

## Research Article

# Euclidean Distance-Based Weighted Prediction for Merge Mode in HEVC

Hongwei Guo <sup>1</sup>, Xiangsuo Fan <sup>2</sup> and Lei Min <sup>3</sup>

<sup>1</sup>School of Engineering, Honghe University, Mengzi, Yunnan 661100, China

<sup>2</sup>School of Electrical and Information Engineering, Guangxi University of Science and Technology, Liuzhou, Guangxi 545006, China

<sup>3</sup>Institute of Optics and Electronics, Chinese Academy of Sciences, Chengdu, Sichuan 610209, China

Correspondence should be addressed to Xiangsuo Fan; [ffanxs@163.com](mailto:ffanxs@163.com)

Received 26 April 2019; Revised 24 July 2019; Accepted 27 August 2019; Published 22 September 2019

Academic Editor: Deepu Rajan

Copyright © 2019 Hongwei Guo et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Merge mode can achieve a considerable coding gain because of reducing the cost of coding motion information in video codecs. However, the simple adoption of the motion information from the neighbouring blocks may not achieve the optimal performance as the motion correlation between the pixels and the neighbouring block decreases with their distance increasing. To address this problem, the paper proposes a Euclidean distance-based weighted prediction algorithm as an additional candidate in the merge mode. First, several predicted blocks are generated by motion compensation prediction (MCP) with the motion information from available neighbouring blocks. Second, an additional predicted block is generated by a weighted average of the predicted blocks above, where the weighted coefficient is related to Euclidean distances from the neighbouring candidate to the pixel points in the current block. Finally, the best merge mode is selected by the rate distortion optimization (RDO) among the original merge candidates and the additional candidate. Experimental results show that, on the joint exploration test model 7.0 (JEM 7.0), the proposed algorithm achieves better coding performance than the original merge mode under all configurations including random access (RA), low delay B (LDB), and low delay P (LDP), with a slight coding complexity increase. Especially for the LDP configuration, the proposed method achieves 1.50% bitrate saving on average.

## 1. Introduction

As one of the main components of multimedia information, the apparent feature of videos is the amount of data. Video coding technologies [1] can effectively achieve data compression by removing various redundancies in videos, which include spatial redundancy, temporal redundancy, statistics redundancy, and visual redundancy. In the current block-based hybrid video coding framework, a picture is partitioned into blocks of variable size and each block is either spatially or temporally predicted with a set of mode or coding parameters. Motion information including motion vectors (MVs) and reference indices is the main parameter for a prediction unit (PU) which adopted interprediction. To remove the motion information redundancy among neighbouring PUs, motion vector coding techniques such as advanced motion vector prediction (AMVP) [2], block merging, and skip mode [3] are introduced in high efficiency

video coding (HEVC) [4]. It is noted that the skip mode can be treated as a special case of the merge mode, where the prediction residual is not transmitted or all discrete cosine transform (DCT) coefficients of the prediction residual are quantized to zeros [5]. The block merging generates a single motion parameter set for a whole region of contiguous motion-compensated blocks, and thus the motion parameters need to be signalled only once for all blocks included in the same merging region [6]. In the mode decision process, rate distortion optimization (RDO) [7] is performed among the merge mode, skip mode, intermode, and intramode [8]. The best mode is selected in terms of the minimization of rate-distortion (R-D) cost, which is expressed as follows:

$$\min \{J\}, \quad \text{where } J = D + \lambda_{\text{mode}} \cdot R, \quad (1)$$

where  $J$  is the R-D cost and  $D$  is the coding distortion, which is measured by sum of squared differences between the reconstructed block and original block;  $R$  is the bits to

encode the corresponding mode and DCT coefficients; and  $\lambda_{\text{mode}}$  is called as a Lagrange multiplier, which is used to adjust the trade-off between distortion and bits.

The block merging technique brings the coding performance improvement significantly in HEVC [9]. Therefore, the merge mode is inherited and extended in the exploration of future video coding which is initiated by the joint video exploration team (JVET) [10], and the new video coding standard has been officially named as versatile video coding (VVC) [11] at the 10th JVET meeting. Recently, some modified versions of the merge mode were proposed for the implementation in video codecs [12–19], and the most of these methods are focused on reducing the computational complexity at the cost of  $R$ - $D$  performance degradation. For example, in [12], based on the all-zero block, the motion estimation information, and the mode selection correlation between the largest coding unit (LCU) and its children coding units (CUs), an early merge mode decision algorithm is proposed to reduce the encoding computational complexity. In [13], an early merge mode decision framework is proposed by identifying smooth/single-motion regions, which achieves the coding time reducing of 46% on average but degrades rate distortion performance to some extent. To implement merge mode module in a hardware-based encoder, in [14, 15], a novel hardware design which can alleviate the computational cost and the requirement of memory accesses was proposed. The coding complexity of three-dimensional- (3D-) HEVC is higher than HEVC, because 3D-HEVC needs to handle the multiple views and depth information. So, in [16], an adaptive merge list construction for 3D-HEVC encoders to fast the coding speed was proposed. In [17], an early merge mode decision for 3D-HEVC encoders is proposed by analysing the inter-view correlation and hierarchical depth correlation of coding modes. In [18], to provide a better balance between computational complexity and coding efficiency, several fast CU encoding schemes are surveyed according to the rate-distortion-complexity characteristics of early merge mode decision. In [19], to accelerate interprediction coding, a merge candidate decision and an early merge termination are proposed based on the sum of absolute transformed difference (SATD) cost, and a merge-based coding unit filtering is proposed to remove the unnecessary CU evaluation process. In addition, [20, 21] focus on improving the  $R$ - $D$  performance rather than reducing coding complexity. In [20], an adjacent-block-based prediction model for more accurately predicting the deformable block is proposed, which is added as an additional candidate in the merge mode. Zhang et al. [21] proposed a merge mode for deformable block motion information derivation in rotation, zoom, and deformation motion, which obtains greater  $R$ - $D$  performance improvement but leads to significant complexities increasing at both encoder and decoder.

In this paper, we proposed a Euclidean distance-based weighted prediction approach as an additional candidate for the merge mode to further improve the  $R$ - $D$  performance in HEVC. It considers the correlation between the motion of

pixels in different locations and the motion of the candidate block and makes full use of the motion information of all available spatial-temporal merge candidates. Experimental results show that the proposed algorithm achieves better coding performance than the original merge mode, with a slight coding complexity increase.

The rest of this paper is organized as follows. Section 2 briefly reviews the merge mode in interframe video coding. Section 3 presents the proposed Euclidean distance-based weighted prediction algorithm. Experimental results are presented in Section 4, and finally, conclusions are drawn in Section 5.

## 2. Merge Mode in HEVC

In order to reduce the number of bits required to encode the motion information, a merge mode [3] was proposed in HEVC. The merge mode utilizes the motion correlation of neighbouring blocks in the spatial and temporal domains. Motion estimation is not required in the merge mode, and the motion information from the neighbouring encoded block is directly adopted to motion compensation prediction (MCP) for the current block. Therefore, encoding the motion information of a merging block only needs a few number of bits to indicate the merge index. The flow chart of the merge mode is shown in Figure 1. A merge candidate list is first constructed, which includes several spatial candidates and one temporal candidate. Then, RDO is employed to choose the best merge candidate from the list. Finally, the predicted block of the current PU is obtained by MCP with the MV of the best merge candidate.

The construction of merge candidate list is given in Figure 2. The output of this process is a list of merge candidates, and each of them has a tuple of motion parameters, which can be used for MCP of the current PU. The flag *NumMergeCands* is a predefined constant, which indicates the number of candidates in the list. In HEVC, the number of merge candidates is five, but the number has been increased to seven in VVC. The whole process of adding candidates will stop as soon as the number of candidates reaches *NumMergeCands* during the following ordered steps: the process starts with the derivation of the initial candidates from spatially neighbouring PUs called spatial candidates. Then, a candidate from a PU in a temporally collocated picture can be included, which is called the temporal candidate. Due to unavailability of certain candidates or some of them being redundant, the number of initial candidates could be less than *NumMergeCands*. In this case, additional candidates are inserted to the list so that the number of candidates in the list is always equal to *NumMergeCands*. For the spatial merge candidate selection, a maximum of four candidates are selected from the neighbouring PUs encoded within the current picture. For the temporal merge candidate derivation, one candidate is generated from the collocated PUs in the previous frame encoded. The specific implementation about the merge mode can be referred from [3].

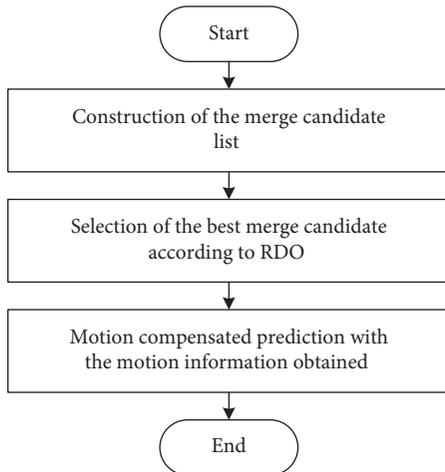


FIGURE 1: Flow chart of the merge mode.

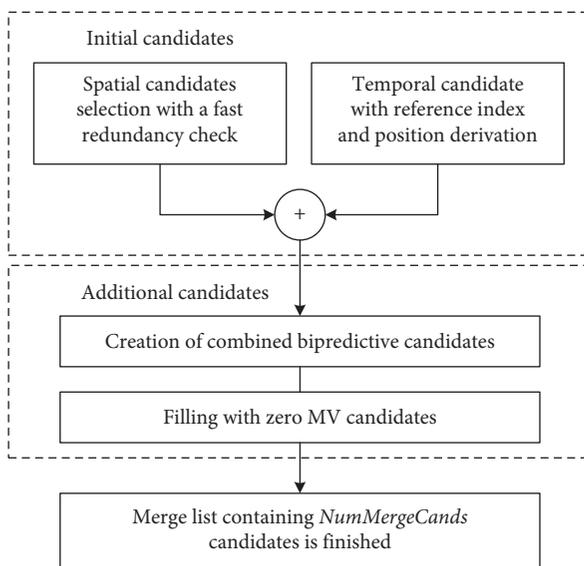


FIGURE 2: Construction of the merge candidate list.

MCP includes uniprediction for  $P$  frames and biprediction for  $B$  frames. Taking the uniprediction as an example, the MCP is expressed as follows:

$$P(x, y) = F_{\text{ref}}(x + i, y + j), \quad (2)$$

where  $P(x, y)$  and  $F_{\text{ref}}(x, y)$  are the predicted picture and reference picture, respectively, and  $(i, j)$  is the MV from the merge candidate. In addition, the biprediction needs two MVs and corresponding reference index for performing weighted prediction.

### 3. Proposed Weighted Prediction for Merge Mode

It has been proved that the merge mode can obtain a considerable  $R$ - $D$  performance improvement because of reducing the cost of encoding motion information. However, the merge mode can be further improved, since the

MCP in the merge mode is not accurate enough. Intuitively, a more accurate prediction can reduce the number of nonzero coefficients after quantization, which directly affects coded bitrate. In this section, we first analyse the characteristic of MCP in the merge mode. Then, a Euclidean distance-based weighted prediction is proposed as an additional candidate in the merge mode to obtain a more accurate predicted block.

#### 3.1. Analysis for Distribution of Predicted Residual.

During the merge mode decision process, MCP is performed with the motion information of each merge candidate to obtain the prediction signal for the current block. Generally, with the distance, between the pixel points in the current block to its neighbouring block which is the merge candidate, increasing, their motion correlation decreases. Consequently, the distribution of the residual after MCP is uneven over the whole block, and the magnitudes of the residual usually become larger with smaller motion correlation. The residual block is the difference between the original block and predicted block as follows:

$$B(x, y) = F(x, y) - P(x, y). \quad (3)$$

If the motion information of the merge candidate at the left top is adopted for MCP, the magnitudes of the residual gradually increase from left to right and from top to bottom. Similar distributions of the residual can be observed when other merge candidates are adopted for MCP. Specifically, the positions of merge candidates and the residual distribution after MCP with different candidates are shown in Figure 3. In Figure 3(a), the positions of left (L), top (T), right top (RT), left bottom (LB), and left top (LT) are in the current picture, and temporal motion vector prediction (TMVP) indicates the position of collocated right bottom in the previous frame. In Figures 3(b)–3(e), the magnitudes of the residual after MCP gradually increase along with the direction of the arrows.

#### 3.2. Euclidean Distance-Based Weighted Prediction.

According to the analysis as above, we can see that the conventional MCP with only one set of motion vector cannot achieve optimal performance for the merge mode because the motion compensation is not accurate enough. Therefore, to obtain a better prediction for the current block, an additional candidate is added to the merge mode decision process. The predicted block of the additional candidate is generated by weighting the predicted blocks obtained from the available merge candidates. Figure 4 illustrates the schematic diagram of adding the additional merge candidate, and the detailed steps are described in the following:

Step 1: for each merge candidate position, the availability is checked according to the order of L, T, RT, LB, LT, and TMVP as in Figure 3(a). It is note that the TMVP only checks the position of co-located right bottom in the temporal candidates. The additional candidate is added as described in the following steps 2

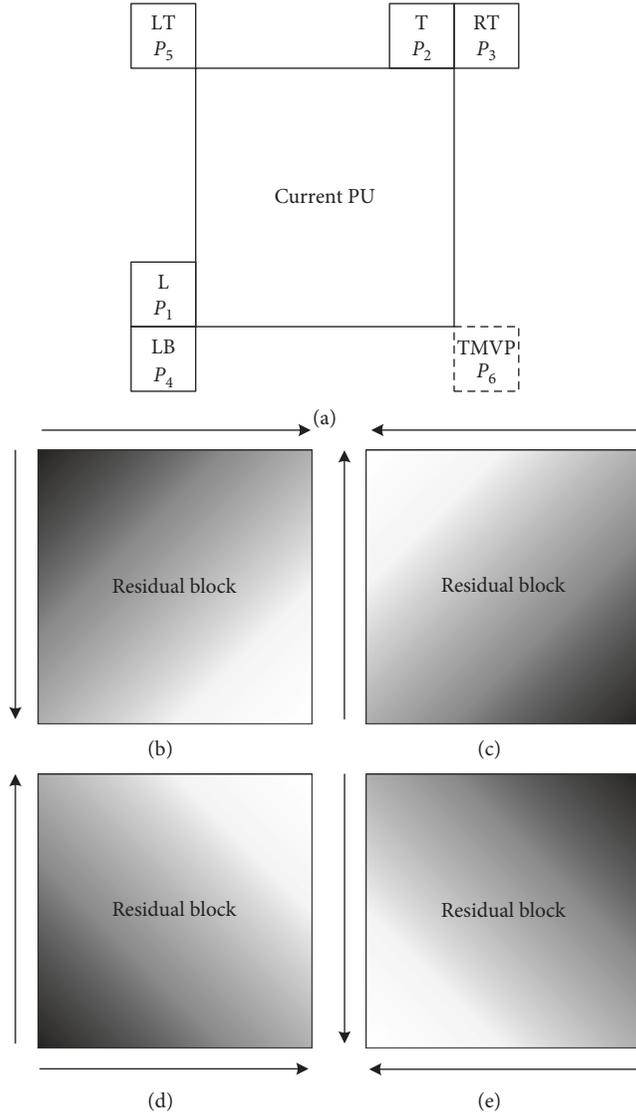


FIGURE 3: Positions of merge candidates and residual distribution after MCP. (a) Positions of spatial-temporal candidates for the merge mode. (b) MCP by LT. (c) MCP by TMVP. (d) MCP by L or LB. (e) MCP by T or RT.

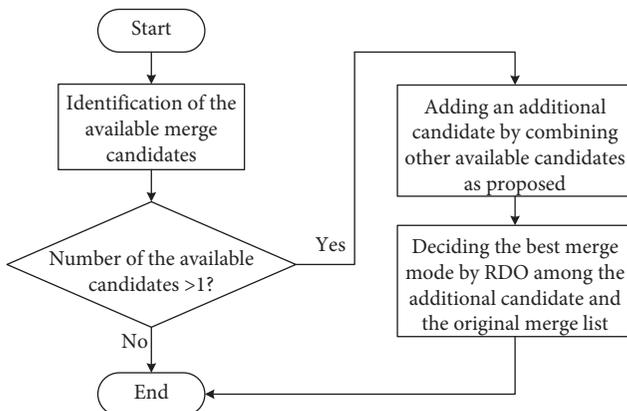


FIGURE 4: Schematic diagram of the proposed merge candidate for the merge mode.

and 3 if the number of the available candidates is larger than one.

Step 2: according to the motion information of the available merge candidates, the corresponding predicted blocks are obtained by (2). Then, the predicted block of the additional candidate is obtained according to (4) with the predefined Euclidean distance-based weights.

$$P(x, y) = \frac{\sum_{i=1}^n W_i(x, y) P_i(x, y)}{\sum_{i=1}^n W_i(x, y)}, \quad n = 2 \dots 6, \quad (4)$$

where  $W_i(x, y)$  is the weighted coefficient which is inversely proportional to the Euclidean distance from the point  $(x, y)$  to the corresponding candidate block and  $n$  is the number of available merge candidates. The definition of the weighted coefficients is illustrated in Table 1, where  $H$  and  $W$  are the height and width of the current PU, respectively.

Step 3: the best merge mode is selected by RDO among the additional candidate and the original merge candidates. If the additional candidate is selected as the best merge mode, motion information needs to be padded for the current PU so that it can be used as a candidate in the coding of the following units. Here, the motion information of the first one of available candidates in Step 1 is used for padding.

In the end, a flag needs to be signalled to decoder side to indicate the additional merge candidate if the proposed Euclidean distance-based weighted prediction method is selected for coding the current PU. Then, the decoder will reconstruct the pixel block using the same weighted algorithm as the encoder side.

## 4. Experimental Results

To evaluate the performance of the proposed Euclidean distance-based weighted prediction for the merge mode, the proposed algorithm is integrated into the joint exploration test model 7.0 (JEM 7.0) [10], which is built up on top of the HEVC test model (HM) by the JVET to evaluate new compression techniques. Several experiments are conducted in comparison with the anchor (the original JEM 7.0) under the JVET common test conditions (CTC) and software reference configurations [22]. The sixteen test sequences from Classes B, C, D, and E are selected for low delay  $B$  (LDB) and low delay  $P$  (LDP) configurations, and the thirteen test sequences from Classes B, C, and D are selected for random access (RA) configuration, according to CTC. The coding performance is measured by the Bjontegaard delta bitrate (BDBR) [23] which calculates the average bitrate reduction under the same peak signal to noise ratio (PSNR) conditions. Note that the positive value of BDBR means performance loss while the negative value means performance improvement. The computational complexities

TABLE 1: Definition of weighted coefficients for each candidate.

Position of candidates	Prediction blocks	$W_i(x, y)$
L	$P_1$	$W - x$
T	$P_2$	$H - y$
RT	$P_3$	$\sqrt{x^2 + (H - y)^2}$
LB	$P_4$	$\sqrt{(W - x)^2 + y^2}$
LT	$P_5$	$\sqrt{(W - x)^2 + (H - y)^2}$
TMVP	$P_6$	$\sqrt{(x + 1)^2 + (y + 1)^2}$

TABLE 2: BDBRs and codec time ratios of the proposed method compared to the JEM 7 anchor.

Class	Sequence	Low delay $P$ main 10			Low delay $B$ main 10			Random access main 10		
		$Y$	$U$	$V$	$Y$	$U$	$V$	$Y$	$U$	$V$
B	Kimono	-0.47%	-0.42%	-0.49%	-0.13%	-0.49%	0.14%	-0.06%	0.27%	-0.02%
	ParkScene	-0.43%	-0.72%	-1.15%	0.03%	0.03%	-0.22%	-0.02%	-0.09%	-0.15%
	Cactus	-1.63%	-3.67%	-2.78%	-0.11%	-0.89%	-0.69%	-0.20%	-0.61%	-0.52%
	BasketballDrive	-1.03%	-2.22%	-1.95%	-0.23%	-0.73%	-0.72%	-0.12%	-0.49%	-0.18%
	BQTerrace	-6.74%	-6.43%	-6.53%	-0.27%	-0.44%	-2.42%	-0.54%	-1.11%	-1.17%
	Average	-2.06%	-2.69%	-2.58%	-0.14%	-0.50%	-0.78%	-0.19%	-0.41%	-0.41%
C	BasketballDrill	-1.49%	-2.79%	-3.18%	-0.22%	-0.11%	-0.38%	-0.15%	0.15%	-0.11%
	BQMall	-0.50%	-0.66%	-1.09%	0.09%	0.46%	-0.17%	-0.07%	0.14%	0.05%
	PartyScene	-1.47%	-1.69%	-1.14%	-0.31%	-0.66%	-0.16%	-0.19%	-0.28%	-0.26%
	RaceHorses	-0.67%	-1.65%	-1.22%	-0.03%	0.04%	-0.18%	-0.07%	0.02%	-0.11%
	Average	-1.03%	-1.70%	-1.65%	-0.12%	-0.30%	-0.22%	-0.12%	0.01%	-0.11%
D	BasketballPass	-0.32%	-0.78%	-0.79%	0.05%	-0.36%	-0.56%	-0.15%	0.12%	0.21%
	BQSquare	-3.24%	-3.20%	-4.92%	-0.25%	0.97%	-2.21%	-0.24%	-0.02%	0.42%
	BlowingBubbles	-0.85%	-0.96%	-1.26%	-0.09%	-0.51%	-0.59%	-0.02%	-0.04%	-0.25%
	RaceHorses	-0.07%	0.92%	0.34%	0.09%	-0.53%	-0.71%	0.01%	0.02%	-0.07%
Average	-1.12%	-1.01%	-1.66%	-0.05%	-0.11%	-1.02%	-0.10%	0.02%	0.08%	
E	FourPeople	-1.32%	-1.47%	-0.80%	-0.31%	-1.50%	-2.04%	—	—	—
	Johnny	-2.82%	-3.84%	-3.66%	0.00%	-0.37%	0.91%	—	—	—
	KristenAndSara	-1.03%	-2.05%	-3.02%	-0.14%	-0.28%	0.39%	—	—	—
	Average	-1.72%	-2.45%	-2.49%	-0.15%	-0.72%	-0.24%	—	—	—
Overall	-1.50%	-1.98%	-2.10%	-0.11%	-0.39%	-0.60%	-0.14%	-0.15%	-0.17%	
EncTR (%)		104			103			103		
DecTR (%)		99			100			100		

are measured by encoding time ratio (EncTR) and decoding time ratio (DecTR), which are defined as follows:

$$TR = \frac{T_{\text{pro}}}{T_{\text{org}}} \times 100\%, \quad (5)$$

where  $T_{\text{pro}}$  and  $T_{\text{org}}$  are the encoding time or decoding time for the proposed algorithm and the original JEM 7.0, respectively. If the time ratios are larger than 100%, it means the computational complexities are increased, and vice versa.

Table 2 gives the experimental results about the coding performance and the computational complexities in the cases of LDP, LDB, and RA configurations. In Table 2, the BDBRs for  $Y$ ,  $U$ , and  $V$  components are given, and the time ratios are calculated by the total time for encoding or decoding. The results show that the proposed algorithm achieves bitrate savings of 1.50%, 0.11%, and 0.14% on average, respectively, under the configurations of LDP, LDB, and RA, with a slight

encoding complexity increase. The encoding complexity increase is due to an additional RDO process in the best merge mode decision. It is worth noting that the average bitrate saving of the LDP configuration is much higher than that of the other two configurations, which is because  $P$  frames have a larger space for improving prediction accuracy than  $B$  frames.

In the case of LDP, some test sequences show significant rate distortion performance improvement and up to 6.74% for test sequence *BQTerrace*. Figure 5 shows some  $R$ - $D$  curve comparisons for four test sequences from different classes under the case of LDP. We can see that the proposed method outperforms the original JEM 7.0 at both high and low bitrates.

In the above-mentioned references about the merge mode, only [20, 21] focus on improving the  $R$ - $D$  performance rather than reducing coding complexity, which is similar to our method. Therefore, the merge mode for deformable block by the bilinear interpolation model in [20]

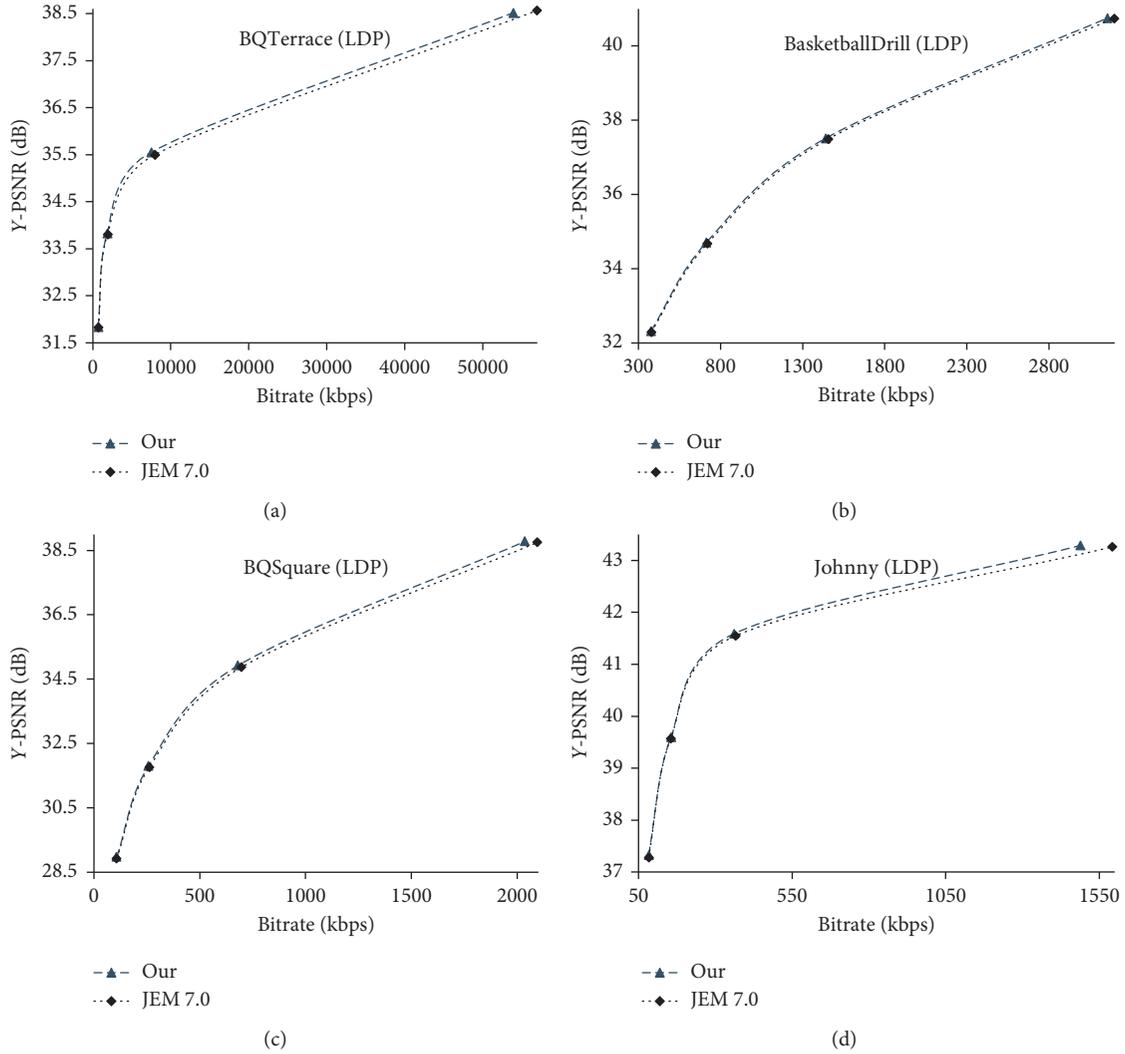
FIGURE 5:  $R$ - $D$  curves comparison under LDP configuration.

TABLE 3: Y BDBRs and codec time ratios of MMD [21], MMD-B [20], and our method.

Sequence	MMD [21]			MMD-B [20]			Our method		
	LDP	LDB	RA	LDP	LDB	RA	LDP	LDB	RA
Class B	-2.3%	-1.7%	-1.8%	-0.3%	-0.4%	0.6%	-2.1%	-0.1%	-0.2%
Class C	-2.2%	-1.2%	-1.0%	-0.1%	-0.2%	-0.2%	-1.0%	-0.1%	-0.1%
Class D	-3.3%	-1.7%	-1.3%	-0.1%	-0.2%	-0.2%	-1.1%	-0.1%	-0.1%
Class E	-3.3%	-2.6%	—	-1.1%	-1.4%	—	-1.7%	-0.2%	—
Overall	-2.7%	-1.7%	-1.3%	-0.4%	-0.5%	-0.3%	-1.5%	-0.1%	-0.1%
EncTR (%)	112	111	116	103	103	104	104	103	103
DecTR (%)	115	114	111	103	103	105	99	100	100

(MMD-B [20]) and the merge mode for deformable block in [21] (MMD [21]) are selected for further performance comparisons. Table 3 shows Y component BDBRs and encode and decode computational complexities for three competitions, where the BDBR, EncTR, and DecTR are calculated by the competition against its corresponding original encoder anchor. We can see that MMD [21] has the best  $R$ - $D$  performance, but it significantly increases encode and decode computational complexities. MMD-B [20] has

the least  $R$ - $D$  performance improvement with a slight encode and decode time increasing. Our method has a moderate  $R$ - $D$  performance improvement with a slight coding complexity increase, but it does not increase any decode time.

## 5. Conclusion

In this paper, we proposed a Euclidean distance-based weighted prediction algorithm for the MCP in the merge

mode to improve the coding performance. It considers the correlation between the motion of pixels in different locations and the motion of the candidate block and makes full use of the motion information of all available spatial-temporal merge candidates. The experimental results demonstrated that, on the platform of JEM 7.0, the proposed method can obtain the rate distortion performance improvement under different coding structures, where up to 1.50% coding gain on average can be obtained under the LDP configuration.

## Data Availability

The data used to support the findings of this study are available from the corresponding author upon request.

## Conflicts of Interest

The authors declare that they have no conflicts of interest.

## Acknowledgments

This work was supported in part by the National Natural Science Foundation of China under grant 41761079 and in part by the Yunnan Local Colleges Applied Basic Research Project under grant 2018FH001-056.

## References

- [1] D. Grois, T. Nguyen, and D. Marpe, "Coding efficiency comparison of AV1/VP9, H.265/MPEG-HEVC, and H.264/MPEG-AVC encoders," in *Proceedings of the Picture Coding Symposium (PCS)*, pp. 1–5, Nuremberg, Germany, December 2016.
- [2] J.-L. Lin, Y.-W. Chen, Y.-W. Huang, and S.-M. Lei, "Motion vector coding in the HEVC standard," *IEEE Journal of Selected Topics in Signal Processing*, vol. 7, no. 6, pp. 957–968, 2013.
- [3] P. Helle, S. Oudin, B. Bross et al., "Block merging for quadtree-based partitioning in HEVC," *IEEE Transactions on Circuits and Systems for Video Technology*, vol. 22, no. 12, pp. 1720–1731, 2012.
- [4] G. J. Sullivan, J.-R. Ohm, W.-J. Han, and T. Wiegand, "Overview of the high efficiency video coding (HEVC) standard," *IEEE Transactions on Circuits and Systems for Video Technology*, vol. 22, no. 12, pp. 1649–1668, 2012.
- [5] Y. Li, G. B. Yang, Y. P. Zhu, X. L. Ding, and X. M. Sun, "Unimodal stopping model-based early SKIP mode decision for high-efficiency video coding," *IEEE Transactions on Multimedia*, vol. 19, no. 7, pp. 1431–1441, 2017.
- [6] Y. Li, G. B. Yang, Y. P. Zhu, X. L. Ding, and R. R. Gong, "Probability model-based early merge mode decision for dependent views coding in 3D-HEVC," *ACM Transactions on Multimedia Computing, Communications, and Applications*, vol. 14, no. 4, pp. 1–15, 2018.
- [7] G. J. Sullivan and T. Wiegand, "Rate-distortion optimization for video compression," *IEEE Signal Processing Magazine*, vol. 15, no. 6, pp. 74–90, 1998.
- [8] S.-h. Jung and H. W. Park, "A fast mode decision method in HEVC using adaptive ordering of modes," *IEEE Transactions on Circuits and Systems for Video Technology*, vol. 26, no. 10, pp. 1846–1858, 2016.
- [9] J.-R. Ohm, G. J. Sullivan, H. Schwarz, T. K. Tan, and T. Wiegand, "Comparison of the coding efficiency of video coding standards-including high efficiency video coding (HEVC)," *IEEE Transactions on Circuits and Systems for Video Technology*, vol. 22, no. 12, pp. 1669–1684, 2012.
- [10] J. Chen, E. Alshina, G. J. Sullivan, J.-R. Ohm, and J. Boyce, "Algorithm description of joint exploration test model 7 (JEM 7)," in *Proceedings of the Joint Video Exploration Team (JVET) of ITU-T SG 16 WP 3 and ISO/IEC JTC 1/SC 29/WG 11 7th Meeting*, Torino, Italy, July 2017.
- [11] A. Kammoun, W. Hamidouche, F. Belghith, J.-F. Nezan, and N. Masmoudi, "Hardware design and implementation of adaptive multiple transforms for the versatile video coding standard," *IEEE Transactions on Consumer Electronics*, vol. 64, no. 4, pp. 424–432, 2018.
- [12] Z. Pan, S. Kwong, M.-T. Sun, and J. Lei, "Early MERGE mode decision based on motion estimation and hierarchical depth correlation for HEVC," *IEEE Transactions on Broadcasting*, vol. 60, no. 2, pp. 405–412, 2014.
- [13] J. Tariq, S. Kwong, and H. Yuan, "Spatial/temporal motion consistency based MERGE mode early decision for HEVC," *Journal of Visual Communication and Image Representation*, vol. 44, pp. 198–213, 2017.
- [14] T. S. Kim, C. E. Rhee, and H.-J. Lee, "Merge mode estimation for a hardware-based HEVC encoder," *IEEE Transactions on Circuits and Systems for Video Technology*, vol. 26, no. 1, pp. 195–209, 2016.
- [15] T. S. Kim, H. J. Lee, and C. E. Rhee, "Highly utilized merge mode estimation for a hardware-based HEVC encoder," in *Proceedings of the IEEE Workshop on Signal Processing Systems (SiPS)*, pp. 1–5, Hangzhou, China, October 2015.
- [16] Y. S. Heo, G. Bang, and G. H. Park, "Adaptive merge list construction for 3D-HEVC fast encoder," *Electronics Letters*, vol. 52, no. 8, pp. 604–605, 2016.
- [17] Y. X. Song and K. B. Jia, "Early merge mode decision for texture coding in 3D-HEVC," *Journal of Visual Communication and Image Representation*, vol. 33, pp. 60–68, 2015.
- [18] J. F. Wu, B. L. Guo, J. Hou, Y. Y. Yan, and J. Jiang, "Fast CU encoding schemes based on merge mode and motion estimation for HEVC inter prediction," *KSII Transactions on Internet and Information Systems*, vol. 10, no. 3, pp. 1195–1211, 2016.
- [19] Z. Cheng, H. Sun, D. Zhou, and S. Kimura, "Accelerating HEVC inter prediction with improved merge mode handling," *IEICE Transactions on Fundamentals of Electronics, Communications and Computer Sciences*, vol. E100.A, no. 2, pp. 546–554, 2017.
- [20] N. Zhang, X. P. Fan, D. B. Zhao, and W. Gao, "Motion vector derivation of deformable block," in *Proceedings of the IEEE International Conference on Image Processing (ICIP)*, pp. 1549–1552, Lake Buena Vista, FL, USA, September 2012.
- [21] N. Zhang, X. Fan, D. Zhao, and W. Gao, "Merge mode for deformable block motion information derivation," *IEEE Transactions on Circuits and Systems for Video Technology*, vol. 27, no. 11, pp. 2437–2449, 2017.
- [22] K. Suehring and X. Li, "JVET common test conditions and software reference configurations," in *Proceedings of the Joint Video Exploration Team (JVET) of ITU-T SG16WP3 and ISO/IEC JTC 1/SC 29/WG11 2nd Meeting*, San Diego, CA, USA, July 2016.
- [23] G. Bjontegaard, "Calculation of average PSNR difference between RDcurves," in *Proceedings of the ITU-T SC16/Q6 13th VCEG Meeting*, Austin, TX, USA, April 2001.

