An Improvement to Merge Mode in ECM With Template Matching

Ru-Ling Liao, Yan Ye, Jie Chen and Xinwei Li

Alibaba Group, China

{ruling.lrl, yan.ye, jiechen.cj, sid.lxw}@alibaba-inc.com

2023 Data Compression Conference (DCC) | 979-8-3503-4795-1/23/$31.00 ©2023 IEEE | DOI: 10.1109/DCC55655.2023.00011

***Abstract**:* In the development of video coding standard, decoder-side motion derivation technology has been proven to provide promising coding efficiency. With this type of technology, the motion information is derived at the decoder instead of being signaled in the bitstream by the encoder, and thus, the number of bits to be sent are reduced. A typical decoder-side derivation technology is template matching, which refines the motion by finding the closest match between neighboring reconstructed samples and corresponding reference samples in the reference pictures. In this paper, the template matching method is extended to temporal motion vector predictor, bi-prediction with CU-level weight and geometric partition modes in order to fully utilize its benefit. Specifically, template matching is used to determine the prediction direction and reference picture of temporal motion vector predictors, and to decide the bi-predicted weight of bi-prediction merge blocks. In addition, the motion of two geometric partitions are individually refined by the template matching mechanism. Simulation results show that on top of Enhanced Compression Model (ECM), which is the software platform for exploring activities beyond versatile video coding established by the joint video experts team (JVET), the three proposed methods achieve 0.20% luma BD-rate savings in random access configuration and 0.33% luma BD-rate saving in low delay B configuration with negligible encoding and decoding runtime impact. It is worth noting that all three proposed methods have been adopted to the ECM software platform.

1. Introduction

In July 2020, the international video coding standard – Versatile Video Coding (VVC) was published. VVC is developed by the Joint Video Experts Team (JVET) of the ITU-T Video Coding Experts Group (ITU-T VCEG) and the ISO/IEC Moving Picture Experts Group (ISO/IEC MPEG), and is aimed at significantly exceeding the compression efficiency of its predecessor, the High Efficiency Video Coding (HEVC/H.265) standard [1]. Similar to HEVC, VVC [2][3] is also based on the hybrid video coding system which is commonly used in modern video compression standards, and includes more coding tools to improve the coding efficiency. In VVC, a coding tree unit (CTU) can be as large as 128×128, and may be partitioned into coding units (CU) with binary tree, ternary tree and quaternary tree partitioning structure [4]. The CU is then predicted using either intra or inter prediction. If intra prediction [5] is used, the CU is predicted using spatial neighboring samples with 67 intra prediction modes (DC mode, planar mode, or 65 angular modes), cross-component linear model prediction, multiple reference line intra prediction, intra sub-partitions or matrix weighted intra prediction modes. If inter prediction [6][7] is used, the CU is predicted using samples of reference pictures with merge mode, merge mode with motion

979-8-3503-4795-1/23/$31.00 ©2023 IEEE
DOI 10.1109/DCC55655.2023.00011

vector difference (MMVD), geometric partition mode (GPM), combined inter and intra prediction (CIIP), bi-prediction with CU-level weight (BCW), adaptive motion vector resolution, affine motion compensation prediction, subblock-based temporal motion vector prediction, bi-directional optical flow, decoder-side motion vector refinement or symmetric motion vector difference coding modes. After either intra or inter prediction, the residues are transformed and quantized. In VVC, multiple transform cores including DCT-II, DCT-VIII and DST-VII are supported. Besides, larger block-size transform with zero-out for high frequency components and subblock transform [8] are used. For quantization [9], dependent quantization and joint coding of chroma residuals are introduced in VVC. Finally, the prediction samples and the residues are added together to generate reconstructed samples, which are filtered using deblocking filter, sample adaptive offset, adaptive loop filter and/or luma mapping with chroma scaling [10].

After VVC was finalized, in order to further improve compression efficiency, the JVET started exploring coding technologies beyond VVC capability and established the Enhanced Compression Model (ECM) software platform to support this exploration activity. In ECM [11], most coding tools in VVC are further improved and more decoder side prediction technologies are added, such as decoder side intra mode derivation, template-based intra mode derivation, convolutional cross-component intra prediction model, local illumination compensation, template matching (TM), etc. It is reported that as compared to VVC test model (VTM) [12], which is the reference software of VVC, the latest version of ECM achieves 18.50% and 15.59% luma BD-rate reduction under random access and low-delay B configurations, respectively [13][14]. In this paper, to fully utilize the benefits of decoder-side derivation coding technology, we propose to combine GPM, BCW and temporal motion vector prediction (TMVP) merge modes with the template matching method. Specifically, the motion vector information of GPM mode, the bi-predictive weight of BCW mode, and the prediction direction and reference pictures of TMVP are derived at decoder side using template matching, instead of being signaled in the bitstream. With these proposed methods, experimental results show that 0.20% luma BD-rate reduction in random access configuration and 0.33% luma BD-rate reduction in low delay B configuration is obtained on top of the ECM software.

The remainder of this paper is organized as follows. Section 2 briefly reviews the GPM, BCW and TMVP in VVC merge mode, and template matching mode in ECM. We then introduce the proposed improvements to merge mode with template matching in Section 3. The experimental results are shown in Section 4. Section 5 concludes this paper.

2. Overview of merge modes

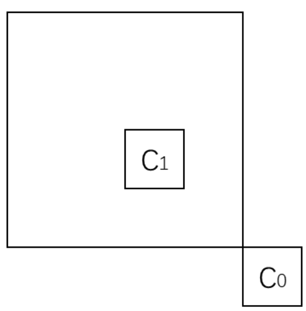
In VVC, a variety of merge modes are introduced to improve the prediction accuracy of inter blocks. In addition to the regular merge mode which is similar to the one used in HEVC, affine motion compensation prediction, MMVD, GPM and CIIP are newly supported. Moreover, the regular merge mode is modified to inherit bi-predicted weight of BCW. In this section, the design of TMVP, BCW and GPM in VVC, and TM in ECM are briefly reviewed.

2.1. Temporal motion vector predictor

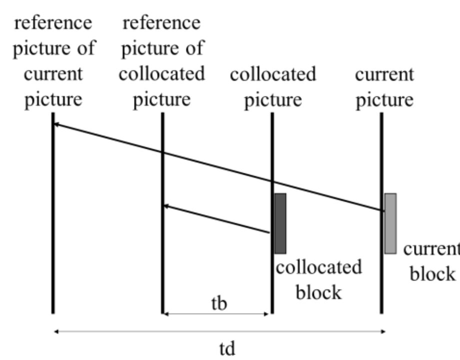
In VVC, temporal motion vector predictor is used as a motion information candidate in merge mode and adaptive motion vector prediction (AMVP) mode. To derive TMVP in merge mode, a collocated picture is selected by the encoder and signaled in the slice header, for example, the collocated picture could be the closest reference picture in temporal distance to the current picture. Then, a collocated block from the collocated picture is firstly obtained from bottom-right (C_0) position or center position (C_1), as illustrated in Figure 1(a). Based on the motion vectors of the collocated block, TMVP is derived based on the list 0 (L0) and/or list 1 (L1) motion vectors of the collocated block using motion vector scaling. It is worth noting that the TVMP construction process in VVC always results in the TMVP representing bi-predicted motion [6]. The motion vector scaling is the same as that in HEVC, that is, the motion vector is scaled according to picture order count (POC) distances between different pictures in Figure 1(b) using the following equation:

$$TMVP = \frac{td}{tb} \times MV\ of\ the\ collocated\ block$$

where *tb* represents the POC difference between the reference picture of the current picture and the current picture, and *td* represents the POC difference between the reference picture of the collocated picture and the collocated picture. Effectively, the motion vector scaling process derives the TMVP from the motion vectors of the collocated block based on corresponding temporal distances.



(a) Collocated block positions.



(b) POC difference in motion vector scaling.

Figure 1. Illustration of collocated blocks and motion vector scaling in TMVP.

2.2. Bi-prediction with CU-level weight

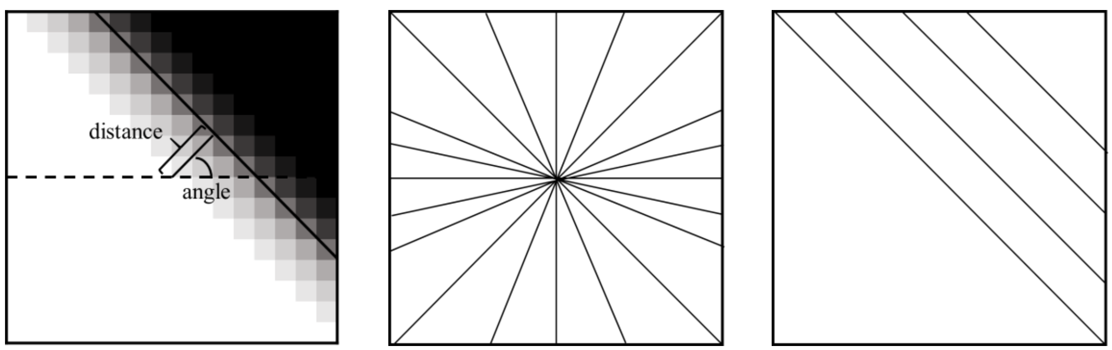
Bi-prediction has been used in video coding standards since H.264/AVC, and is effective for improving inter prediction accuracy. In VVC, a new coding mode called bi-prediction with CU-level weight or BCW is introduced to allow weighted averaging in addition to simple averaging of the two prediction signals used in HEVC. In BCW mode [15][16], the bi-prediction weight is determined at CU-level. Five weights {-2/8, 3/8, 4/8, 5/8, 10/8} and three weights {3/8, 4/8, 5/8} are allowed for low delay pictures and non-low delay pictures, respectively. Herein, a low delay picture is defined as an inter-coded picture predicted from only reference pictures in the past (i.e. the reference pictures precede the current picture in display order); and a non-low delay picture is an inter-coded picture without such restriction. Specifically, when BCW is enabled, the bi-predicted signal is generated as follows:

$$P_{bi-pred} = (8 \times (1 - w) \times P_0 + 8 \times w \times P_1 + 4) \gg 3$$

where P_0 and P_1 are the prediction signals from list 0 and list 1 reference picture, respectively. For non-merge coded blocks, the BCW weights are determined at encoder and are explicitly signaled in the bitstream, whereas for merge coded blocks, the BCW weights are inherited from the corresponding neighboring block based on the merge index.

2.3. Geometric partition mode

In GPM [18], a CU is firstly split into two partitions along a straight line (Figure 2(a)) that is defined by an angle and a distance from the center of the CU and each partition is predicted using its own motion information. The motion information is derived from a uni-predicted candidate list which is derived from spatial neighboring blocks, temporal motion vector predictor from collocated blocks, history-based MVP table, pairwise average MVP and zero motion. After predicting each partition with its own motion information, a blending process is applied to the samples along the partition edge. The blending weight for each sample is determined according to mathematical equations which are described in detail in [17]. In VVC, 64 partitioning modes are supported, which includes 20 partitioning angles (Figure 2(b)) with 4 partitioning distance (Figure 2(c)).



(a) GPM (b) 20 partitioning angles (c) 4 partitioning distance

Figure 2. Illustration of geometric partition mode.

2.4. Template matching

Template matching is a decoder-side motion vector derivation technology, which refines the motion vector by finding the minimum template matching cost of a template and its corresponding reference samples. As shown in Figure 3, a template is constructed using the left and the above neighboring reconstructed samples of a current block. Within a pre-defined search area, template matching costs are calculated and are measured by the sum of absolute difference between the template and its corresponding reference samples. In ECM, template matching is applied to both merge and AMVP modes to refine the motion vector of merge candidates and motion vector predictors. Additionally, template matching is also applied to inter modes to reorder candidates such as adaptive reordering of merge candidates with template matching (ARMC-TM) [19], or to decide the best set of modes such as template matching based OBMC [20] which determines the weights for blending in OBMC according to template matching cost.

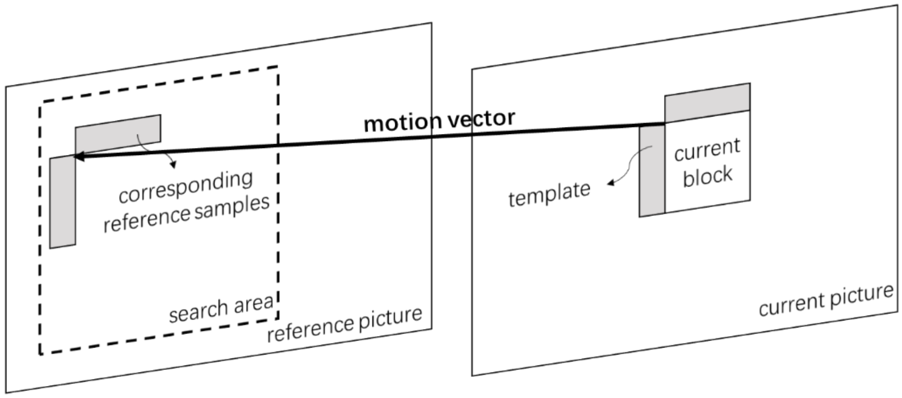


Figure 3. Illustration of template matching technology.

3. Proposed improvements to merge mode with template matching

In this paper, we propose to apply template matching to TMVP, BCW and GPM modes in order to fully utilize the benefits of decoder side motion information derivation technology. The proposed algorithm is described in detail in this section.

3.1. Improvement to TMVP

In VVC, the TMVP is always derived as bi-predicted motion no matter whether the collocated block is uni-predicted or bi-predicted motion. It is observed that scaling an uni-predicted motion to be a bi-predicted motion is not always suitable since the motion vector scaling may not be accurate. Therefore, we propose to decide the prediction direction of TMVP using the template matching technology. More specifically, template matching costs are calculated for L0-predicted, L1-predicted and bi-predicted TMVP, and the one with minimum template matching cost is used as TMVP for merge mode.

Moreover, we observe that the reference picture index of the current picture is always set to 0 in VVC for the TMVP. This is not optimal. In the proposed modification, the scaling factor of each reference picture in the reference picture list of the current picture is calculated, and the one with scaling factor closest to 1 is used as the reference picture of the TMVP such that the impact due to temporal scaling can be reduced. As shown in Figure 4, POC 2 is selected to be the reference picture of TMVP instead of being fixed to POC 0.

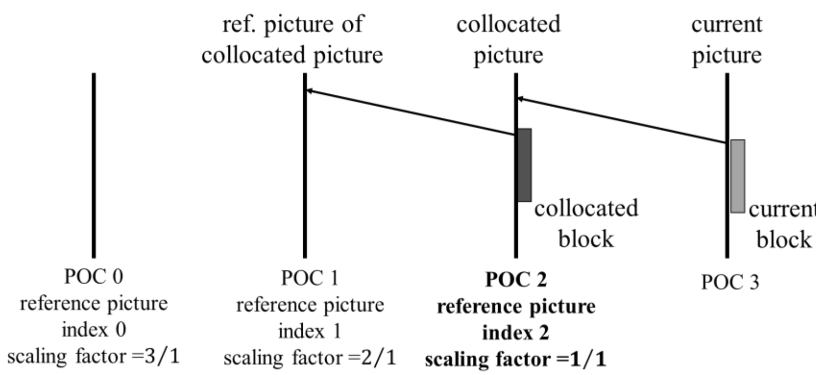


Figure 4. Example of reference picture selection of the proposed method.

3.2. Improvement to BCW

As mentioned in section 2.2, for merge coded blocks, the bi-predicted weight is directly inherited from the neighboring block specified by the merge index. However, the inherited bi-predicted weight is not always optimal for all merge blocks. In this paper, we propose to adjust bi-predicted weights according to template matching cost. Firstly, the bi-predicted weights for merge coded blocks are extended from {-2, 3, 4, 5, 10}/8 to {-4, -3, -2, -1, 1, 2, 3, 4, 5, 6, 7, 9, 10, 11, 12}/8. For non-merge coded blocks, in order to control signaling bits overhead, the weights remain the same as those allowed in VVC. Secondly, for a merge coded block, its bi-predicted weight is derived based on template matching cost, and the one with minimum template matching cost is selected to predict the merge coded block. When calculating template matching cost for bi-predicted weights, the following rules are applied:

- Since the inherited bi-predicted weight is likely to have higher accuracy than others, only the two neighboring bi-predicted weights (i.e. ±1) and the inherited bi-predicted weight are considered. For example, the inherited bi-predicted weight is 4/8, only three weights {3, 4, 5}/8 are involved in template matching cost calculation.
- The template matching cost of the inherited bi-predicted weight is reduced by 3/32 to slightly bias towards the inherited weight.
- Similarly, the template matching cost of the equal weight (i.e. the bi-predicted weight is equal to 4/8) is also reduced by 3/32 to slightly bias toward the case of equal weight, because such CU's can benefit from bi-directional optical flow which is only applied to blocks predicted with equal weight.

3.3. Improvement to GPM

To improve the prediction accuracy of GPM, we propose to refine the GPM motion information using template matching technology. When a block is coded in GPM, a CU-level flag is signaled to indicate whether the motions are refined by template matching. When it is chosen, two motions, one each for the two geometric partitions are firstly selected from the uni-predicted candidate list which is derived in the same way as that in VVC standard. Then, a template is constructed using left, above or left and above neighboring samples based on GPM partition angle, as shown in Figure 5. The basic principle for the template selection is that, if a partition has only above neighboring samples, then only above template is used, if a partition has only left neighboring samples, then only left template is used, and if a partition has both left and above neighboring samples, then both left and above templates are used. The motion is thus refined by minimizing the difference between the current template and the template in the reference picture with the following search pattern in order: full-pel diamond search, full-pel cross search, half-pel cross search, quarter-pel cross search and 1/8-pel cross search. The refined motion is used to predict each part of the geometric partition, and the blending is applied to the partition edge in the same way as that in VVC standard.

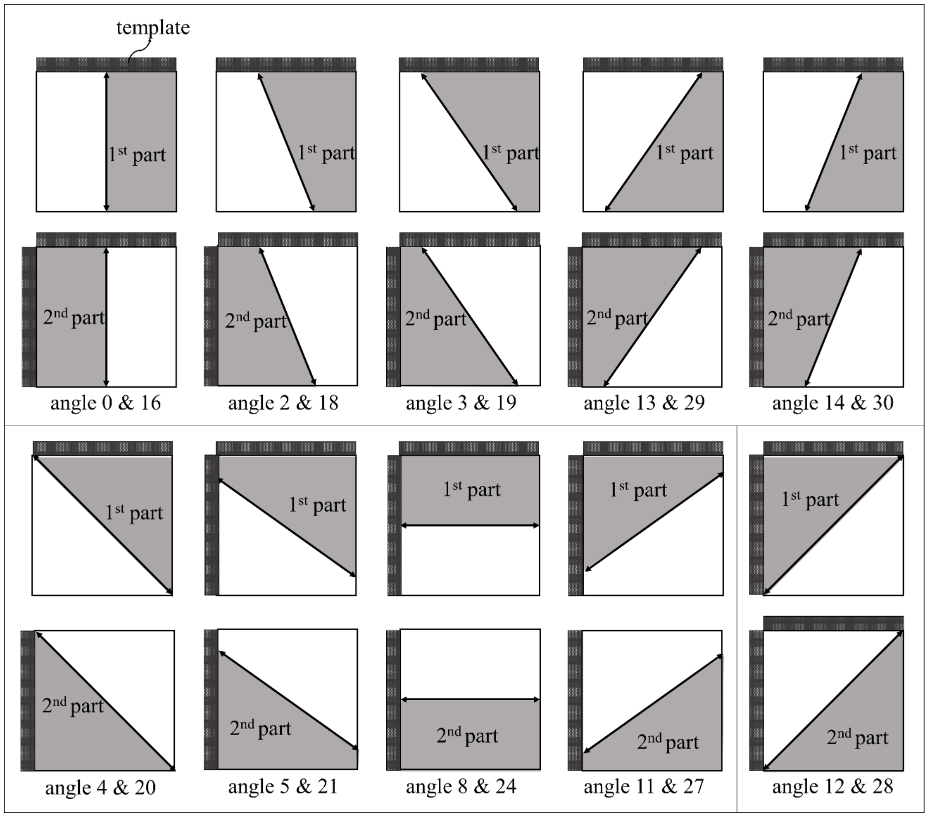


Figure 5. Example of selecting template based on partitioning angle.

4. Experimental results

The proposed improvements are implemented on top of ECM-6.0. Simulation results are conducted under the common test condition of ECM [21]. The random access (RA) and low delay B (LB) configurations with 4 QP values (22, 27, 32, 37) are tested. BD-rate is used to measure the rate distortion coding performance of the proposed methods. It is noted that since two proposed methods, improving TMVP and GPM with template matching [22][23], have been adopted into ECM-6.0 software platform, we disable these two methods in the anchor.

As shown in Table 1, compared to the anchor, the overall BD-rate saving for {Y, U, V} components are {0.20%, 0.35%, 0.34%} for RA and {0.33%, 0.57%, 0.43%} for LB with all three proposed methods enabled. It is also observed that the BD-rate saving in LB is higher than that in RA. It is reasonable since TMVP has higher coding efficiency in LB. The encoding and decoding runtime remains the same as the anchor. The results verify that the proposed methods provide a good trade-off between coding efficiency and runtime complexity.

Furthermore, to verify the effectiveness of individual improvement, the three proposed methods are enabled separately and the results are summarized in Table 2 to Table 4. Only class C sequences are evaluated in the tests. Table 2 shows that the improvement to TMVP mainly benefits the LB configuration since TMVP has higher probability of being selected in LB. Table 3 shows the results of improvement to BCW, which provides coding gain in both RA and LB. It is noted that the improvement to BCW achieves higher coding gain in larger resolution sequences [24]. The results of improvement to GPM are shown in Table

4, which achieves more than 0.15% BD-rate reduction in both RA and LB. In addition, it is also observed that the coding gain of three proposed methods are largely additive with minor overlap in coding gain from each proposed method. As shown in Table 1, the BD-rate saving of class C are 0.20% for RA and 0.55% for LB, which are closed to the sum of BD-rate saving of each proposed method (i.e. 0.20% for RA and 0.60% for LB).

Table 1. Simulation results of proposed methods.

Class (Resolution)	Sequence	Random Access Main10			Low delay B Main10		
		Y	U	V	Y	U	V
A1 (3840×2160)	Tango2	-0.31%	-0.78%	-0.78%			
	FoodMarket4	-0.27%	-0.46%	-0.41%			
	Campfire	-0.21%	-0.22%	-0.42%			
A2 (3840×2160)	CatRobot1	-0.14%	-0.26%	-0.31%			
	DaylightRoad2	-0.13%	0.04%	-0.21%			
	ParkRunning3	-0.14%	-0.34%	-0.36%			
B (1920×1080)	MarketPlace	-0.26%	-0.33%	-0.45%	-0.24%	-0.74%	-0.57%
	RitualDance	-0.25%	-0.40%	-0.41%	-0.21%	-0.40%	-0.50%
	Cactus	-0.02%	-0.21%	-0.19%	-0.56%	-0.35%	-0.53%
	BasketballDrive	-0.28%	-0.78%	-0.69%	-0.27%	-0.18%	-0.50%
	BQTerrace	-0.14%	-0.27%	-0.17%	-0.47%	-1.20%	-0.81%
C (832×480)	BasketballDrill	-0.18%	-0.30%	-0.12%	-0.54%	-0.59%	-0.90%
	BQMall	-0.07%	-0.10%	-0.11%	-0.53%	-0.01%	-0.21%
	PartyScene	-0.25%	-0.37%	-0.13%	-0.77%	-0.28%	-1.03%
	RaceHorses	-0.29%	-0.40%	-0.39%	-0.36%	-0.98%	0.19%
E (1280×720)	FourPeople				-0.11%	0.35%	-0.29%
	Johnny				0.00%	-3.23%	-1.02%
	KristenAndSara				0.08%	0.73%	0.99%
Class Summary	Class A1	-0.26%	-0.49%	-0.54%			
	Class A2	-0.14%	-0.19%	-0.29%			
	Class B	-0.19%	-0.40%	-0.38%			
	Class C	-0.20%	-0.29%	-0.19%	-0.55%	-0.46%	-0.49%
	Class E				-0.01%	-0.71%	-0.11%
Overall Summary	Average	-0.20%	-0.35%	-0.34%	-0.33%	-0.57%	-0.43%
	EncT		101%			103%	
	DecT		99%			100%	

Table 2. Simulation results of improvement to TMVP.

Class (Resolution)	Sequence	Random Access Main10			Low delay B Main10		
		Y	U	V	Y	U	V
C (832×480)	BasketballDrill	0.05%	-0.14%	0.08%	-0.31%	-0.39%	-0.26%
	BQMall	0.13%	0.30%	0.17%	-0.36%	0.56%	-0.24%
	PartyScene	-0.17%	0.03%	-0.06%	-0.67%	-0.35%	-0.64%
	RaceHorses	-0.03%	-0.14%	-0.12%	-0.15%	-0.52%	-0.02%
Class C Summary		0.00%	0.01%	0.02%	-0.37%	-0.18%	-0.29%

Table 3. Simulation results of improvement to BCW.

Class (Resolution)	Sequence	Random Access Main10			Low delay B Main10		
		Y	U	V	Y	U	V
C (832×480)	BasketballDrill	-0.01%	-0.09%	0.17%	-0.10%	-0.18%	-0.44%
	BQMall	-0.07%	0.29%	0.03%	-0.01%	1.06%	-0.53%
	PartyScene	-0.04%	-0.31%	-0.04%	0.01%	0.52%	0.14%
	RaceHorses	-0.06%	-0.01%	0.08%	-0.09%	-0.60%	0.66%
Class C Summary		-0.04%	-0.03%	0.06%	-0.05%	0.20%	-0.04%

Table 4. Simulation results of improvement to GPM.

Class (Resolution)	Sequence	Random Access Main10			Low delay B Main10		
		Y	U	V	Y	U	V
C (832×480)	BasketballDrill	-0.17%	-0.26%	0.05%	-0.22%	-0.01%	-0.83%
	BQMall	-0.15%	0.15%	-0.23%	-0.16%	0.73%	-0.44%
	PartyScene	-0.05%	-0.16%	0.12%	-0.17%	0.42%	-0.09%
	RaceHorses	-0.26%	-0.35%	-0.04%	-0.15%	0.10%	0.05%
Class C Summary		-0.16%	-0.15%	-0.02%	-0.18%	0.31%	-0.33%

5. Conclusion

In this paper, we propose to apply template matching mechanism to merge tools to improve the prediction accuracy of temporal motion vector predictor, bi-prediction with CU-level weight and geometric partition modes. The prediction direction and reference picture of TMVP, bi-predicted weights for BCW and motion of GPM are determined and refined using the template matching method. Experimental results show that on top of ECM, the proposed schemes achieve 0.20% and 0.33% luma BD-rate saving in random access and low delay B configurations, respectively. In addition, the impact on the encoding and decoding runtime is negligible. It is worth noting that all the three proposed methods have been adopted into the ECM software platform due to the good trade-off between coding efficiency and runtime complexity.

References

[1] G. J. Sullivan, J.-R. Ohm, W.-J. Han and T. Wiegand, “Overview of the high efficiency video coding (HEVC) standard,” IEEE Trans. Circuits and Systems for Video Technology, vol. 22, no. 12, pp. 1649-1668, Dec. 2012

[2] B. Bross, Y.-K. Wang, Y. Ye, S. Liu, J. Chen, G. J. Sullivan and J.-R. Ohm, “Overview of the Versatile Video Coding (VVC) Standard and its Applications,” IEEE Trans. Circuits and Systems for Video Technology, vol 31, no. 10, pp. 3736-3764, Oct. 2021.

[3] A. Browne, Y. Ye and S. Kim, “Algorithm description for Versatile Video Coding and Test Model 17 (VTM 17),” JVET-Z2002, July 2022.

[4] Y.-W. Huang, J. An, H. Huang, X. Li, S.-T. Hsiang, K. Zhang, H. Gao, J. Ma and O. Chubach, “Block Partitioning Structure in the VVC Standard,” IEEE Trans. Circuits and Systems for Video Technology, vol 31, no. 10, pp. 3818-3833, Oct. 2021.

[5] J. Pfaff, A. Filippov, S. Liu, X. Zhao, J. Chen, S. Hernandez, T. Wiegand, V. Rufitskiy, A. K. Ramasubramonian and G. Van der Auwera, “Intra Prediction and Mode Coding in VVC,” IEEE Trans. Circuits and Systems for Video Technology, vol 31, no. 10, pp. 3834-3847, Oct. 2021.

[6] W.-J. Chien, L. Zhang, M. Winken, X. Li, R.-L. Liao, H. Gao, C.-W. Hsu, H. Liu and C.-C. Chen, “Motion Vector Coding and Block Merging in the Versatile Video Coding Standard,” IEEE Trans. Circuits and Systems for Video Technology, vol 31, no. 10, pp. 3848-3861, Oct. 2021.

[7] H. Yang, H. Chen, J. Chen, S. Esenlik, S. Sethuraman, X. Xiu, E. Alshina and J. Luo, “Subblock-Based Motion Derivation and Inter Prediction Refinement in the Versatile Video Coding Standard,” IEEE Trans. Circuits and Systems for Video Technology, vol 31, no. 10, pp. 3862-3877, Oct. 2021.

[8] X. Zhao, S. Kim, Y. Zhao, H. E. Egilmez, M. Koo, S. Liu, J. Lainema and M. Karczewicz, "Transform Coding in the VVC Standard," IEEE Trans. Circuits and Systems for Video Technology, vol 31, no. 10, pp. 3878-3890, Oct. 2021.
[9] H. Schwarz, M. Coban, M. Karczewicz, T.-D. Chuang, F. Bossen, A. Alshin, J. Lainema, C. R. Helmrich and T. Wiegand, "Quantization and Entropy Coding in the Versatile Video Coding (VVC) Standard," IEEE Trans. Circuits and Systems for Video Technology, vol 31, no. 10, pp. 3891-3906, Oct. 2021.
[10] M. Karczewicz, N. Hu, J. Taquet, C.-Y. Chen, K. Misra, K. Andersson, P. Yin, T. Lu, E. Francois and J. Chen, "VVC In-Loop Filters," IEEE Trans. Circuits and Systems for Video Technology, vol 31, no. 10, pp. 3907-3925, Oct. 2021.
[11] M. Coban, F. Le Léannec, K. Naser, J. Ström and L. Zhang, "Algorithm description of Enhanced Compression Model 6 (ECM 6)," JVET-AA2025, July 2022.
[12] VTM software, https://vcgit.hhi.fraunhofer.de/jvet/VVCSoftware_VTM.git.
[13] V. Seregin, J. Chen, F. Le Léannec and K. Zhang, "JVET AHG report: ECM software development (AHG6)," JVET-AB0006, Oct. 2022.
[14] ECM software, https://vcgit.hhi.fraunhofer.de/ecm/ECM.git.
[15] C.-C. Chen, X. Xiu, Y. He and Y. Ye, "Generalized bi-prediction method for future video coding," Proc. Picture Coding Symposium (PCS), Dec. 2016.
[16] Y.-C. Su, C.-Y. Chen, Y.-W. Huang, S.-M. Lei, Y. He, J. Luo, X. Xiu and Y. Ye, "CE4-related: Generalized bi-prediction improvements combined from JVET-L0197 and JVET-L0296," JVET-L0646, Oct. 2018.
[17] H. Gao, R.-L. Liao, K. Reuze, S. Esenlik, E. Alshina, Y. Ye, J. Chen, J. Luo, C.-C. Chen, H. Huang, W.-J. Chien, V. Seregin and M. Karczewicz, "Advanced Geometric-Based Inter Prediction for Versatile Video Coding," Proc. Data Compression Conference (DCC), Mar. 2020.
[18] H. Gao et al., "Integrated Text for GEO," JVET-Q0806, Jan. 2020.
[19] N. Zhang, K. Zhang, L. Zhang, H. Liu, Z. Deng and Y. Wang, "EE2-3.1/EE2-3.2: Adaptive Reordering of Merge Candidates with Template/Bilateral Matching," JVET-W0090, July 2021.
[20] Z. Lv, C. Zhou and J. Zhang, "EE2-2.3: Template matching based OBMC", JVET-Z0061, Apr. 2022.
[21] M. Karczewicz and Y. Ye, "Common Test Conditions and evaluation procedures for enhanced compression tool testing," JVET-Y2017, Jan. 2022.
[22] Y.-J. Chang, H. Huang, V. Seregin, C.-C. Chen, M. Karczewicz, R.-L. Liao, J. Chen, Y. Ye, X. Li, L. Zhao, K. Zhang, N. Zhang, L. Zhang, G. Laroche, P. Onno and R. Bellessort, "EE2-3.4, EE2-3.5, EE2-3.6: Experimental results of the MV candidates reordering in candidate types based on template matching costs," JVET-Y0134, Jan. 2022.
[23] X. Xiu, C.-W. Kuo, X. Wang, R.-L. Liao, Y. Ye, X. Li, J. Chen, Z. Deng, K. Zhang, L. Zhang, N. Zhang, Y. Wang, Y.-J. Chang, H. Huang, V. Seregin, C.-C. Chen and M. Karczewicz, "EE2-related: Combination of EE2-3.3, EE2-3.4 and EE2-3.5," JVET-W0097, July 2021.
[24] R.-L. Liao, J. Chen, Y. Ye and X. Li, "EE2-2.2: Template matching based BCW index derivation for merge mode," JVET-AB0079, Oct. 2022.