https://doi.org/10.2352/ISSN.2470-1173.2017.2.VIPC-407

Improved Diamond Half-pel Hexagon Search Algorithm for Block-Matching Motion Estimation

Abdelrahman Abdelazim^A, Ahmed M. Hamza^B, Bassam Noaman^A , Djamel Ait-Boudaoud^B
American University of the Middle East, Kuwait City, Kuwait ^{A}
University of Portsmouth, Portsmouth, UK ^{B}

Abstract

The paper presents an efficient algorithm that reduces the time complexity of video coding in the H.265/HEVC encoder, towards an implementation employable in real-time video coding and transmission applications. The optimization targets the motion estimation search procedure, which occupies a large part of the compute time per Coding Unit. Experimental results demonstrate extensive processing time savings while maintaining similar compression quality and bit rate as the standard.

Introduction

The modern era of video, particularly live streaming applications, relies on advancements made since the early hybrid block-coding standards of H.120. Since then, H.261 and H.262 have been developed, which were advanced versions compared to H.120 [1]. These were followed by H.263 and H.264/AVC, which is the Advanced Video Coding. These were developed incrementally the past decade with several extensions for scalable and multicast formats, allowing for higher quality, but still limited to fixed 16×16-pixel coding blocks. A persistent problem has been the extensive processing time of encoding video sequences, which increased with the complexity of these encoders. HEVC stands for the High Efficiency Video Coding (also known as H.265) and was first published in January 2013 [2]. Despite the HEVC design to fit High Definition videos it shares the same complexity bottleneck with pervious standards, with as much as ten-fold increase in compute times due to its architecture. This work presents an efficient algorithm to help the realization of the standard in real time applications.

The paper is organized as follows: Section II provides a background on video coding, and the concept of the motion estimation method in inter-frame coding is outlined. Similar research in the area is presented in Section III. Our algorithm is discussed and tested against the standard in sections IV and V respectively. Section VI concludes the paper.

Background

Video Coding

A video codec is a piece of software that works to compress or decompress digital videos, which means transforming uncompressed video into a compressed one or vice-versa.

H.265

HEVC/H.265 aims to significantly improve the compression efficiency compared with the previous standards and in particular the H.264/AVC. It supports larger encoding blocks, also it has a more flexible partitioning structure to allow smaller blocks to be used for more textured regions. Large blocks will generally work better for flat regions of a picture, the probabilities are that HD videos contain bigger smooth regions, which can be encoded more effectively when large block sizes are used. HEVC encoder flexibility stems from the fact that it contains a highly configurable encoding setup parameters, and large number of coding tools, beyond those provided by earlier video coding standards [3, 4] This added flexibility allows an encoder to determine block dependent parameters in terms of:

i. Coding unit (CU) quad-tree structure, prediction unit (PU), partition modes and transform unit (TU) structure
ii. Intra PU prediction mode
iii. Inter PU motion parameters and reference list index or indices, for motion estimation
iv. Rate distortion optimized quantization

Motion Estimation

Nowadays, due to the large size of videos, video communication faces a big challenge. Video compression techniques are used to reduce redundancy in video data without affecting visual quality [3]. Motion estimation [4] is considered as a main component of video compression techniques [5]. It is one of the most computational intensive operations in video compression [6]. The reason of using motion estimation in the H.265 is that it helps to eliminate redundancy between frames within a video sequence [7]. It is used to illustrate the transformation from one image to another in 2D form. Moreover, the concept of motion estimation means finding the motion vector pointing to the best prediction macroblock in a reference frame. It is used to detect non-changeable blocks, and motion vectors are saved in place of blocks. In reality, motion estimation based encoders are the most widely used in video compression techniques [8]. Also, the goal of the motion estimation is to find the relative motion between two images in order to eliminate temporal redundancy. Different motion estimation algorithms have been invented to reduce the complexity of motion estimation [9]. In the following subsection, further clarifications are presented to explain the block-matching technique of the motion estimation and Rate Distortion (RD) which is the main factor of the block matching.

Block Matching

The Block Matching Algorithm is a method for finding matching macroblocks (called Coding Units in HEVC) in a sequence of digital video frames for the purposes of motion estimation. It works by dividing the current frame of a video sequence into macroblocks, then comparing each of the macroblocks with a corresponding block and its adjacent neighbors in a near frame of the video. The rate distortion optimization procedure involves multiple calculations of the distortion between the current and reference frames. Discovering the finest coding parameters is executed in a rate distortion (RD) optimization procedure, since it allows tradeoffs between the numbers of bits used to compress a block of the image vs. the resulting distortion that is formed by using that number of bits. The RD optimization problem here can be expressed as:

$$J = \text{Distortion(SAD)} + \lambda.\text{MV}$$

which is used as the cost function adapted to various stages in the RDO.

Low bit rate can be accomplished in video coding at the cost of reduced quality (high distortion). This generic cost function in different modes of prediction is used to find the 'best' matched block. This is done by iterating through the entire CTU at all depths, recursively, with the final mode decision being made using the similar function

$$J = \text{Distortion (SSE)} + \lambda.MV$$

SSE denotes the Square of Sum Differences function instead of the simple sum – a measure more closely aligned to the PSNR quality calculation.

In a particular instance, the process of subdividing the 64-pixel by 64-pixel block starts in current frame. A subdivision of the reference frame into 128-pixel by 128-pixel (or other research area) in each CU instance is then done, which considered as the search region. To populate the Prediction Unit parameters, the cost function is employed to find the 64×64 block best match in the reference frame, this carried out pixel by pixel is raster manner. The best match is the one with the minimum cost. This movement can also be sub-pixel movement; either half-pixel or quarter-pixel, though all information is expressed at the quarter pixel level. Additionally, the block is divided to the smaller CU/PU blocks to find a better match. While using the sub blocks, the sum of all of the differences is carried out in order to get the overall cost to be able to compare it to the larger-block cost. Each cost calculation involves finding the minimum value of the respective J as the first possibility for the perfect block match. The extensive list of subdivision modes (symmetric and asymmetric in HEVC) includes carrying out the same process for the left and the right 32×64 block and finding the minimum J value. The process is recursively repeated, subdividing the 64×64 block in the current frame into four 32×32 blocks and finding the best J value for all the blocks, and so on till it reaches 8×8 blocks. The encoder chooses the mode configuration that for CU, TU, and PU that gives the lowest total cost for constituent partitions.

Existing Research

Over the past 20 years, many algorithms were put forth to reach a permissably low computational complexity while preserving quality with the same low bit rate as the various standards. We group them together to discuss common factors. The are three main factors that effect the output encoding of raw video: objective quality, bitrate, and complexity of the encoder. The proposals in the Video Coding Experts Group (VCEG) have highlighted different techniques of motion estimation, most of them geared towards having high efficiency leading to a high quality of significantly reduced bitrate counts at the same resolution and frame rate of encoded video. The overall complexity of the encoding process tends to rise as these were sucessively added to the standards. With speed as an important real-time encoding consideration, the full search could use up to 70% of the complete computation of the video compression process. [10].

Alternatives in the literature include the three step search (low quality), the four step search and the DGDS[11]. The four step search has significance improvements in motion detection over the three step search pattern[12]. The DGDS stands for Directional Gradient Descent Search. Under the low quality parameter, there is the three step search and under the medium quality parameter, there is the four step search, beneath them the fixed shape patterns such as cross search, diamond search and hexagon search. Within these there are two methods used, either integer pixel or sub-integer pixel level comparisons.

For the sub-integer pixel there exist granularity levels of half and quarter-integer pixel, where half-pixel is typically used. A third step in motion search is the high quality geared, higher complexity combination of the fixed shape patterns such as MCDH (Modified Cross-Diamond-Hexagonal) and HEXFS [12]. HEXFS stands for Hexagonal pattern search.

Proposed Algorithm

This proposal is a variation of Modified Cross Hexagon Diamond Search Algorithm, aiming for a faster motion estimation procedure, with some modifications that target the time complexity. The proposed search, which we call Improved Diamond Half-Pixel Hexagon Search (IDHHS), relies on only two stages. For testing purposes we are applying this modification to only the current/latest standard, HEVC. The IDHHS is considered a block matching optimization technique under the spatial domain.

Step 1: The search will begin by mapping the search region then creating around the central point of the search region eight-point Diamond pattern with the center point. By using Sum Absolute Difference (SAD) each point of the diamond's shape will have a minimum cost; we will choose the lowest one. If the minimum cost is found to be at center point, then we will stop the search process. This stage will be repeated three times; each time choosing the new diamond's central point depends on the previous diamond's best matched point to preform another new eight-point diamond. The idea behind implementing this is; three diamonds provides more accuracy as shown in Fig. 3, the first and second iterations, and Fig. 4 represents the remaining steps.

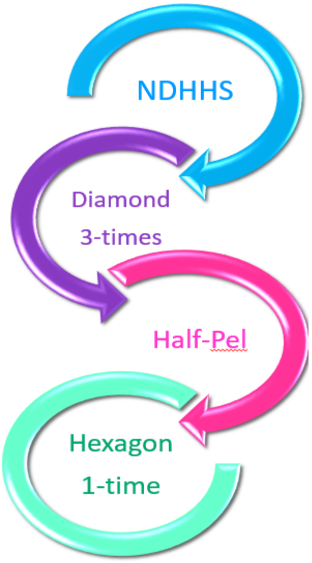

Figure 1. *Overview of NDHHS*

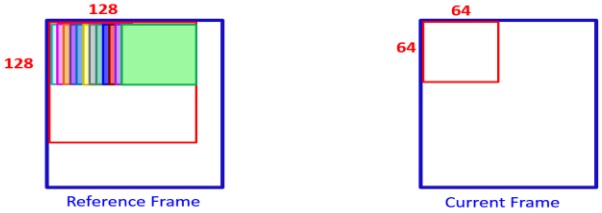

Figure 2*. Block Motion search for a single CU depth*

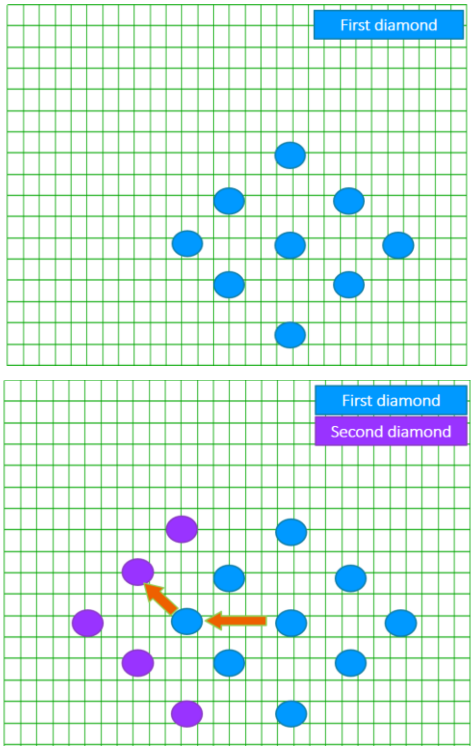

Figure 3. *Initial Search Steps in NDHHS*

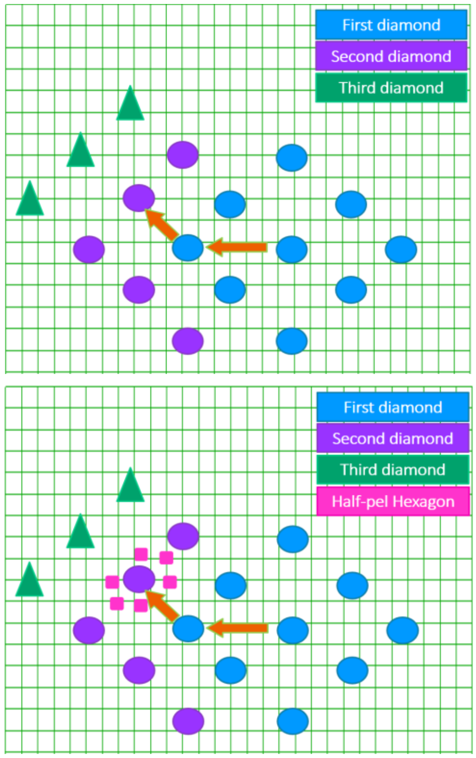

Figure 4. *Half-PEL Hexagon Search in NDHHS*

Step 2: After finishing from stage 1 with the three iterations, the half-pel six-point Hexagon is performed as the second search. This stage is considered as a fourth and final stage. Lastly, the algorithm will end up having the best matched block. The reason behind choosing a half-pel is because the resolution will remain high as shown in Figure 5 the final/fourth step. This algorithm is done by bunch of if-statements in order to check; first, whether the used point is out of the search region or not. Second, if the step was in the last stage to decide whether performing the half-pel hexagon or diamond. The main advantage of this proposal is that it has one constraint, and it can be implemented easily in HEVC (H.265). This proposal introduced the Improved Diamond Half-Pel Hexagon (IDHHS) algorithm for fast Motion Estimation, which has improvement in quality and reduction in time.

Experimental Results

Implementation consisted of modifying the reference H.265 code base to add the new search modules. The NDHHS algorithm is summarized in Fig 1.

Classification of video sequence, according to the Joint Collaborative Team on Video Coding (JCT-VC) organization, provides categories of six classes of video resolution and frame rate: A, B, C, D, E and F. Our set of encoding tests shown in Table 1 are comparing the optimization gains of FastSearch and the improved diamond search presented in this work. This basic non-averaged result was done at the standard QP values of 22, 27, 32 and 37 for all sequences.

The rest of this section will detail the test environment and comprehensive results across the range of video sequences.

Testing was implemented on a PC running Windows 10 with processor: Intel® Core™ i7-4720HQ CPU @ 2.60 GHz. Also, it has an installed memory (RAM) 16.0 GB, and the system type of the computer is 64-bit Operating System x64-based processor. The Full Search and Fast search are existing standard modules in HEVC; but fall short of the time savings recorded under the proposed method. The average figures for NDHHS show execution time decreased by (7.5%) compared with the Fast Search, and compared with Full search it decreased by (21%).

The comprehensive test results in Table 1, below, show similar patterns across video sequences of all classes. We used a slightly different test setup, utilizing the BD-rate and BD-PSNR versions of the measures, to be standards compatible in testing environment. The results show clear performance gains over the FastSearch optimization, with slight penalties in quality (below 0.1 dB on average) and compression (around 2%). Comparison against the basic Full Search shows the expectedly larger gains and slightly higher penalties, at bitrate increase for a 2K (above HD) video. This is, for most applications, an acceptable tradeoff.

We note that FastSearch is enabled by default. We also note that the full search gives by definition an optimal HEVC encoding for the block, with no optimization for complexity applied. The best overall performance was shown by the proposed algorithm.

Table 1. Comprehensive test runs of proposed method against Full and Fast Search for various video sequences

Sequence	**Full Search**			**Fast Search**		
	BD-Bitrate (%)	***BD-PSNR (dB)***	***Time Saving (%)***	***BD-Bitrate (%)***	***BD-PSNR (dB)***	***Time Saving (%)***
Traffic 2560x1600	2.79	-0.21	15.81	1.46	-0.04	6.54
BOMall 832x480	3.22	-0.39	23.56	1.98	-0.07	8.69
Racehorses 832x480	2.96	-0.056	18.79	1.02	-0.02	7.31
BasketballPass 416x240	3.64	-0.26	21.42	2.07	-0.11	7.74
KristenAndSara	2.55	-0.078	17.92	1.53	-0.03	5.33
ChinaSpeed 1204x768	4.01	-0.41	28.4	2.61	-0.04	9.10
Average	**3.2**	**-0.23**	**20.98**	**1.78**	**-0.05**	**7.45**

Conclusion

This work improves on standard optimizations in motion estimation search, combining ideas from previous work on search patterns. HEVC reference tools provide excellent solutions for high bitrate encoding, but the features remain complex and the computational overhead high. We proposed and tested an alternative block search method aiming to reduce encoding time in the motion estimation procedure, with results being promising, limited penalties in quality and bitrate. We believe speed optimization in the high time complexity processes to be important in ensuring the broad success of HEVC as it emerges in a wide range of products and real-time streaming applications.

References

[1] Efficiency and V. Coding, “High Efficiency Video Coding,” 2009.J.

[2] G. J. Sullivan, J. R. Ohm, W. J. Han, and T. Wiegand, “Overview of the high efficiency video coding (HEVC) standard,” IEEE Trans. Circuits Syst. Video Technol., vol. 22, no. 12, pp. 1649– 1668, 2012.

[3] D. Salomon and G. Motta, “Video Compression,” Handb. Data Compression, pp. 855–952, 2010.

[4] P. Davis and S. Marikkannan, “Implementation of Motion Estimation Algorithm for H . 265 / HEVC,” vol. 3, no. 3, 2014.

[5] S. Ashwin, S. J. Sree, and S. A. Kumar, “Study of the Contemporary Motion Estimation Techniques for Video Coding,” Int. J. Recent Technol. Eng. (IJRTE), ISSN 2277-3878, vol. 2, no. 1, pp. 190–194, 2013.

[6] S. Zuo, M. Wang, and L. Xiao, “Motion Estimation Algorithm for HEVC Suitable for Hardware Implementation motion of HEVC estimation,” vol. 10, pp. 611–618, 2014.

[7] T. Aziz and D. R. J. Dolly, “Motion Estimation and Motion Compensated Video Compression Using DCT And DWT,” vol. 2, no. 12, pp. 667–671, 2012.

[8] I. Engineering, “Review and Comparative Study of Motion Estimation Techniques To Reduce Complexity in Video,” pp. 3574–3584, 2013

[9] J. R. Ohm, G. J. Sullivan, H. Schwarz, T. K. Tan, and T. Wiegand, “Comparison of the coding efficiency of video coding standards-including high efficiency video coding (HEVC),” *IEEE Trans. Circuits Syst. Video Technol.*, vol. 22, no. 12, pp. 1669–1684, 2012.

[10] L.-M. Po, K.-H. Ng, K.-W. Cheung, K.-M. Wong, Y. M. S. Uddin, and C.-W. Ting, “Novel Directional Gradient Descent Searches for Fast Block Motion Estimation,” *IEEE Trans. Circuits Syst. Video Technol.*, vol. 19, no. 8, pp. 1189–1195, 2009.

[11] Lai-man Po and Wing-chung Ma, “A novel four-step search algorithm for fast block motion estimation,” *IEEE Trans. Circuits Syst. Video Technol.*, vol. 6, no. 3, pp. 313–317, 1996.

[12] H. Jia and L. Zhang, “Directional Cross Diamond Search Algorithm for Fast Block Motion Estimation,” *arXiv Prepr. arXiv0806.0689*, pp. 1–23, 2008.

Author Biography

Abdelrahman Abdelazim *is an associate professor of Computer Engineering at the American University of the Middle East (AUM) in Kuwait. He holds a BEng (Hons) degree in Digital Communication and a PhD degree in Engineering, both from the University of Central Lancashire (UCLAN), Preston, UK. Between 2008 and 2012 he worked as Lecturer in Electronics within the School of Computing, Engineering and Physical Sciences (CEPS). His research interests are in the area of reducing the complexity of Video Coding Encoders in real-time Scalable and Multi-view applications and the area of teaching and learning in higher education. Abdelrahman is member of the IET since 2006. He is a chartered engineer and a Fellow of the Higher Education Academy.*

Ahmed Hamza *is a PhD student in the School of Computing at Portsmouth University, where he is pursuing research in real-time video encoding optimization by deep learning networks. He obtained a M.S in Computer Science from Georgetown University in 2010, with a thesis on structural similarity algorithms in chem-informatics. His interests are in Video Coding, algorithms, information theory, and natural systems. Ahmed is a member of IET and IEEE.*

Bassam Noaman *is the vice dean of the college of engineering and technology at the American University of the Middle East (AUM) in Kuwait Before AUM he worked as a patent examiner for United States Patent and Trademark Office (USPTO), and as an assistant professor at the Department of Electrical and Computer Engineering, Catholic University*
He received his B.Sc. - 2001, M.Sc. -2003, and Ph.D. -2011 degrees from George Washington University (GWU). His research interests are predominantly focused on microelectronics, numerical modeling and noise characterization of semiconductor devices, micro and nano-electro-mechanical systems (MEMS/NEMS) and nanotechnologies. He is also active researcher in the areas of Active learning.

Djamel Ait-Boudaoud *joined the Portsmouth University in 2010 and is currently Dean of the Faculty of Technology. Before Portsmouth he was the head of the School of Computing, Engineering and Physical Sciences at the University of Central Lancashire for close to 10years. His research interests are predominantly focused on the problems of optimization with applications in 3D computer vision, video standards (H264) and solving combinatorial (ordered-sequence) problems using evolutionary algorithms. He gained a PhD in 1991 from the Department of Electrical and Electronic Engineering at University of Nottingham, UK., is a Chartered Engineer (CEng) and a fellow of the Institution of Engineering and Technology (FIET). He is currently a committee member of the Engineering Professors Council (EPC) and serve on the reader’s panel for the Queens Anniversary Prize for Higher and Further Education, UK*