



The H.264/MPEG4 Advances Video Coding Standards, Advantages and Review on the Latest H.264 Motion Estimation Techniques

Jasmeet Kaur*

Assistant Professor, Department of Computer Science & Engineering
A P Goyal Shimla University, H.P., India

Abstract— *H.264/MPEG4-AVC is the latest video coding standard of the ITU-T video coding experts group (VCEG) and the ISO/IEC moving picture experts group (MPEG).H.264/MPEG4-AVC has recently become the most widely accepted video coding standard since the deployment of MPEG2 at the dawn of digital television. 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. Onthoroughly studyingthe UMHEXAGONS algorithm, there are numerous methods to improve the UMHEXAGONS speed whilst maintaining the quality of the encoding process. In this review paper several techniques are proposed 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.This work provides an overview of the technical features of H.264/AVC describes profiles, outlines the history of the standardization process and discusses its advantages.*

Keywords-AVC, H.263, H264, UMHEXAGONS, JVT, MPEG-2, MPEG-4.

I. INTRODUCTION

Video technologies are generating ever new applications with high resolution, wide range of basic video characteristics and requirements such as spatiotemporal chroma format and sample accuracy. Application areas today range from video conferencing over mobile TV to broadcasting of standard high-definition. TV content up to very high quality applications such as professional digital video recording or digital cinema/ large-screen digital imagery. Prior video coding standards such as MPEG2/H.262 [1], H.263 [2], and MPEG4 Part2 [3] are already established in their application domains.

But with the proliferation of digital video into new application spaces such as mobile TV or high-definition TV broadcasting, the requirements for efficient representation of video have increased up to operation points where previously standardized video coding technology can hardly keep pace. Besides, more cost-efficient solutions in terms of bit rate vs. end-to-end reproduction quality are increasingly required in traditional application areas of digital video.

Considering the mentioned challenges in H.264/MPEG4, Advanced Video Coding (AVC) [4] that is the latest entry of international video coding standards has demonstrated significantly improved coding efficiency, substantially enhanced error robustness, increased flexibility and scope of applicability relative to its predecessors [5]. A recently added amendment to H.264/ MPEG4 -AVC, so-called fidelity range extensions (FRExt) [6], further broaden the application domain of the new standard toward areas like professional contribution, distribution, or studio/post production. Another set of extensions for scalable video coding (SVC) has been designed [7, 8], that allows the reconstruction of video signals with lower spatiotemporal resolution or lower quality from parts of the coded video representation (i.e., from partial bit streams). Also, multi-view video coding (MVC) capability has been successfully demonstrated using H.264/MPEG4-AVC [9], requiring almost no change to the technical content of the standard.

In early 1998, the *Video Coding Experts Group* (VCEG) ITU-T SG16 Q.6 issued a call for proposals on a project calledH.26L, with the target to double the coding efficiency (which means halving the bit rate necessary for a given level of fidelity) in comparison to any other existing video coding standards for a broad variety of applications. The first draft design for that new standard was adopted in October of 1999. In December of 2001, VCEG and the *Moving Picture Experts Group* (MPEG) ISO/IEC JTC 1/SC 29/WG 11 formed a *Joint Video Team* (JVT), with the charter to finalize the draft new video codingstandard for formal approval submission as H.264/AVC [11] in March 2003.

The scope of the standardization is illustrated in Fig. 1, whichshows the typical video coding/decoding chain (excluding thetransport or storage of the video signal).

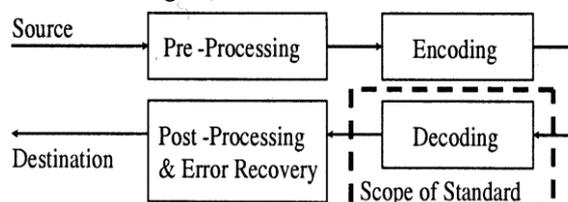


Fig.1.scope of video coding standardization

As has been the case for all ITU-T and ISO/IEC video coding standards, only the central decoder is standardized, by imposing restrictions on the bit stream and syntax and defining the decoding process of the syntax elements such that every decoder conforming to the standard will produce similar output when given an encoded bit stream that conforms to the constraints of the standard.

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, in which consume 70% to 90% of the total encoding time [13]. 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) [15], new three step search (NTSS)[16], four step search (FSS)[17], diamond search (DS)[18] and hexagon-based search (HEXBS)[19].

II. ADVANTAGES

The latest advancements in H.264 compression technology now deliver the best of both worlds - high quality video and a superb compression ratio. Specifically related to security DVRs, here is how H.264 technology advantages:

Longer record times: As mentioned earlier, H.264 provides superb compression, in most cases more than doubling record times over previously popular compression methods. Here is a quick chart (i.e. table 1) comparing H.264 and other compression technologies, using a 160GB hard drive, recording video at 30 frames per second, at a resolution setting of 720 x 480.

Table 1: Comparison of different compression techniques

H.264	records 77 Hours
MPEG4	records 44 Hours
JPEG2000	records 13 Hours

The advantages of H.264 are impressive! The ability to record longer periods of time on the same sized hard drive will save you money, and greatly increase the amount of video evidence you can archive.

Quality and speed does not suffer: Previous compression methods have always allowed you to record excellent picture quality at real-time frame rates - but you were eating up valuable hard drive space. Similar to MPEG4, H.264 uses predictive technology to drastically reduce the amount of redundant video recorded. In simple terms, instead of constantly recording say a room with no motion, H.264 will use previously recordings images. However, if someone walks into the room, H.264 will record that person and continue to use the previously recorded background.

Improved remote monitoring: Saving hard drive space is not the only benefit of H.264. The ability to bring together high quality and low memory sizes allows for seamless presentations of video when transmitted. Coupled with low frame rate recordings, previous compressions made watching video from remote video security DVR units a trying experience - especially when activities in question happened fast. Video would be jumpy, and audio (if used) wasn't likely to be synchronized with the video. Watching real-time H.264 is a completely different experience, and the reason so many broadcast companies around the globe have rapidly adopted it.

A word of caution: Although H.264 compression is without question a giant leap in video technology for all industries, just having that feature in a DVR does not mean that it's a good DVR. To make all the magic happen, it takes more processing horsepower. So be sure to use a manufacturer with a proven track record, who stands behind their products, and one who will be there for you should you need a hand.

Encoding and decoding complexity: H.264 encoding and decoding is more computationally complex than some other codec's such as MPEG-4 Part 2 (DivX, XviD). However, the compression performance of H.264 is significantly better than these so it depends on what is more important to you. This is becoming less of a problem as more devices are including hardware support for H.264.

Error Resiliency: There are some things in H.264 to deal with bit errors, but often they are not used and a single bit error can still have a catastrophic effect. From what we have seen in my study of video codec's, error resiliency seems to being pushed to another layer in most systems. That is, the video codec is designed for maximum compression, and another layer is added on top of the video data to take care of bit errors. That way those who don't need the error resiliency don't pay for it with lower compression rates.

III. UMHEXAGONS

UMHEXAGONS has been adopted as fast search algorithm for H.264 video compression standard since it performs well in both small and large video motion. The UMHEXAGONS accuracy and rate distortion performance is very close to FS while maintaining the computational complexity low, up to one tenth of the FS [14]. 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 [14].

In the first step, initial search point prediction is performed using five different types of prediction. These are median prediction, (0,0) prediction, up layer prediction, corresponding block prediction and neighboring reference picture

prediction. The best match amongst these predictions is chosen as the initial search point. The second step performs unsymmetrical cross search where the horizontal is twice of vertical search as shown in Fig.2. For the third step, a small rectangular full search is done using a 5x5 search area from 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 as yellow points in Fig.5. 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 candidate lies in the center of the hexagon. This will be followed by small diamond search pattern for final refinement.

IV. IMPROVEMENTED UMHEXAGON

A number of improvements have been proposed for UMHEXAGONS algorithm, each of these focusing on certain step in the algorithm. Most of the proposed algorithms focus on increasing the speed by reducing the number of search points

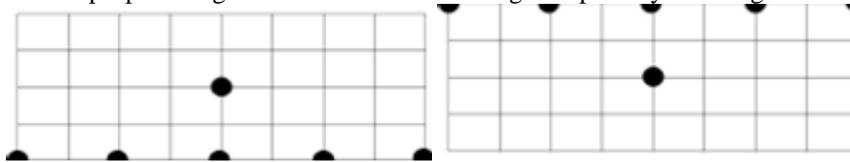


Fig. 2 Unsymmetrical-cross search pattern

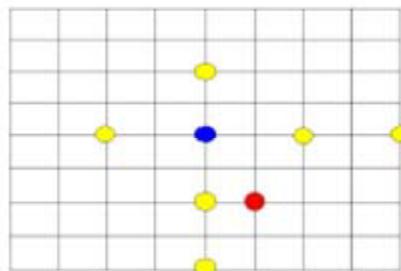


Fig.3. Irregular-cross search pattern: Case 1

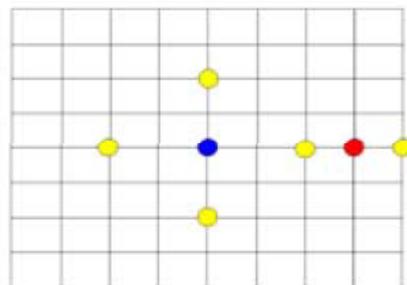


Fig.4. Irregular-cross search pattern: Case 2

A. Irregular-Cross Search Pattern

As discussed in the previous section the unsymmetrical cross search is performed within the UMHEXAGONS algorithm with the horizontal search range is set twice of the vertical search range as shown in Fig.2. The search is bias towards horizontal based on the assumption that most of the motion vector (MV) in aggressive video sequence are on the horizontal plane.

However, 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 [14] where it considers the motion trend to predict the possible area of the motion vector. In other words, the technique analyses the initial search point and the current location of the block before deciding the search pattern. In this irregular-cross step search, the authors divide the cross into four quadrants as shown in Fig.3 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.

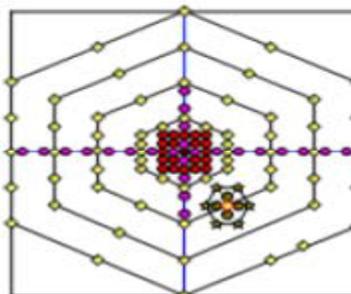


Fig.5. Hexagon search pattern

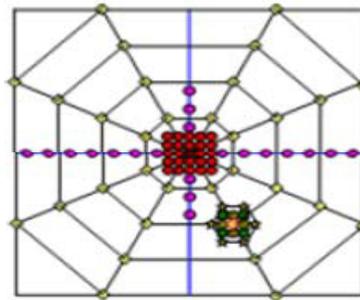


Fig.6. Octagon search pattern

As the current block is located in the fourth quadrant, the search range will be biased towards this quadrant as shown in Fig.3. Thus, the best match will more likely to be located between the current block and the initial search location. Another possibility that might occur using this method is that the current block may be located exactly between the two quadrants. In this case, only the search point that is biased toward this location is done as shown in Fig.4. This will result in less search points since only one side of the search is done.

B. New Square Search Pattern

As mentioned in the previous section, small full search will be performed within 5x5 search area relative to search centre which makes a total of 25 search points. More than 80% of the motion vector is located within this region and more than 70% is located within the 3x3 region [22]. With only approximately 10% difference, the 3x3 search area is a better choice as it 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 within the four hexagon rings as shown in Fig.5. To reduce the search point, an octagon-grid search is proposed in [21]. Using this method, the search points are reduced to 32 points. This is a 50% reduction compared to the original algorithm. The points are hereby distributed evenly (as shown in Fig.6) and it is particularly suitable for high or aggressive video motions.

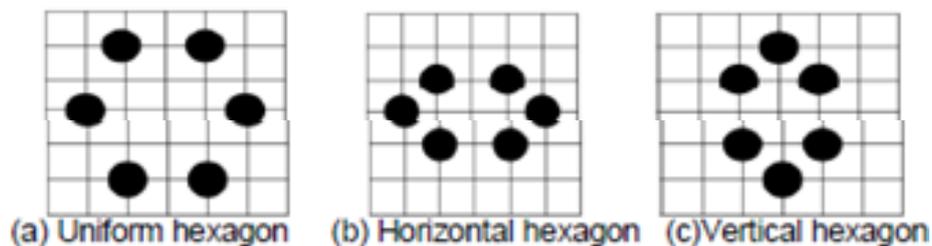


Fig.7. Three hexagonal patterns

D. Horizontal and Vertical Hexagon Search

For the last step of the UMHEXAGONS algorithm, an improvement proposed in [22] 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 Fig.4.6. When the current block size is 16x16 or 8x8, the uniform hexagon pattern as shown in Fig.7 (a) is used. For 16x8 or 8x4 macro blocks, the horizontal hexagon pattern as shown in Fig.7 (b) is adopted. On the other hand, vertical hexagon as shown in Fig.7(c) will be used for 8x16 or 4x8 block size. If a 4x4 macro block is used, the block will be finalized with small diamond search as the last step for UMHEXAGONS algorithm.

V. CONCLUSION

The emerging H.264/AVC video coding standard has been developed and standardized collaboratively by both the ITU-T VCEG and ISO/IEC MPEG organizations. H.264/AVC represents a number of advances in standard video coding technology, in terms of both coding efficiency enhancement and flexibility for effective use over a broad variety of network types. It has so many advantages and in this Technique, 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 Fig.4.6. Three hexagonal patterns provide only a small improvement.

REFERENCES

- [1] ITU-T Rec. H.262 and ISO/IEC 13818-2 (MPEG2), *Generic Coding of Moving Pictures and Associated Audio Information — Part 2: Video*, Nov. 1994.
- [2] ITU-T Rec. H.263, *Video Coding for Low Bit Rate Communication*, v1, Nov. 1995; v2, Jan. 1998; v3, Nov. 2000.

- [3] ISO/IEC JTC 1, *Coding of Audio-Visual Objects — Part2: Visual,*” ISO/IEC 14496-2 (MPEG4 Visual Version 1), April 1999; Amendment 1 (Version 2), Feb. 2000;Amendment 4(streaming profile), Jan. 2001.
- [4] ITU-T Rec. H.264 and ISO/IEC 14496-10(MPEG4-AVC),*Advanced Video Coding for Generic Audiovisual Services*,v1, May, 2003; v2, Jan. 2004; v3 (with FRExt), Sept. 2004; v4, July 2005.
- [5] A. Luthra, G. J. Sullivan and T. Wiegand, Eds., Special issue on the “*H.264/AVC Video Coding Standard,*” *IEEETrans. Circuits and Sys. for Video Tech.*, vol. 13, no. 7,July 2003.
- [6] G. J. Sullivan *et al.*, *Draft text of H.264/AVC Fidelity Range Extensions Amendment*, ISO/IEC MPEG and ITUT VCEG, JVT-L047, Redmond, WA, July 2004.
- [7] T. Wiegand *et al.*,*Joint Draft 5: Scalable Video Coding*, ISO/IEC MPEG and ITU-T VCEG, Doc. JVT-R201, Bangkok, Thailand, Jan. 2006.
- [8] H. Schwarz, D. Marpe and T. Wiegand, “Overview of the Scalable H.264/MPEG4-AVC Extension,” *IEEE Int’l. Conf. Image Processing*, Atlanta, GA,Oct. 2006.
- [9] K. Müller *et al.*, “*Multi-View Video Coding Based on H.264/AVC Using Hierarchical B-Frames,*” *Proc. PCS2006*, Beijing, China, Apr. 2006.
- [10] A. M. Tourapis *et al.*, “*Direct Mode Coding for Bipredictive Slices in the H.264 Standard,*” *IEEE Trans. Circuits and Sys.for Video Tech.*, vol. 15, no. 1, Jan. 2005, pp. 119–26.
- [11] Draft ITU-T recommendation and final draft international standard of joint video specification (ITU-T Rec. H.264/ISO/IEC 14 496-10 AVC, in Joint Video Team (JVT) of ISO/IEC MPEG and ITU-T VCEG, JVTG050,2003.
- [12] *Generic Coding of Moving Pictures and Associated Audio Information- Part 2: Video*, ITU-T and ISO/IEC JTC 1, ITU-T Recommendation H.262 and ISO/IEC 13 818-2 (MPEG-2), 1994.
- [13] 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
- [14] Peng Huang, and Cui-hua Li, “*Irregularity -cross multi-hexagon-gridSearch Algorithm for Fast Motion Estimation on H.264*” in 2nd International Conference on Computer Engineering and Technology 2010 (IC CET), vol. 3, pp. 587-592.
- [15] 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.
- [16] C. Karthikeyan. “*Performance Analysis of Block Matching Algorithmsfor Highly Scalable Video Compression*”, International Symposium on Ad Hoc and Ubiquitous Computing 2006 (ISAUHC ‘06), pp. 179-182, Dec 2006
- [17] 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.
- [18] 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
- [19] 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.
- [20] Zhibo Chen, JianfengXu, Yun He and JunliZheng, “*Fast integer-pel and fractional -pel motion estimation for H.246/AVC*” *Journal of Visual Communication & Image Representation*, April 2006, pp. 264-290.
- [21] Zhu Weixing, Chen Xianyong and Li Xincheng, “*A New Search Algorithm Based On Multi-Octagon-Grid*” in *CISP '09*. 2ndInternational Congress Image and Signal Processing, 2009,pp.1-5.
- [22] 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.