Motion Vector Predictor Selection and Adaptive Refinement Process for H.264/AVC Enhanced Predictive Zonal Search

Ji-Hee Moon and Yo-Sung Ho
Gwangju Institute of Science and Technology (GIST)
261 Cheomdan-gwagiro, Buk-gu, Gwangju, 500-712, South Korea
{jhmoon, hoyo}@gist.ac.kr

Abstract

Key technology area (KTA) proposed larger motion partition sizes than the macroblock of 16× 16 pixels to improve the performance of H.264/AVC in high resolution video coding. Even if coding efficiency is improved, it has high complexity due to increased motion partition blocks. Especially, the motion estimation part takes a large portion of complexity. Enhanced predictive zonal search (EPZS) is recommended as the motion vector search method for high resolution video due to its low computational complexity. In this paper, we propose a motion vector predictor selection method and adaptive refinement process for EPZS. The proposed method reduces the computational complexity of EPZS by 15.92% without any visible performance degradations.

Keywords--- **Motion vector, EPZS, Motion estimation**

1. Introduction

H.264/AVC is the latest video coding standard that provides high coding efficiency. It contains a number of new features that allow it to compress the video much more effectively than previous video coding standards. For higher compression efficiency, H.264/AVC has adapted several powerful coding techniques, such as variable block-size macroblock modes, multiple reference frames, integer discrete cosine transform (DCT), and efficient entropy coding techniques.

However, these coding tools cannot provide sufficient performance in the high resolution video sequences. In general, the motion estimation part is not suitable for real-time implementation due to high complexity. Especially, motion estimation of H.264/AVC takes approximately 50~60 % of the total encoding time [1]. Several researches on fast motion estimation have been proposed to reduce the complexity of motion estimation.

A new standardization initiative is referred to as high efficiency video coding (HEVC) and its development has been undertaken by joint collaborative team on video coding (JCT-VC) [2]. HEVC is intended to provide significantly better compression capability than existing standards.

In order to achieve higher coding efficiency in high resolution sequences, they have proposed several video coding tools using motion partitions larger than 16×16 [3]. It has been demonstrated that video coding with blocks bigger than 16×16 pixels can provide substantial coding gains compared to H.264/AVC, especially when applied to high resolution video sequences. The correlation between neighboring pixels is increased in high resolution contents. Therefore, when we apply the video coding tools using blocks bigger than 16×16 pixels to high resolution video sequence, the residual error can be reduced effectively.

Unfortunately, motion estimation using larger block sizes requires a significant increase of the encoding time. There are lots of integer pixel values to be calculated for motion estimation. It is a disadvantage of video coding using large block sizes. Thus, enhanced predictive zonal search (EPZS) for high resolution video coding is recommended as a motion estimation method.

In this paper, we propose new motion vector predictor selection and adaptive refinement processes to enhance the predictive zonal search. The search range is determined by the spatial correlation of neighboring motion vectors. If the motion vector predictor exceeds a determined search range, we do not calculate the motion cost of the motion vector predictor. Besides, we apply the refinement process adaptively to reference frames except for the first reference frame. Our experimental results demonstrate that more than 15.92% of the motion estimation time is reduced

2. Predictive zonal search in H.264/AVC

In general, most fast motion estimation methods are characterized by three different operations: 1) early termination, 2) generation of predictor set, and 3) various search patterns. The predictor set is used to start a search point. Motion vector predictors have high probability to be the best motion vector. The early termination condition determines to terminate the search at any checking point. Various search patterns are used for optimizing and improving the search. In this section, we explain each part of the EPZS algorithm [4] in detail.

2.1. Early termination

The threshold value for early termination determines whether the search process should be terminated or not. The threshold value affects the performance of motion estimation and its complexity. Even if a large threshold value can achieve high speed up of motion estimation, it reduces the performance of motion estimation. A small threshold value improves the performance of motion estimation. However, it spends more time of motion estimation compared to that of a large threshold value. Thus, it is important to select a proper threshold value.

The threshold value for early termination considers the smallest sum of absolute differences (SAD) in neighboring blocks of the current macroblock. The median value among the smallest SAD and predefined threshold values is calculated. We can modify the predefined threshold value. The final early termination threshold value is calculated by weighted averaging.

2.2. Generation of predictor set

Each motion vector predictor is important for the refinement process. The median motion vector predictor is checked at the first time. The median motion vector predictor has higher probability to be the best motion vector. If the cost of the median motion vector predictor is smaller than the early termination threshold value, motion estimation is terminated. Otherwise, the predictor set is generated. All predictors are based on the spatial and temporal correlation of neighboring motion vectors and the macroblock type.

After the predictor set is generated, the motion cost of each motion vector predictor is calculated. Among all motion vector predictors, two motion vector predictors that have small motion cost are obtained. If the smallest motion cost is larger than the early termination threshold value, the refinement process is applied to the two motion vector predictors. Otherwise, the motion vector predictor is set as the best motion vector, and the motion estimation process is completed.

2.3. Refinement process

The refinement process tries to avoid the local minimum value. Two motion vector predictors are important for the refinement process. Figure 1 shows several search patterns for the refinement process.

These patterns are determined by the macroblock type and the distance between the current macroblock and reference frames. The refinement process is performed with two motion vector predictors. The refinement process is continued when the best matched point is located at the center point of the search pattern. After the refinement process is finished, the motion vector that has the smallest motion cost is decided as the best motion vector of the current macrocblock.

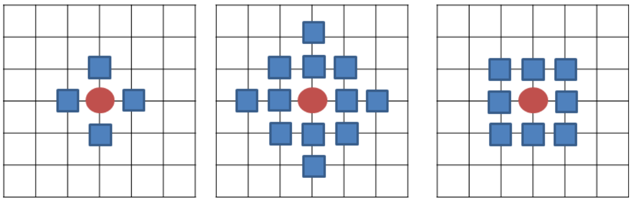


Figure 1. Refinement process pattern

3. Proposed algorithm

The proposed algorithm consists of two parts: 1) reduction of motion vector predictors and 2) adaptive refinement process. In order to reduce the motion estimation time, we remove unnecessary search points in each reference frame based on the spatial correlation. Then, we apply the refinement process adaptively based on the temporal correlation.

3.1. Reduction of motion vector predictors

In order to remove unnecessary motion vector predictors, we use the spatial correlation between neighboring motion vectors. Considering the variance of neighboring motion vectors, we assign a new search range in each direction. We can determine whether each motion vector predictor is included in the new search range or not. We do not calculate the motion cost of the motion vector predictor that exceeds the determined search range. Figure 2 explains the reduction of motion vector predictors.

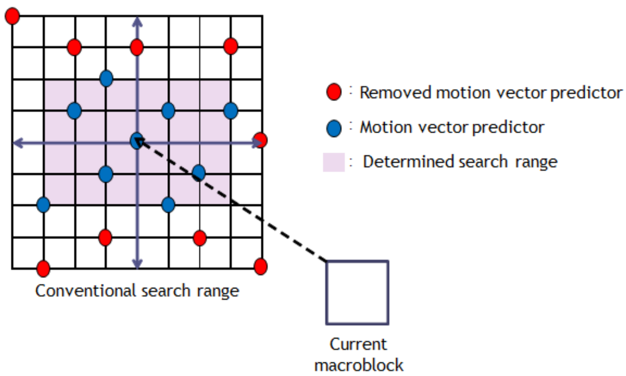


Figure 2. Reduction of motion vector predictors

When the search range is reduced, we may remove the best matching point. In order to guarantee improved coding efficiency over EPZS, we apply a nonlinear function. In order to avoid a situation that the determined search range is larger than the initial search range, we modify the search range by the scaling factor. The motion search range is determined by the spatial correlation.

$$R = max\{1, \alpha\sqrt{standard\ deviation(MVs)}\} \quad (1)$$

where R is the determined search range based on the spatial correlation using the standard deviation of neighboring motion vectors. *MVs* are motion vectors of the neighboring macroblocks.

3.2. Adaptive refinement process

The refinement process is applied over all the multiple reference frames. We can observe that the nearest reference frame is likely to be selected as the best reference frame for the current macroblock, as shown in Fig. 3. The probability that the best reference frame is selected from the nearest reference frame to the third reference frame is approximately 95%. It means that the computation cost is increased, especially when we apply the refinement process to farther reference frames.

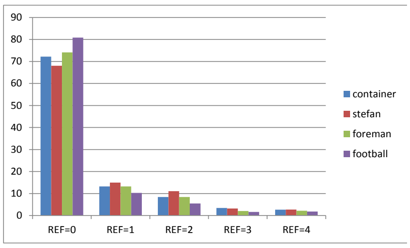


Figure 3. Probability of reference frames

In order to reduce the computation cost except for the first reference frame, the smallest motion cost of the first reference frame is the threshold value that determines whether the refinement process is applied or not in other reference frames. If the smallest motion cost of each reference frame is smaller than a new threshold value, the refinement process is applied. Otherwise, we do not apply the refinement process. Since the reference frame without the first frame is rarely chosen as the best reference frame we can reduce the computation cost through an adaptive refinement process. Figure 4 explains the proposed adaptive refinement process.

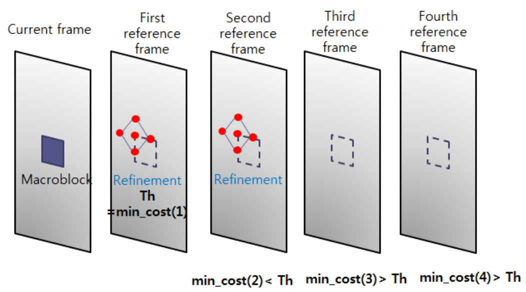


Figure 4. Proposed refinement process

As shown in Fig. 4, min_cost(i) is the minimum cost of the best motion vector in each reference frame. We compare the smallest cost in each frame with a new threshold value. Since the probability that a farther reference frame is selected as the best frame is low, it is not necessary to apply the refinement process to all the reference frames. Figure 5 shows the flow chart of our adaptive refinement process.

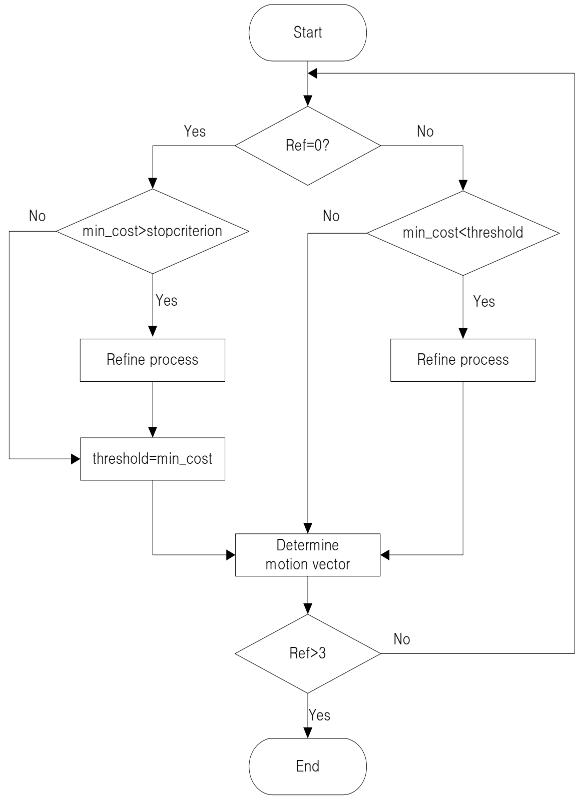


Figure 5. Adaptive refinement process

4. Experimental results

We have implemented the proposed algorithm on KTA 2.7 [5] and tested several test sequences of HD resolution. Table 1 summarizes the encoding parameters for the reference software.

Table 1. Encoding parameters

Profile	100 (high profile)
QP	22, 27, 32, 37
SearchRange	96
SymbolMode	1 (CABAC)
FrameStructure	IPPP…P
UseIntraMDDT	1
UseExtMB	2 (64×64)
Transform8×8mode	1
SearchMode	3 (EPZS)
NumberReferenceframes	4

We have compared percentage time savings relative to KTA by

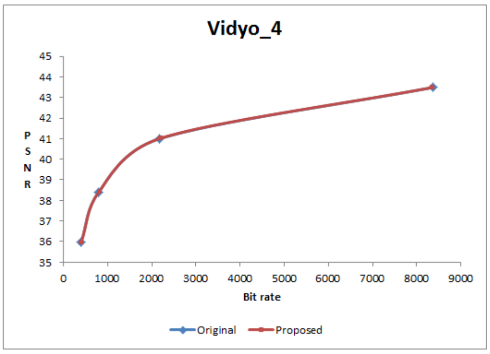
$$\Delta Time\,saving(\%) = \frac{Time_{KTA} - Time_{proposed}}{Time_{KTA}} \times 100 \qquad (2)$$

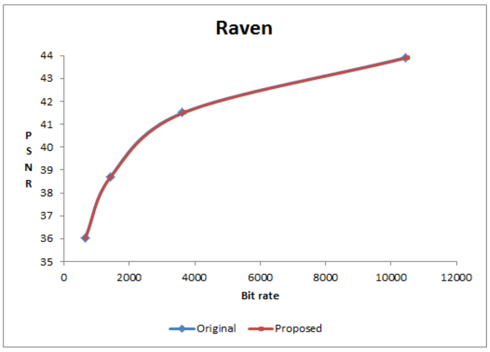
Table 2 shows that the proposed algorithm provides approximately 15.92% time savings, compared to the conventional KTA EPZS method.

Table 2. Comparison of processing time savings

Sequence	QP	Original	Proposed	△Time (%)
Night	22	47791.38	41094.92	-14.01
	27	41057.20	34547.12	-15.86
	32	35688.50	29536.75	-17.24
	37	31103.24	25667.72	-17.48
Crew	22	45969.38	39287.40	-14.54
	27	36089.86	30024.73	-16.81
	32	32471.74	26532.78	-18.30
	37	29387.82	23877.96	-18.75
Raven	22	36224.93	28972.75	-20.02
	27	29918.33	23900.96	-20.11
	32	27251.87	22010.67	-19.23
	37	25243.53	20986.38	-16.86
Vidyo_4	22	27153.30	24616.21	-9.34
	27	22774.60	20283.27	-10.94
	32	21255.46	18492.76	-12.98
	37	20041.87	17572.50	-12.32
Average				-15.92

Figure 6 shows rate distortion curves. As shown in Fig. 6, there is no significant degradation of PSNR values and bit rates.





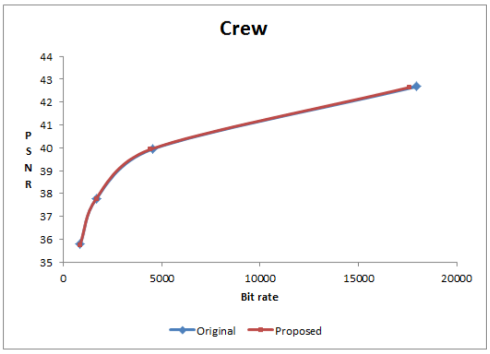


Figure 6. Rate-distortion curves

5. Conclusions

We have proposed adaptive motion vector predictor selection and refinement processes for the enhanced predictive zonal search (EPZS) in KTA. After we determine a new search range based on the variance of neighboring motion vectors, we remove the motion vector predictors that exceed the new search range. In addition, the smallest motion cost of the first reference frame is selected as a threshold value that determines the refinement process for the rest of reference frames. We can reduce the search time by 15.92% without any significant degradation.

Acknowledgements

This research was supported by MKE under the ITRC support program supervised by NIPA (NIPA-2011-(C1090-1111-0003)).

References

[1] J.H. Moon and Y.S. Ho, Fast Motion Estimation with Extended Block Sizes for High-definition Video Coding. *Asia-Pacific Signal and Information Processing Association (APSIPA)* 2010, pp. 619-622, Dec. 2010.

[2] Joint Call for Proposals on Video Compression Technology, ISO/IEC MPEG 91st Meeting, Kyoto, Japan, Document N11113, April 2010.

[3] "Technical consideration for Ad Hoc Group on New Challenges in Video Coding Standardization," ISO/IECJTC/SC29/WG11Document M1558, July 2008, Hannover, Germany.

[4] A.M. Tourapis, Enhanced Predictive Zonal Search for Single and Multiple Frame Motion Estimation. *Visual Communications and Image Processing*, pp.1069-1079, Jan. 2002.

[5] http://iphome.hhi.de/suehring/tml/download/KTA/jm11.0kta2.7