Adaptive Search Range Determination for H.264 based Disparity and Motion Estimation of MVC

Narasak Boonthep

***Abstract*— Multi-view video consists of a set of multiple video sequences from multiple viewpoints or view directions in the same scene. It contains extremely a large amount of data which has a high degree of correlation with both spatial and temporal redundancy. The redundant data can be eliminated with multi-view video coding (MVC) by using motion estimation and disparity estimation. I proposed fast disparity and motion estimation algorithm to reduce computational complexity. The redundancy reduction problem can be solve by determine the optimize search range. The search range can vary area depended on object moving rate. Disparity estimation can eliminate mismatched points from data set. The experimental results show that the proposed method can reduce the encoding time, so that the encoder complexity can be reduced but achieves almost the same compress quality. In case that the distance is zero**

***Index Terms*—A Multi-view video coding(MVC), motion estimation, disparity estimation, H.264/AVC, Search range**

I. Introduction

Nowadays, the world's video broadcasting technologies are developed and changed all the time. In the past, users were limited to see only 2D so that they could access to the image only one side. However, recently emerged 3D video technology allows them to access the image in more freedom view so that it seems more realistic video to users. The viewer can interactively choose viewpoint in 3D space to observe the action of a dynamic real-world. In developing 3D video broadcasting technology, researchers presented multi-view video (MVV) as a key technology for various application [1].

MVV [2- 3] basically consists of a set of video sequences capturing the same scene simultaneously from cameras at different view directions as shown in Fig. 1. It contains of huge amount of data which has a high degree of correlation with both spatial dependencies between the camera views (inter- view) and temporal redundancies (inter-frame) as among the frames each view. The redundant data should be eliminated with multi-view video coding (MVC) by using motion estimation (ME) and disparity estimation (DE). ME has been used to reduce temporal redundancy, and DE has been employed to reduce inter-view redundancy.

Manuscript received February 2, 2016. This work was supported in part by the University of Phayao.

Narasak Boonthep is with Department of Computer Engineering, University of Phayao, Phayao 56000 Thailand (corresponding author to provide phone: 66-54-466-666; e-mail: narasak.bo@ up.ac.th).

The ideal of ME and DE is to track every pixel from frame to frame. However, it requires very expensive computational cost. Since this system has serious limitations on data distribution applications, such as broadcasting, multimedia streaming services, and other commercial applications, I need to compress the multi-view sequence efficiently without sacrificing its visual quality significantly [4-6].

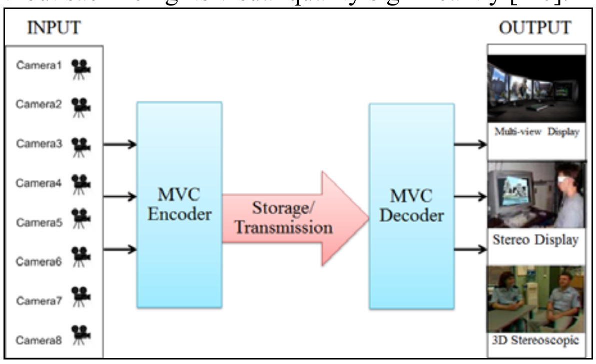

Fig 1 Overall of multi-view video coding.

Many algorithms have been proposed for saving the ME and DE computational to reduce processing time and more accurate. FS (Full Search) is the best one in terms of the quality of the predicted image and its resolution of the motion vector it is very computationally intensive. This algorithm find the optimal motion vector for a given block in the anchor frame by comparing it with all candidate blocks in the target frame within a predefined search range and finding the one with the minimum error, but it requires very expensive computational cost [9]. For solving the complexity problem of motion estimation, many fast motion estimation algorithms.

Fast ME algorithms have been proposed as alternatives to the exhaustive full search technique. These algorithms reduce the search positions through the efficient search pattern. Some of these algorithms have been widely utilized in video coding applications. Examples include the three step search (TSS) [10], the new three step search (NTSS) [11], the four step search (4SS) [12], the diamond search (DS) [13], Zhu et al. [14] suggested a diamond search (DS) method based on the characteristic of the center-biased motion vector (MV) distribution typically existed in the real-world video sequences, the gradient search, hierarchical search, and the cross search (CS). These algorithms reduce their computation by using techniques such as sub-sampling search area and by restricting the search window. Due to their uniform search strategy without differentiating the types of motion, these algorithms perform well only for certain video sequences.

The authors set up the research target to develop a coding algorithm with low complexity and low computational cost. I proposed fast disparity and motion estimation for reduce complexity, to solve the redundancy reduction problem by considering the information of regions, and set up priority among region based on information in the pre-processing.

This paper proposes algorithm to reduce computational complexity of ME and DE for MVC. I utilize multi-view camera geometry and the relationship between the disparity and motion vectors, and then determined a search range and reduced the searching points within the limited window.

The rest of this paper is organized as follows. Proposed Method of Adaptive Search Range Determination is presented in Section 2. The experimental results are shown in Section 3.. Finally, the conclusion is drawn in Section 4.

II. Proposed Method

Disparity estimation computed in every group of pictures enables the application of virtual view construction. By knowing the mapping between the left and right images, an intermediate view can be reconstructed through interpolation. In the case of our MVC model, a mapping between

A. Initialization of Disparity

I define the epipolar geometry between a pair of images. Let C and C' be a pair of pinhole cameras in 3D space. Let m and m' be the projections through cameras C and C' of a 3D point Min images I and I' respectively. The geometry of these definitions is shown in Fig. 2. The epipolar constraint is defined as follows.

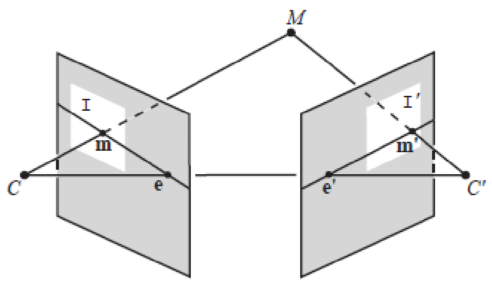

Fig 2 Epipolar geometry between a pair of images.

Image rectification can be view as the process of transforming the epipolar geometry of a pair of images into a canonical form. This is accomplished by applying a homography to each image that maps the epipole to a predetermined point as shown in Fig. 3.

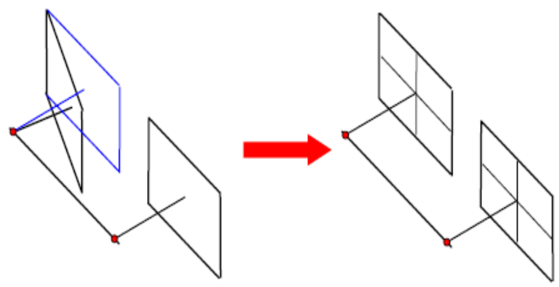

Fig 3 Image rectify.

B. Motion Estimation and Search Range Determination

In order to track a moving object on an image efficiently, position of moving object in the next timing frame should be properly estimated, and its search range also should be then determined in appropriate size. The proposed motion estimation method of moving object on the next timing frame and the proposed search range determination method are described in below.

- Motion Estimation

The position of moving object in the next timing frame can be estimated by using the distance of search point (r) of a previous pair frames which can be simply expressed by the following equation

$$r = \sqrt{|\Delta x|^2 + |\Delta y|^2} \quad (1)$$

where Δx and Δy stand for distance in x and y direction respectively. In case of the distance is zero, It probably means some errors occur, and full search as same as initial state needs to be performed for finding moving object.

- Search Range Determination

Suppose the search point representing object position in the next frame is already estimate in the previous process, search range determination is then start. In this paper, I apply mode complexity (MDC) and global disparity vector (GDV) for determining searches range located on the search point on the next frame. I decide the tested modes of each MB based as located the corresponding position in the next frame. The mode size is derived from the corresponding MB and use to compute MDC shown in Fig. 4. I can find the new search range by using dynamically adjusted. It can divide 4 region which depend on movement frame rate.

$$GDV = \begin{cases} Low: \frac{SR}{4} \times \frac{SR}{4} \\ Medium: \frac{SR}{2} \times \frac{SR}{2} \\ High: SR \times SR \\ Otherwise: FS \end{cases} \quad (2)$$

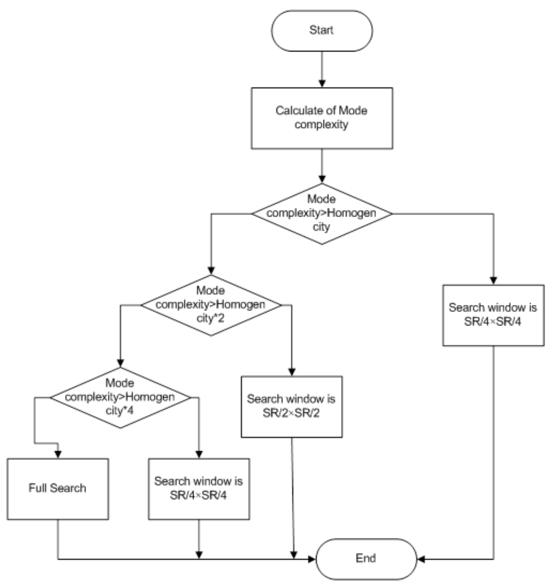

Fig 4 Flowchart of the proposed adaptive motion search range.

ISBN: 978-988-19253-8-1
ISSN: 2078-0958 (Print); ISSN: 2078-0966 (Online)

In most cases, the predicted search range works well. However, there are exceptions. To avoid the over limited search range being set, I propose two additional methods, early termination of ME and adaptive search range modification. Early termination is used to reduce the search range for homogeneous motion MBs that are misclassified as complex motion. A large search range should be unnecessary if the resulted block-matching residuals are already quite small with all-zeros after quantization. Calculate the mode complexity MDC by use global disparity vector to located the corresponding picture of the neighbor view and the information of mode size is derived from the corresponding MB and use to compute MDC

$$MDC = \frac{1}{N} \times \sum_{i=0}^{n-1} W_i \qquad (3)$$

Where W_i is is the MB mode factor defines;

III. Experiments

A. Experimental Setup

I evaluated the coding complexity and performance of the proposed algorithms using various multi-view sequences. The test sequences Flamenco1 were provided by the KDDI. In addition, the test sequence Exit came from MERL (Mitsubishi Electric Research Laboratories) and “Breakdancers” and “Ballet” came from Microsoft Research. The test sequences are shown in Fig. 5. Table I describes the properties of the various test data sets.

Table I Data Set

Sequences	Resolution No. of Horizontal	Views	Camera Spacing
Breakdancers	1024x768	8	20 cm
Ballet	1024x768	8	20 cm
Flamenco1	640x480	8	20 cm
Exit	640x480	8	20 cm

The coding parameters for the motion search and view order in JMVM are listed in Table II. According to our experiment, if the search mode is full search or in a large search range like 96, In addition, since the motion prediction part in H.264/AVC was modified according to our method, reduction of bits for motion vectors is compared with fast search algorithm.

(a) Breakdancers

(b) Ballet

(c) Flamenco1

(d) Exit

Fig 5 Data set for Experiment

Table II Parameters in JMVM

Parameters	Value	Note
Search Mode	4	(0:BlockSearch, 4:FastSearch)
Search Range	8/96	Search Range
View Order	0-2-1-4-3-6-5-7	View Order

B. Evaluation

Evaluation of video coding algorithms can be done using objective and subjective measures and processing time all the process. The most widely used objective measure is the peak signal-to-noise-ratio (PSNR) of the signal which is given as

$$PSNR = 10\log\frac{255^2}{MSE} \qquad (4)$$

with being the mean squared error (MSE) between the original and decoded video samples.

C. Experimental Results

This section includes the performance of efficiency MVC. I have simulate Four MVC test sequences which the JMVM software. I compare the proposed method with full search algorithm, motion estimation without vary the search range, and conventional method in term of processing time and PSNR. The PSNR results are shown in Fig.6 According to the PSNR figures, the PSNR can improve from another method. When I compare with FS algorithm, PSNR can improve 0.2-0.3 dB.

The result of processing time is shown in table III. The proposed method uses processing time less than another method except fixed search range method. When I compare with fixed search range method, the processing time do not too difference but PSNR can improve ostensibly. The proposed method can reduce the search candidate by determine the new search range so that it can reduces estimation time and speed up all process.

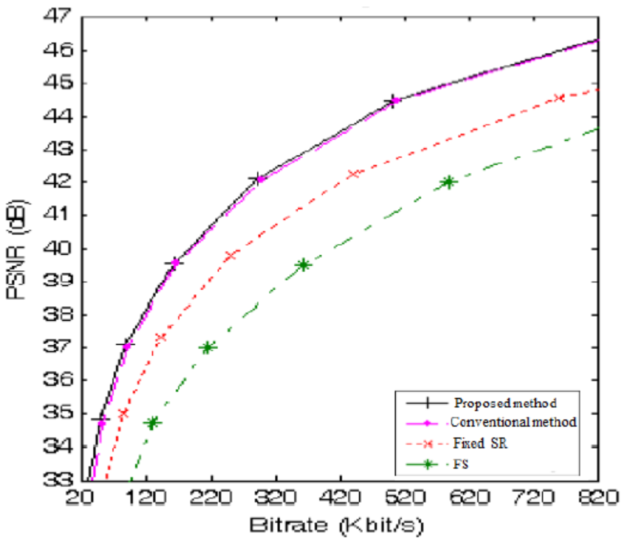

Fig 6 Compares the rate-distortion performance

ISBN: 978-988-19253-8-1
ISSN: 2078-0958 (Print); ISSN: 2078-0966 (Online)

Table III Comparison of Processing Time

Methods	Test Data		
	Ballet	*Breakdancer*	*Flamenco1*
Full Search Algorithm	18,378 ms	17,414 ms	17,890 ms
Proposed Method	13,900 ms	14,904 ms	14,250 ms
Fixed Search Range	13,500 ms	14,700 ms	14,200 ms
Conventional Method	16,790 ms	16,700 ms	16,650 ms

This section includes the performance of efficiency MVC. I have simulate Four MVC test sequences which the JMVM software. I compare the proposed method with full search algorithm, motion estimation without vary the search range, and conventional method in term of processing time and PSNR. The PSNR results are shown in Fig.6 According to the PSNR figures, the PSNR can improve from another method. When I compare with FS algorithm, PSNR can improve 0.2-0.3 dB.

The result of processing time is shown in table III. The proposed method uses processing time less than another method except fixed search range method. When I compare with fixed search range method, the processing time do not too difference but PSNR can improve ostensibly. The proposed method can reduce the search candidate by determine the new search range so that it can reduces estimation time and speed up all process.

IV. Conclusion

This paper proposed fast disparity and motion estimation for reduce complexity to solve the redundant reduction problem by determining the optimize search range. The new search range is calculated from global disparity vector. This proposed tools aim to reduce the search region and the number of reference frames of searching candidate. The search range can vary area depended on object moving rate. The multi-view video coding based on the H.264 by using geometry based to find disparity estimation. The experimental results show that the proposed method can reduce the encoding time, so that the encoder complexity can be reduced but achieves almost the same compress quality. The encoding time reduce up to 25.4%.

References

[1] A. Smolic and P. Kauff, ”Interactive 3-D Video Representation and Coding Technologies,” Proceedings of IEEE , vol.93, No. 1, pp. 98-110.

[2] T.Wiegand, J. Sullivan, G. Bjontegaard, and A. Luthra, ”Ovrelated overview of the H.264 / AVC Video Coding Standard,” IEEE Transactions on Circuits and Systems for Video Technology, 2003.

[3] G. J. Sullivan, P. Topiwala, and A. Luthra, ”The H.264/AVC Advanced

Video Coding Standard: Overview and Introduction to the Fidelity Range Extensions,” SPIE Conference on Applications of Digital Image Processing XXVII, Colorado, USA., 2004.

[4] H. Schwarz, D. Marpe, and T. Wiegand, ”Overview of the Scalable

H.264/MPEG4-AVC Extension,” Proceedings of the IEEE International

Conference on Image Processing,Georgia, USA., 2006.

[5] P. Merkle, K. Muller, A. Smolic, and T. Wiegand, ”Efficient Compression of Multi-view Video Exploiting Inter-view Dependencies based on h.264/mpeg4-avc,” IEEE International Conference on Multimedia and Expo, pp. 1717 - 1720, 2002.

[6] Masayuki Tanimoto, ”FTV (Free viewpoint TV) and Creation of Ray-Based Image Engineering,” ECTI Transaction on Electrical Engineering, Electronics and Communications, Vol. 6, No. 1, pp.3-14, 2008.

[7] K. Muller, P. Merkle, and T. Wiegand, ”Compressing Time-Varying Visual Content,” IEEE Signal Processing Magazine, Vol. 24, Issue 6, pp. 58 - 65, 2007.

[8] Iain E.Richardson, ”The H.264 Advanced Video Compression Standard,” John Wiley & Sons, UK., 2010

[9] Yu-Wen Huang, Ching-Yeh Chen, Chen-Han Tsai, Chun-Fu Shen and Liang-Gee Chen, ”Survey on Block Matching Motion Estimation Algorithms and Architectures with New Results,” Journal of VLSI Signal Processing , Springer, Vol.42, No.7, pp.297-320, 2006.

[10] Alan Conrad Bovik, ”The Essential Guide to Video processing,” Elsevier, USA., 2006 , (2009). , , 978-0-12-374456- 2, USA.

[11] Renxiang Li, Bing Zeng, and Liou, M.L. ”A new three-step search algorithm for block motion estimation,” IEEE Transactions on Circuits

and Systems for Video Technology, Vol. 4, Issue 4, pp. 438 - 442, 2002.

[12] Lai-Man Po, and Wing-Chung Ma. ”A Novel Four-Step Search Algorithm for Fast Block Motion Estimation,” IEEE Transactions Circuits Systytem for Video Technology, vol. 6, No. 3, pp. 313-317, Jun.

1996.

[13] J. Y. Tham, S. Ranganath, M. Ranganath, and A. A. Kassim, ”A novel unrestricted center-biased diamond search algorithm for block motion estimation,” IEEE Transactions on Circuits and Systems for Video Technology, Vol. 8, no. 4, pp. 369377, 1998.

[14] S. Zhu and K. K. Ma, ”A new diamond search algorithm for fast block matching,” IEEE Transactions on Circuits and Systems for Video Technology, Vol. 9, no. 2, pp. 287290, 2000.