PAPER Special Section on Picture Coding and Image Media Processing

Geometric Partitioning Mode with Inter and Intra Prediction for Beyond Versatile Video Coding

Yoshitaka KIDANI[†,††a)], **Haruhisa KATO**[†], **Kei KAWAMURA**[†], ***and*** **Hiroshi WATANABE**[††], ***Members***

SUMMARY Geometric partitioning mode (GPM) is a new inter prediction tool adopted in versatile video coding (VVC), which is the latest video coding of international standard developed by joint video expert team in 2020. Different from the regular inter prediction performed on rectangular blocks, GPM separates a coding block into two regions by the predefined 64 types of straight lines, generates inter predicted samples for each separated region, and then blends them to obtain the final inter predicted samples. With this feature, GPM improves the prediction accuracy at the boundary between the foreground and background with different motions. However, GPM has room to further improve the prediction accuracy if the final predicted samples can be generated using not only inter prediction but also intra prediction. In this paper, we propose a GPM with inter and intra prediction to achieve further enhanced compression capability beyond VVC. To maximize the coding performance of the proposed method, we also propose the restriction of the applicable intra prediction mode number and the prohibition of applying the intra prediction to both GPM-separated regions. The experimental results show that the proposed method improves the coding performance gain by the conventional GPM method of VVC by 1.3 times, and provides an additional coding performance gain of 1% bitrate savings in one of the coding structures for low-latency video transmission where the conventional GPM method cannot be utilized.

key words: *video coding, geometric partitioning mode (GPM), inter prediction, intra prediction, versatile video coding (VVC)*

1. Introduction

Video coding is a fundamental technology for video streaming and broadcast services. High efficiency video coding (HEVC) [1] is widely used for ultra high definition television (UHDTV) distribution worldwide. The coding performance of HEVC is not yet sufficient for UHDTV services over the mobile network due to the lack of transmission capacity. To address this situation, versatile video coding (VVC) [2], which is the state-of-the-art video coding of international standard with the coding performance gain of about 30–40% bitrate savings compared to that of HEVC, was developed by the joint video expert team (JVET) in 2020 [3]. The VVC coding performance gain is realized by a variety of newly adopted coding tools: non-square block partitioning, extended intra predictions, various inter predictions, enhanced transforms, and several new in-loop filters [3].

Manuscript received January 9, 2022.
Manuscript revised April 25, 2022.
Manuscript publicized June 21, 2022.

[†]The authors are with KDDI Research, Inc., Fujimino-shi, 356–8502 Japan.
[††]The authors are with Waseda University, Tokyo, 169–0072 Japan.
a) E-mail: yo-kidani@kddi.com

DOI: 10.1587/transinf.2022PCP0005

Geometric partitioning mode (GPM) is a newly VVC adopted inter prediction that contributes to the coding performance gain of VVC. The overview of the GPM in VVC is shown in Fig. 1 (a) [4], [5]. Different from the regular inter prediction performed on rectangular blocks in VVC, the GPM separates a coding block into two regions by the predefined 64 types of straight lines, generates inter predicted sampling values of luma and chroma component signals ("samples" hereinafter) for each GPM-separated region (P_0, P_1) with different motion vectors (MV_0, MV_1), and then blends them to obtain the final inter predicted samples (P_G). With this feature, the GPM improves the prediction accuracy at the boundary between the foreground and background with different motions. Especially, the GPM highly contributes to the coding performance gain of the coding structure for the low-latency video transmission, called low delay B (LB) configuration in JVET. The coding performance gain is up to 1.54% bitrate savings compared to the whole bitrate savings (= 30%) in the LB configuration [5], [6].

The GPM in VVC organized by two different inter predictions ("GPM-Inter/Inter" hereinafter) has room to further improve the prediction accuracy if the final predicted samples can be generated using the intra prediction as well, which generates the predicted samples by fetching the reconstructed samples adjacent to a coding block in the same picture. Furthermore, if the application of intra prediction to GPM is realized, GPM could be applied to for the lower-latency coding structure, called low delay P (LP) configuration, since the inter prediction with two different motion vectors ("bi-prediction" hereinafter) is prohibited in LP configuration. However, efficient methods to apply the intra prediction to GPM for the improvement of GPM have not been proposed.

Geometric partitioning of the coding blocks before the prediction stage ("GEO" hereinafter) [7]–[9] is a potential solution, while GEO significantly increase encoder and decoder complexities, such as the need for the additional transforms with adaptive shapes. Due to the high level of the complexities, GEO is not adopted in VVC. Combined inter-intra prediction (CIIP) with triangular partitions (CIIP-TP) [10] is another potential solution. Here, CIIP [5] is another new inter prediction in VVC, which generates the final predicted samples by blending inter and intra predicted samples without the additional partitioning at the prediction stage such as GPM. CIIP-TP further extends CIIP such that rectangular coding blocks are diagonally split and the predicted samples in the two split regions are generated using

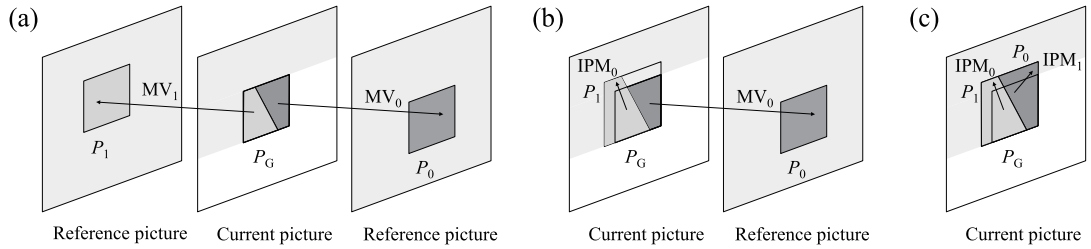


Fig. 1 Overview of the generation process for predicted samples by GPM. (a) GPM-Inter/Inter in VVC, (b) GPM-Inter/Intra, and (c) GPM-Intra/Intra. The shaded region of the current picture and the reference picture indicates the reconstructed sample areas available for inter and intra predictions.

inter or intra prediction, and finally combined. On the other hand, CIIP-TP is not adopted in VVC due to its small performance improvement over VVC reference software, since CIIP-TP is restricted to two types of splitting shapes and only one intra-prediction mode (i.e., Planar mode).

To tackle the problem, we propose GPM with inter and intra prediction methods as shown in Fig. 1 (b) ("GPM-Inter/Intra" hereinafter), which generates P_G by blending the inter and intra predicted samples with one motion vector (MV) and one intra prediction mode (IPM). To maximize the coding performance of the proposed method, we restrict applicable IPMs to only Planar, DC, and/or angular modes in the proposed method. These modes have simpler algorithms than those of the VVC newly adopted IPMs, while the prediction accuracy improvement can be expected by utilizing them appropriately according to the GPM block boundary shapes. Furthermore, a variation of GPM could be GPM with two different intra predictions as shown in Fig. 1 (c) ("GPM-Intra/Intra" hereinafter), while we prohibit GPM-Intra/Intra to save its signaling overhead. The application of GPM-Intra/Intra can be estimated lower than the other GPM since the inter prediction is sufficient to predict the background areas with no motion or the foreground areas including large flat regions.

The experimental results show the proposed method improves the coding performance gain by the GPM-Inter/Inter by 1.3 times, and provides an additional coding performance gain with 1% bitrate savings for the LP configurations. In addition, a subjective improvement by the proposed GPM can be newly observed in the LP configuration for test sequences where GPM are highly applied.

The rest of this paper is organized as follows. Related work and the corresponding problems are explained in Sect. 2. Section 3 presents the details of the proposed method. Section 4 describes the experimental results and discussion. Finally, we conclude the paper in Sect. 5.

2. Related Work

2.1 Block Partitioning in VVC

The maximum coding block size of VVC, i.e., coding tree block size, is extended from 64 × 64 samples of HEVC to 128 × 128 samples [11]. Furthermore, a recursive quad-tree plus binary-ternary tree (QTBTT) block partitioning is introduced as shown in Fig. 2 [11]. The recursive QTBTT block partitioning generates various sizes and shapes of rectangular blocks including non-square blocks, making uniform the sizes of the coding blocks, prediction blocks, and transform blocks. It improves the coding performance while increasing the encoder and decoder complexity compared to those of HEVC [11].

In addition, dual tree, which allows the coding tree block of the luma and chroma components to separate coding tree structure, is newly introduced in intra slice (i.e., I slice) where only the intra prediction can be utilized, while not in inter slice (i.e., B or P slice) where the inter and intra prediction can be utilized [11]. The dual tree improves the coding performance because the block partitioning size of the chroma components can be enlarged when the change of the sample values of the chroma components are relatively small compared to the luma component, for instance.

2.2 Intra Prediction in VVC

For the intra prediction for a luma component, Planar, DC, and angular modes similar to HEVC can be utilized in VVC. The Planar and DC modes are effective for predicting the picture with flat characteristics, whereas the angular mode is effective for predicting the object edges. In particular, the number of angular modes is increased from 32 to 64 as shown in the solid arrows of Fig. 3 to improve the prediction accuracy for a coding block [12]. Moreover, wide angular modes which replace partial angular modes described as dashed arrows (= No. -1–No. -14 and No. 67–No. 80) in Fig. 3 are also newly adopted to improve the prediction accuracy for non-square blocks [12]. This means that the total number of modes is the same as the regular angular modes. Whether the wide angular modes is applied or not is determined by the block sizes and shapes. For the intra prediction for chroma components, the direct mode, which applies the same IPM as the luma component to the chroma components, in order to save the signaling overhead for the chroma component, is available in VVC the same as HEVC.

In VVC, the IPM derivation is designed based on an IPM candidate list, which is similar to HEVC. In fact, the candidate list size (the maximum number of candidates) is extended from three of HEVC to six in VVC. Planar mode

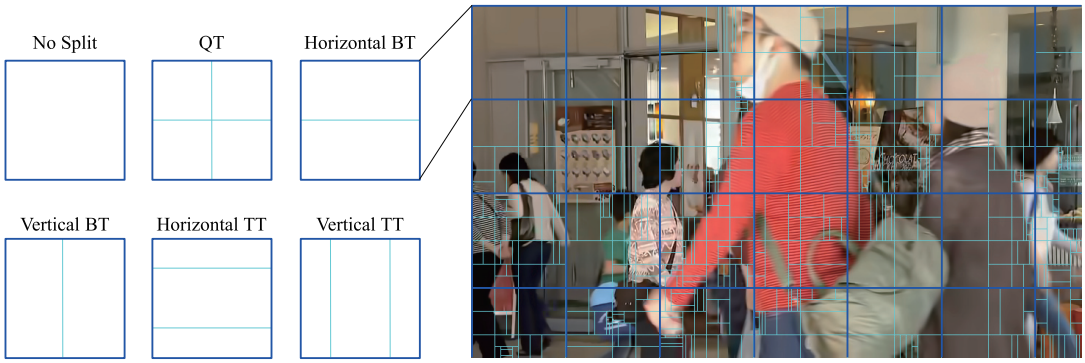


Fig. 2 Quad-tree (QT), binary-tree (BT), and ternary-tree (TT) block partitioning in VVC and an example of the recursive QTBTT block partitioning. Blue grids denote the coding tree blocks and sky-blue lines indicate the QT, BT, or TT splitting lines.

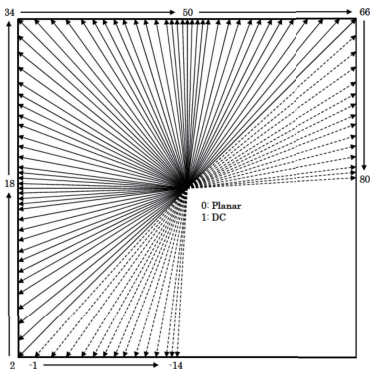


Fig. 3 Intra prediction modes in VVC.

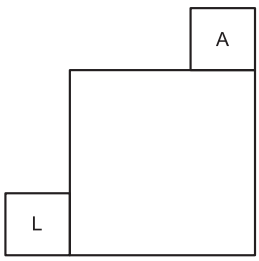


Fig. 4 Available neighboring blocks for derivation of IPM candidates in VVC. A and L denote the neighboring blocks located in the above and left sides of a coding block.

is registered preferentially, and the remaining candidates are determined depending on the IPMs of the two neighboring blocks, A and L, as shown in Fig. 3. Specifically, when these two modes are Planar or DC, the remaining candidates are registered as DC, angular No. 50 (i.e., Vertical mode), angular No. 18 (i.e., Horizontal mode), angular No. 46, and angular No. 54. Otherwise, A and L angular modes and also the three angular modes close to the direction of the A and L angular modes are registered. Moreover, the list already includes the target IPM candidate, it is not registered to reduce the signaling overhead for the duplicated IPM candidate. This is called the pruning process in VVC. After the list construction, the actual IPM can be uniquely identified by the index signaled from the encoder.

2.3 Inter Prediction in VVC

In VVC, the inter prediction is basically organized by an adaptive motion vector prediction (AMVP) mode and a merge mode, the same as HEVC [5]. AMVP mode derives the MV by adding the MV prediction from the decoded region and the MV difference signaled from the encoder. On the other hand, merge mode derives directly the MV from the merge candidate list including a maximum of six sets of merge candidates (i.e., six sets of MV_0 and MV_1). These merge candidates are registered from the spatially and/or temporally neighboring inter prediction blocks. Because of these different MV derivation processes, AMVP and merge modes are effective for predicting the picture with non-uniform and uniform motions, respectively.

The maximum number of MVs applicable to a block is two, the same as HEVC. The minimum block sizes are defined as 4×8 or 8×4 for the inter prediction with only one motion vector ("uni-prediction", hereinafter), and 8×8 samples for the bi-prediction, respectively, due to the worst-case memory bandwidth requirement.

Various new technologies have been developed to improve the prediction accuracy of the AMVP and merge modes in VVC [5]. In this paper, we focus on the GPM among them since it has potentiality for enhanced compression beyond VVC.

2.4 Geometric Partitioning Mode in VVC

As described in Sect. 1, P_G is derived by blending P_0 and P_1 with the integer blending matrices, W_0 and W_1, which contain weights in the value range of [0, 8]. The blending process is expressed by the following formula,

$$P_G = (W_0 \circ P_0 + W_1 \circ P_1 + 4) \gg 3 \quad (1)$$

with

$$W_0 + W_1 = 8\, J_{w,h}, \quad (2)$$

where "$\circ$" in Eq. (2) represents the Hadamard product and $J_{w,h}$ is an all-ones matrix with the coding block size, $w \times h$.

The weights of W_0 and W_1 depend on the displacement between the sample to be predicted and the GPM block boundary, $d(x_c, y_c)$, where x_c and y_c denote the individual sample position within a coding block. In fact, the one of the W_0 and W_1 is given by a ramp function $\gamma_{x_c,\, y_c}$ as,

$$\gamma_{x_c,y_c} = \begin{cases} 0 & d(x_c, y_c) \le -\tau, \\ \frac{8}{2\tau}(d(x_c, y_c) + \tau) & -\tau < d(x_c, y_c) < \tau, \\ 8 & d(x_c, y_c) \ge \tau, \end{cases} \tag{3}$$

and the other blending matrix is derived from Eq. (2). Here, τ indicates the width of the soft blending area which contains non-maximum or non-minimum weights within the matrices, and two samples are selected for τ of GPM-Inter/Inter in VVC. Examples for W_0 and W_1 are shown in Fig. 5. For the chroma components, the same matrices as that for luma are utilized, and the matrices are downsampled when the color format is 4:2:0, which has chroma planes with half the width and height of the luma plane.

Here, the shape of the GPM block boundary required for $d(x_c, y_c)$ is defined by the Hessian normal form with the combination of the angle φ and Euclidean distance ρ between the GPM block boundary and the center position of the coding block as shown in Fig. 6 (a). The angle φ is quantized into 20 discrete angles φ_i shown in Fig. 6 (b), with the range of $[0, 2\pi)$ symmetrically divided. The φ_i is desigined with fixed $\tan(\varphi_i)$ values corresponding to the aspect ratio of the coding block, i.e., $\{0, \pm 1/4, \pm 1/2, \pm 1, \pm 2, \pm\infty\}$ and can be uniquely identified by the angleIdx signaled from the encoder. Similarly, the ρ is quantized into 4 discrete distance ρ_j shown in Fig. 6 (c). Starting from ρ_0, which passes through the center of the block, the distance between ρ_0 and ρ_3 is calculated at equal intervals based on the height h, width w and angle φ of the block. In VVC, $d(x_c, y_c)$ is rounded into the integer precision sample position to avoid the additional interpolation in the GPM prediction process when calculating W_0 and W_1. The shapes of the GPM block boundary are restricted to a total of 64 combinations of 20 φ_i and 4 ρ_j, minus 16 redundant offsets. Here, the 16 redundant offsets are the 10 angles that overlap due to the 180-degree rotation and the 6 overlapping split lines due to the BT and TT splitting. An index to specify the shape of the GPM block boundary among the 64 candidates, *gpm_partition_idx*, is defined in VVC [2].

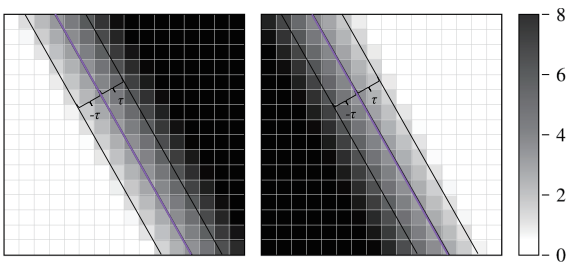


Fig. 5 An example of the GPM blending matrix for each GPM-separated region based on Fig. 5 of Gao et al., 2021 [4]. The purple lines indicate the GPM block boundary.

The MV_0 and MV_1 for the two GPM-separated regions are derived by GPM specific merge indices, *gpm_merge_idx*0 and *gpm_merge_idx*1, and the same merge candidate list as the regular merge mode [2]. *gpm_merge_idx*0 and *gpm_merge_idx*1 indicate the different merge candidates within the list. Hence, assuming the maximum merge candidates for GPM are six, the variation of *gpm_merge_idx*0 and *gpm_merge_idx*1 becomes 30 (= *gpm_merge_idx*0×*gpm_merge_idx*1 = 6×(6−1)) at most. It means that the encoder needs to select the best combination of the GPM block boundary shape and the MVs among the 1920 (= $64 \times 6 \times (6-1)$) candidates.

To reduce the encoder complexity, the encoder of the VVC reference software, VTM version 11 [13], has an early termination method based on full rate-distortion optimization [14] such as the following three-step cost comparisons; In the first step, the sum of absolute difference (SAD) costs for all 1920 candidates are compared and 60 candidates with smaller SAD costs are selected. In the second step, the sum of absolute transformed difference (SATD) costs for the 60 candidates are compared and 8 candidates with smaller SATD costs are selected. In the last step, the rate distortion (RD) costs for the 8 candidates are compared and the best combination with the smallest RD cost is determined.

Furthermore, to refrain the decoder complexity, the range of GPM applicable block sizes is restricted from 8×8 to 64×64 (i.e., $8 \le h \le 64$ and $8 \le w \le 64$) in VVC. In addition, GPM is prohibited for blocks with an aspect ratio

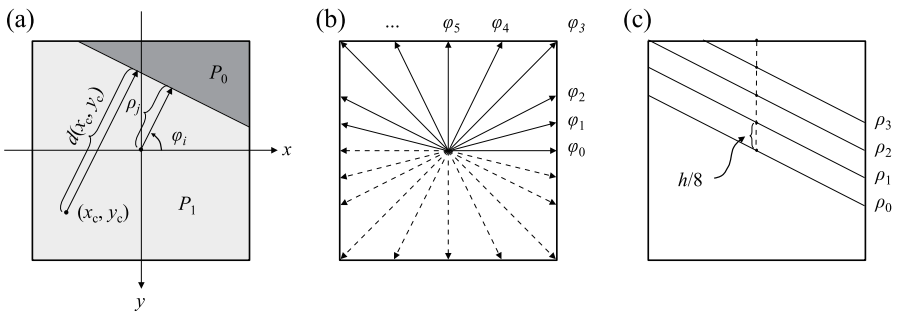


Fig. 6 (a) Example of a Hessian normal form-based GPM block boundary; (b) quantized angle parameters φ; (c) quantized distance parameters ρ.

of 1 : 8 / 8 : 1 among the applicable block size range (i.e., 8×64 / 8×64) because it was confirmed that the improvements by GPM for these block sizes is relatively small in the VVC standardization process [15].

With the feature and algorithm, GPM-Inter/Inter can generate the highly accurate predicted samples around the block boundaries between foregrounds and backgrounds with different motions. Especially, GPM contributes to increasing the coding performance of the coding structure without significant increments of the encoder complexity for the LB configuration as described in Sect. 1

However, GPM has room to further improve the prediction accuracy if the final predicted samples can be generated using intra prediction as well. Moreover, GPM-Inter/Inter cannot be utilized for the LP configuration where bi-prediction is prohibited.

3. Proposed Method

3.1 GPM with Inter and Intra Prediction

We propose a GPM with inter and intra prediction methods to improve the coding performance of GPM. Furthermore, we also propose the following restriction of IPMs to maximize coding performance of the proposed method. In this paper, the detail of the proposed methods [16], [17] is described as follows.

The increment of the IPMs number improves the intra prediction accuracy while increasing the signaling overhead and encoding time to determine the best IPM candidate. Hence, we propose to restrict the number of applicable IPM candidates to four at most; parallel angular mode to the GPM boundary as shown in Fig. 7 (a) ("Parallel mode" hereinafter), perpendicular angular mode to the GPM boundary as shown in Fig. 7 (b) ("Perpendicular mode" hereinafter), Planar mode, and IPM of the neighboring blocks ("Neighbor mode" hereinafter). The decoder can derive the actual IPM from the IPM candidate list similar to the merge candidate list and the index for the IPM of each GPM-separated region signaled from the encoder.

The IPM number shown in Fig. 3 corresponding to the Parallel and Perpendicular modes can be identified with the index for specifying the shape of the GPM block boundary. This is because all the angles of the GPM block boundary are covered by the angles of the angular modes in VVC. When the angle of the GPM block boundary is above right or bottom left in the 45-degree direction, Parallel mode has two IPM candidates, i.e., No. 2 or No. 66, but No. 66 is always utilized in this proposed method.

For Neighbor mode, a maximum of two candidate can be derived from up to five neighboring blocks as shown in Fig. 8. A new Neighbor mode is registered in the IPM candidate list only when the neighboring block is intra predicted and the IPM is not yet registered in the list. That is, the pruning process similar to the merge candidate list is also introduced in the IPM candidate list to avoid IPM duplication. In order to register more effective Neighbor modes, the maximum number of neighbor blocks that can be referenced is increased from 2 in HEVC to 5 in Fig. 8 while restricting the available neighboring blocks according to the shape of the GPM block boundary. The restriction method is represented as Table 1.

The registration order of IPM candidates is designed such that IPM candidates expected to be more effective against GPM are registered earlier. First, the Parallel mode, registered first in the proposed method, can be estimated to be effective when the GPM block boundary is extended over

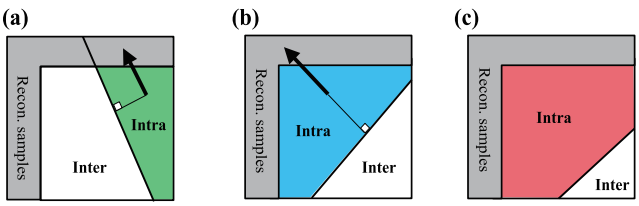


Fig. 7 Examples of the GPM-Inter/Intra block applied by the proposed IPM candidates. (a) Parallel mode, (b) Perpendicular mode, and (c) Planar mode. Gray shaded regions indicate the reconstructed sample areas.

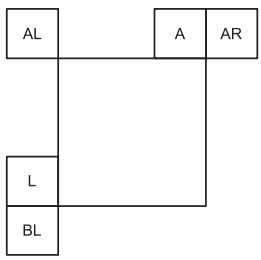


Fig. 8 Available neighboring blocks for the Neighbor mode in the proposed method. AL, A, AR, L, and BL indicate the positions of the neighboring block; above left, above, above right, left, and bottom left, respectively.

Table 1 The restriction method of available neighboring blocks to derive the Neighbor mode for each GPM-separated region. GPM angleIdx corresponds to the angleIdx of GPM in VVC. A, L, and AL indicate the positions of the applicable neighbor blocks; A includes AL, A, and AR of Fig. 8; L includes AL, L, and BL of Fig. 8; L+A includes all the positions of Fig. 8

GPM angleIdx	0	2	3	4	5	8	11	12	13	14
P_0	A	A	A	A	L+A	L+A	L+A	L+A	A	A
P_1	L+A	L+A	L+A	L	L	L	L	L+A	L+A	L+A
GPM angleIdx	16	18	19	20	21	24	27	28	29	30
P_0	A	A	A	A	L+A	L+A	L+A	L+A	A	A
P_1	L+A	L+A	L+A	L	L	L	L	L+A	L+A	L+A

the reconstructed samples as shown in Fig. 7 (a). Then, the Neighbor mode is registered next to the Parallel mode, and is expected to be as effective as the Parallel mode since the IPM candidate of the neighboring blocks considered for the GPM block boundary can be applied. Next, the Perpendicular mode, registered next to the Neighbor mode, is considered to be effective when the total distances between the samples in the GPM-separated regions and the reconstructed samples are closer than that of the Parallel mode as shown in Fig. 7 (b). After that, the Planar mode, which is registered next to the Perpendicular mode, is assumed to be effective when the GPM-separated region size is large as shown in Fig. 7 (c). Finally, the DC mode, which is usually applied preferentially in the regular intra prediction as well as Planar mode, can be registered as the final candidate in the case that the list is not filled due to the absence of the Neighbor mode that can be caused by the available check or pruning process for Neighbor mode.

The same blending matrices as GPM-Inter/Inter are utilized for generating the final inter predicted samples in the proposed method (i.e., both GPM-Inter/Intra and GPM-Intra/Intra). In other words, the derivation method for the shape of the GPM block boundary is the same as GPM-Inter/Inter.

3.2 Prohibition of GPM-Intra/Intra

We also propose a prohibition of GPM-Intra/Intra to maximize the coding performance of the proposed method. The prohibition of GPM-Intra/Intra can be realized by the flag, $gpm\_intra\_enabled\_flag$, signaled from the encoder for specifying whether inter or intra prediction is applied for each GPM-separated region. Specifically, the signaling is designed such as when the intra prediction is applied to one region, the intra prediction cannot be applied to the other region.

The reason that the prohibition of GPM-Intra/Intra improves the coding performance can be explained as follows. First, GPM is easily applied to the boundary the between foreground and background with different motions as described in Sect. 1. This means that inter prediction is sufficient to predict the background areas with no motion or the foreground areas including large flat regions. Second, the intra prediction accuracy for the bottom right region within the prediction block is lower than that of the above left region. This is because the distance between the sample to be predicted and the reconstructed sample is large. With these features of the GPM and intra prediction, the application rate of GPM-Intra/Intra can be estimated lower than the other GPM. In other words, the prohibition of GPM-Intra/Intra can be expected to have no impact other than saving the signaling overhead and the GPM-Inter/Intra will be effective even for the LP configuration.

The prohibition of GPM-Intra/Intra will not reduce the encoder complexity but will reduce the decoder complexity as follows. Regarding the encoder, this is because all SAD costs to apply each IPM candidate to the two GPM-separated regions in all GPM shapes need to be calculated for GPM-Inter/Intra at the first early termination stage as described in 2.4, even without prohibiting GPM-Intra/Intra. On the other hand, as for the decoder, the prohibition can avoid an increase in the circuit size for the intra prediction, which is a critical problem for hardware-based video decoders.

3.3 The Other Specification of Signaling

In the proposed method, GPM-Inter/Intra is applied only for inter slices. This makes that the dual tree in VVC does not need to be considered. Therefore, the direct mode is always utilized for deriving the chroma component IPMs to further save the signaling overhead.

The block size ranges to which the proposed method can be applied to are not changed from those of GPM-Inter/Inter as described in Sect. 2.4. The reasons are as follows. For small size blocks, the application of intra prediction reduces the worst-case memory bandwidth requirement, making GPM applicable. On the other hand, GPM for small size blocks increases the overhead more than improvements of the prediction accuracy. For larger size blocks, the intra prediction accuracy becomes lower as described in Sect. 3.2. Furthermore, the calculation and storage of the blending mask for larger size blocks become a burden especially for the decoder.

In addition to the original signals in GPM-Inter/Inter, two additional block-level signals, $gpm\_intra\_enabled\_flag$ as described in Sect. 3.2 and gpm_intra_idx, are introduced in the proposed method. gpm_intra_idx is signaled to identify the IPM candidate for GPM only when the number of IPM candidates are larger than one. $gpm\_intra\_enabled\_flag$ is coded with fixed-length code, whereas is coded with a zero-order exponential Golomb the same as that for signaling merge candidates.

4. Experimental Results and Discussion

4.1 Test Conditions

1) Software Settings: The VVC reference software VTM version 11 (VTM-11) [13] was used for the simulation software and the proposed method was implemented in the VTM-11. To evaluate the effect by GPM-Inter/Intra compared to that by GPM-Inter/Inter, the coding performance and the complexity of VTM-11, disabling GPM-Inter/Inter, were evaluated as the baseline (i.e., anchor). In addition, a total of nine different proposed methods as shown in Table 2 were conducted to verify the effects by IPM candidate list sizes, IPM candidate variations, and the prohibition of GPM-Intra/Intra.

2) Encoder Configurations: The coding conditions were followed with the VTM Common Test Condition (CTC) [18]. Random access (RA), low delay B (LB), and low delay P (LP) configurations defined in the VTM CTC were used since GPM-Inter/Inter and the proposed methods

Table 2 Details of the proposed methods categorized by IPM candidate list size, registrable IPM candidates and their registering order, and prohibition of GPM-Intra/Intra. The arrow in the third column indicates the registering order of registrable IPM candidates in the list. The bracketed Planar and DC modes in Prop. 5 and Prop. 6 are registrable when Neighbor modes is not registered. "On" and "Off" within the fourth column denote the existence and absence of the prohibition of GPM-Intra/Intra in the proposed method.

Method	IPM candidate list size	Registrable IPM candidates and their registering order	Prohibition of GPM-Intra/Intra
Prop. 1	1	Parallel	On (Default)
Prop. 2 (Phbt.Off)	2	Parallel → Planar	Off
Prop. 2	2	Parallel → Planar	On
Prop. 3 (Phbt.Off)	2	Parallel → Perpendicular	Off
Prop. 3	2	Parallel → Perpendicular	On
Prop. 4 (Phbt.Off)	3	Parallel → Perpendicular → Planar	Off
Prop. 4	3	Parallel → Perpendicular → Planar	On
Prop. 5	3	Parallel → Neighbor → Perpendicular (→ Planar)	On
Prop. 6	4	Parallel → Neighbor → Perpendicular → Planar (→ DC)	On

Table 3 Details of the VTM CTC test sequences from class A to F categorized by resolutions, frame rates, and video content.

Class	Resolutions [pixels × lines]	Frame rates [fps]	Video content
A1	3840 × 2160	30–60	Camera-captured content (Natural scene)
A2	3840 × 2160	50–60	Camera-captured content (Natural scene)
B	1920 × 1080	50–60	Camera-captured content (Natural scene)
C	832 × 480	30–60	Camera-captured content (Natural scene)
D	416 × 240	30–60	Camera-captured content (Natural scene)
E	1280 × 720	60	Camera-captured content (Conversation scene)
F	832 × 480–1920 × 1080	20–60	Pure screen content (SCC) and mixed SCC and camera-captured content

Table 4 Overall performance of each method including the existing GPM-Inter/Inter compared to the anchor in RA, LB, and LP configurations, which is evaluated by BDY, BDY, BDV, EncT, and DecT.

Method	RA [%]					LB [%]					LP [%]				
	BDY	BDU	BDV	EncT	DecT	BDY	BDU	BDV	EncT	DecT	BDY	BDU	BDV	EncT	DecT
GPM-Inter/Inter	−0.74	−1.08	−1.18	103	99	−1.54	−1.86	−1.80	105	98	0.00	0.00	0.00	100	100
Prop. 1	−0.88	−1.29	−1.38	104	100	−1.81	−2.58	−2.55	106	99	−0.78	−1.34	−1.24	107	101
Prop. 2 (Phbt.Off)	−0.89	−1.48	−1.47	104	100	−1.89	−2.75	−2.74	106	99	−0.87	−1.64	−1.54	109	102
Prop. 2	−0.94	−1.42	−1.49	104	100	−1.94	−2.86	−2.88	106	99	−0.95	−1.63	−1.54	108	101
Prop. 3 (Phbt.Off)	−0.89	−1.52	−1.49	104	100	−1.90	−2.70	−2.82	107	99	−0.91	−1.84	−1.76	108	101
Prop. 3	−0.93	−1.47	−1.48	104	100	−1.91	−2.62	−2.72	106	99	−0.95	−1.84	−1.59	108	101
Prop. 4 (Phbt.Off)	−0.91	−1.58	−1.54	105	100	−1.97	−2.84	−2.77	107	99	−1.01	−1.84	−1.81	109	102
Prop. 4	−0.96	−1.56	−1.55	105	100	−1.96	−2.94	−2.76	107	99	−1.09	−1.92	−1.71	110	102
Prop. 5	−0.97	−1.58	−1.62	107	100	−1.99	−3.16	−2.81	108	99	−1.15	−2.10	−1.96	111	102
Prop. 6	−0.96	−1.66	−1.67	107	100	−2.01	−3.10	−2.70	110	99	−1.12	−2.19	−1.91	113	101

can be applied only to B and P slices. RA is often used for general video transmission while LB and LP are utilized for low-latency video transmission. Only the performance and complexity of the proposed methods in LP configuration were evaluated since GPM-Inter/Inter is disabled by default in the LP configuration.

3) Test Sequences: The test sequences from classes A to F, as listed in VTM CTC, were used. They are categorized with different resolutions, frame rates, and video content as shown in Table 3. For each test sequence, four quantization parameter (QP) values 22, 27, 32, and 37 defined in the VTM CTC were used to generate the different rate points.

4) Evaluation Metrics: The coding performance was evaluated by the BD-rate of the luma (BDY) and two chroma (BDU, BDV) components [19], [20]. The BD-rate is the evaluation index used to quantify the difference of the generated bitrate for the identical level of peak signal-to-noise ratio (PSNR) between two coding methods. The negative BD-rate values indicate the coding performance gain with bitrate savings. In other words, the positive value is coding performance loss. The complexity was evaluated by the relative encoding time (EncT) and decoding time (DecT) of the two coding methods measured on a homogenous cluster PC. Note that the results of Class D and F are not included in the overall results, which is the average of BD-rate, EncT, and DecT for all test sequences except for these classes, in accordance with the VTM CTC.

4.2 Comparison of Overall Results

The overall results of each method compared to the anchor in RA, LB, and LP configurations, which is evaluated by BDY, BDU, BDV, EncT, and DecT, are described in Table 4.

First, regarding the coding performance, the GPM-Inter/Inter and all the proposed methods provide the coding performance gain against the anchor in RA, LB, and LP as observed in Table 4. The values of LB is the largest among the three coding conditions. The proposed meth-

ods give additional coding performance gains compared to GPM-Inter/Inter. From the comparison of GPM-Inter/Inter and Prop. 5 or Prop. 6, the coding performance gains of RA and LB become 1.3 times of that of the GPM-Inter/Inter at most, and the additional coding performance gain with a maximum of 1% BD-rate savings is newly produced in LP. They are derived from an increase of GPM applied samples, which will be clarified in Sect. 4.4.

Second, regarding the complexity, the GPM-Inter/Inter and all the proposed methods increase EncT compared to the anchor. Similarly, the proposed methods provide additional EncT increments against GPM-Inter/Inter. The increment of the proposed method in LP is significant since the entire functionality of GPM is introduced into the anchor in LP, whereas only the functionality regarding the intra prediction of the GPM-Inter/Intra is added in RA and LB. On the other hand, there is almost no DecT increment of GPM-Inter/Inter and all the proposed methods over the anchor. This is because the decoder can specify the GPM block boundary shape, motion vectors, and IPMs signaled from the encoder without any burden.

Third, the coding performance and complexity of the proposed methods compared by different IPM candidates, corresponding to the results of Prop. 1–6 of Table 4, are described as follows.

Regarding the performance, more coding performance gains are achieved as the number of IPM candidates (i.e., IPM candidate list sizes) is increased. The comparison of Prop. 1–3 proves the addition of Planar or Perpendicular modes to Parallel mode yields additional coding performance gain. The further coding performance gain of Prop. 4 against Prop. 2 and 3 indicates an additive effect of Planar and Perpendicular modes. Compared with Prop. 4 and 5, more coding performance gain by introducing Neighbor mode can be confirmed. The same level of BD-rate between Prop. 5 and 6, which have different IPM candidate list sizes, implies the coding performance gain is saturated around three IPM candidates. This is because the PSNR improvements and bitrate increments by extending IPM candidate list sizes from 3 to 4 are counterbalanced. The above observation will be clarified with a rate distortion curve and an analysis of GPM intra predicted samples described in Sect. 4.4.

Regarding complexity, EncT is similarly increased as the increment of IPM candidates. This is because it raises the number of rate distortion optimization (RDO) processing of the encoder side to derive the minimum RD cost corresponding to the best combination of the GPM block boundary shape, motion vectors, and IPMs. In contrast, DecT is not increased since the decoder can identify the best combination by their indices signaled from the encoder.

Fourth, the coding performance and complexity of the proposed methods without and wit the prohibition of GPM-Intra/Intra, i.e., Prop. 2–4 (Phbt.Off) vs. Prop. 2–4 in Table 4, are described as follows. Regarding the coding performance, the proposed method with the prohibition gives further coding performance gains compared to that without the prohibition, which suggests the bitrate reduction by the prohibition contributes to the gains. Regarding EncT, there is no difference between the proposed methods with and without the prohibition since only the addition processing of the two SAD costs for GPM-Intra/Intra (i.e., two different IPMs) is different, as described in Sect. 3.1. DecT is also not different because the reduction of the GPM-Intra/Intra process is minor due to the lower application rate of the GPM-Intra/Intra, which will be clarified in Sect. 4.4.

4.3 Comparison of Sequence-Level Results

In this section, the coding performance and complexity of GPM-Inter/Inter and the proposed method are compared by sequence level with different resolutions and content. The sequence-level results of GPM-Inter/Inter and Prop. 5 to the anchor in RA, LB, and LP, which is evaluated by BDY, EncT, and DecT, are described in Table 5.

Regarding the coding performance, no tendency on the resolutions can be confirmed from the comparison of classes A1/A2, B, C, and D results. In contrast, the following tendency can be found with the content in both GPM-Inter/Inter and Prop. 5, which is common to the RA, LB, and LP. Specifically, the coding performance gains are relatively large for the sequence with differently moving foregrounds and backgrounds across classes, e.g., RitualDance, BQMall, and RaceHorses. GPM is originally easy to apply in these sequences so that Prop. 5 provides additional coding performance gains against GPM-Inter/Inter. On the other hand, the coding performance gains of the sequence without foregrounds and backgrounds, e.g., BQTerrace and BQSquare, are relatively small. GPM is seldom applied in these sequences so that the bitrate increments for GPM-Inter/Intra raise the BD-rate of Prop. 5 compared to that of GPM-Inter/Inter. In the pure SCC sequences, i.e., SlideEditing and SlideShow, very minor coding performance gains or coding performance losses are observed. This is because these sequences often include moving objects with sharp edges and the gradated blending matrices described in Sect. 2.1 overly smooth the sharp edges, and rather raise residuals.

In association with these characteristics, the EncT increments of the test sequences with minor coding performance gains or losses are smaller than those with larger coding performance gains since GPM are terminated early in the RDO processing of the encoder. On the other hand, DecT of the sequences with larger coding performance gains are smaller than those with minor coding performance gain or losses because GPM increases the number of large size blocks and reduces the decoding processing on QTBTT partitioning compared to the anchor.

4.4 Picture-Level Analysis

As a picture-level analysis, the rate-distortion (RD) curves of the anchor and all the proposed methods for BQMall and BQSquare in the LP configuration are shown in Fig. 9. To investigate the effect of the proposed method, we se-

Table 5 Sequence-level performance of GPM-Inter/Inter and Prop. 5 compared to the anchor (VTM-11 without GPM) in RA, LB, and LP configurations, which is evaluated by BDY, EncT, and DecT. Note that the performance of GPM-Inter/Inter in LP is not written since GPM-Inter/Inter is disabled in the LP configuration.

Sequense	GPM-Inter/Inter						Prop. 5								
	RA [%]			LB [%]			RA [%]			LB [%]			LP [%]		
	BDY	EncT	DecT	BDY	EncT	DecT	BDY	EncT	DecT	BDY	EncT	DecT	BDY	EncT	DecT
Tango2	−0.64	103	99	–	–	–	−0.75	108	100	–	–	–	–	–	–
FoodMarket4	−0.40	103	100	–	–	–	−0.54	108	101	–	–	–	–	–	–
Campfire	−0.20	103	99	–	–	–	−0.40	110	99	–	–	–	–	–	–
Average Class A1	**−0.41**	**103**	**99**	–	–	–	**−0.56**	**107**	**100**	–	–	–	–	–	–
CatRoad	−0.61	102	100	–	–	–	−0.80	107	99	–	–	–	–	–	–
DaylightRoad2	−0.46	102	100	–	–	–	−0.44	107	100	–	–	–	–	–	–
ParkRunning3	−0.56	104	100	–	–	–	−0.69	109	100	–	–	–	–	–	–
Average Class A2	**−0.55**	**103**	**100**	–	–	–	**−0.64**	**106**	**100**	–	–	–	–	–	–
MarketPlace	−0.44	103	100	−0.96	104	99	−0.60	108	100	−1.19	107	100	−0.67	110	101
RitualDance	−0.57	103	100	−1.04	105	98	−0.83	108	100	−1.87	110	99	−1.36	113	102
Cactus	−0.75	103	99	−1.40	105	95	−1.12	107	100	−2.00	108	96	−1.14	111	102
BasketballDrive	−0.32	103	100	−0.80	105	99	−0.49	108	101	−1.31	110	99	−0.93	112	101
BQTerrace	−0.38	103	100	−0.43	104	100	−0.41	106	101	−0.54	107	100	−0.49	109	102
Average Class B	**−0.49**	**103**	**100**	**−0.93**	**105**	**98**	**−0.69**	**106**	**100**	**−1.38**	**108**	**99**	**−0.92**	**111**	**102**
BasketballDrill	−1.20	103	98	−2.23	106	98	−1.85	108	101	−2.94	110	99	−1.92	115	102
BQMall	−2.30	102	98	−3.06	106	96	−2.61	107	101	−3.81	110	97	−2.28	111	101
PartyScene	−0.67	104	100	−1.08	107	98	−1.01	108	101	−1.79	111	100	−1.14	111	101
RaceHorses	−1.52	104	98	−1.83	107	98	−2.01	108	100	−2.47	111	99	−1.79	112	100
Average Class C	**−1.42**	**103**	**99**	**−2.05**	**106**	**98**	**−1.87**	**107**	**99**	**−2.75**	**111**	**99**	**−1.78**	**112**	**101**
BasketballPass	−0.79	103	99	−1.61	107	98	−1.20	108	101	−2.05	111	99	−1.29	110	102
BQSquare	−0.08	103	100	−0.96	105	99	−0.19	106	103	−0.87	107	101	−0.07	108	104
BlowingBubbles	−0.68	104	99	−1.78	107	97	−0.84	109	102	−2.01	111	98	−0.75	111	103
RaceHorses	−1.29	104	99	−2.17	106	96	−1.67	108	100	−3.03	110	98	−1.76	109	101
Average Class D	**−0.71**	**104**	**100**	**−1.63**	**106**	**98**	**−0.98**	**107**	**100**	**−1.99**	**110**	**99**	**−0.97**	**109**	**102**
FourPeople	–	–	–	−1.73	103	97	–	–	–	−2.05	107	99	−1.07	110	103
Johnny	–	–	–	−2.05	102	98	–	–	–	−2.01	104	100	−0.39	109	103
KristenAndSara	–	–	–	−1.92	102	97	–	–	–	−1.86	105	99	−0.66	108	102
Average Class E	–	–	–	**−1.90**	**102**	**97**	–	–	–	**−1.98**	**105**	**99**	**−0.71**	**109**	**102**
BasketballDrillText	−1.21	102	99	−1.88	104	98	−1.82	106	101	−2.66	108	98	−1.76	110	102
ArenaOfValor	−0.90	101	99	−1.56	103	97	−1.14	104	100	−2.13	107	98	−1.08	109	101
SlideEditing	−0.02	100	100	0.21	101	101	−0.04	102	102	−0.14	104	101	0.30	107	102
SlideShow	−0.12	101	101	0.04	103	100	0.16	103	103	0.05	106	101	−0.13	108	102
Average Class F	**−0.56**	**102**	**100**	**−0.80**	**103**	**94**	**−0.71**	**103**	**101**	**−1.22**	**106**	**99**	**−0.67**	**108**	**102**
Overall	**−0.74**	**101**	**99**	**−1.54**	**103**	**99**	**−0.97**	**107**	**100**	**−1.99**	**108**	**99**	**−1.15**	**111**	**102**

lected BQMall and BQSquare which produce the largest and smallest coding performance gains among all the test sequences, respectively.

First, regarding BQMall, the comparison by the same QP value shows that coding performance gains of the proposed method mainly come from the bitrate reduction as shown in Fig. 9 (b) and (c). The reason is that GPM-Inter/Intra improves the prediction accuracy so that reduces residuals around the boundaries of the foreground and background with different motions. Not only a bitrate reduction, but also the PSNR improvement can be observed in the high QP range as shown in Fig. 9 (c). This is because the effect of GPM on the qualities of the reconstructed samples, which determines the size of PSNR, became more apparent in the high QP range. The quantization step of the high QP range is larger than that of the low QP range so that most of the residual coefficients become zero. Therefore, the qualities of the reconstructed samples highly depend on those of the predicted samples. In this sense, the additional PSNR improvement can be observed as the number of the IPMs is increased in the high QP range. Moreover, the counterbalance between Prop. 5 and Prop. 6, and the additional bitrate reduction by the prohibition of GPM-Intra/Intra, i.e., Prop. 2–4 vs. Prop. 2–4 (Phbt.Off) can be confirmed, as described in Sect. 4.3.

On the other hand, regarding BQSquare, the bitrate reductions are very minor in both high and low QP ranges while only small PSNR improvements are seen in Prop. 4 (Phbt.Off) and Prop. 5 in the high QP range as shown in Fig. 9 (e) and (f). The small coding performance gain of Prop. 5 in Table 5 comes from the small PSNR improvements. The same tendencies as BQMall and BQSquare can be observed in the other test sequences with larger and smaller coding performance gains, respectively.

To clarify the evidence of the difference of coding performance gains among RA, LB, and LP and the effect by the prohibition of GPM-Intra/Intra, an analysis of the total GPM applied samples organized by GPM-Inter/Inter, GPM-Inter/Intra, and GPM-Intra/Intra with Prop.4 (Phbt.Off) for BQMall and BQSquare is shown in Fig. 10. They are normalized by the total samples of the encoded test sequences. The difference of coding performance gains among RA, LB,

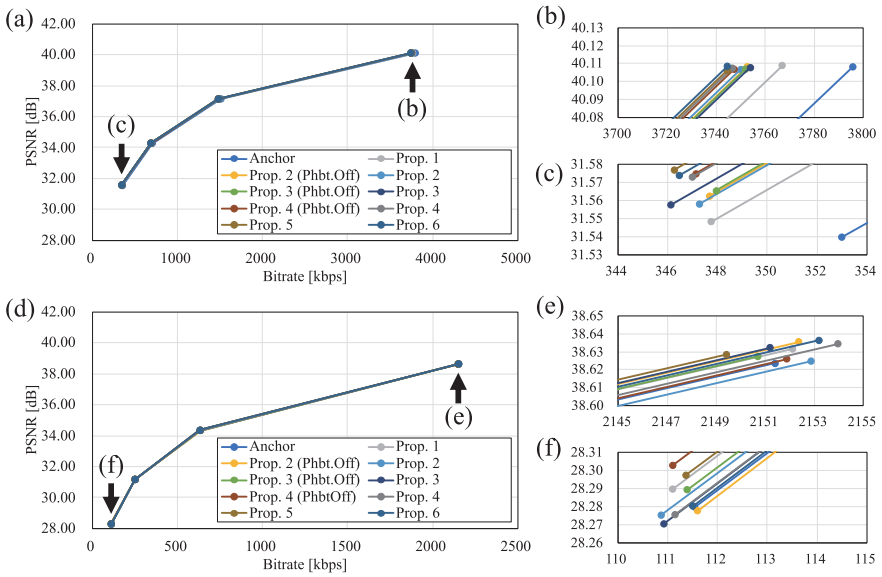


Fig. 9 Rate distortion curves of the anchor and all the proposed methods for BQMall and BQSquare in the LP configuration. (a)–(c) BQMall, (d)–(f) BQSquare.

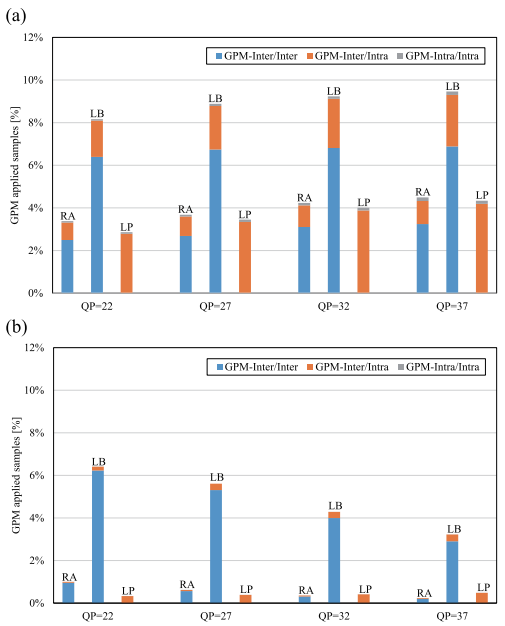


Fig. 10 Analysis of the total GPM applied samples organized by GPM-Inter/Inter, GPM-Inter/Intra, and GPM-Intra/Intra with Prop. 4 (Phbt.Off) in RA, LB, and LP configurations. (a) BQMall, (b) BQSquare.

and LP corresponds to the difference of their GPM applied samples as follows. Figure 10 (a) shows that the total GPM applied samples in the RA and LB configurations become around 1.3 times of that of only the GPM-Inter/Inter samples, which also matches the additional coding performance gain of Prop. 4 (Phbt.Off) as observed in BQMall. In contrast, Fig. 10 (b) shows the total GPM applied samples in the RA and LB configurations have not increased much compared with only the GPM-Inter/Inter samples, which matches the coding performance loss of Prop. 4 (Phbt.Off) against the GPM-Inter/Inter as observed in BQSquare. In these two sequences, the total GPM-applied samples of RA and LP are the same level, which corresponds to the coding performance gains of the Prop. 4 (Phbt.Off) in the RA and LP configurations. The GPM-Intra/Intra samples are significantly smaller than the GPM-Inter/Inter or GPM-Inter/Intra samples. It means that the prohibition of the GPM-Intra/Intra does not affect the PSNR improvement and contributes to the bitrate reduction.

An analysis of the GPM intra predicted samples of Prop. 6 in each GPM applicable block size for two different test sequences and QPs in the LB configuration is shown in Fig. 11. A certain number of the GPM intra predicted samples in Figs. 11 (a) and (b) is observed while not in Figs. 11 (c) and (d), which corresponds to the coding performance gain of BQMall or the coding performance losses of BQSquare of the proposed method against the GPM-Inter/Inter in the LB configuration. The ratios of the GPM-Inter/Intra samples to all GPM applied samples (i.e., GPM-Inter/Intra and GPM-Inter/Inter predicted samples) in Figs. 11 (a), (b), and (d) are decreased as the GPM applicable block size becomes larger. This is because the accuracy of the intra prediction degrades for the large block size where the prediction target samples are far from the reconstructed samples. Compared to the GPM intra predicted samples in Figs. 11 (a) and (b), the majority of the samples shifts from the smaller size blocks to the larger size blocks becauese the number of larger size blocks increases in the high QP range.

In addition, the ratio of GPM-Inter/Intra of each block size in high QP range is larger than that in small QP range from the comparison of Figs. 11 (a) and (b). The reasons of this tendency are as follows. First, the inter prediction accuracy is relatively higher for low QP than for high QP. This is because the coding noise on the reference frame re-

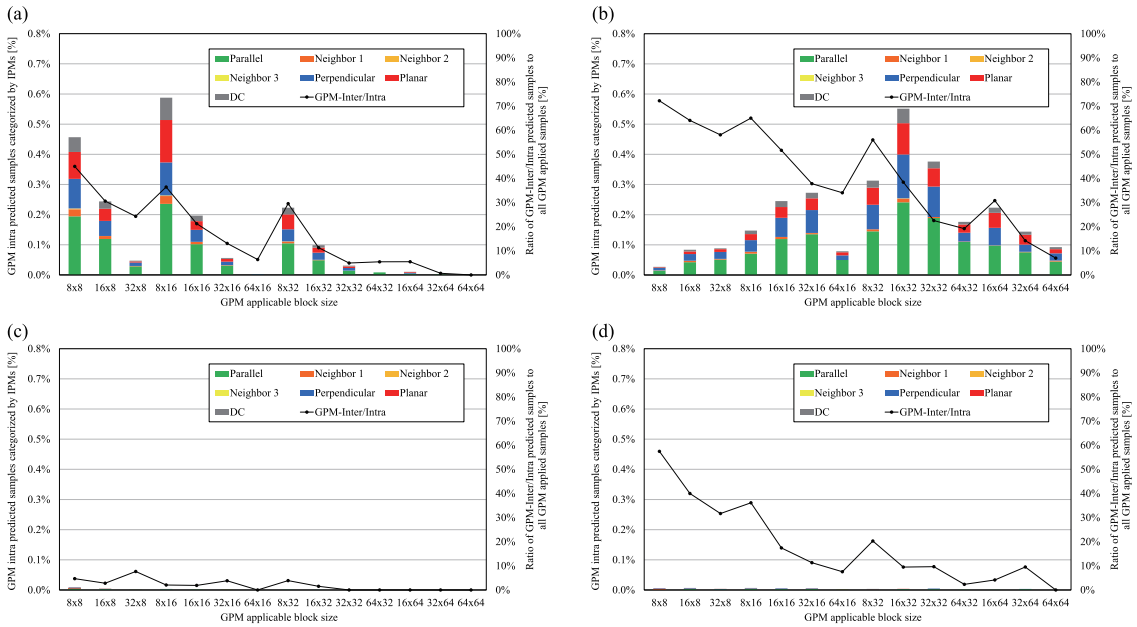


Fig. 11 An analysis of the GPM intra predicted samples by Prop. 6 in each GPM applicable block size for two different test sequences and QPs in the LB configuration; (a) BQMall, QP = 22, (b) BQMall, QP = 37, (c) BQSquare, QP = 22, and (d) BQSquare, QP = 37. The left axis indicates GPM intra predicted samples categorized by IPMs. The right axis denotes the ratio of GPM-Inter/Intra predicted samples to all GPM applied samples (i.e., GPM-Inter/Intra and GPM-Inter/Inter predicted samples). Both values are normalized by the total encoded samples of each test sequence.

quired for inter prediction is smaller for low QP than for high QP. On the other hand, the intra prediction accuracy is not always higher for low QP than for high QP. This is because intra prediction is highly dependent on the similarity of the reference samples themselves, as well as the presence or absence of coding noise on the reference samples adjacent to the current block. These characteristics based on the quantization parameters of inter and intra predictions are also corresponded to GPM. Hence, at high QP, the application rate of GPM-Inter/Inter decreases, while those of GPM-Inter/Intra increases, which causes the increment in the ratio of GPM-Inter/Intra applied samples in Fig. 11.

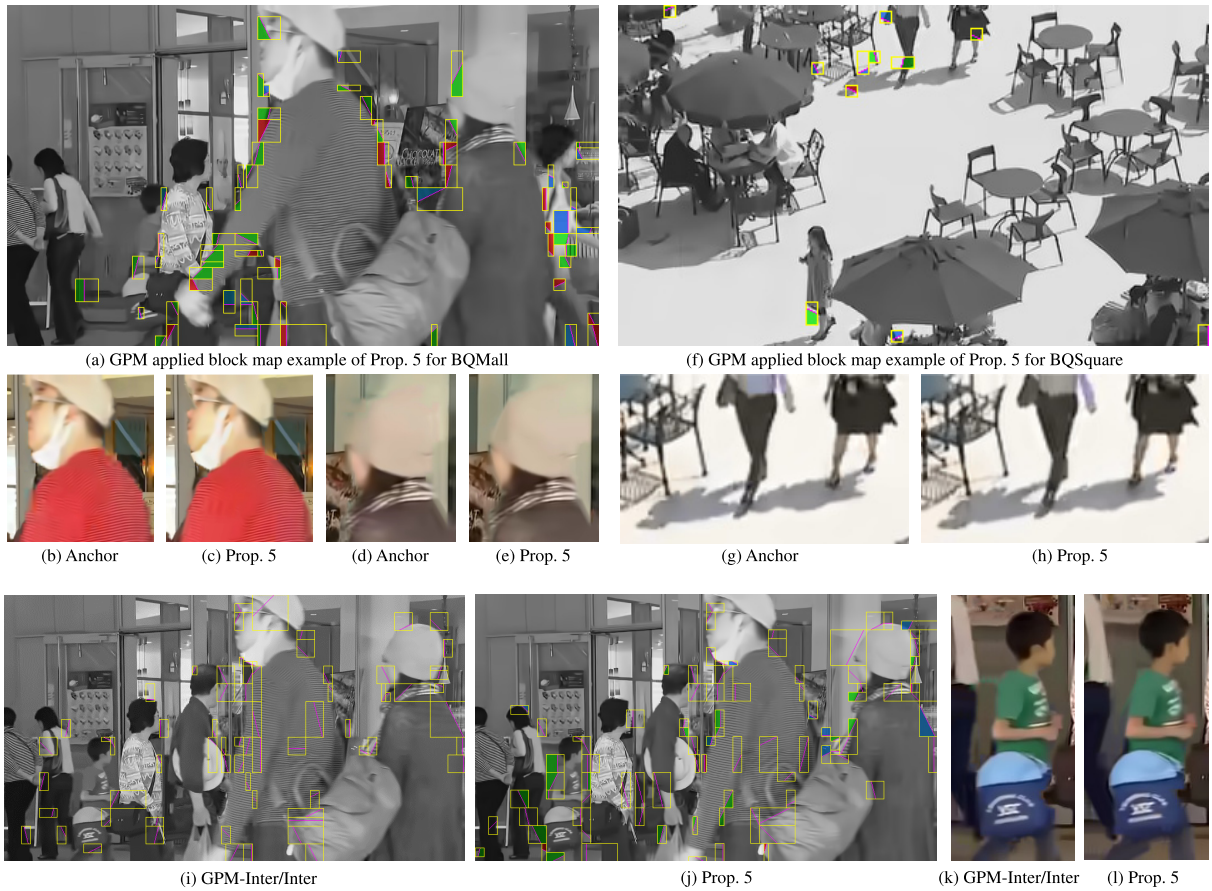
In terms of IPMs, the number of GPM intra predicted samples applied by Parallel, Perpendicular, and Planar modes dominates. Especially, those of the Parallel mode are the largest, which corresponds to our expectation described in Sect. 3.1. In contrast, those of Neighbor and DC mode are smaller. This is because Neighbor mode is only registered in the IPM candidate list only when the neighbor block is intra block, whereas DC mode is registered in the IPM candidate list only when the list is not filled due to the absence of the Neighbor mode. In this sense, the further coding performance gains for the Neighbor mode can be expected if the neighboring block is an inter predicted block but stores the IPM of the reference block, which has being studied for enhanced compression beyond VVC in JVET.

4.5 Subjective Evaluation

In addition to the improvement of the coding performance, subjective quality improvements are provided by the proposed method. First, the significant improvements can be seen in GPM-Intra/Inter applied block areas of the LP configuration for BQMall from the comparison of Fig. 12, which are example of the GPM applied block map of Prop. 5 and decoded image of the anchor and the Prop. 5. For example, the block noises around the man's shoulder in Fig. 12 (b) are removed in Fig. 12 (c). As another example, the blur of the woman's cap in Fig. 12 (e) is sharpened in Fig. 12 (d). These improvements correspond to the PSNR improvements in the high QP range, described in the rate distortion curve of Sect. 4.4. In addition, Parallel, Perpendicular, and Planar modes were applied as expected in areas with the image characteristics described in Sect. 3.1. Note that the application of Neighbor mode was not observed in Fig. 12 (a) due to its lower application rate, but it was confirmed in another frame.

Second, the any of the subjective improvement and degradation by the proposed GPM cannot be seen in BQSquare as shown in Figs. 12 (g) and (h) due to the lower application of the GPM-Inter/Intra blocks, which matches the objective evaluation results and analyses for BQSquare.

Third, an additional improvement was also confirmed in the LB configuration even where the conventional GPM (GPM-Inter/Inter) can be applied from the comparison of the block maps and decoded images of GPM-Inter/Inter and Prop. 5. An example is shown in Figs. 12 (i)–(l). Specifically, the artifacts behind the boy and the blur near the boy's feet seen in Fig. 12 (k) were removed in the GPM-Inter/Intra applied block area in Fig. 12 (l). Other than this example,



(a) GPM applied block map example of Prop. 5 for BQMall
(f) GPM applied block map example of Prop. 5 for BQSquare
(b) Anchor
(c) Prop. 5
(d) Anchor
(e) Prop. 5
(g) Anchor
(h) Prop. 5
(i) GPM-Inter/Inter
(j) Prop. 5
(k) GPM-Inter/Inter
(l) Prop. 5

Fig. 12 Subjective evaluation results of Anchor, GPM-Inter/Inter, Prop. 5 for BQMall and BQSquare in the LP and LB condition at QP = 37. (a), (f), (i) and (j) are GPM applied block map examples. (a) Prop. 5 for the 77th frame of BQMall in LB; (f) Prop. 5 for the 32nd frame of BQSquare in LP; (i) and (j) GPM-Inter/Inter and Prop. 5 for the 68th frame of BQMall in LB. In these block maps, the area surrounded by the yellow grid is the GPM applied area while the other areas are the non-GPM areas. A purple line indicates the GPM block boundary. Gray, green, blue, and red areas within the yellow grid denotes the areas to which inter prediction, Parallel mode, Perpendicular mode, and Planar mode are applied, respectively. (b), (c), (d) (e), (k) and (l) are decoded images corresponding to the block maps for each method and each test condition.

additional improvements have been identified, although not as much as in the LP configuration.

5. Conclusion

In this paper, we proposed a geometric partitioning mode (GPM) with inter and intra prediction to achieve further enhanced compression beyond VVC. To maximize the coding performance of the proposed method, the restrictions of the intra prediction mode number and the prohibition of GPM with two different intra predictions were proposed. The experimental results show that the proposed method improved the coding performance gain by the existing GPM method in VVC, realized by two different inter prediction, by 1.3 times in RA and LB configuration, and provided an additional coding performance gain with 1% bitrate savings in LP configuration. Moreover, subjective evaluation for the existing GPM in VVC and the proposed GPM revealed that the proposed method further improves the quality around the picture boundary of foregrounds and backgrounds with different motions.

Acknowledgments

This work was supported by Ministry of Internal Affairs and Communications (MIC) of Japan (Grant no. JPJ000595).

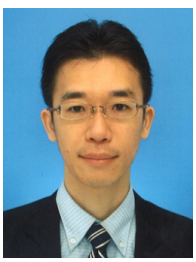
References

[1] Recommendation ITU-T H.265, “High efficiency video coding,” Edition 1.0, April 2013.
[2] Recommendation ITU-T H.266, “Versatile video coding,” Edition 1.0, Aug. 2020.
[3] 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 Transactions on Circuits and Systems for Video Technology, vol.31, no.10, pp.3736–3764, Oct. 2021, doi: 10.1109/TCSVT.2021.3101953.
[4] H. Gao, S. Esenlik, E. Alshina, and E. Steinbach, “Geometric partitioning mode in versatile video coding: algorithm review and analy-

sis," IEEE Transactions on Circuits and Systems for Video Technology, vol.31, no.9, pp.3603–3617, Sept. 2021, doi: 10.1109/TCSVT.2020.3040291.
[5] 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 Transactions on Circuits and Systems for Video Technology, vol.31, no.10, pp.3848–3861, Oct. 2021, doi: 10.1109/TCSVT.2021.3101212.
[6] W.-J. Chien, J. Boyce, Y.-W. Chen, R. Chernyak, K. Choi, R. Hashimoto, Y.-W. Huang, H. Jang, R.-L. Liao, and S. Liu, "JVET AHG report: tool reporting procedure and testing (AHG13)," Joint Video Experts Team (JVET) of ITU-T SG 16 WP 3 and ISO/IEC JTC 1/SC 29, by teleconference, JVET-T0013, Oct. 2020.
[7] M. Bläser, J. Sauer, and M. Wien, "Description of SDR and 360° video coding technology proposal by RWTH Aachen University," Joint Video Experts Team (JVET) of ITU-T SG 16 WP 3 and ISO/IEC JTC 1/SC 29/WG 11, San Diego, JVET-J0023, April 2018.
[8] M. Bläser, J. Schneider, J. Sauer, and M. Wien, "Geometry-based partitioning for predictive video coding with transform adaptation," Proc. Picture Coding Symp. (PCS), pp.134–138, June 2018, doi: 10.1109/PCS.2018.8456238.
[9] B. Bross, K. Andersson, M. Blaser, V. Drugeon, S.-H. Kim, J. Lainema, J. Li, S. Liu, J.-R. Ohm, G.J. Sullivan, and R. Yu, "General video coding technology in responses to the joint call for proposals on video compression with capability beyond HEVC," IEEE Trans. Circuits Syst. Video Technol., vol.30, no.5, pp.1226–1240, May 2020, doi: 10.1109/TCSVT.2019.2949619.
[10] S. Blasi, A. Seixas Dias, and G. Kulupana, "Non-CE4: CIIP using triangular partitions," Joint Video Experts Team (JVET) of ITU-T SG 16 WP 3 and ISO/IEC JTC 1/SC 29/WG 11, Gothenburg, JVET-O0522, July 2019.
[11] 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 Transactions on Circuits and Systems for Video Technology, vol.31, no.10, pp.3818–3833, Oct. 2021, doi: 10.1109/TCSVT.2021.3088134.
[12] J. Pfaff, A. Filippov, S. Liu, X. Zhao, J. Chen, S. De-Luxan-Hernandez, T. Wiegand, V. Rufitskiy, A.K. Ramasubramonian, and G.V. Auwera, "Intra prediction and mode coding in VVC," IEEE Transactions on Circuits and Systems for Video Technology, vol.31, no.10, pp.3834–3847, Oct. 2021, doi: 10.1109/TCSVT.2021.3072430.
[13] J. Chen, Y. Ye, and S. Kim, "Algorithm description for versatile video coding and test model 11 (VTM 11)," in Joint Video Experts Team (JVET) of ITU-T SG 16 WP 3 and ISO/IEC JTC 1/SC 29, by teleconference, JVET-T2002, Oct. 2020.
[14] G.J. Sullivan and T. Wiegand, "Rate-distortion optimization for video compression," IEEE Signal Processing Magazine, vol.15, no.6, pp.74–90, Nov. 1998, doi:10.1109/79.733497.
[15] Z. Deng, L. Zhang, H. Liu, K. Zhang, and Y. Wang, "CE4-related: further constraints on block shapes for GEO," Joint Video Experts Team (JVET) of ITU-T SG 16 WP 3 and ISO/IEC JTC 1/SC 29/WG 11, Brusells, JVET-Q0309, Jan. 2020.
[16] Y. Kidani, H. Kato, and K. Kawamura, "AHG12: GPM with inter and intra prediction," Joint Video Experts Team (JVET) of ITU-T SG 16 WP 3 and ISO/IEC JTC 1/SC 29, by teleconference, JVET-W0110, July 2021.
[17] Y. Kidani, H. Kato, and K. Kawamura, "EE2-related: Modified GPM with inter and intra prediction," Joint Video Experts Team (JVET) of ITU-T SG 16 WP 3 and ISO/IEC JTC 1/SC 29, by teleconference, JVET-X0078, Oct. 2021.
[18] F. Bossen, J. Boyce, X. Li, V. Seregin, and K. Sühring, "VTM common test conditions and software reference configurations for SDR video," Joint Video Experts Team (JVET) of ITU-T SG 16 WP 3 and ISO/IEC JTC 1/SC 29, by teleconference, JVET-T2010, Oct. 2020.
[19] G. Bjøntegaard, "Calculation of average PSNR differences between RD-curves," ITU-T SG 16 WP 3, Texas, VCEG-M33, April 2001.
[20] K. Andersson et al., "Summary information on BD-rate experiment evaluation practices," Joint Video Experts Team (JVET) of ITU-T SG 16 WP 3 and ISO/IEC JTC 1/SC 29/WG 11, Brussels, JVET-Q2016, Jan. 2020.

Yoshitaka Kidani received B.E. and M.E. degrees in Engineering from Nagoya University, in 2010 and 2012. He joined KDDI Corporation in 2012. He is currently a research engineer at KDDI Research, Inc. He has been involved with the development of VVC standards under JVET. His research interests include video coding and multimedia distribution. He is a member of ITE.

Haruhisa Kato received B.E. and M.E. degrees in electronical and electronic engineering from Kobe University in 1997 and 1999, respectively. He received a Ph.D. from Graduate School of Engineering, Tokyo Institute of Technology in 2009. In 1999, he joined KDD Co. Ltd. His research interests include image processing and augmented reality. He is a Senior Expert of the XR Space Creation Laboratory at KDDI Research, Inc.



Kei Kawamura received his B.E., M.Sc., and Ph.D. degrees in Global Information and Telecommunication Studies from Waseda University, Japan, in 2004, 2005, and 2013, respectively. He joined KDDI in 2010. He has been involved with the development of HEVC and VVC standards under JCT-VC and JVET. He is currently engaged in the research and development of a video coding system at KDDI Research, Inc. His research interests include image and video processing, video coding, and multimedia distribution. He is a member of a steering committee of PCSJ/IMPS. He is a member of IEEE and ITE.

Hiroshi Watanabe received the B.E., M.E. and Ph.D. degrees in Engineering from Hokkaido University in 1980, 1982, and 1985, respectively. He joined Nippon Telegraph and Telephone Corporation (NTT) in 1985, and was engaged in research and development of image and video coding systems at NTT Human Interface Labs. and NTT Cyber Space Labs. until 2000. He has also been involved with the development of JPEG, MPEG standards under JTC 1/SC 29. He is currently a professor at the Department of Communications and Computer Engineering, School of Fundamental Science and Engineering, Waseda University. His research interests include object recognition, deep learning, image processing, video coding, and multimedia distribution. He is a member of IEEE, IPSJ, ITE, and IIEEJ.