Adaptive Planar Weighting Strategy for Enhanced DIMD Performance

Jeonghyun Lee[a] and Kiho Choi[a]‡

Abstract

This paper addresses the structural limitation of the Decoder-side Intra Mode Derivation (DIMD) tool in ECM-15.0, where the Planar mode is always included with a fixed weight. While the fixed participation of the Planar mode in blending provides stability in prediction, it can reduce prediction accuracy for blocks with strong directional characteristics. To overcome this limitation, two adaptive weight adjustment methods for planar blending are proposed. The first method determines the Planar weight by evaluating the amplitude distribution of the top candidate modes derived from the Histogram of Gradient. The second method determines the Planar weight based on the number of directional modes associated with non-square block shapes. Experimental results showed that Adaptive Planar Weighting based on Amplitude (APWA) consistently improved performance in the luma component, while Adaptive Planar Weighting based on Shape (APWS) provided gains in chroma components. These results demonstrate that adaptively adjusting the participation of the Planar mode according to conditions can enhance coding efficiency without increasing complexity.

Keyword : ECM, DIMD, Intra prediction, Adaptive Planar weight

a) Department of Electronics and Information Convergence Engineering, Kyung Hee University
‡ Corresponding Author : Kiho Choi
　　　　　　E-mail: aikiho@khu.ac.kr
　　　　　　Tel: +82-31-750-5799
　　　　　　ORCID: https://orcid.org/0000-0003-3686-4786

Ⅰ. Introduction

In the recent digital media market, the demand for ultra-high-definition video such as 4K and 8K, as well as immersive content including VR and AR, has been rapidly increasing. Consequently, the need for high-performance compression technology to efficiently transmit and store video data has become more critical. However, due to limited network bandwidth and storage capacity, it is increasingly difficult to meet these requirements with conventional

coding techniques alone. To address these challenges, the Versatile Video Coding (VVC) standard was established as an international standard in 2020 and has since been adopted in industry. VVC achieves approximately 50% higher compression efficiency compared to High Efficiency Video Coding (HEVC) while maintaining flexibility across a wide range of application scenarios[1,2]. This standardization effort was led by the Joint Video Experts Team (JVET), a collaboration between ITU-T's VCEG (Video Coding Experts Group) and ISO/IEC's MPEG (Moving Picture Experts Group)[3].

Currently, JVET is conducting research toward the next-generation video coding standard, for which it has developed a software known as the Enhanced Compression Model (ECM)[4]. The ECM integrates a variety of advanced coding tools not only for intra coding but also for inter coding. In particular, the intra coding stage has been significantly enhanced through the adoption of tools such as Template-based Intra Mode Derivation (TIMD), Decoder-side Intra Mode Derivation (DIMD), Occurrence-Based Intra Coding (OBIC), Spatial Geometric Partition Mode (SGPM), Template-based Multiple Reference Line (TMRL), IntraTMP, Convolutional Cross-Component Model (CCCM), Gradient-based Linear Model (GLM), and Cross-Component Prediction (CCP)[5-12]. As a result, ECM-15.0 demonstrated bitrate savings of approximately 18% in intra coding and 28% in inter coding compared to VVC[13].

This paper focuses on DIMD, one of the intra coding tools included in ECM. Specifically, the study aims to analyze the structure and operational principles of DIMD, and to explore methods for improving its performance.

II. DIMD Description and Limitations

DIMD is one of the intra prediction techniques in which the decoder derives the prediction mode. In the VVC standard, intra prediction modes are explicitly encoded by the encoder and signaled in the bitstream, which causes significant overhead and consequently reduces coding efficiency. To overcome this limitation, DIMD does not transmit explicit mode information; instead, it derives candidate modes according to rules that can be identically reproduced at both the encoder and the decoder.

The DIMD adopted in the ECM derives prediction modes at the decoder side. Therefore, it extracts directional information from the already reconstructed reference samples around the current prediction block to identify the dominant angle. Specifically, as shown in Figure 1, a 3x3 horizontal Sobel filter and a 3x3 vertical Sobel filter are applied to the samples in the left, top, and top-left regions of the current prediction block to obtain their horizontal and vertical gradients. The each of Sobel filters are as follows:

$$Sobel_{Hor} = \begin{bmatrix} 1 & 0 & -1 \\ 2 & 0 & -2 \\ 1 & 0 & -1 \end{bmatrix}$$
$$Sobel_{Ver} = \begin{bmatrix} -1 & -2 & -1 \\ 0 & 0 & 0 \\ 1 & 2 & 1 \end{bmatrix}. \tag{1}$$

Based on the horizontal gradient (i.e., iDx) and the vertical gradient (i.e., iDy) calculated from the Sobel filters, the angle is obtained and mapped to the closest prediction mode among the 65 directional modes. In addition, the amplitude of the derived angle is computed as follows:

$$angle = arctan\left(\frac{iDx}{iDy}\right) \tag{2}$$

$$amp = |iDx| + |iDy|. \tag{3}$$

After obtaining the prediction mode and the amplitude of the angle for the samples in the reference area, the amplitude values are accumulated for each prediction mode
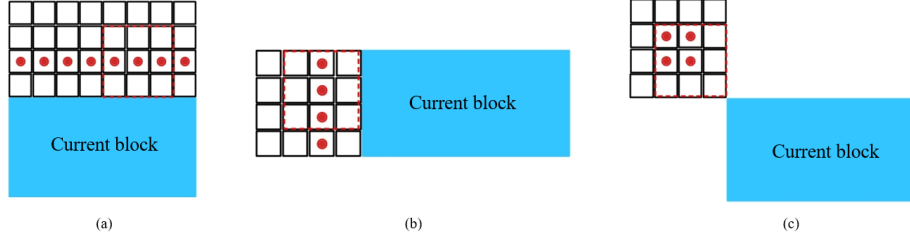
Fig. 1. Each of reference area : (a) Top (b) Left (c) Top-Left

to form a Histogram of Gradient (HoG). Finally, the HoGs constructed from each reference area are linearly combined to generate the final HoG.

After the HoG is generated, DIMD performs a blending process to construct the final prediction block. First, the five prediction modes with the highest accumulated amplitude in the HoG are selected, and a corresponding prediction block is generated for each of them. Then, weights are assigned to the prediction blocks according to their accumulated amplitude values, and the final prediction block is produced as follows:

$$Pred_{blending} = weight_{Planar} * Pred_{Planar} + \\ - \sum_{i=0}^{4} weight_{mode_i} * Pred_{mode_i}$$

(4)

where $Pred_{blending}$ is the final predicted block obtained after blending, $Pred_{Planar}$ is the planar-mode-based prediction block and $Pred_{mode_i}$ indicates the prediction block generated from the ith directional mode. $weight_{Planar}$ denotes the blending weight assigned to the planar mode and $weight_{mode_i}$ denotes the weight assigned to the ith candidate directional mode.

In the blending process, the Planar mode is always included. To provide stability for the final prediction block, it is always assigned a fixed weight ratio of 16 to 64, while the remaining prediction modes are weighted in proportion to their accumulated amplitudes. However, including the

Planar mode in all cases with a fixed contribution introduces certain limitations.

First, it becomes unsuitable for blocks with strong directional characteristics. When a block in the image exhibits a distinct boundary or a dominant edge in a single direction, the Planar mode may unnecessarily mix interpolated components, thereby reducing prediction accuracy. In addition, since a portion of the total blending weight budget is always reserved for the Planar mode, it restricts the ability to allocate sufficient weights to the most effective directional modes for that block. Such inefficiency ultimately prevents adaptive optimization according to the content characteristics, which can degrade prediction performance. Therefore, this paper proposes a method to adaptively adjust the participation of the Planar mode.

Ⅲ. Proposed Methods for Improving DIMD Performance

After the candidate intra prediction modes are derived through the HoG, the directional similarity of the selected modes is evaluated in preparation for the blending process. The selected modes are considered the most dominant prediction modes across all reference regions, and thus are expected to have a significant impact on the current prediction block. However, only the top three modes with the highest accumulated strengths are evaluated. The similarity measurement is performed by using the strongest mode

(i.e., mode[0]) as the reference and evaluating how similar the second (i.e., mode[1]) and third (i.e., mode[2]) ranked modes are to it. First, the distances between mode[0] and each of the mode[1] and mode[2] are calculated. As shown in Figure 2, the distance between modes is defined as the numerical difference between their mode indices. The distance is calculated as follows:

$$distance = |mode[0] - mode[i]| \ (i = 1,2). \tag{5}$$
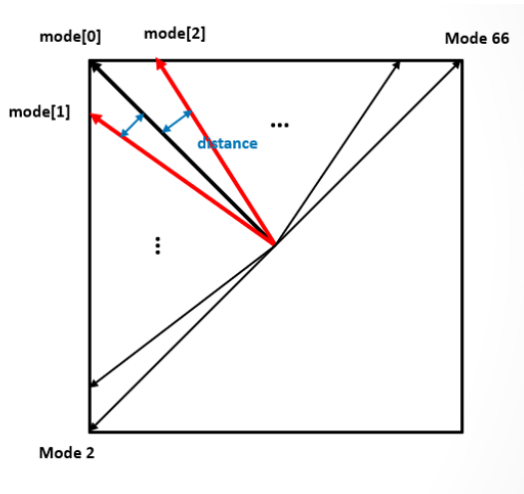

Fig. 2. Angular distance between mode[0] and other intra prediction modes

If each calculated distance is smaller than a predefined threshold, the mode is considered close to the mode[0] and thus regarded as directionally similar. In this paper, the threshold is set to 2. If the distances between the mode[0] and the mode[1] and mode[2] exceed this threshold, for example, the modes are not directionally similar, it is determined that the dominant modes do not exhibit directional consistency. In such cases, the Planar mode is deemed necessary to play a mediating role and is therefore included in the blending process. Conversely, if the distances do not exceed the threshold, meaning the modes are directionally similar, the involvement of the Planar mode may interfere with better directional prediction, and thus

its participation is restricted.

1. Adaptive Planar Weighting based on Amplitude (APWA)

The first approach determines the Planar weight by comparing the amplitude values of each mode[14]. The amplitudes obtained from the HoG are not only crucial for assessing the influence of each directional mode on the prediction of the current block, but also serve as essential parameters for calculating the weights assigned to each mode during the final blending process. Hence, the amplitude plays a critical role in generating the final prediction block. The algorithm for determining the Planar weight according to the amplitude is presented, as shown in Figure 3.
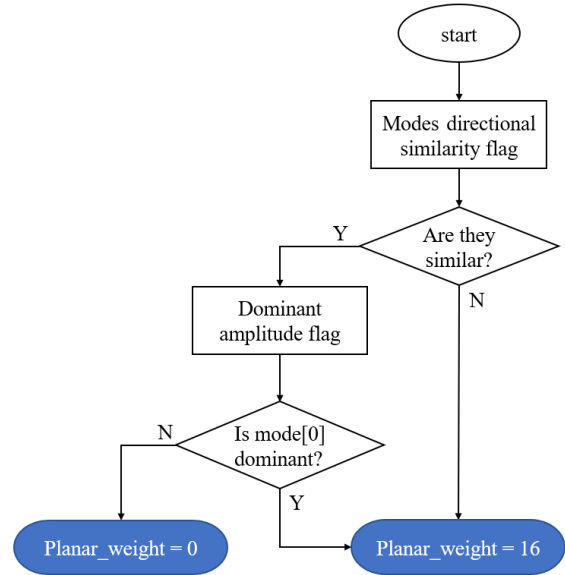

Fig. 3. Adaptive Planar weight decision based on directional similarity and dominance of mode[0]

First, the similarity among the top three modes described earlier is examined. If the three modes are found to be similar, the process proceeds to the next step, which is the amplitude comparison. The comparison of the amplitudes among the three modes is given by the follows:

$$amp[0] \geq \alpha \times amp[i] \; (i = 1,2) \quad (6)$$

where $amp[0]$ denotes the amplitude of mode[0], and $amp[i]$ represents the amplitude of mode[i], where i = 1 or 2. α is a predefined constant, which is set to 1.5 in this paper.

It is first examined whether the amplitude of the top-ranked mode (i.e., mode[0]) is greater than or equal to α times the amplitudes of the other modes (i.e., mode[1] and mode[2]). If this condition is not satisfied, the amplitudes of the three modes are considered to be evenly distributed. In such a case, the Planar mode is regarded as unnecessary in the blending process, since it is assumed that the prediction block already has a consistent directional tendency defined by the top three similar modes, and the evenly distributed amplitudes are sufficient to ensure stability. Therefore, in this case, the Planar weight is set to zero and the Planar mode is excluded from the blending process.

2. Adaptive Planar Weighting based on Shape (APWS)

The second proposed method evaluates the presence of dominant directional modes according to the shape of the block[15]. In the Intra Prediction of ECM, not only square blocks but also non-square blocks are utilized during prediction. The emergence of these non-square blocks is a result of applying various block partitioning methods to represent finer details of the video more accurately. Based on this observation, it is assumed that non-square blocks are more likely to exhibit dominant directional tendencies depending on their shape.

In VVC, wide-angle intra prediction adaptively adjusts the set of angular prediction modes according to a block's width-to-height ratio, and for non-square blocks assigns additional wide-angle directions toward the longer side[18]. In other words, for horizontal blocks, wide-angle modes that more densely sample the horizontal direction (e.g., modes 67‐80) are selectively enabled, whereas for vertical blocks, wide-angle modes that more densely sample the vertical direction (e.g., modes −1 to −14) are activated. Accordingly, in this paper, it is assumed that prediction with a horizontal mode is advantageous for horizontal blocks, while prediction with a vertical mode is advantageous for vertical blocks. Therefore, it is necessary to distinguish between the intra modes used in the ECM that exhibit horizontal characteristics and those that exhibit vertical characteristics. In the current HoG process of DIMD, prediction is carried out using a total of 65 modes, ranging
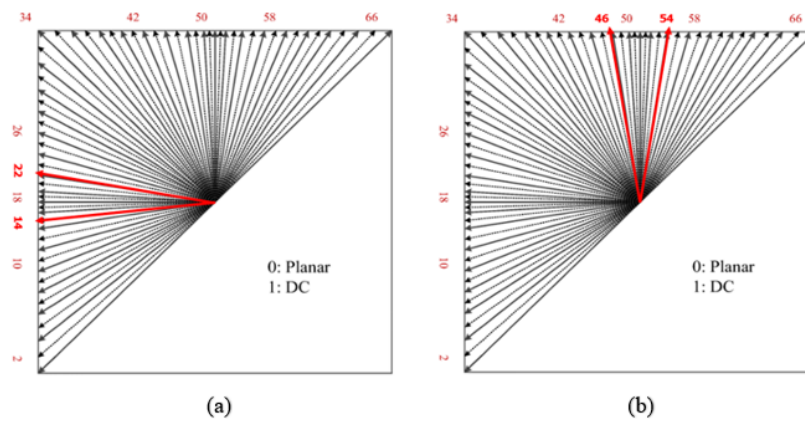

Fig. 4. (a) Horizontally oriented intra prediction modes (b) Vertically oriented intra prediction modes

from mode 2 to mode 66, as shown in Figure 4.

As shown in Figure 4, mode 18, which represents a clear horizontal prediction, and mode 50, which represents a distinct vertical prediction, are taken as references. Based on these, the ranges from -4 to +4 around each, namely mode 14 to mode 22 and mode 46 to mode 54, are defined as horizontally oriented prediction modes and vertically oriented prediction modes, respectively. The related process can be found in the following question:

$$mode_{k-4} \leq mode_k \leq mode_{k+4} \quad \begin{cases} k = 18 \\ k = 50 \end{cases} \quad (7)$$

where, $mode_k$ denotes the intra prediction mode index, k is the reference mode index (18 for horizontal and 50 for vertical). Therefore, the algorithm for determining the Planar weight according to the degree of dominance of the prediction modes depending on the block shape is shown in Figure 5.

First, as in the previous method, the directional similarity of the top three modes is evaluated. If the directions are similar, a Planar_weight = 0 is assigned, while APWS is applied when the directions are not similar. The shape of the current prediction block is then determined by comparing its width and height, classifying it as a horizontal block, a vertical block, or a square block. In particular, a block is categorized as horizontal when its width is larger than its height, as vertical when its height is larger than its width, and as square when its width and height are the same. In this paper, in order to focus on evaluating the performance for non-square blocks, square blocks are always assigned a Planar weight of 16.

After determining the block shape, the modes that exhibit the corresponding directional tendency are identified. For horizontal blocks, modes 14 to mode 22 are considered favorable, while for vertical blocks, modes 46 to mode 54 are advantageous. Next, among the top modes selected after the HoG process, the presence of such tendency modes is checked. In this work, only the top three modes with the strongest influence are considered. If all three modes belong to the tendency set, Planar_weight = 0 is assigned. If two modes belong, Planar_weight = 4 is assigned, and if only one mode belongs, Planar_weight = 8 is assigned. If none of the top three modes fall into the tendency set, Planar_weight = 16.

The reason for decreasing the value of the Planar_weight as the number of tendency modes among the top three increases is that the more tendency modes are present, the stronger the directional consistency of the current prediction block is judged to be.
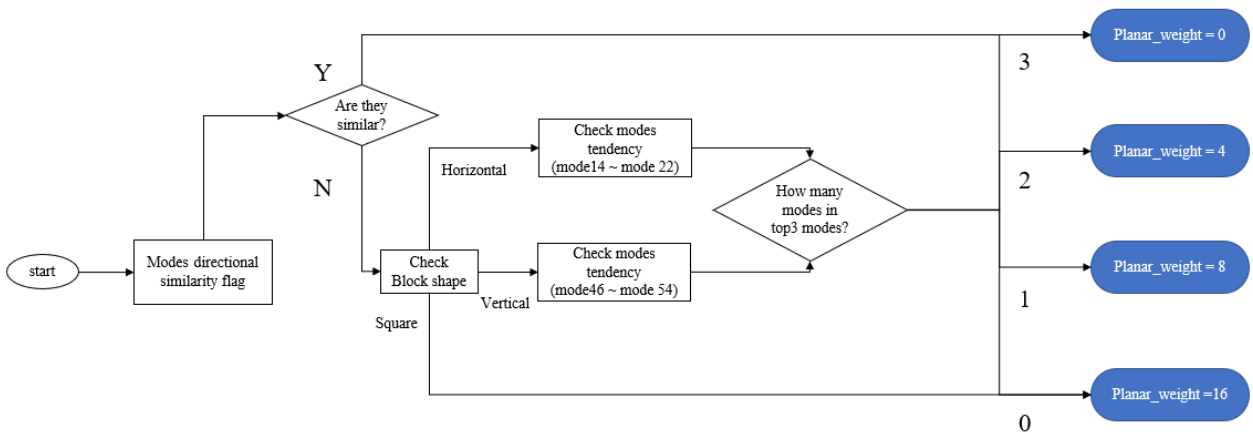

Fig. 5. Adaptive Planar weight decision based on block shape and directional tendency

Ⅳ. Testing results

To further evaluate the proposed adaptive Planar weight adjustment method, the implementation was carried out on the ECM-15.0 reference source code[16], and the evaluation followed the JVET Common Test Conditions (CTC)[17] under the All Intra (AI) configuration. In this setup, the complete sequences of Classes A1, A2, B, C, and E were encoded for performance analysis.

1. Results of APWA

This proposed method was also tested by applying it to the 1sec test sequences. Table 1 presents the experimental results the proposed APWA.

Table 1 shows the experimental results when the proposed method was applied to the DIMD of luma blocks. In the Class A2 sequences, coding performance improvements of 0.12% BD-rate reduction for the luma component

were observed, and in the Class E sequences, 0.02% BD-rate reduction for the luma component was achieved. On average, coding performance improvements of 0.02%, 0.28%, and 0.15% were observed for Y, Cb, and Cr, respectively.

2. Results of APWS

The algorithm was also tested by applying it to the full test sequences. Table 2 summarizes the experimental results obtained with APWS.

The proposed APWS was applied to the DIMD of luma blocks, and the results are reported in Table 2. For the Class A1 sequences, the Chroma (i.e., U) component achieved a BD-rate reduction of 0.14%, while for the Class E sequences, a 0.01% reduction was observed in the luma component. When averaged across all test classes, the proposed method yielded overall coding gains of 0.00%, 0.02%, and 0.01% for Y, Cb, and Cr, respectively.

Table 1. Results of APWA

Class	Sequence	BD-rate		
		Y	U	V
A1	Tango2	0.05%	-2.78%	-0.17%
	FoodMarket4	0.12%	-0.83%	-1.34%
	Campfire	-0.15%	0.10%	-0.08%
A2	CatRobot	-0.02%	-0.78%	-0.73%
	DaylightRoad2	-0.11%	-0.74%	-0.32%
	ParkRunning3	-0.22%	-0.18%	0.18%
B	MarketPlace	0.01%	0.07%	-0.28%
	RitualDance	-0.03%	0.69%	-0.02%
	Cactus	0.02%	0.21%	0.26%
	BasketballDrive	0.09%	-0.49%	-0.72%
	BQTerrace	-0.03%	0.29%	-0.51%
C	BasketballDrill	0.02%	-0.40%	-0.33%
	BQMall	0.03%	-0.60%	0.31%
	PartyScene	0.03%	-0.05%	-0.30%
	RaceHorses	-0.02%	-0.09%	1.59%
E	FourPeople	-0.04%	0.38%	0.19%
	Johnny	0.01%	0.64%	-0.07%
	KristenAndSara	-0.02%	-0.52%	-0.31%
Average		-0.02%	-0.28%	-0.15%

Table 2. Results of APWS

Class	Sequence	BD-rate		
		Y	U	V
A1	Tango2	0.02%	-0.26%	-0.14%
	FoodMarket4	0.05%	-0.11%	-0.01%
	Campfire	0.01%	-0.05%	0.15%
A2	CatRobot	0.01%	-0.09%	-0.07%
	DaylightRoad2	0.01%	-0.22%	-0.17%
	ParkRunning3	0.01%	-0.02%	0.01%
B	MarketPlace	0.01%	-0.18%	-0.09%
	RitualDance	0.01%	0.23%	0.00%
	Cactus	0.01%	0.05%	0.04%
	BasketballDrive	0.00%	-0.15%	-0.23%
	BQTerrace	0.00%	0.03%	-0.35%
C	BasketballDrill	-0.02%	0.05%	0.18%
	BQMall	0.00%	-0.08%	-0.05%
	PartyScene	0.02%	-0.03%	-0.08%
	RaceHorses	0.00%	0.04%	0.13%
E	FourPeople	-0.02%	0.10%	0.10%
	Johnny	-0.02%	-0.31%	-0.04%
	KristenAndSara	0.01%	0.37%	0.06%
Average		0.00%	-0.03%	-0.02%

Ⅴ. Discussion

This study verified the effectiveness of an adaptive Planar weight adjustment method to address the structural limitation in DIMD tool of ECM-15.0, where the Planar mode is always included with a fixed weight. Among the proposed approaches, APWA evaluates the amplitude distribution of the top three candidate modes derived from HoG. The experimental results indicate that, considering ECM-15.0 is already a highly optimized environment, even such modest improvements are meaningful and demonstrate that fine-grained blending adjustment can contribute to enhancing intra prediction performance.

APWS, on the other hand, utilizes block-shape information. Based on the aspect ratio of the block, horizontal blocks are associated with horizontal tendency modes (e.g., 14‑22), vertical blocks with vertical tendency modes (e.g., 46‑54), and square blocks are classified with mandatory Planar participation. This approach is based on the idea that the uneven shape of non-square blocks makes them more suitable for certain directional modes. While certain cases showed performance improvements, performance degradation was observed in the luma component in particular. This suggests that block shape alone is insufficient to fully characterize directional tendencies and that other factors such as texture complexity and edge continuity must also be considered. An interesting observation is that improvements were relatively larger in the chroma components. This can be interpreted as the reduced resolution of chroma in the 4:2:0 sampling format making directional prediction inherently more difficult, where suppressing unnecessary Planar involvement allows for more efficient utilization of chroma samples.

Ⅵ. Conclusion

In this paper, we proposed adaptive Planar weight adjust-

ment methods to address the limitation of the fixed Planar mode weighting in DIMD tool of ECM-15.0. The proposed method consists of APWA, which utilizes the amplitude distribution of HoG, and APWS which incorporates block-shape information. Experimental results showed that APWA achieved stable improvements in the luma component, while APWS yielded partial gains in the chroma components.

These findings demonstrate that adaptively adjusting the involvement of the Planar mode according to block characteristics can improve efficiency compared to the conventional fixed approach. Although the overall coding gains were limited, even small BD-rate reductions are meaningful in a highly optimized environment such as ECM, indicating that fine-grained refinements in intra prediction tools remain an important research direction for future video coding standards. Future work will extend the evaluation beyond the All Intra configuration to include Random Access and Low-Delay B conditions.

Furthermore, a multidimensional Planar weight estimation method that accounts for not only amplitude and block shape but also texture complexity, edge continuity, and block size should be investigated. In addition, integrating learning-based approaches to automatically derive optimal block-level weights could provide both efficiency and adaptability. Such research is expected to contribute not only to ECM but also to the development of next-generation video coding standards by further enhancing intra prediction performance.

References

[1] B. Bross, J. Chen, J.-R. Ohm, G. J. Sullivan, and Y.-K. Wang, "Developments in International Video Coding Standardization After AVC, With an Overview of Versatile Video Coding (VVC)," Proceedings of the IEEE, Vol. 109, No. 9, pp. 1463-1493, Sep.2021. doi : https://doi.org/10.1109/JPROC.2020.3043399

[2] B. Bross, Y.-K. Wang, Y. Ye, S. Liu, J. Chen, G. J. Sullivan, and J.-R. Ohm, "Overview of the Versatile Video Coding (VVC) Standard and

Its Applications," IEEE Transactions on Circuits and Systems for Video Technology, Vol. 31, No. 10, pp. 3736 – 3764, Oct. 2021. doi: https://doi.org/10.1109/TCSVT.2021.3101953

[3] Joint Video Experts Team (JVET). "JVET Official Website." https://jvet.hhi.fraunhofer.de (accessed Nov. 25, 2025).

[4] Xu, J., et al. "Enhanced Compression Model (ECM): Beyond VVC." JVET Document, 2023.

[5] Y. Wang, L. Zhang, et al. "EE2-related:Template-based intra mode derivation using MPMs." in Document JVET-V0098, Teleconference, Apr. 2021.

[6] E. Mora, A. Nasrallah, et al. "CE3-related:Decoder-side Intra Mode Derivation" in Document JVET-L0164, Macao, Oct. 2018.

[7] R. G. Youvalari, M.Abdoli, et al. "EE2-2.2:Occurrence-based intra coding (OBIC)." in Document JVET-AH0076, Rennes, Apr.2024.

[8] F. Wang, Y. Yu, et al. "Non-EE2:Spatial GPM." in Document JVET-Z0124, Teleconference, Apr. 2022.

[9] L. Xu, Y. Yu, et al. "EE2-1.10:Template-based multiple reference line intra prediction." in Document JVET-AB0156, Mainz, Oct. 2022.

[10] X. Li, Y. Ye, et al. "EE2-related:On Gradient Linear Model (GLM)." in Document JVET-AA0138, Teleconference, Jul. 2022.

[11] P. Astola, J. Lainema, et al. "AHG12:Convolutional cross-component model (CCCM) for intra prediction." in Document JVET-Z0064, Teleconference, Apr. 2022.

[12] H. Huang, V. Seregin, et al."AHG12:On the cross-component merge mode." in Document JVET-AE0097, Geneva, Jul. 2023.

[13] JVET. "ECM-15.0 Test Results." JVET-Q reports, 2024.

[14] J. Lee and K. Choi, "Improving DIMD Performance Through Adaptive Planar Weight Adjustment," Proceedings of the Korean Society of Broadcast and Media Engineers Conference, Jeju, Korea, pp. 222-224, June 2025.

[15] J.-H. Lee, K. Choi (KHU), C. W. Ryu (Kaon Group). "Non-EE2: Adaptive Planar Weight for DIMD." in Document JVET-AM0141, Daejeon, Jun. 2025.

[16] ECM-15.0, https://vcgit.hhi.fraunhofer.de/ecm/ECM/-/tree/ECM-15.0

[17] M. Karczewicz, Y. Ye, "Common test conditions and evaluation procedures for enhanced compression tool testing, Joint Video Experts Team of ITU-T and ISO/IEC, in Document JVET-AI2017, Sapporo, Oct. 2024.

[18] M. Lee, H. Song, J. Park, B. Jeon, J. Kang, J.-G. Kim, Y.-L. Lee, J.-W. Kang, and D. Sim, "Overview of Versatile Video Coding (H.266/ VVC) and Its Coding Performance Analysis," IEIE Transactions on Smart Processing and Computing, vol. 12, no. 2, pp. 122 – 151, Apr. 2023. doi: https://doi.org/10.5573/IEIESPC.2023.12.2.122

Introduction Authors

Jeonghyun Lee

- Feb. 2025 : B.S. in Electronic Engineering, Kyung Hee University, Suwon, Korea.
- Mar. 2025 ~ Present : M.S. student (2nd semester) in Electronics and Information Convergence Engineering, Kyung Hee University, Korea.
- ORCID : https://orcid.org/0009-0003-5103-7292
- Research interests : Video coding and compression standards (JVET ECM), Intra prediction algorithms.

Kiho Choi

- 2012. : Electronics and computer engineering, Hanyang University, B.S.(2008), Ph.D.(2012)
- 2012. 09. ~ 2014. 02. : Lecturer, Post Doc., Hanyang University
- 2014. 03. ~ 2021. 02. : Senior Engineer, Visual Technology Lab. of Samsung Research
- 2021. 03. ~ 2023. 02. : Professor, School of Computing, Gachon University
- 2023. 03. ~ Present : Professor, Electronic Engineering, Kyung Hee University
- ORCID : https://orcid.org/0000-0002-2869-0440
- Research interests : Image and video coding, multimedia data compression, multimedia streaming, AI-based multimedia technology, and multimedia standardization