일반논문-11-16-2-16

움직임 특성을 이용한 적응적 교차 움직임 벡터 부-복호화

원 광 현[a)], 양 정 엽[a)], 박 대 윤[a)], 전 병 우[a)‡]

Adaptive Interleaved Motion Vector Coding using Motion Characteristics

Kwanghyun Won[a)], Jungyoup Yang[a)], Daeyun Park[a)], and Byeungwoo Jeon[a)‡]

요 약

본 논문은 영상의 움직임 특성을 참조하여 움직임 벡터 수평-수직 성분중 하나를 최적의 예측 움직임 벡터 성분으로 선택하여 사용할 수 있는 개선된 교차 움직임벡터 부호화 방법을 제시한다. 제안방법은 부-복호화를 수행할 움직임 벡터 성분의 순서를 복호화기에 알리는 신호비트를 별도로 보낼 필요 없이, 시공간적 움직임필드의 특성을 참조하여 적절한 움직임벡터 성분 예측치를 적응적으로 사용할 수 있다. 기존 H.264/AVC 방법과 비교한 실험 결과, 평균적으로 전체 비트량을 약 1.99% (최대 8.71%) 감소시킬 수 있음을 보여준다.

Abstract

This paper proposes an improved design of an interleaved motion vector coding scheme with flexibility in predictive motion vector component by exploiting motion characteristics. It can use component-wise adaptive motion vector predictor based on the utility of spatial and temporal motion field without any signaling bit for indicating decoder of the selected predictive motion vector component. Experiment with test video data shows about 1.99% (max 8.71%) bit rate reduction compared to the conventional H.264/AVC method.

Keyword : H.264/AVC, interleaved motion vector coding, motion vector component, compression, motion characteristics

Ⅰ. Introduction

The H.264/AVC video coding standard has achieved substantially high compression performance compared to its predecessors thanks to variety of tools such as 1/4-pel motion compensated inter prediction, directional intra prediction, variable block sizes, multiple reference pictures, context adaptive entropy coding, in-loop deblocking filtering, and so on[1][2][3]. All those coding tools contribute to improving coded video quality with spending minimum coded data for mode, motion, and textural information. Especially, superior coding gain is brought by the motion compensated inter-picture prediction with 1/4-pel motion accuracy in variable block sizes. The data for textural information are effectively reduced by exploiting temporal

a) 성균관대학교 정보통신공학부 디지털미디어 연구실
Digital Media Lab., School of Information and Communication Engineering, Sungkyunkwan University
‡ 교신저자 : 전병우 (bjeon@skku.edu)
※ 이 논문은 2010년도 정부(교육과학기술부)의 재원으로 한국연구재단의 지원을 받아 수행된 기초연구사업 연구임(No. 20100008030).
· 접수일(2011년3월4일),수정일(2011년3월22일),게재확정일(2011년3월22일)

correlation using multiple reference pictures. However, the 1/4-pel accuracy in representing motion displacement leads to increasing motion information whose amount can reach up to 40% of the total bit-rate[4].

After the completion of H.264/AVC standardization, the ITU-T video coding experts group (VCEG) challenges the coding tools of the H.264/AVC for even higher compression, and various coding tools have been proposed and some of them are included in the key technical area (KTA) software[5][6]. Most coding tools proposed for better compressing motion vector (MV) require using additional signaling overhead bits which set some limit in the coding gain.

The Joint Collaborative Team on Video Coding (JCT-VC) of ITU-T VCEG and ISO/IEC MPEG has recently launched a standardization activity known as the high efficiency video coding (HEVC) with a goal of achieving at least 50% of compression gain against the existing H.264/AVC at equivalent quality[7]. In order to fulfill the goal, more precise and effective MV coding is necessary due to non-trivial bit portion of compressed MV data.

This paper presents a new way of MV coding in which one selected motion vector component is predicted first by the conventional median predictive motion vector while the other component is predicted by a corresponding component of a neighboring motion vector which is nearest to the current motion vector in its first selected component. The component selection is flexibly determined by measuring motion property of objects in picture. A criterion determining which motion vector component is predicted by median predictor is designed, and its implementation on KTA software is tested for evaluation of its performance gain against H.264/AVC.

The remainder of this paper is organized as follows. Prior works on MV coding relevant to the proposed method are presented in Section II. The proposed method of the flexible MV coding is described in Section III. Finally, Section IV presents simulation results using KTA software and discussion. Section V makes concluding remarks.

II. Related Works

1. MV Coding

The MV coding has been one of the main areas of research in the video coding, thus lots of idea already have been addressed much in the literature[8][9][10][11][12][13]. However, almost all those methods have exploited the basic concept of predictive coding as following:

$$DMV = MV - PMV \tag{1}$$

where MV denotes motion vector of a current block, PMV denotes its predictive MV (PMV), and DMV denotes differential MV (DMV) between MV of a current block and its PMV. Hence, selection of a proper PMV most similar to MV of a current block is the key to MV coding. The H.264/AVC uses, as the PMV, the median of MVs of three spatial neighboring blocks - left (A0), upper (B0), and upper-right (C) blocks shown in Fig. 1. In case of using the median as the PMV, the encoder does not need to transmit any signaling bits to inform decoder of the selected PMV since the decoder can generate the PMV for itself. However, the median may give ineffective coding efficiency of MV when the other PMV is more similar to the current MV.

In this context, a MV coding scheme based on competition has been researched to choose the optimal PMV[14]. The competition scheme selects a PMV which produces the minimum bit-rate among several MVs located in spatially and temporally neighboring blocks. The scheme is effective to reduce motion information because a PMV giving the minimum bit-rate is always used in Eq.(1). However, it comes with increment of overhead bits in order to notify

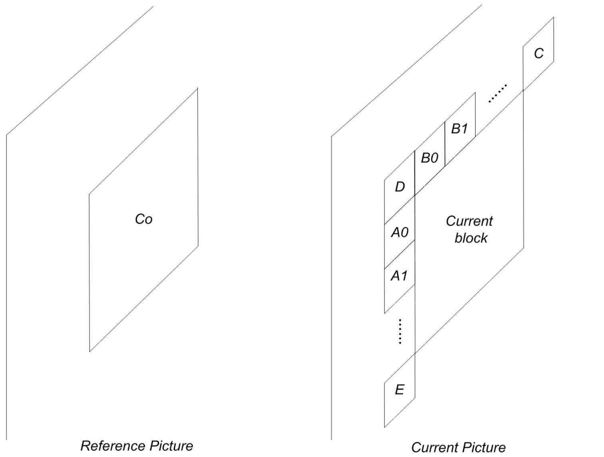

그림. 1. 현재 블록과 공간적으로 이웃한 블록들(A0, A1, …, B0, B1, …, C, D and E)과 시간적으로 이웃한 동일위치 블록(Co)
Fig. 1. Current and spatially neighboring blocks (A0, A1, …, B0, B1, …, C, D and E), and temporally co-located block (Co)

a decoder of the selected PMV. Furthermore, if the number of candidate PMVs is increased, the amount of signaling bits also goes up.

There has been also research on MV coding which does not require such indicating overhead[15]. The method chooses a PMV which can be predicted at the decoder by matching the reconstructed pixel data of current and reference blocks as indicated by the recovered motion vector. In this scheme, however, since the decoder may not identify the same PMV chosen at the encoder all the time, the encoder transmits signaling information to let the decoder differentiate the case in which the decoder cannot identify the same PMV of encoder’s choice.

2. Interleaved MV Coding

Of being closely relevant to the proposed design, there has been a report on interleaved MV coding which also does not require any signaling bit for the selected PMV[16]. As shown in Fig. 2, the scheme compresses a horizontal MV component using a predictor determined after finishing coding the vertical component first. The vertical PMV component of a current block is selected by using Eq.(2) as the vertical component of the median calculated in the same way as the MV coding of the H.264/AVC.

$$PMV_y = median(MV_y^{A0}, MV_y^{B0}, MV_y^{C}) \quad (2)$$

where y denotes the vertical direction, PMV_y denotes a vertical component of PMV, $median(\bullet)$ denotes a median function, MV_y^{A0} , MV_y^{B0} , and MV_y^{C} denote vertical components of corresponding MVs in spatially neighboring blocks as described in Fig. 1. After finishing coding the vertical component first, the horizontal component of PMV is chosen as that of a neighboring MV which minimizes absolute difference between the vertical MV components of current block and its spatially neighboring blocks of left (A0), upper (B0) and upper-right (C) blocks shown in

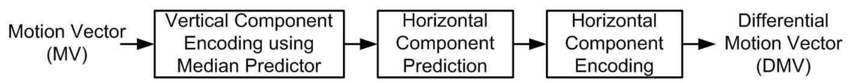

그림. 2. 교차 움직임 벡터 부-복호화 방법의 부호화 과정[16]
Fig. 2. Encoding in interleaved MV coding[16]

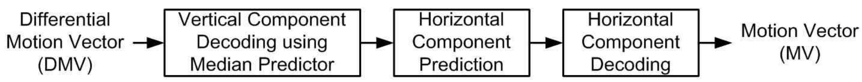

그림. 3. 교차 움직임 벡터 부-복호화 방법의 복호화 과정[16]
Fig. 3. Decoding in interleaved MV coding[16]

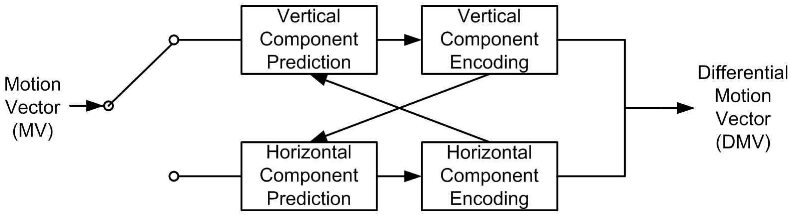

그림. 4. 제안한 움직임 벡터 부-복호화 방법의 부호화 과정
Fig. 4. Encoding in the proposed MV coding design

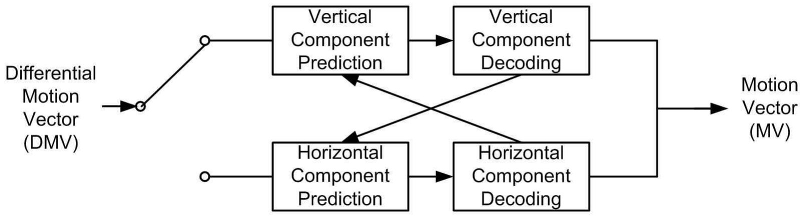

그림. 5. 제안한 움직임 벡터 부-복호화 방법의 복호화 과정
Fig. 5. Decoding in the proposed MV coding design

Fig.1. The selection is made by Eq.(3).

$$PMV_x = f\left(\underset{MV_y^{CS} \in CS}{\arg min} \left| MV_y^{Curr} - MV_y^{CS} \right|\right) \quad (3)$$

where x denotes the horizontal direction, PMV_x denotes a horizontal component of PMV, MV_y^{curr} denotes a vertical MV component of the current block, CS denotes a candidate set for vertical components of MVs in left (A0), upper (B0) and upper-right (C) blocks, $\arg min(\bullet)$ finds a vertical MV component among CS which minimizes the absolute difference, and the function $f(\bullet)$ identifies the horizontal component of a motion vector corresponding to the selected vertical MV component by $\arg min(\bullet)$. The horizontal component is coded using the identified PMV_x.

Its decoding process is almost identical to the encoding process as illustrated in Fig. 3. When DMV of a current block is parsed, the vertical MV component is first decoded using median PMV in Eq.(2). Following, the other PMV component is found using Eq.(3). Although this interleaved MV coding scheme is effective, its coding of vertical component always using median predictor regardless of the direction of object motion or camera panning may reduce coding efficiency when the motion characteristic does not match with the fixed scheme.

III. Proposed Motion Vector Coding

The proposed flexible interleaved MV coding approach seeks to enhance MV coding gain by adaptively coding each MV component using appropriate predictor. It allows one MV component better interleaved with the other MV component. Furthermore, it requires no signaling bit for indicating selected MV predictor. The flexibility according to sequence characteristics helps to improve coding efficiency. The design concept is particularly useful for supporting various directions of camera panning, object action, and zooming motion.

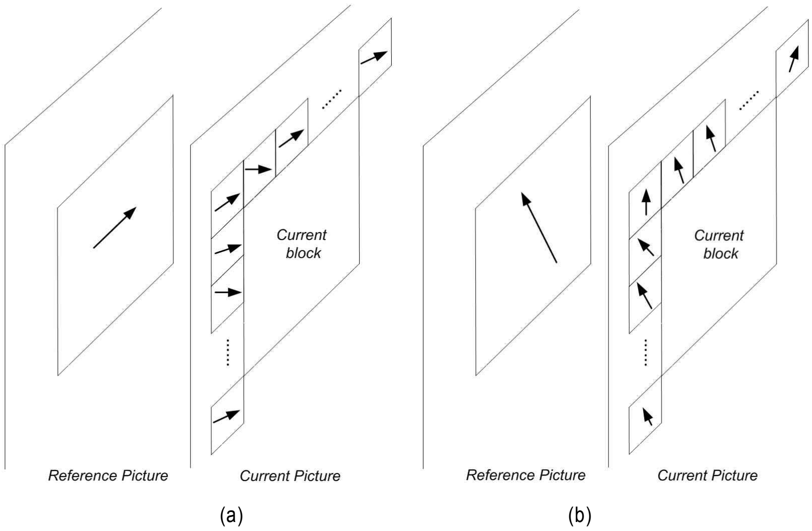

그림. 6. 수평 움직임(a) 또는 수직 움직임(b)을 나타내는 주변 블록들
Fig. 6. Neighboring blocks with (a) horizontal or (b) vertical motion displacement

1. Coding Procedure

As described in Fig. 4, there are two ways of selecting a predictor for each MV component. The first encoded component is predicted by the same component of the median PMV of the H.264/AVC.

$$PMV_{1st} = median(MV_{1st}^{A0}, MV_{1st}^{B0}, MV_{1st}^{C}) \quad (4)$$

Using Eq.(5), the predictor of the other MV component of a current block is decided based on SAD between the first component of motion vector of a current block in Eq.(4) and those neighboring blocks spatially and temporally located in Fig. 1.

$$PMV_{2nd} = f\left(\underset{MV_{1st}^{PropCS} \in CS}{\arg min} \left| MV_{1st}^{Curr} - MV_{1st}^{PropCS} \right|\right) \quad (5)$$

where PropCS denotes a candidate set for the first MV component of left blocks (A0, A1,⋯), upper blocks (B0, B1,⋯), diagonal blocks (C, D, E), and temporally co-located block (Co) in Fig. 1. The function $f(\bullet)$ identifies the second component of a motion vector corresponding to the selected first MV component by $\arg min(\bullet)$. The first MV component can be selected either horizontal or vertical component depending on the motion characteristics. Its decoding process is almost identical to the proposed encoding process as illustrated in Fig. 5. If the horizontal MV component is firstly decoded, Eq.(4) and Eq.(5) can be utilized as the above-mentioned method. Otherwise, Eq.(2) and Eq.(3) can be employed in the identical manner as explained in Section II-2.

2. Predictor Decision of MV Components

For making component-wise decision on motion vector predictor, motion vector disparity of a current block with respect to the spatio-temporally neighboring blocks is measured. The measurement is carried out using the recovered motion field so that the decoder can do the exactly same procedure. Thus the encoder does not need to code any overhead signaling bit. The motion vector disparity is computed as the difference between vertical and horizontal components of motion vectors in the identical blocks given by Eq. (6).

$$SCORE = \sum_{blk \in PropBlk} g(|MV_x^{blk}| - |MV_y^{blk}|)$$
$$where\ g(x) = \begin{cases} 1 & \text{if } x > 0 \\ 0 & \text{if } x = 0 \\ -1 & \text{if } x < 0 \end{cases} \quad (6)$$

where PropBlk denotes a subset of neighboring blocks in the candidate set indicated by PropCS in Eq.(5). In Eq.(5), non-existing motion vector of a neighboring block (for example, due to being intra block or out of the picture boundary) is set to (0,0). A non-negative value of $SCORE$ means that the object action or camera panning is more in the horizontal direction than in the vertical direction as illustrated in Fig. 6-(a). Consequently a range of vertical motion displacement is expected to be smaller than a range of horizontal one. Thus a possibility that variance of vertical motion displacement is decreased highly rises. Hence, it should be effective to firstly code vertical motion vector using median predictor. On the contrary, its negative value means that the activity is mainly in the vertical direction as shown in Fig. 6-(b), thus coding the horizontal one in advance using the median predictor can achieve more coding efficiency.

IV. Simulation and Results

1. Simulation Condition

In order to evaluate the performance of the proposed method, it is implemented in the KTA reference software version 2.0[5], and then applied to all possible macroblock modes (SKIP, P16x16, P16x8, P8x16, P8x8, and all sub-macroblocks). The experiment is simulated under the test condition recommended by the JCT-VC for evaluating

표 1. 실험 조건
Table 1. SIMULATION CONDITION

Spatial Resolution	416x240	832x480	1280x720	1920x1080	2560x1600
Sequence (Frame Rate in Hz)	Flowervase (30) RaceHorses (30) BlowingBubbles (50) BasketballPass (50) BQSquare (60)	Flowervase (30) RaceHorses (30) PartyScene (50) BasketballDrill (50) BQMall (60)	vidyo1 (60) vidyo3 (60) vidyo4 (60) Jets (60) ShuttleStart (60)	ParkScene (24) Kimono1 (24) Toy&Calendar (25) Cactus (50) BasketballDrive(50) BQTerrace (60)	PeopleOn Street(30) Traffic (30)
Number of Encoded Pictures	300 except for (ParkScene: 240) , (Kimono1: 240) , (Toy and Calendar: 250) , (PeopleOnStreet: 150)				
Motion Estimation	EPZS Motion Estimation with search range ± 64				
GOP Structure	IPPP				
Quantization Parameter	QPI : 22, 26, 30, 34, QPP: QPI + 1				
Number of Reference Pictures	3				
Profile	High Profile				
Other Coding Options	Weighted Prediction on, 8x8 transform on, RD on, CABAC on, 1/4-pel ME accuracy				
Performance Comparison	1. Anchor (MV coding in H.264/AVC scheme) 2. Interleaved MV coding[16] 3. Flexible interleaved MV Coding (Proposed Design)				
Implementation	KTA2.0 software based on JM11.0 software				

HEVC standardization as depicted in Table 1. The proposed flexible interleaved MV coding is compared against the MV coding in the H.264/AVC standard (as an anchor) and the interleaved MV coding scheme[16] in terms of Bjøntegaard Delta Bit Rate (BDBR)[17]. Since MV coding is lossless, no comparison is made in PSNR.

2. Analysis of Experimental Results

Table 2 summarizes experimental results of the proposed method and its comparison to existing methods. Its vertical-axis indicates percentage bit-saving. Table 2 shows that the existing interleaved MV coding[16] achieves high coding gain compared to the anchor because of its capability of appropriately matching vertical MV components of current and its neighboring blocks. It also shows that the proposed flexible interleaved MV coding gives better coding performance in all test sequences than the fixed interleaved method[16] because of properly exploiting the spatial and the temporal motion fields with no consumption of signaling bits. The results show that the proposed one and the fixed interleaved one[16] respectively achieve BDBR gain

표. 2. 부호화 성능 비교
Table. 2. Coding performance comparison against H.264/AVC

Resolution	Sequences	Anchor VS Interleaved MV Coding[16]		Anchor VS Proposed Design	
		BDPSNR[dB]	BDBR[%]	BDPSNR[dB]	BDBR[%]
416x240	Flowervase	0.17	-3.84	0.21	-4.82
	RaceHorses	-0.02	0.46	0.00	-0.08
	BlowingBubbles	-0.01	0.18	0.03	-0.63
	BasketballPass	0.01	-0.17	0.03	-0.77
	BQSquare	-0.01	0.25	0.01	-0.28
832x480	Flowervase	0.08	-2.14	0.12	-3.10
	RaceHorses	-0.01	0.26	0.00	-0.03
	PartyScene	0.00	0.06	0.02	-0.42
	BasketballDrill	0.02	-0.37	0.03	-0.72
	BQMall	0.02	-0.45	0.05	-1.19
1280x720	Vidyo1	0.09	-2.91	0.12	-3.74
	Vidyo3	0.00	-0.12	0.01	-0.16
	Vidyo4	0.04	-1.51	0.07	-2.32
	Jets	0.11	-5.76	0.17	-8.71
	ShuttleStart	0.04	-1.52	0.07	-2.52
1920x1080	Toys and Calendar	0.08	-3.49	0.11	-5.07
	ParkScene	0.05	-1.57	0.07	-1.97
	Kimono1	0.05	-1.46	0.06	-1.74
	Cactus	0.00	-0.15	0.02	-0.99
	BasketballDrive	0.01	-0.29	0.02	-0.68
	BQTerrace	0.01	-0.35	0.05	-2.50
2560x1600	PeopleOnStreet	0.00	-0.08	0.03	-0.71
	Traffic	0.05	-1.44	0.10	-2.72
Overall average		0.03	-1.15	0.06	-1.99

of 1.99% and 1.15% on average compared to the H.264/AVC standard. Especially, in the Jets sequence of 1280x720, the proposed one performs very well by attaining bit-saving of 8.71% against the anchor. For sequenceshaving mainly horizontal motions such as ParkScene, Kimono1, Vidyo1, and BQMall, the performance of the proposed and the interleaved ones hardly makes any difference even though both methods obtain quite good codinggain compared to anchor as represented in Fig. 7. However, for sequences with diverse directional motions such as Traffic, Jets, ShuttleStart, Flowervase, the proposed one overwhelmingly gains more coding efficiency compared to the other methods as illustrated in Fig. 8.

Fig. 9 shows number of selected coding modes. It shows

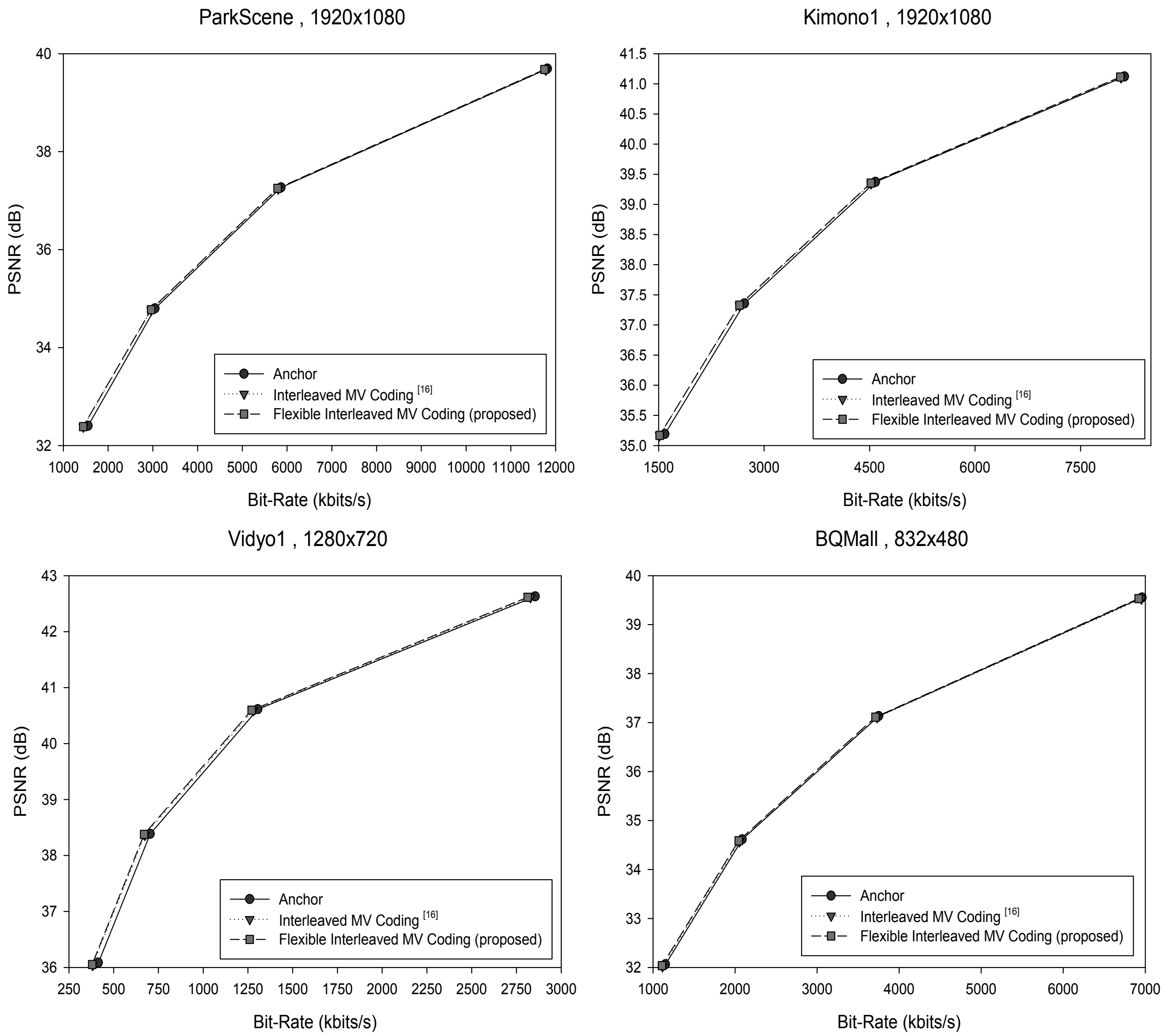

그림. 7. 율-왜곡 그래프 (수평방향의 움직임이 발생하는 영상)
Fig. 7. R-D curve for sequences having mainly horizontal motions

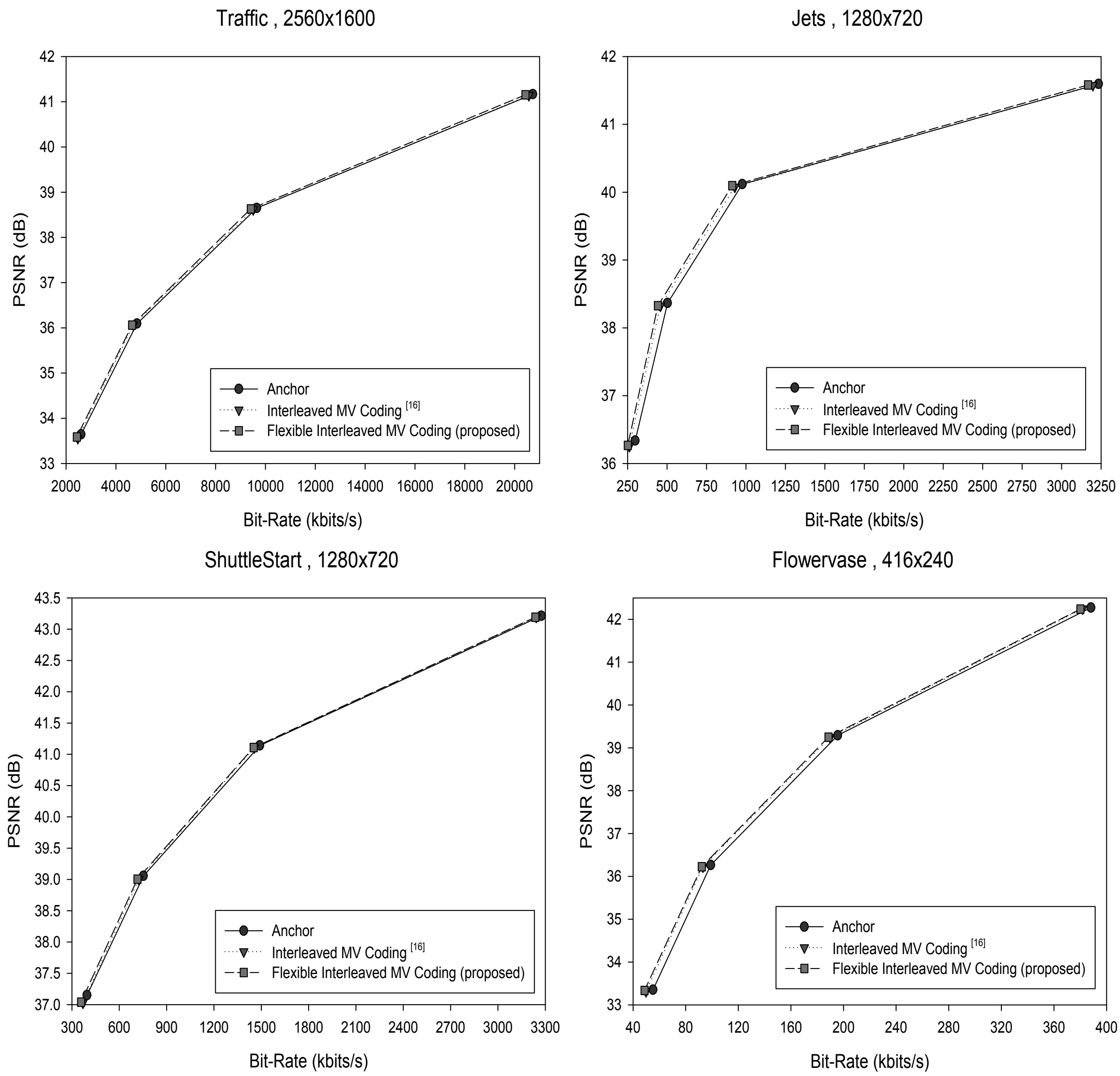

그림. 8. 율-왜곡 그래프 (다양한 움직임이 발생하는 영상)
Fig. 8. R-D curve for sequences having diverse directional motions

that the proposed method generates more SKIP modes than the anchor and the interleaved method[16] because the proposed one selects the proper PMV which is considerably close to a current MV so that it produces more zero DMV which is one condition of being a skip mode. This can explain why the proposed one attains highly effective compression. Moreover, its adaptive usage of component-wise predictor operated without any signaling bit also contributes to coding gain.

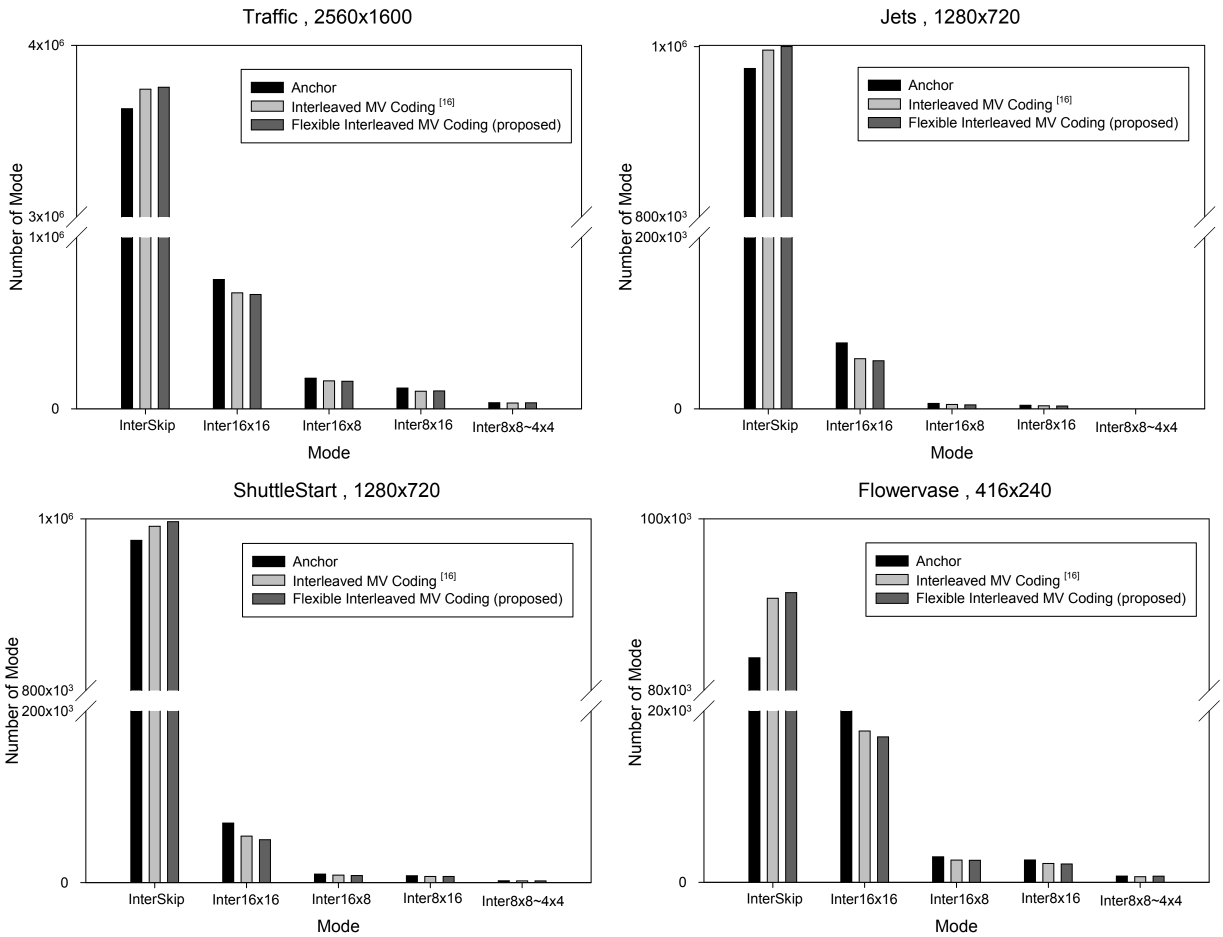

그림. 9. 모드 결정 과정에서 결정된 블록 모드의 개수 (다양한 움직임이 발생하는 영상)
Fig. 9. Coding modes used for each block in sequences having diverse directional motions

V. Conclusion

In this paper, a new MV coding with flexible usage of predictive motion vector components is proposed to improve the coding efficiency of motion vector. To reduce motion information data, this design allows each MV component to be flexibly interleaved with the other. Indeed, the proposed design promotes many zero DMVs and the approach with the flexibility according to sequence characteristics is effective to enhance coding gain. Compared to the anchor and the previous interleaving-based method[16], the proposed one provides improvement in coding efficiency by allowing flexibility in choosing a coding procedure.

참 고 문 헌

[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] T. Wiegand, G. J. Sullivan, G. Bjøntegaard, and A. Luthra, “Overview of the H.264/AVC Video Coding Standard,” IEEE Trans. on Circuits and Syst. for Video Tech., vol.13, no.7, pp.560-576, Jul. 2003.
[3] J. Ostermann, J. Bormans, P. List, D. Marpe, M. Narroschke, F. Pereira, T. Stockhammer, and T. Wedi, “Video coding with

H.264/AVC: tools, performance, and complexity," IEEE Circuits Syst. Mag., vol. 4, no. 1, pp. 7–28, 1Q, 2004.
[4] 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.
[5] Key Technology Area Reference Software, Version 2.0(KTA2.0). [Online]. Available: http://www.tnt.uni-hannover.de/~vatis/kta/.
[6] T. Wedi, "1/8-pel Motion Vector Resolution for H.26L," ITU-T VCEG, Proposal Q15-K-21, Aug. 2000.
[7] "Joint Call for Proposals on Video Compression Technology," ISO/IEC MPEG/ITU-T SG16 VCGE Doc. N11113/VCEG-AM91, Jan. 2010.
[8] L. A. Da Silva Cruz, and J. W. Woods, "Adaptive motion vector quantization for video coding," in Proc. IEEE ICIP, Oct. 2000, vol. 2, pp. 867–870.
[9] A. Deever and S. S. Hemami, "Dense motion field reduction for motion estimation," Signals, Syst. Comput., vol. 2, pp. 944–948, Nov. 1998.
[10] M. H. Chan, Y. B. Yu, and A. G. Constantinides, "Variable size block matching motion compensation with applications to video coding," Proc. Inst. Elect. Eng., vol. 137, pp. 205–212, Aug. 1990.
[11] J. Yeh, M. Vetterli, and M. Khansari, "Motion compensation of motion vectors," Proc. IEEE Inter. Conf. Image Processing (ICIP), Oct. 1995, vol. 1, pp. 574–577.
[12] M. C. Chen and A. N. Willson, "A spatial and temporal motion vector coding algorithm for low-bit-rate video coding," Proc. IEEE Inter. Conf. Image Processing (ICIP), Oct. 1997, vol. 2, pp. 791–794.
[13] S. D. Kim and J. B. Ra, "An efficient motion vector coding scheme based on minimum bit rate prediction," IEEE Trans. Image Process., vol. 8, no. 8, pp. 1117–1120, 1999.
[14] G. Laroche, J. Jung, and B. Pesquet-Popescu, "RD Optimized Coding for Motion Vector Predictor Selection," IEEE Trans. on Circuits and Syst. for Video Tech., vol.18, no.12, pp.1681-1691, Dec. 2008.
[15] J. Yang, K. Won, and B. Jeon, "Motion vector coding with selection of an optimal predictive motion vector" SPIE Opt. Eng., vol. 48, no.1, pp.010501, Jan. 2009.
[16] JCT-VC, "Description of video coding technology proposal by Fraunhofer HHI", Joint Collaborative Team on Video Coding (JCT-VC) of ITU-T SG16 WP3 and ISO/IEC JTC1/SC29/WG11, document JCTVC-A116, Dresden, DE, 15-23 April, 2010.
[17] G. Bjøntegaard, "Calculation of average PSNR differences between RD-curves," ITU-T SG16/Q6 Doc. VCEG-M13, April 2001.

저 자 소 개

원 광 현

- 2007년 : 광운대학교 컴퓨터공학부 학사 졸업
- 2009년 : 성균관대학교 휴대폰학과 석사 졸업
- 2009년 ~ 현재 : 성균관대학교 휴대폰학과 박사 과정
- 주관심분야 : 멀티미디어 영상압축

양 정 엽

- 2004년 : 성균관대학교 정보통신공학부 학사 졸업
- 2006년 : 성균관대학교 전자전기공학과 석사 졸업
- 2006년 ~ 현재 : 성균관대학교 전자전기공학과 박사 과정
- 주관심분야 : 멀티미디어 영상압축

저 자 소 개

박 대 윤

- 2010년 : 성균관대학교 전자전기공학과 학사 졸업
- 2011년 ~ 현재 : 성균관대학교 휴대폰학과 석사 과정
- 주관심분야 : 비디오 영상압축

전 병 우

- 1985년 : 서울대학교 전자공학과 학사 졸업
- 1987년 : 서울대학교 전자공학과 석사 졸업
- 1992년 : Purdue Univ, School of Elec. 공학박사 졸업
- 1993년 ~ 1997년 : 삼성전자 신호처리연구소 선임/수석연구원
- 1997년 ~ 현재 : 성균관대학교 정보통신공학부 교수
- 주관심분야 : 멀티미디어 영상압축, 영상인식, 신호처리