



# Implementation of Motion Estimation Algorithm for H.265/HEVC

Davis P, Sangeetha Marikkannan

M.E. Embedded System Technologies, Department of ECE, KarpagaVinayaga College of Engineering and Technology Chennai, India

Professor and Head, Department of ECE, KarpagaVinayaga College of Engineering and Technology Chennai, India.

**Abstract**—The most essential part of almost all video coding standards such as H.262/MPEG-2, H.264/MPEG-4 and newly emerged High Efficiency Video Coding (H.265/HEVC) is Motion Estimation. HEVC provides better performance in compression and efficiency in coding for Beyond HD and UHD videos. The main drawback is its computation complexity. The complexity is mainly due to Motion Estimation (ME), as it consumes more than half of time to encode. Many fast motion estimation algorithms have been developed so as to reduce the complexity in encoding. This paper proposes the implementation of New Combined Three Step Search pattern for ME algorithm. The simulation result shows average time saving of 50.07% and slight improvement in quality compared with existing patterns. A System is designed using Altera System on a Programmable Chip (SOC) builder and the algorithm is being implemented using Nios-II software build tools to get optimal performance.

**Keywords**—Motion estimation; HEVC; search pattern; Combined Three Step Search; ME time

## I. INTRODUCTION

Due to the growing popularity of High Definition (HD) Video and beyond HD videos, the traffic caused by video applications in today's network is high. Personal computers, Tablets and even mobile devices need to receive and display HD videos, and this become a severe challenge on today's network. Network traffic can be minimized by increasing video compression. HEVC has been designed to address essentially all existing applications and issues of H.264/MPEG-4 AVC. It particularly focuses on two key issues (1) increased video resolution (2) increased use of parallel processing architectures [3].

High Efficiency Video Coding standard was jointly developed by ISO/ICE Moving Picture Experts Group (MPEG) and ITU-T Video Coding Experts Group (VCEG) as ISO/ICE 23008-2 MPEG-H Part-2 and ITU-T H.265. The development of ITU-T and ISO/IEC

organizations are the main cause for the evolution of video coding standards. H.261 and H.263 are developed by the ITU-T, MPEG-1 and MPEG-4 Visual is produced by ISO/ICE, and the two organizations jointly produced H.262/MPEG-2 Video and H.264/MPEG-4 Advanced Video Coding (AVC) standards, which are widely used in variety of product in our daily life [7]. HEVC thus emerged by the joint effort of both ITU-T and ISO/ICE.

HEVC has various improvements related to its predecessor H.264/MPEG-4 (AVC) standard. Table I shows some of the key improvements of HEVC/H.265 standard related to H.264/MPEG-4 (AVC).

Table I Key Improvements in HEVC standards

Category	HEVC
Bit rate	40-50% bit rate reduction with same visual quality compared to H.264
Specification	Supports up to 8k (8192x4320) Supports up to 300 fps
Compression model	Enhanced Hybrid Spatial-temporal prediction model
Block structure	CTU supporting large block structure of 64x64
Intra prediction directional modes	35 directional modes
Improvements	Potential to realize UHD, 2k, 4k for broadcast and online
Other Names	MPEG H, H.265

HEVC standard uses Quadtree based partitioning as shown in Figure 1(a). Each frame in the video sequence is split into Coding Tree Unit (CTU) which comprises of a luma CTB and two chroma CTB's. The size of the CTU is selected by encoder as represented in Figure 1(b). The CTB can be of 16x16 or 32x32 or 64x64, where largest size provides an efficient compression. A CTB may be split to form multiple Coding Units (CU) or contain a single Coding Unit. Luma CB can be small as

8x8 and chroma CB's can be small as 4x4. CU further splits into Prediction Units (PU) depending on the prediction type decision. Transform Units (TU) have their root at CU where prediction residuals are coded using block transforms [5]. This approach has a problem of transmitting redundant set of motion parameters which can be effectively removed by Block merging algorithm [6].

HEVC encoders are more complex compared to H.264/AVC. The motion related complexity aspects were considered in standardization in HEVC [4]. Inter-Picture prediction process consists of choosing motion data comprising the selected reference picture and motion vector to be applied for predicting the samples of each block. Identical inter-picture prediction signals are generated by both encoder and decoder by motion compensation. Inter-Prediction demands a heavy complexity burden up to 80% in the entire encoding process [2]. This complexity is mainly caused due to Motion Estimation (ME). So, it is essential to reduce the heavy complexity of ME. Motion Estimation itself consumes more than 50% of coding complexity or time to encode. Block based motion estimation is widely used in various video coding standards including HEVC which is efficient for parallel translation between frames. For other motion models such as zooming, rotating tilting and wrapping, various algorithms are been discussed in [10]. To reduce the computational complexity, the time required for Motion Estimation should be reduced. This paper discusses various existing search patterns and proposes the New Combined Three Step Search Pattern.

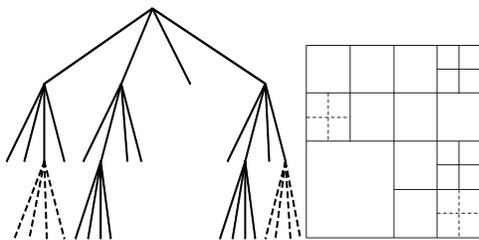


Figure 1 (a) Quad tree (b) Coding Tree Unit (CTU)

The rest of this paper is organized as follows. Section II presents the Existing Search Methodologies. In Section III, the New Combined Three Step Search Algorithm for Motion Estimation is discussed. Section IV presents the performance measures in terms of Time saving for New Combined Three Step Search Algorithm with various test sequences and compared with existing algorithms. Finally, Conclusion is made in Section V.

## II. EXISTING SEARCH METHODOLOGIES

Block based motion estimation removes the temporal redundancy effectively and so it is used widely in various video coding standard. In order to get motion vector for each PU, block matching algorithm (BMA) is performed at the encoder. To generate half-pel and quarter-pel accuracy samples, interpolation filtering is performed for reference picture samples. Integer-pel accuracy motion vector is obtained at first, instead of searching all the positions for quarter-pel accuracy motion. Sum of Absolute Difference (SAD) is used for integer-pel motion search and Sum of Absolute Transformed difference (SATD) is used for half-pel and quarter-pel motion search in HEVC.

Motion Estimation is the process of estimating the best matched block of current frame in the search window of reference frame so as to find the motion vectors. Block based motion estimation removes the temporal redundancy effectively and so it is used widely in various video coding standard. In order to get motion vector for each PU, block matching algorithm (BMA) is performed at the encoder.

### A. Integer pel accuracy Motion search

A three step motion search strategy is used so as to reduce the search point for integer-pel motion search. The start position of the search is selected first. Motion vector predictor (PMV) obtained by motion vector predictor derivation process is used as default start position of integer-pel search and optionally, motion vectors of neighbor positions and zero motion can be checked. After selecting the start position, the second step is to perform the first search. Figure 2 illustrates the three search patterns used for the first search. First search uses diamond search pattern or square search pattern. Currently, diamond search pattern is used as default. Raster search is performed when there is too much difference between obtained motion vector and start position. Currently, search range is set as 64 in integer-pel accuracy. Centre circle represents current position and squares represent candidate search positions for each pattern. The square of same color represents positions having same distance from the start position. Finally refinement search is performed by changing the start position to the best position from the second step.

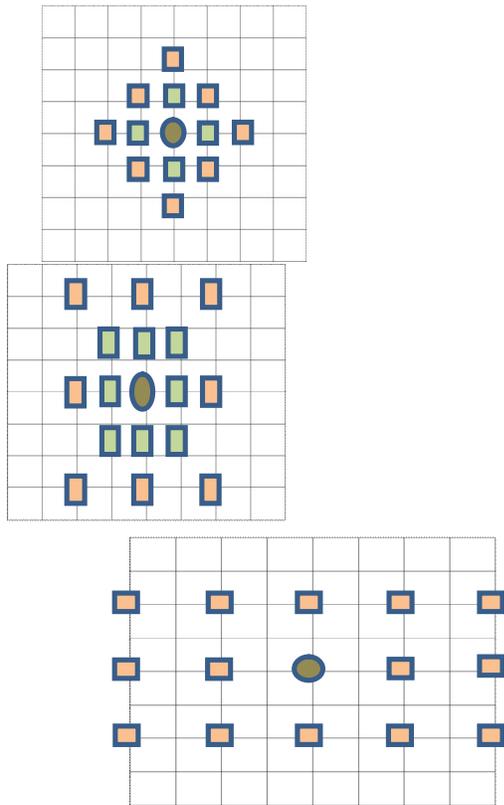


Figure 2(a)Diamond (b)Square and (c)Raster search patterns

Three step search became very popular because of its simplicity and also robust and near optimal performance [1]. It searches for the best motion vectors in a coarse to fine search pattern. The algorithm may be described as:

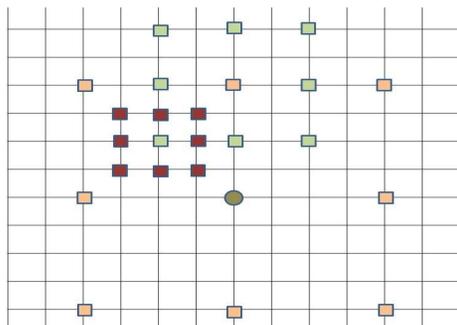


Figure 3 Three Step Search

Step 1: An initial step size is picked. Eight blocks at a distance of step size from the centre (around the centre block) are picked for comparison.

Step 2: The step size is halved. The centre is moved to the point with the minimum distortion.

Steps 1 and 2 are repeated till the step size becomes smaller than one. A particular path for the convergence of this algorithm is shown in Figure 3.

### B. Bipredictive Search

Bi-prediction enhances the inter-frame coding performance and allows the encoder to search for a better reference block from forward and backward pictures. It is efficient for video sequence having scene or illumination changes, camera panning, zoom in/out, and abrupt object. In order to reduce the bi-prediction time, Sum of absolute Difference (SAD) based selective bi-prediction method is discussed in [2]. This method determines whether or not to perform bi-prediction for a current CU block. Bi-prediction of current block is performed as in (1), only if the SAD of uni-prediction is larger than the average SAD of previously coded CU blocks by uni-prediction.

Bi-prediction= on, if

(1)

$$\left. \begin{array}{l} J_{L0} \\ J_{L1} \end{array} \right\} \begin{array}{l} L0 \\ L1 \end{array}$$

Here,  $J_{L0}$  and  $J_{L1}$  are the SAD's computed by forward and backward prediction mode for current CU, where  $\bar{J}_{L0}$  and  $\bar{J}_{L1}$  are average SAD for previously coded CU by uni-prediction respectively.

### C. Fast Search

The ME algorithm which perform search on all blocks in reference frame search window are called full search algorithms and the algorithms which perform search only on blocks which are likely to produce the sub optimal MV's are called fast motion estimation algorithm. Test Zone Search (TZS) is the fast ME algorithm that is mostly preferred in HEVC.

There are four stages in searching the best matched block in fast ME algorithms. Prediction is the initial stage, where the algorithm uses the motion vectors of previously coded CUs to predict the initial search block. The search patterns are employed to find the global minimum point in the second stage. The third stage uses a threshold to terminate the search process called early termination stage [9]. If the motion vector does not

satisfy the early termination criteria, the ME algorithm refines the motion vector in the final stage.

The global minimum points can be estimated by using search patterns such as diamond, square or hexagon grid patterns. Hexagon pattern save around 23% of search window when compared to diamond pattern. There are two basic hexagon search pattern. Horizontal hexagons is good for horizontal motion and have poor performance for vertically moving objects, whereas Vertical hexagons is good for vertical motion and have poor performance for horizontally moving objects. In order to cope with these patterns without loss of performance we consider rotating hexagon patterns. Rotating hexagonal patterns as shown in Figure 4 is used, as it incorporates both vertical and horizontal motion. Type-1 Rotating hexagon pattern is slightly better than Type-2 Rotating hexagon pattern [8].

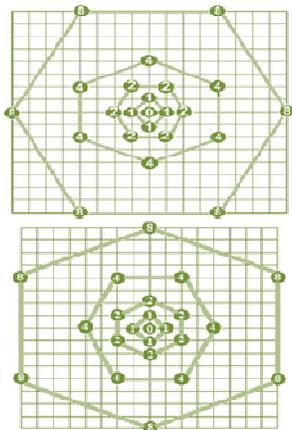


Figure 4 (a) Rotating Hexagon type-1  
(b) Rotating Hexagon type-2

### III. NEW COMBINED THREE STEP SEARCH

Hexagon pattern helps reducing the computational complexity to a great extent, but there is still complexity in coding. Taking this in account, a new search pattern is designed so as to reduce the computational complexity. It performs a three step search with three different patterns at each step. The algorithm may be defined as:

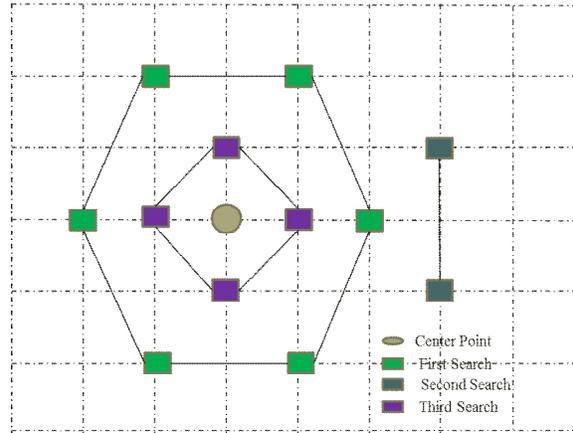


Figure 5 New Combined Three Step Search

Step 1: The first search is done using Hexagon pattern with a step size of two from the center point and six points are taken for comparison at the initial step. If the point with minimum distortion is at the center, then go to step 3.

Step 2: The center point is moved to the point with minimum distortion. The second search is done using linear pattern with a distance one. For minimum distortion along the horizontal points, linear pattern along vertical indices is used and for distortion along vertical points, linear pattern along horizontal indices are used.

Step 3: The center point is moved to the point with minimum distortion. The final search is done using small diamond search with distance of one from the center.

The point with minimum distortion at the final stage is taken as the best match for the current block. Skip mode is imposed in the first step. Figure 5 shows the pattern for New combined Three Step Search. Snake eye scanning order is used for scanning.

### IV. PERFORMANCE MEASURE

Performance measure is analyzed based on average computational time savings and PSNR. Encoding is done for different test sequences with different patterns available. As motion estimation consumes more time in total encoding process, the time taken for motion estimation to predict the motion vector is analyzed. The algorithm is implemented on MATLAB 7.10.0.499 (R2010a). The maximum CU size of 64 with maximum partition of depth 2 is taken. Number of frames taken for each sequence is 30.



Table II Simulation Results of ME time

Sequence	Depth level	ME Time (sec)			
		TSS	DS	HS	NCTSS
BQMALL 832x480	1	25.47689	16.40981	13.21377	13.00017
	2	32.54606	23.97998	18.80197	18.40985
MOBISO DE 832x480	1	24.74718	18.19181	12.91023	12.32081
	2	32.80111	28.66917	19.79492	19.34664
PEOPLE ON STREET 2560x1600	1	303.02874	223.3282	136.1478	135.6764
	2	370.8152	311.356	186.7926	185.6142
TRAFFIC 2560x1600	1	289.8494	188.7696	139.6987	139.9564
	2	374.7651	269.2973	201.2912	201.5436
<b>Average</b>		<b>181.75371</b>	<b>135.0002</b>	<b>91.08139</b>	<b>90.73384</b>
<b>Average Time Saving Compared with TSS</b>			<b>25.72%</b>	<b>49.88%</b>	<b>50.07%</b>

Simulations are carried out on Windows 8.1 OS platform with Intel I3-2310M @ 2.10 Ghz CPU and 4GB RAM. Table II shows the computational results of ME time for different patterns Three Step Search (TSS), Diamond search (DS), Hexagon Search (HS) and proposed New Combined Three Step Search (NCTSS). The simulation results reveal that the proposed algorithm reduces around 45 -55% of time for motion estimation depends on motion in video sequences and depth levels compared to TSS. A slight reduction of time for motion estimation is also achieved compared to the existing Hexagon Search in HEVC. Table III shows the improvement in PSNR for Diamond search (DS), Hexagon Search (HS) and New Combined Three Step Search (NCTSS) patterns with TSS as reference.

Table III Improvement in PSNR with TSS as Reference

Sequence	Depth level	PSNR Improvement (db) compared to TSS		
		DS	HS	NCTSS
BQMALL 832x480	1	-0.23276	0.27705	0.04092
	2	0.13086	0.08098	0.23582
MOBISODE 832x480	1	1.1086	0.82732	0.9011
	2	-0.60223	0.51061	0.68056
PEOPLE ON STREET 2560x1600	1	0.31395	0.02099	0.13663
	2	0.31346	0.25937	0.30735
TRAFFIC 2560x1600	1	-0.43395	0.19148	0.24657
	2	-0.07319	0.11079	0.17016
<b>Average improvement in quality (db)</b>		<b>0.06559</b>	<b>0.28482</b>	<b>0.33989</b>

This New Combined Three Step Search is being implemented in Altera Cyclone-IV FPGA to get further optimal results. Altera SOPC (System on a Programmable Chip) builder is used to connect the soft-hardware components to create a complete computer system using Nios-II softcore processor that runs on various FPGA chips. Nios-II software build tools are then used for implementing the algorithm in the designed system.

## V. CONCLUSION

The proposed algorithm saves about 50.07% of motion estimation time compared with Three Step Search and 0.38% of motion estimation time compared to existing Hexagon pattern in HEVC with a slight improvement in PSNR. The algorithm is implemented for HEVC and it can also be used for older standards such as H.264/MPEG-4 AVC. The implementation result is expected to have better optimal performance.

## REFERENCES

- [1] Jong-Nam Kim, Tae-Sun Choi, "A fast Three-Step search algorithm with minimum checking points using unimodal error surface assumption," in IEEE Transaction on Consumer Electronics, vol. 44, no. 3, pp. 638-648, Aug. 1998.
- [2] J. Kim, D. S. Jun, S. Jeong, S. Cho, J. S. Choi, J. Kim, C. Ahn, "An SAD based selective bi-prediction method for fast motion estimation in high efficiency video coding," in ETRI Journal, vol. 34, no. 5, pp. 753-758, Oct. 2012.
- [3] G. J. Sullivan, J.R. Ohm, W.J. Han, T. Wiegand, "Overview of the high efficiency video coding (HEVC) standard," in IEEE Trans. on circuits and syst. for video technol., vol. 22, no. 12, pp. 1649-1668, Dec. 2012.
- [4] F. Bossen, B. Bross, K. Sühling, D. Flynn, "HEVC complexity and implementation analysis," in IEEE Trans. on circuits and syst for video technol, vol. 22, no. 12, pp. 1685-1696, Dec. 2012.
- [5] I. K. Kim, J. Min, T. Lee, W. J. Han, J. H. Park, "Block partitioning structure in the HEVC standard," in IEEE Trans. on circuits and syst. for video technol., vol. 22, no. 12, pp. 1697-1706, Dec. 2012.
- [6] P. Helle, S. Oudin, B. Bross, D. Marpe, M. O. Bici, K. Ugur, J. Jung, G. Clare, T. Wiegand, "Block merging for quadtree-based partitioning in HEVC," in IEEE Trans. on circuits and syst. for video technol., vol. 22, no. 12, pp. 1720-1731, Dec. 2012.
- [7] J. R. Ohm, G. J. Sullivan, "High efficiency video coding: The next frontier in video compression," in IEEE Signal Processing Magazine, pp. 152-158, Jan. 2013.
- [8] N. Purnachand, L. N. Alves, A. Navarro, "Fast motion estimation algorithm for HEVC," IEEE, ICCE, Berlin, 2012.
- [9] Z. Pan, Y. Zhang, S. Kwong, X. Wang, L.Xu, "Early termination for TZSearch in HEVC motion estimation," Int. Conf. ASSP 2013, pp 1389-1393.
- [10] H. S. kim, J. H. Lee, C. K. Kim, B. G. Kim, "Zoom motion estimation using block-based fast local area scaling," in IEEE Trans. on circuits and syst. for video technol., vol. 22, no. 9, pp. 1280-1291, Sep. 2012.