3 and ISO/IEC JTC1/SC29/WG11 (JCTVC-A200)[C], Joint Collaborative Team on Video Coding (JCT-VC),Dresden,2010.

[3] Pourazad D M T, Douter C, Azimi M. HEVC: the new gold standard for video compression: how does HEVC compare with H.264/AVC? [J]. IEEE Consumer Electronics Magazine, 2012, 1(3): 36-34.

[4] Liang Hong, Weifeng He, Hui Zhu, Zhigang Mao. A full-pipelined 2-D IDCT/IDST VLSI architecture with adaptive block-size for HEVC standard[J], Institute of Electronics Information and Communication Eng, 2013,10(9):1-11.

[5] Liquan Shen, Zhi Liu, Xinpeng Zhang, Wenqiang Zhao, Zhaoyang Zhang. An Effective CU Size Decision Method for HEVC Encoders [J]. IEEE Transactions On Multimedia, 2012, 15(2): 465-470.

[6] Pan Zhaoqing, Zhang Yun, Kwong Sam; Wang Xu, Xu Long. Early termination for TZSearch in HEVC Motion Estimation [C], 2013 38th IEEE International Conference on Acoustics, Speech, and Signal Processing. Canada,2013.

[7] JinWen Hu,Guowei Xi,Yilong Cheng,Ke Yan,Guoping Li. Fast Inter-Frame Predietion Algorithm for HEVC Based on Motion Features [J]. Shanghai University Journal ,2013, 19(3): 246-478.

[8] Peng Jinhu, Cen Feng. Inter-motion merge techniques research of HEVC [J]. Shanghai University journal ,2013, 32(1): 40-44.

[9] N.Purnachand, Luis Nero Alves, Antonio Navarro Fast Motion Estimation Algorithm for HEVC [C], 2012 IEEE Second International Conference on Consumer Electronics, Berlin, 2012, 981.

[10] Sinang Mahmut E. Chandrakasan Anantha P. Sze Vivienne. Zhou Minhua. Hardware-aware motion estimation search algorithm development for high-efficiency video coding (HEVC) standard [C], 2012 19th IEEE International Conference on Image Processing, Lake Buena Vista , 2012, 1529-1532.

[11] Dai Wei, Au Oscar, Pang Chao, Sun Lin, Zou Ruobing, Li Sijin. A novel fast two step sub-pixel motion estimation algorithm in HEVC [C], 2012 IEEE International Conference on Acoustics, Speech, and Signal Processing, Kyoto, 2012, 1197-1200.

[12] HM Reference Software [OL]. Available: https://hevc.hhi.fraunhofer.de/svn/svn_HEVCSoftware.2012.12

[13] Sinangil Mahmut E, Chandrakasan Anantha P, Sze Vivienne, Zhou Minhua. Memory cost vs. coding efficiency trade-offs for HEVC motion estimation engine[C], 2012 19th IEEE International Conference on Image Processing, Lake Buena Vista, 2012, 1533-1536.

[14] Lin Jian-Liang, Chen Yi-Wen, Huang Yu-Wen, Lei Shaw-Min. Motion vector coding in the HEVC Standard[J], Institute of Electrical and Electronics Engineers, 2013,7(6): 957-968.

[15] Matsuo Shohei, Takamura Seishi. Standardization trends for the HEVC next-generation video coding standard[J], Nippon Telegraph and Telephone Corp, 2014,12(2): 67-68.

[16] Saab Farah, Elhajj, Imad H, Chehab Ali, Kayssi Ayman. Energy-efficient HEVC using stochastic processing[C], 2013 4th Annual International Conference on Energy Aware Computing Systems and Applications, Istanbul, 2013, 152-157.

[17] Viitanen Marko, Vanne Jarno, Hamalainen Timo D, Gabbouj Moncef, Lainema Jani. Complexity analysis of next-generation HEVC decoder[C], 2012 IEEE International Symposium on Circuits and Systems, Seoul, 2012, 882-885.

[18] Sullivan Gary J, Ohm Jens-Rainer, Han Woo-Jin, Wiegand Thomas. Overview of the high efficiency video coding (HEVC) standard[J], Institute of Electrical and Electronics Engineers, 2013,22(12): 1649-1668.

[19] ITU-T, SERIES H: Audiovisual and Multimedia systems: Infrastructure of audiovisual services - Coding of moving video, *ITU-T Rec*. 2003.

[20] L. Deng, W. Gao, M. Z. Hu, and Z. Z. Ji, An efficient hardware implementation for motion estimation of AVC standard, *IEEE Transactions Consumer Electron*, Vol.51, No.4, 2005, pp. 1360–1366.

[21] Purnachand N, Alves, Luis Nero, Navarro, Antonio. Fast motion estimation algorithm for HEVC[C], 2012 IEEE 2nd International Conference on Consumer Electronics, Berlin, 2012, 34-37.

[22] Pan Zhaoqing, Zhang Yun, Kwong Sam; Wang Xu, Xu Long. Early termination for TZSearch in HEVC Motion Estimation [C], 2013 38th IEEE International Conference on Acoustics, Speech, and Signal Processing. Canada,2013.

[23] Xiong Jian, Li Hongliang, Wu Qingbo, Meng Fanman. A fast HEVC inter CU selection method based on pyramid motion divergence [J]. IEEE Transactions on Multimedia,2014 16(2): 559-564.