특집논문 (Special Paper)
방송공학회논문지 제24권 제1호, 2019년 1월 (JBE Vol. 24, No. 1, January 2019)
https://doi.org/10.5909/JBE.2019.24.1.41
ISSN 2287-9137 (Online) ISSN 1226-7953 (Print)

대체 병합 후보를 이용한 병합 후보 리스트 구성 기법

박 도 현[a)], 윤 용 욱[a)], 도 지 훈[a)], 김 재 곤[a)‡]

A Method of Merge Candidate List Construction using an Alternative Merge Candidate

Do-Hyeon Park[a)], Yong-Uk Yoon[a)], Ji-Hoon Do[a)] and Jae-Gon Kim[a)‡]

요 약

HEVC(High Efficiency Video Coding) 보다 뛰어난 압축 성능을 갖는 차세대 비디오 압축 표준으로 진행중인 VVC(Versatile Video Coding)에서는 보다 향상된 움직임 병합(Merge) 모드를 위한 기법들이 다루어지고 있다. 기존의 병합 모드에서는 다수의 주변 블록에서 움직임 정보가 존재하지 않으면 현재 블록의 움직임 정보로 사용할 병합 후보 리스트에 제로(zero) 움직임 벡터가 후보에 삽입될 수 있으며 이는 부호화 효율을 저하시킬 수 있다. 본 논문에서는 병합 후보 리스트에 제로 움직임 후보가 삽입되는 경우를 줄이기 위해 대체 움직임 정보를 생성하여 병합 후보 리스트에 삽입하는 효율적인 병합 모드 후보 리스트 구성 기법을 제시한다. 실험결과 제안 기법은 VTM1.0 대비 3% 복호화 복잡도 증가와 함께 0.2%의 부호화 성능 향상을 보였다.

Abstract

Recently, enhanced methods on the inter merging have been being investigated in Versatile Video Coding (VVC) standardization which will be a next generation video coding standard with capability beyond the High Efficiency Video Coding (HEVC). If there is not enough motion information available in the neighboring blocks in the merge mode, zero motion candidate is inserted into the merge candidate list, which could make the coding efficiency decreased. In this paper, we propose an efficient method of constructing the merge mode candidate list to reduce the case that the zero motion is used as a candidate by generating an alternative merge candidate. Experimental results show that the proposed method gives the average BD-rate gain of 0.2% with the decoding time increase of 3% in the comparison with VTM 1.0.

Keywords : VVC, JVET, VTM, Merge mode, Inter prediction

a) 한국항공대학교 항공전자정보공학부(Korea Aerospace University, School of Electronics and Information Engineering)
‡ Corresponding Author : 김재곤(Jae-Gon Kim)
E-mail: jgkim@kau.ac.kr
Tel: +82-2-300-0414
ORCID: https://orcid.org/0000-0003-3686-4786
※이 논문의 연구결과 중 일부는 "한국방송·미디어공학회 2018년 하계학술대회"에서 발표한 바 있음.
※이 논문은 과학기술정보통신부의 재원으로 정보통신산업진흥원의 지원을 받아 수행된 연구임(No. 2016-0-00572).
※This work was supported by IITP grant funded by Korea Government (MSIT) (No. 2016-0-00572).
※Parts of this work have been published in the 2018 Summer Conference of the Korean Institute of Broadcasting and Media Engineers.
・Manuscript received November 13, 2018; Revised January 4, 2019; Accepted January 4, 2019.

I. Introduction

More efficient video coding technologies are still required in the emerging immersive media services such as 4K/8K Ultra High-Definition (UHD) videos and 360-degree/Virtual Reality (VR) videos. Accordingly, Joint Video Exploration Team (JVET) was founded by ISO/IEC MPEG and ITU-T VCEG to explore the future video coding technologies beyond High Efficient Video Coding (HEVC). JVET conducted the search and verification new coding technologies and released a reference software codec, Joint Exploration Model (JEM) [3]. After the exploration stage about two years, Joint Video Experts Team (JVET) released the Call for Proposal (CfP) for future video coding standard. At the 10th JVET meeting in April 2018 [4], JVET started to develop a new video coding standard named as Versatile Video Coding (VVC) and released software codec, Versatile Test Model (VTM), based on the evaluation results of the response to the CfP.

Merge mode is a conventional inter-predicted coding mode which allows reuse of the motion information of neighboring blocks. In this way, the coding efficiency can be enhanced by reducing the signaling overhead of motion information of the current coding unit (CU). For the merge mode, firstly, the merge candidate list is constructed. Each merge candidate includes a motion vector, reference index, and a prediction direction. Then the best merge candidate minimizing the rate-distortion (RD) cost is selected for the inter prediction of the current block. A merge flag and a merge index are signaled to indicate that merge mode is selected as the best coding mode of the current CU and which candidate is selected in the merge mode, respectively.

In VVC, some methods have been proposed to enhance the merge mode. In [5], predefined motion pairs in the merge candidate list are averaged and it is inserted to the merge candidate list. In [6], a history-based method that use the motion information of previously coded blocks in the merge is proposed.

In this paper, a method of merge candidate construction which uses an alternative merge candidate (AMC) as a merge candidate is proposed to enhance inter merging prediction. In the proposed merge candidate list construction, an AMC is obtained as a weighted combination of neighboring motion information available, and the computed AMC is assigned to the neighboring block in which motion information is not available. In this way, the coding efficiency of merge mode could be enhanced by reducing the case that uses zero motion as a candidate when sufficient motion information is not available in the neighboring blocks.

The rest of the paper is organized as follows. Section 2 presents the merge mode coding and the procedure of merge candidate construction used in VTM1.0. In Section 3, the details of proposed algorithm are described. Section 4 presents the experimental results with analysis on the results. Finally, Section 5 concludes the paper.

II. Merge mode coding

Each inter coded block estimates, derives, and stores motion information. In the merge mode, motion information in the current CU is derived from spatial/temporal neighboring positions which allows avoidance of signaling of full motion information. In VTM1.0, the same as HEVC [7], the merge candidate list consists of 5 motion candidates, each of which consists of prediction direction, reference index, and motion vector. Motion information in the merge candidate list can be an uni-predictive or a bi-predictive one.

In VTM1.0 [8], the merge candidates are derived among neighboring blocks located at the positions shown in Fig. 1: five spatial neighboring candidate positions and two temporal neighboring candidate positions. The co-located block is in a previously decoded picture that is the closest

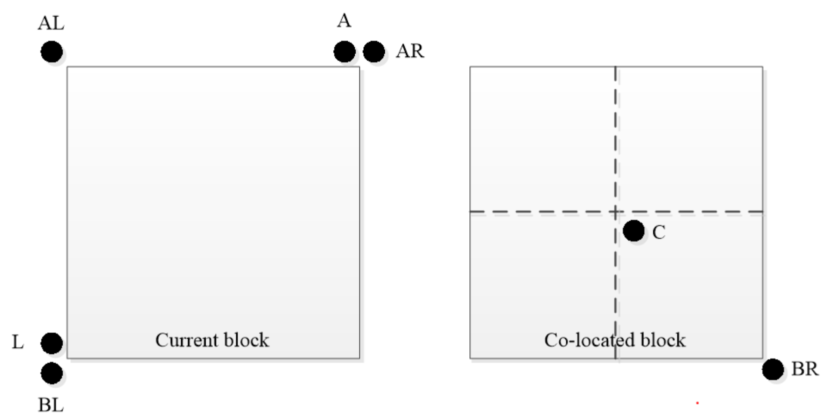

그림 1. 현재 부호화 블록의 주변 부호화 블록 위치
Fig. 1. Neighboring blocks' positions of the current block

to the current picture temporally. The process of the merge candidate list construction consists of the following four steps. First, spatial neighboring candidates are registered into the merge candidate list up to 4 according to the following order: (L, A, AR, BL, AL). Second, temporal neighboring candidates can be registered up to 1 in the order: (BR, C). Third, combined merge candidates are inserted into the merge candidate list, which are only bi-predictive motion candidates and results of pre-defined pair-wise combination of uni-predictive motion candidates in the merge list. Table 1 shows the pre-defined pair position in the merge list. Finally, the zero-motion candidate is filled in the remaining array in the merge list. Additionally, when the searched motion candidate is not available in the defined positions or redundant in the process of the merge candidate list construction, it is pruned. Also, if all elements of the merge candidate list are filled, the process is completed.

표 1. 혼합 운직임 벡터 생성을 위한 움직임 후보 인덱스 쌍
Table 1. Pairs for combined merge candidates

Order	0	1	2	3	4	5	6	7	8	9	10	11
Candidiate index (L0)	0	1	0	2	1	2	0	6	1	3	2	3
Candidiate index (L1)	1	0	2	0	2	1	3	0	3	1	3	2

Ⅲ. Proposed merge mode

In general, the motion information of current CU is inherited from blocks on neighboring positions in the merge mode. In VTM1.0, if the number of valid neighboring blocks having motion information is not sufficient, the merge candidate list is filled with the zero-motion candidate. For example, if only two neighboring blocks are valid, the zero-motion candidate is included in the merge candidate list unconditionally. A valid motion candidate means the motion information is available and not duplicated during the construction of the merge candidate list. However, it is likely that the zero-motion candidate in the merge candidate list does not represent the motion of the current CU well, which causes the decrease of the coding performance.

The proposed AMC is generated by combining the left and above motion candidates of adjacent neighboring positions so that the generated motion candidate represents the current CU well, and the details are presented in sub section 3.1. Then, the details on inserting the AMC into the merge candidate list are described in section 3.2.

1. Alternative merge candidate

The AMC is calculated by weighted averaging of two motion candidates, each of which is selected from the left side and the above side of the current block, respectively. Therefore, there must be at least one motion candidate on each side to generate the AMC. That is, one motion candidate of the block on the L or BL position and the other motion candidate of the block on the A or AR position should be available. If there are two valid motion candidates on one side, one motion information having a reference picture temporally closer to the current picture is selected.

Then selected two motion candidates from both sides are averaged based on block shape of the current CU. Since a block partitioning called Multi Type Tree (MTT) in VTM1.0 makes various rectangular shape of the CU possible, a correlation between the motion information of the current CU and the motion information on one side may be different depending on the width and height of current CU. Table 2 shows the correlations between the motion vector of the current CU and the motion vector on one side according to block conditions. The motion correlations are calculated by

$$c_x = \frac{\Sigma(CMV_x - \overline{CMV_x})(NMV_x - \overline{NMV_x})}{\sqrt{\Sigma(CMV_x - \overline{CMV_x})^2 \times \Sigma(NMV_x - \overline{NMV_x})^2}}$$
$$c_y = \frac{\Sigma(CMV_y - \overline{CMV_y})(NMV_y - \overline{NMV_y})}{\sqrt{\Sigma(CMV_y - \overline{CMV_y})^2 \times \Sigma(NMV_y - \overline{NMV_y})^2}} \quad (1)$$
$$c = 0.5 \times c_x + 0.5 \times c_y$$

Where the CMV is the motion vector of current CU; the NMV is the motion vector of neighboring CU. The c is the correlation coefficient between the CMV and the NMV. The closer the value of the coefficient is to 1, the stronger the correlation. From the table 2, it is shown that the correlation of motion information of the longer side is higher than the opposite side. Thus, based on this observation, the motion vector of the AMC is determined by

$$MV_{AMC} = \frac{MV_L \times Width + MV_A \times Height}{Width + Height} \quad (2)$$

If the reference pictures of two motion vectors to be averaged are different, scaling for one motion vector is used to make two motion vectors have the same reference picture that is closer to the current picture before the motion combination. Then, the reference index of the AMC is inherited from the unscaled motion candidate among the two motion candidates. The motion vector of the AMC is calculated separately for each reference list.

2. Proposed merge candidate list construction

The concept of proposed merge candidate list construction is to assign the generated candidate motion information to the block of AL when there is no available motion information in there. Since the AMC generation needs a process of searching four spatial neighboring positions (L, A, AR, and BL), the AMC can be used as an alternative motion information when the block on the AL position is invalid, as shown in Fig. 2. The process of the merge candidate list construction except the proposed substitution with AMC is basically the same as that of VTM1.0.

표 2. 현재 블록과 주변 블록의 움직임 상관 계수
Table 2. Motion correlations between the neighboring and the current block

Sequence name	RaceHorses		BlowingBubbles		BQSquare	
Block condition	W > H	W< H	W > H	W< H	W > H	W< H
Correlation coefficient (L and Current)	0.84	0.66	0.87	0.38	0.67	0.31
Correlation coefficient (A and Current)	0.65	0.76	0.44	0.72	0.29	0.45

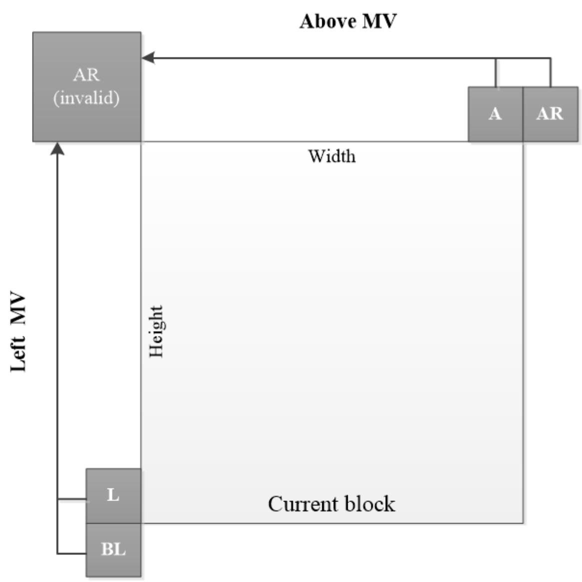

그림 2. 병합 후보 리스트에 대한 대체 병합 움직임 후보 삽입
Fig. 2. Insertion of the AMC into merge candidate list

Ⅳ. Experimental results

To evaluate the overall performance of the proposed alternative merge candidate, the proposed method is implemented in the VVC reference software (VTM1.0). The simulations are performed following the JVET Common Test Conditions (CTC) [9]. The quantization parameters (QPs) of {22, 27, 32, 37} are used for all sequences, and the random access (RA) encoding configuration is used. The Bjøntegaard Delta (BD)-rate measurement with piece-wise cubic interpolation is employed for the objective coding performance comparison [10]. Table 3 shows the average BD-rate for each class and Table 4 shows the details on the results for each sequence.

표 3. 클래스 별 실험 결과(Anchor: VTM1.0)
Table 3. Experimental result for each class (Anchor: VTM1.0)

Class	BD-rate (piecewise cubic)		
	Y	U	V
Class A1	-0.37%	-0.54%	-0.30%
Class A2	-0.22%	-0.07%	-0.10%
Class B	-0.19%	-0.06%	-0.13%
Class C	-0.05%	-0.10%	0.01%
Overall	-0.20%	-0.17%	-0.12%
Class D	-0.16%	-0.08%	-0.21%
Class F (optional)	-0.13%	-0.26%	-0.17%
Enc. time (s)	100%		
Dec. time(s)	103%		

표 4. 시퀀스 별 실험 결과(Anchor: VTM1.0)
Table 4. Experimental result for each sequence (Anchor: VTM1.0)

Class	Sequences	BD-rate (piecewise cubic)		
		Y	U	V
Class A1 (4K)	Tango2	-0.45%	-0.40%	-0.36%
	FoodMarket4	-0.65%	-1.06%	-0.41%
	Campfire	-0.03%	-0.16%	-0.12%
Class A2 (4K)	CatRobot1	-0.27%	-0.18%	0.02%
	DaylightRoad2	-0.23%	0.14%	-0.06%
	ParkRunning3	-0.17%	-0.18%	-0.25%
Class B (1080p)	MarketPlace	-0.38%	-0.04%	-0.20%
	RitualDance	-0.21%	-0.29%	-0.18%
	Cactus	-0.18%	-0.06%	-0.37%
	BasketballDrive	-0.15%	0.51%	0.31%
	BQTerrace	-0.02%	-0.44%	-0.22%
Class C (WVGA)	BasketballDrill	0.00%	0.13%	0.71%
	BQMall	0.00%	-0.18%	-0.56%
	PartyScene	-0.02%	-0.40%	0.01%
	RaceHorses	-0.17%	0.06%	-0.12%
Class D (WQVGA)	BasketballPass	-0.25%	-1.49%	0.35%
	BQSquare	-0.21%	0.05%	0.12%
	BlowingBubbles	-0.16%	-0.40%	-0.29%
	RaceHorses	-0.03%	1.52%	-1.00%
ClassF	BasketballDrillText	0.12%	-0.05%	0.05%
	ChinaSpeed	0.04%	0.09%	-0.10%
	SlideEditing	-0.02%	0.10%	0.21%
	SlideShow	-0.65%	-1.19%	-0.85%

As a result, from Table 3, it is noted that the proposed method achieves on average 0.20%, 0.17% and 0.12% BD-rate reductions for the Y, U and V components, respectively. The encoding time of proposed method has not changed but, the decoding time is increased by 3%.

Ⅴ. Conclusions

This paper proposed merge candidate construction with an alternative merge candidate to enhance the coding performance. By inserting the AMC to the merge candidate list instead of zero-motion candidate, the current block can be represented better in terms of motion information. Thus, the experimental result shows that proposed method can obtain a remarkable coding gain increase on average 0.2% BD-rate gain for luma component over VTM1.0. However, the decoding time is increased by 3% due to the additional complexity of motion combination process. Thus, a simplified method for motion combination will be studied in further work. In VTM3.0, the technique called the pairwise average candidates similar to the proposed method was adopted [11]. The pairwise method improves the coding performance by combining various motion vector pairs. But, the proposed AMC obtains the coding gain from the precisely combination of one motion vector pair. Therefore, further study on the harmonization of the pairwise method and the proposed AMC is also needed.

참 고 문 헌 (References)

[1] Rec, ITU-T H.264, ISO/IEC 14496-10 AVC, May 2003.
[2] High Efficiency Video Coding, Version 1, Rec. ITU-T H.265, ISO/IEC 23008-2, Jan. 2013.
[3] J. Chen, E. Alshina, G. Sullivan, J. Ohm and J. Boyce, "Algorithm Description of Joint Exploration Test Model 7," Joint Video Exploration Team of ITU-T SG 16 WP 3 and ISO/IEC JTC 1/SC 29/WG 11, JVET-G1001, Jul. 2017.
[4] A. Segall, V. Baroncini, J. Boyce, J. Chen and T. Suzuki, "Joint Call for Proposals on Video Compression with Capability beyond HEVC," Joint Video Exploration Team of ITU-T SG 16 WP 3 and ISO/IEC JTC 1/SC 29/WG 11, JVET-H1002, Oct. 2017.
[5] C. Hsu, C. Chen, T. Chuang, H. Huang and S. Hsiang, "Description of SDR Video Coding Technology Proposal by MediaTek," Joint Video Exploration Team of ITU-T SG 16 WP 3 and ISO/IEC JTC 1/SC 29/WG 11, JVET-J0018, Apr. 2018.
[6] L. Zhang, K. Zhang, H. Liu, Y. Wang, P. Zhao and D. Hong, "CE4-realted: History-based Motion Vector Prediction," Joint Video Exploration Team of ITU-T SG 16 WP 3 and ISO/IEC JTC 1/SC 29/WG 11, JVET-K0104, Jul. 2018.
[7] K. Kim, S. Kim, G. Park, H. Kim, S. Lim and J. Lee, "CU-based Merge Candidate List Construction Method for HEVC," Journal of Broadcast Engineering, Vol. 17, No. 2, Mar. 2013.
[8] [Online]. Available at: https://jvet.hhi.fraunhofer.de/svn/svn_VVC Software_VTM/tags/
[9] K. Suehring and X. Li, "JVET common test conditions and software reference configurations," Joint Video Exploration Team of ITU-T SG 16 WP 3 and ISO/IEC JTC 1/SC 29/WG 11, JVET-G1010, Aug. 2017.
[10] G. Bjøntegaard, "Calculation of average PSNR differences between RD curves," ITU-T SG.16 Q.6, VCEG-M33, Apr. 2001.
[11] Y. Hsiao, T. Chuang, C. Chen, C. Hsu, Y. Huang and S. Lei, "CE4.4.12: Pairwise average candidates," Joint Video Experts Team of ITU-T SG 16 WP 3 and ISO/IEC JTC 1/SC 29/WG 11, JVET-L0090, Oct. 2018.

저 자 소 개

박 도 현

- 2016년 2월 : 국립한밭대학교 멀티미디어공학과 학사
- 2018년 2월 : 한국항공대학교 항공전자정보공학과 석사
- 2018년 3월 ~ 현재 : 한국항공대학교 항공전자정보공학과 박사과정
- ORCID : https://orcid.org/0000-0002-5873-0132
- 주관심분야 : 비디오 부호화, 영상처리, 360 VR 비디오

저 자 소 개

윤 용 욱

- 2017년 2월 : 한국항공대학교 항공전자정보공학과 학사
- 2017년 3월 ~ 현재 : 한국항공대학교 항공전자정보공학과 석사과정
- ORCID : https://orcid.org/0000-0002-5105-5437
- 주관심분야 : 비디오 부호화, 영상처리, 360 VR 비디오

도 지 훈

- 2018년 2월 : 한국항공대학교 항공전자정보공학과 학사
- 2018년 3월 ~ 현재 : 한국항공대학교 항공전자정보공학과 석사과정
- ORCID : https://orcid.org/0000-0002-8254-2481
- 주관심분야 : 비디오 부호화, 영상처리, 360 VR 비디오

김 재 곤

- 1990년 2월 : 경북대학교 전자공학과 학사
- 1992년 2월 : KAIST 전기 및 전자공학과 석사
- 2005년 2월 : KAIST 전기 및 전자공학과 박사
- 1992년 3월 ~ 2007년 2월 : 한국전자통신연구원(ETRI) 선임연구원/팀장
- 2001년 9월 ~ 2002년 11월 : 뉴욕 콜롬비아대학교 연구원
- 2007년 9월 ~ 현재 : 한국항공대학교 항공전자정보공학부 교수
- 2015년 12월 ~ 2016년 1월 : UCSD 방문교수
- ORCID : https://orcid.org/0000-0003-3686-4786
- 주관심분야 : 비디오 신호처리, 비디오 부호화 표준, UHD/Immersive Media, 영상통신