Motion Vector Coding using Decoder-side Estimation of Motion Vector

Kwanghyun Won, Jungyoup Yang, and Byeungwoo Jeon, *Senior Member, IEEE*

***Abstract*— The H.264/AVC standard employs the predictive motion vector coding technique using the median predictor of spatially neighboring three motion vectors. Although the median is effective in reducing redundancy, it is not always optimal in minimizing bits. To solve the matter, a new motion vector coding scheme, known as, MV competition in which decoder is signaled on the selected optimal PMV, has been reported. Though it can use the optimal PMV(Predicted Motion Vector), the bits consumed to indicating the optimal PMV to the decoder increases bit-rate. In this paper, we propose a new motion vector coding scheme that allows usage of an optimal PMV without consuming additional bits to inform the choice of PMV to decoder. Simulation results show that the proposed method gains in BDBR by 3.22% on average, and in BDPSNR by 0.13dB compared to the H.264/AVC.**

***Index Terms*— motion vector coding, predictive coding, predicted motion vector, template matching.**

I. Introduction

The H.264/AVC standard has superior compression performance compared to its predecessors due to improved coding tools such as intra prediction, 1/4-pel motion compensated inter prediction with variable block sizes, multiple reference pictures, context adaptive entropy coding, in-loop deblocking filtering, and so on [1]. All these coding tools contribute to enhancing video quality at minimum coded data of mode information, textual information, and MV (motion vector) information. Particularly, the motion compensated inter-picture prediction technique achieves high compression by reducing data for textual information by effective exploitation of temporal correlation using reference pictures. However, the amount of MV information increases in order to represent accurate displacement. Note that considerable portion of compressed bits is consumed in encoding MV data - they take about 40%~50% of the total bit-stream at low bit-rate case [2], for example.

The H.264/AVC standard applies predictive coding technique using the PMV which is selected as a median value of MVs of three spatially neighboring blocks - upper, left, and upper-right blocks. The median PMV is likely to be similar to the current MV, however, it does not necessarily lead to minimal MV bit-rate always. If the other PMV being different from the median is more similar to the current MV, there does exist chance to better compress the MV by using the other PMV.

To utilize this opportunity, a method called the MV competition has been proposed at ITU-T VCEG(Video Coding Experts Group) and adopted to KTA(Key Technology Area) software which amasses state-of-the-art video coding tools for future video coding standardization effort [3]. It selects an optimal PMV among spatial as well as temporal candidate PMVs using a competing scheme. It is always capable of selecting a PMV optimal in the sense of minimum bitrate of MV information. However, the transmission of index information to inform decoder of the predictor does not always guarantee the minimal bitrate even though the optimal PMV with minimal MV bitrate is selected. It is because the bit increment by sending index of the PMV may make it overwhelm the bit-saving in MV coding. If the number of candidate PMVs is increased, the additional information for index also increases in proportion to the number of candidate PMVs.

In this paper, an MV coding scheme using an optimal PMV giving minimal bitrate is proposed. It requires no additional information of signaling the optimal PMV. In the proposed method, the encoder makes a list of possible candidate PMVs using spatio-temporally neighboring MVs, and then encoder classifies specific PMVs in the list which can be decided at the decoder by matching the reconstructed neighboring pixel data of current block with those of the block in the reference picture as indicated by the recovered motion vector [4]. Lastly, the encoder selects the optimal PMV giving minimal bitrate in the specific PMVs. If the optimal PMV cannot be uniquely found at the decoder using the matching criterion, the encoder selects the median as a default PMV.

II. Conventional Motion Vector Coding

A. Motion Vector Coding in H.264/AVC Standard

The coding scheme of motion vector in H.264/AVC standard is predictive coding. The notation DMV indicates a difference vector between the current MV and the predicted MV. As shown in Fig.1, where $\overrightarrow{mv^D}$ denotes MV of current block, $\overrightarrow{mv^A}$, $\overrightarrow{mv^B}$ and $\overrightarrow{mv^C}$, respectively denote MVs in spatially neighboring blocks of the current block, and $\overrightarrow{pmv^D}$ denotes MV which both the encoder and the decoder can predict. The H.264/AVC standard uses median PMV expressed as:

Manuscript received April 09, 2009.

K. Won, J. Yang, and B. Jeon is with School of Information and Computer Engineering, Sungkyunkwan University 300 Chunchun-dong, Jangan-gu, Suwon, 440-746, Korea (e-mail: wkh12345@skku.edu; binbak1@ece.skku.ac.kr; bjeon@skku.edu).

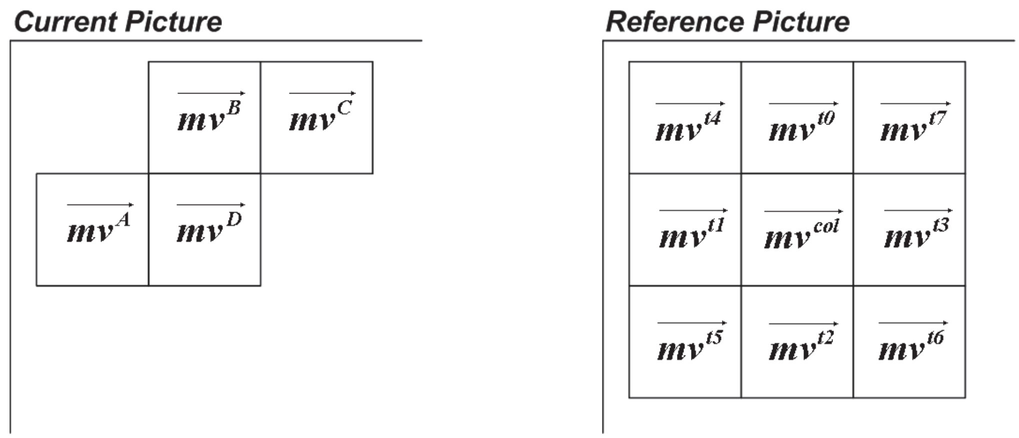

Fig. 1. Motion vector of current block(D) and spatially neighboring blocks(A, B, and C) in current picture, and temporally co-located block(Col) and others(t0~t7) in reference picture.

$$\overrightarrow{pmv^{D(H.264/AVC)}} = \begin{bmatrix} Median(mv_x^A, mv_x^B, mv_x^C) \\ Median(mv_y^A, mv_y^B, mv_y^C) \end{bmatrix} \quad (1)$$

where $Median(\bullet)$ is a function giving median value of its arguments. In case of using the median as PMV, the encoder does not need to transmit additional information to indicate median to decoder since the decoder can generate the same median value for itself. However, the median PMV gives ineffective MV bitrate when the other PMV being different from the median is more similar to the current MV.

B. *Motion Vector Competition Method in KTA Software*

To take advantage of the potential problem mentioned above, MV competition method has been proposed at ITU-T VCEG and adopted to KTA software. It employs not only spatially neighboring candidate PMVs, but also temporally co-located candidate PMVs as illustrated in Fig.1 and Eq. (2):

$$\overrightarrow{pmv^{D(Conv.method)}} = f\left(\overrightarrow{mv^A}, \ldots, \overrightarrow{mv^{Col}}, \ldots\right) \quad (2)$$

The function $f(\bullet)$ selects an MV among its arguments $\left(\overrightarrow{mv^A}, \ldots, \overrightarrow{mv^{Col}}, \ldots\right)$ which gives minimum bitrate. The MV competition needs to inform decoder of the PMV selection. If the number of candidate PMVs is increased, the number of signaling bits is increased as in Eq. (3).

$$signaling\ bit = \left\lceil \log_2^N \right\rceil (N = number\ of\ candidate\ PMVs) \quad (3)$$

III. Proposed Motion Vector Coding

The proposed algorithm selects the optimal PMV giving maximal bit-saving without spending additional information indicating PMV selection. The proposed method selects a PMV among spatio-temporally candidate PMVs using following equation:

$$\overrightarrow{pmv^{D(Prop.method)}} = g\left(\overrightarrow{mv^A}, \ldots, \overrightarrow{mv^{Col}}, \ldots\right) \quad (4)$$

Where $g(\bullet)$ denotes a function which selects an MV among its arguments $\left(\overrightarrow{mv^A}, \ldots, \overrightarrow{mv^{Col}}, \ldots\right)$ which gives minimum template matching error by comparing reconstructed neighboring pixel data adjacent to the current block at current picture with those in a block in the reference picture indicated by the recovered motion vector as described in Fig.2 [4]. The reconstructed neighboring block in reference picture is derived from tentatively decoded MV for each candidate PMVs as seen in Fig.2 as follows:

$$\overrightarrow{mv_n^{D(temp)}} = \overrightarrow{dmv^D} + \overrightarrow{pmv_n^{D(temp)}} \quad (5)$$

The matching procedure makes it possible to take the same action taken at the encoder also at the decoder because it operates on just a set of reconstructed pixel data which are already available at decoder. The encoder categorizes particular PMVs with minimal matching error. Lastly, the encoder chooses an optimal PMV which generates minimal number of coded bits of the DMV which the encoder transmits to decoder. However, the proposed algorithm cannot guarantee that the decoder always estimate the same optimal PMV using the matching criterion because it is possible to find other candidate PMV at the decoder. If the situation happens, the encoder chooses the median as a PMV. Therefore, the encoder transmits a signaling flag to let the decoder to differentiate these two cases.

IV. Experimental Results

A. *Experimental conditions*

In order to evaluate the performance of the proposed method, the proposed scheme is implemented in KTA software version 2.0 reference software, and then applied to all possible

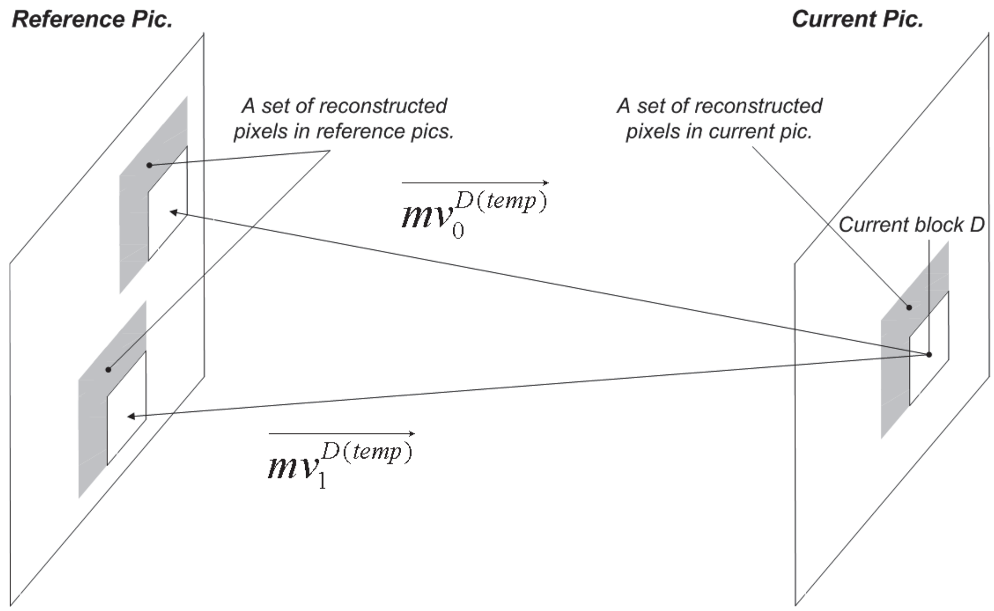

Fig. 2. Optimal PMV Estimation at Decoder using Template Matching

macroblock modes (SKIP, P16x16, P16x8, P8x16, P8x8 and all the sub-macroblocks) [5]. The experiment is simulated under the test condition depicted in Table I. The comparison of coding performances of the H.264/AVC standard (*Anchor*), the MV competition scheme (*Conventional method*), and the proposed method (*Proposed*) are carried out in terms of BDBR (Bjøntegaard Delta Bit Rate) and BDPSNR (Bjøntegaard Delta Peak Signal to Noise Ratio) [6]. The proposed method is compared with the conventional method which exploits 2 candidate PMVs. Note that the MV competition achieves the best coding performance with 2 candidate PMVs.

B. Experimental results

Table II shows experimental results of the proposed and MV competition method compared to the anchor. As described in Table II, the first column of the simulated results shows that the conventional method achieves high coding efficiency compared to the anchor because of its capability of making appropriate trade-off between mode information and MV information. The second column of the simulated results shows that the proposed method fulfills the best coding performance on average because there is no consumption of bits for the indexing information. The results show that the proposed method and the conventional method achieve BDBR gain of respectively, 3.22% and 3.05% on average compared to the H.264/AVC standard. For the crew sequence of 720p, the proposed method is performing very well by showing maximum PSNR gain of 0.15 dB and bit savings of 5.24% against the anchor. Note also that, the conventional method sometimes performs better then the proposed method for the Coastguard, the Bigship, and the City sequences which have slow and linear motion.

Fig.3 further verifies the results of Table II. It can be observed that the proposed method generates more SKIP modes than both the H.264/AVC standard and the conventional method in all resolutions and bitrates. It is because the proposed MV coding scheme can use the most appropriate PMV so that it promotes more zero DMVs, thus satisfying one condition of being a SKIP mode (that is, DMV=(0,0)). This explains why the proposed method attains higher compression of MV information than the H.264/AVC standard and the conventional method.

Since the proposed method does not transmit indicative information of selected PMV, it is expected to have more gain

TABLE I
EXPERIMENTAL CONDITIONS

Sequence size	QCIF	CIF	720p
Sequence name	Carphone Foreman Stefan	Coastguard Foreman Stefan	Bigship City Crew
Encoded picture	150	300	150
Frame rate (Frame skip)	30hz (1)	30hz (0)	60hz (0)
Motion Estimation	Full Search (±32)	Full Search (±32)	EPZ Search (±64)
GOP structure	IPPP…		
QP	QPI : 22, 27, 32, 37, QPP: QPI + 1		
Reference picture	4		
Other coding option	Adaptive rounding with factor 8, RDO on, 1/4-pel ME accuracy, 1 picture = 1 slice, CAVLC		
Comparison	*Anchor* (H.264/AVC standard scheme)		
	Conventional method (MV competition scheme with 2 candidate PMVs)		
	Proposed method (The proposed scheme with 5 candidate PMVs)		

TABLE II
EXPERIMENTAL RESULTS

Sequence		Conventional method vs.Anchor		Proposed method vs.Anchor	
		BDPSNR [dB]	BDBR [%]	BDPSNR [dB]	BDBR [%]
QCIF	*Carphone*	0.04	-0.70	0.07	-1.35
	Foreman	0.09	-1.70	0.14	-2.46
	Stefan	0.07	-1.01	0.12	-1.72
CIF	*Coastguard*	0.09	-2.20	0.07	-1.66
	Foreman	0.16	-3.79	0.17	-4.16
	Stefan	0.13	-2.36	0.14	-2.56
720p	*Bigship*	0.15	-5.52	0.14	-5.06
	City	0.20	-6.13	0.16	-4.80
	Crew	0.00	-4.02	0.15	-5.24
Average of QCIF		**0.07**	**-1.13**	**0.11**	**-1.84**
Average of CIF		**0.12**	**-2.78**	**0.13**	**-2.79**
Average of 720p		**0.12**	**-5.22**	**0.15**	**-5.03**
Average of overall		**0.10**	**-3.05**	**0.13**	**-3.22**

*BDPSNR (+ is gain against the anchor)
BDBR (- is gain against the anchor)

in contrast to the conventional method of motion vector competition scheme if the number of candidate of PMV is increased.

V. CONCLUSION

In this paper, we proposed an MV coding scheme capable of using optimal PMV in the sense of minimal bitrate without its indexing overhead. The proposed method shows the coding performance gain about 3.22% bit-saving on average compared to the H.264/AVC standard. Especially, compared to the MV Competition method, the proposed method provides improvement in coding efficiency by allowing flexibility in choosing PMV to encode.

ACKNOWLEDGMENT

This work was supported by the Korea Science and Engineering Foundation (KOSEF) NRL Program grant funded by the Korea government (MEST) (ROA-2006-000-10826-0 (2008))

REFERENCES

[1] JVT of ISO/IEC MPEG and ITU-T VCEG, Draft ITU-T Recommendation and Final Draft International Standard of Joint Video Specification, ITU-T Rec. H.264, ISO/IEC 14496-10 AVC, 2003.
[2] G. Laroche, J. Jung, and B. Pesquet-Popescu, "A spatio-temporal competing scheme for the rate-distortion optimized selection and coding of motion vectors," Proc. 14th European Signal Processing Conf., Italy, 2006.
[3] J. Jung and G. Laroche, "Competition-Based Scheme for Motion Vector Selection and Coding," ITU-T SG16/Q6 Doc. VCEG-AC06, Hangzhou, October 2006.
[4] K. Sugimoto, M. Kobayashi, Y. Suzuki, S. Kato, and Choong Seng Boon, "Inter frame coding with template matching spatio-temporal prediction," Proc. IEEE Int. Conf. Image Processing, vol. 1, pp. 465-468, October, 2004.
[5] Key Technology Area Reference Software, Version 2.0(KTA2.0). [Online]. Available: http://www.tnt.uni-hannover.de/~vatis/kta/
[6] G. Bjøntegaard, "Calculation of average PSNR differences between RD-curves," ITU-T SG16/Q6 Doc. VCEG-M13, April 2001.

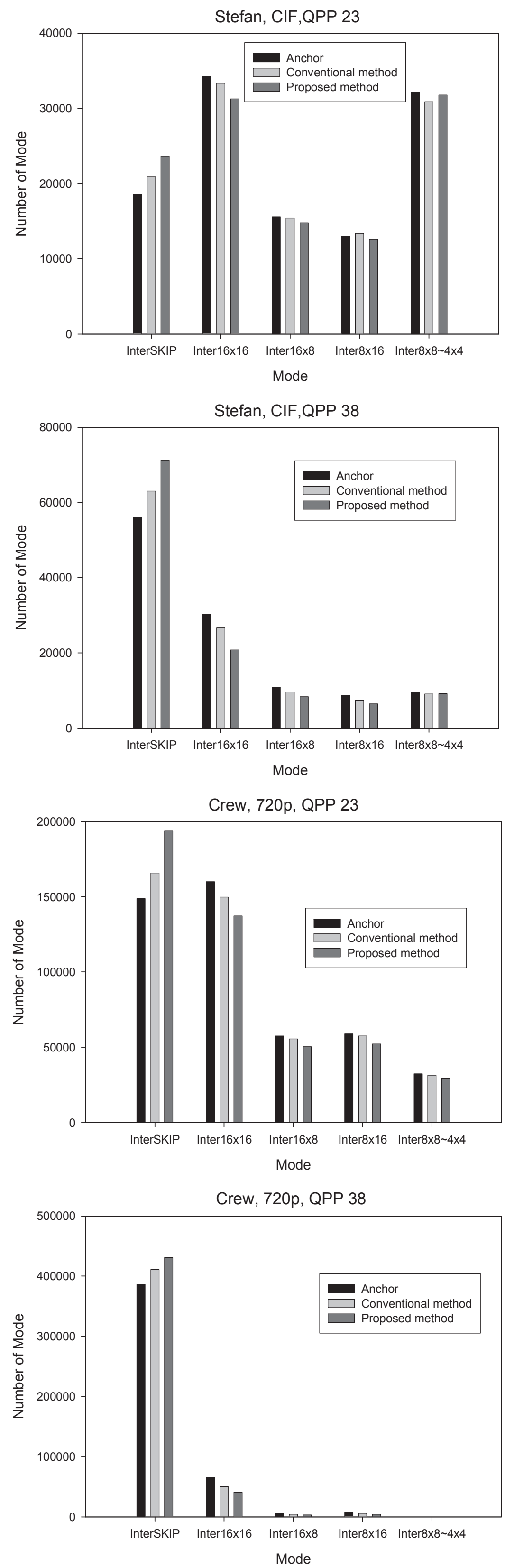

Fig. 3. Mode used for each block mode in sequence