The following publication C. -H. Fu, Y. -W. Zhao, H. -B. Zhang, Y. -L. Chan and W. -C. Siu, "Depth modelling mode decision for depth intra coding via good feature," 2017 IEEE International Conference on Image Processing (ICIP), Beijing, China, 2017, pp. 4018-4022 is available at https://doi.org/10.1109/ICIP.2017.8297037.

DEPTH MODELLING MODE DECISION FOR DEPTH INTRA CODING VIA GOOD FEATURE

Chang-Hong Fu[1,2], *Ya-Wen Zhao*[1*], *Hong-Bin Zhang*[1,2], *Yui-Lam Chan*[2], *Wan-Chi Siu*[2]

[1]School of Electronic and Optical Engineering,
Nanjing University of Science and Technology, Nanjing, China
[2]Department of Electronic and Information Engineering,
The Hong Kong Polytechnic University, Hong Kong

ABSTRACT

The depth modelling modes (DMM) and 35 conventional intra modes (CHIMs) introduced in 3D-HEVC results in unacceptable huge complexity of depth intra coding. However, some redundancy between DMM and CHIMs could be avoided to accelerate the process. In this paper, a good feature-corner point (CP) is proposed to evaluate the orientation of edge in a given prediction unit (PU), by which a binary classifier is created. We further investigate the probability distribution of DMM, which is selected as the optimal intra mode in each category. According to the statistical analysis, the skipping of DMM decision is proposed to eliminate the cases which have been predicted well by CHIMs. The experimental results show that, compared with the test model HTM-13.0 of 3D-HEVC, the proposed algorithm can yield about 17% time reduction for depth intra coding with almost no degradation in coding performance.

Index Terms— Multi-view video plus depth, 3D-HEVC, Corner point, Depth map, Intra prediction.

1. INTRODUCTION

In the last decade, supporting geometric information, depth map has being a very hot topic in computer vision, such as stereo reconstruction, 3D modelling, layer separation, and view interpolation [1]-[5]. In this context, depth map is introduced into 3D video to break the limitation of conventional glass-based 3D product and realize the free view video. In order to promote the industrialization of depth-based 3D video, a new generation coding standard has been developed by the JCT-3V, the official name of which is 3D-HEVC [6]. Assisted by depth map, the required views for 3D scene reconstruction can be synthesized via the depth image based render (DIBR) technique [7].

Depth map is very different from natural videos, which is employed in synthesis process instead of directly being viewed by audiences. In order to avoid the possible confusion between background and foreground in rendering, preserving sharp edge is the most important task in depth coding. The prediction of conventional HEVC intra modes (CHIMs) [8] just copy the neighboring reference pixels into the current block along a direction. This may generate large residuals along the sharp edge. When the high frequency coefficients in large amplitude are truncated in the conventional transform-quantization-based coding scheme, the around sharp edges may be degraded significantly. Therefore, the partition-based depth modelling mode (DMM) [9] is proposed as an alternative of CHIMs to keep the sharp edges. When DMM is applied, the current PU can be segmented into two region by a straight line. By describing each segment using a constant pixel value, DMM can prevent the distortion of sharp edge and yield a better quality of synthesized views. However, calculating the optimal pattern of DMM requires tremendous computational power for exhaustively evaluating the rate distortion cost (RD-cost) function.

To accelerate the depth intra coding, some research works have been developed to skip DMMs in unnecessary cases [11-13, 17-19]. Gu *et al.* checked the variance or best rough mode in [10]-[11] to skip DMM evaluation for smooth PUs. Park skipped unnecessary DMM by a pair of edge detectors in Hadamard transform domain in [12]. However, there are still redundant DMM evaluations and the time consumption is still high for depth intra mode decision. Hence, we raise an outstanding feature-corner point in this paper to skip DMM in more blocks by checking the local monotonicity of edges and finding out the joint case suitable for both DMM and CHIMs.

The organization of this paper is as follows. In Section 2, we make the problem formulation and introduce the basic concept of corner detection. Then, the feature extraction for the proposed classifier is presented in Section 3, together with the DMM skipping strategy. Finally, simulation results are presented and discussed in Section 4. Some concluding remarks are provided in Section 5.

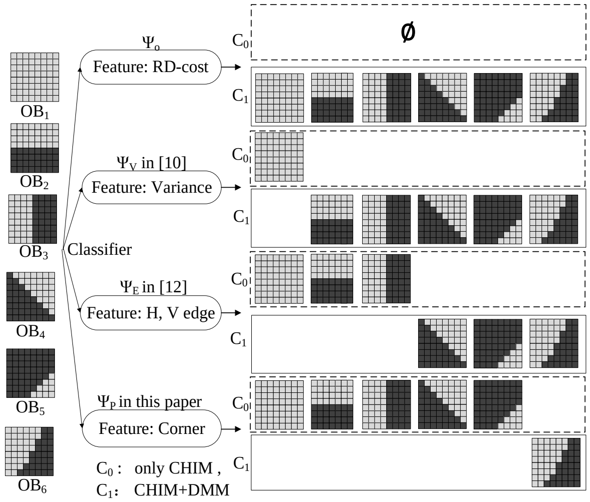

Fig. 1. Several DMM decision algorithms are formulated as classification problem with different features.

2. PROBLEM FORMULAITON AND MOTIVATION

For the simplicity of illustration, we could address the problem of skipping DMM as a classification problem. All PUs are classified into two categories by a classifier: skip the DMM evaluation or not. The input of classifier is the original block (OB) and the output is class label C_0 or C_1. Here, C_0 denotes that the case to skip DMM evaluation and C_1 presents the case where DMM decision is performed as usual. Fig. 1 presents the classifiers in the existing algorithms [10]-[12] together with our proposed one. Each OB in the Fig. 1 stands for one type of blocks. OB_1 presents smooth blocks. OB_2 and OB_3 have horizontal or vertical edges inside correspondingly. OB_4 and OB_5 contain other orientation of edges, where the edges are straight and monotonic. And OB_6 denotes the remaining blocks. Some of them contain edges with slight variation, which could be coded well by DMM but not CHIMs as shown in Fig. 1. In the original 3D-HEVC, all OBs are classified as C_1 by the classifier Ψ_o. The searching of wedgelet pattern is performed for all the PUs by exhaustively evaluating the RD cost function. The numbers of wedgelet pattern candidates for different PU sizes are illustrated to show the complexity of DMM in Table 1. The Ψ_V proposed in [10] classifies OB_1 with smaller variance into C_0 category to avoid DMM in smooth blocks. Furthermore, since OB_2/OB_3 could be predicted well by both angular(10)/angular(26) modes in CHIMs and DMM, the DMM decision could also be skipped in these two cases. The Ψ_E [12] hereby further classifies OB_2 and OB_3 into C_0 category by checking the horizontal and vertical edges. In this way, DMM evaluations could be skipped for $OB_{1\text{-}3}$.

It is observed that OB_4/OB_5 could also be predicted well by both of the corresponding CHIMs and DMM, as illustrated by the example in Fig. 2. This observation motivated us to seek a better classifier, which could distinguish the OB_4 and OB_5 from the mixture case OB_6. In the next section, a classifier Ψ_P based on corner point is raised to describe whether the orientation of edge is straight and monotonic and skip DMMs in OB_4 and OB_5 furthermore.

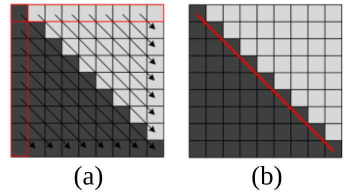

Fig. 2 The example of the OB_4, which could be predicted by (a) angular(18) mode, (b) wedgelet pattern in DMM. The red rectangle in (a) denotes the reference pixels of intra prediction.

Table 1. Number of Wedgelet Patterns in DMM

PU	Full Search	Fast Algorithm
4x4	86	58
8x8	766	310
16x16	510	384
32x32	510	384

3. PROPOSED CORNER POINT BASED DMM SKIPPING ALGORITHM

In this paper, we use corner point as the feature of the proposed Ψ_P. It is due to the fact that corner point has two dominant directions around its neighborhood and can reflect the level of consistent orientation.

3.1. Feature Extraction

The corner point is defined as the intersection of two edges. Smooth areas have no edges, while single straight edges only have one direction. The variant (Shi-Tomasi) of the classical corner detection algorithm Harris is selected from the various existing corner detection algorithms. In Shi-Tomasi, an image I(x,y) can be regarded as multiple random variables and the two main directions are selected by a classical Principal Component Analysis (PCA) method. The covariance matrix of a sample pixel p(x,y) and its neighboring random variance is given by

$$M = \sum_{x,y} w(x,y) \begin{bmatrix} I_x^2 & I_x I_y \\ I_x I_y & I_y^2 \end{bmatrix} \qquad (1)$$

where w(x,y) is the weight function, and I_x and I_y are the horizontal and vertical gradients, respectively. Based on the PCA theory, the two eigenvectors of M (quadric form) are the two principal directions of p(x,y) and their eigenvalues can reflect the degrees of change in their directions. Since each real corner point should have at least two dominant directions. The smaller one of the two eigenvalues is noted as $\lambda_{M(x,y)}$, which is employed to sort the candidates of corner point.

A two-step strategy is adopted in the extraction of corner points. First, all points with zero or near-zero $\lambda_{M(x,y)}$

are filtered out. The remaining candidates are sorted by the magnitude of $\lambda_{M(x,y)}$ in the descending order. Then, quantization parameter (QP) is employed to pick up the first Th_{Num} of corner points in the sorting list for each frame.

$$Th_{Num} = \begin{cases} N_c & ,QP \le 36 \\ N_c \times \dfrac{5-((QP-37)\%3}{3\times 2^{([((QP-37)/3)]+1)}} & ,Otherwise \end{cases} \qquad (2)$$

where N_c is the total number of candidates after the first round filtering. And Th_{Num} is the final number of corner points after the second thresholding. After that, these first Th_{Num} of samples in the list are designate as corner point for further utilization. The concept of corner points is also employed in our previous work on quadtree decision in [13]. Readers who are interested in this feature are encouraged to read that. Besides, researchers who are interested to verify the results or utilize corner points in other fields are suggested to read the function "GoodFeaturesToTrack()" in the OpenCV library.

3.2. Statistical analysis

Since corner point represent two dominant directions in gradient changing, it is utilized for evaluate whether the edge is strictly straight and monotonic in our classification. The classifier Ψ_P could be defined as whether the input block OB contains any corner points (CPs). In order to verify the accuracy of classifier by corner points, the conditional probability of DMM utilization is calculated for each class. The results for different block sizes are summarized in Table 2, where $p(DMM|C_0)$ and $p(DMM|C_1)$ represent the probability of DMM to be the best intra mode for these OBs classified as C_0 and C_1 , respectively. As shown in Table 2, when the OBs are classified as C_0, the average probability of DMM to be the optimal mode is only 1.62%, 0.69%, 0.38% and 0.48% for different block sizes correspondingly. On the other side, about 24.95%, 19.50%, 14.01% and 11.01% of the OBs classified as C_1 choose the DMM as the optimal intra mode. The results confirm that almost all OBs classified as C_0 could be predicted well by CHIMs. And the time consuming DMM could be skipped safely in these OBs with almost no influence on the coding performance.

3.3. Proposed algorithm

There are three steps for the original depth intra mode decision in 3D-HEVC. First, a few effective CHIMs are selected and inserted into the candidate list by evaluating the low-complexity rough mode RD cost. Then, the optimal wedgelet pattern by DMM decision is also added into the candidates list. Finally, the full RD calculations are performed for all the candidates. In the current HTM, two features from Gu' work [10]-[11] have already been adopted to skip the DMM. The corresponding skipping conditions are as follows: (i) the variance of the OB is smaller than a threshold [10], (ii) the rough mode with minimal low complexity RD cost is Planar Mode[11]. Our proposed classifier by corner points is implemented in the current together with [10-11]. Since [10-11] are dealing with the same problem as our proposed method, in the following illustration, they are also included as step (c) in the flowchart for more comprehensive understanding of the final coding structure. As demonstrated in Fig. 3, the flowchart presents intra mode decision process as follows:

Table 2. The conditional probabilities of DMM to be the best intra mode for the OBs classified as C_0 and C_1 with different block sizes.

Test seq.		32x32	16x16	8x8	4x4
Kendo	p(DMM|C0)	1.82%	1.05%	0.34%	0.33%
	p(DMM|C1)	32.48%	22.79%	12.92%	9.08%
Newspaper	p(DMM|C0)	0.95%	0.86%	0.87%	1.25%
	p(DMM|C1)	30.51%	24.43%	15.75%	11.41%
Balloons	p(DMM|C0)	0.99%	0.79%	0.44%	0.55%
	p(DMM|C1)	31.43%	25.09%	14.65%	10.20%
Poznan_Hall2	p(DMM|C0)	1.24%	0.37%	0.13%	0.08%
	p(DMM|C1)	16.56%	9.59%	6.60%	4.74%
Poznan_Street	p(DMM|C0)	0.00%	0.38%	0.38%	0.66%
	p(DMM|C1)	17.81%	9.57%	4.85%	3.84%
GT_Fly	p(DMM|C0)	0.22%	0.14%	0.14%	0.24%
	p(DMM|C1)	21.56%	17.77%	14.62%	12.41%
Undo_Dancer	p(DMM|C0)	6.14%	1.23%	0.36%	0.23%
	p(DMM|C1)	24.33%	27.28%	28.69%	25.38%
Average	p(DMM|C0)	1.62%	0.69%	0.38%	0.48%
	p(DMM|C1)	24.95%	19.50%	14.01%	11.01%

(a) For a given depth frame, adaptive corner point extraction is performed to find the corner points and store their positions.

(b) For a given OB, a few of CHIMs are selected by rough mode decision and added into the candidate list.

(c) In the DMM evaluation process, if the conditions of [10-[11] as illustrated above are satisfied, the current OB is classified as smooth OB_1 as shown in Fig. 1. Skip DMM decision and go to step (f); otherwise, go to step (d).

(d) The proposed Ψ_P is applied. If the current OB is classified as C_0, it is assumed to be one of the cases from OB_2 to OB_5 in Fig. 1. The DMM decision is also skipped and the process goes to (f). Otherwise, the current OB is classified as C_1 go to step (e) for DMM decision.

(e) The optimal wedgelet pattern is searched for DMM and added into the list of candidate mode.

(f) The full RD calculation is performed to find the optimal depth intra mode from the candidate list.

It is noted that in the above process and Figure 4, step (a) and (d) are the main contribution of this paper. Besides, the corner point extraction in step (a) is very fast compared to the DMM mode decision steps. The overhead in time consumption caused by step (a) could be neglected.

4. EXPERIMENTAL RESUTLS

In the experiments, the reference software HTM-13.0 of 3D-HEVC [14] was used to implement the state-of-the-art method [12] and the proposed one. All the experiments foll-

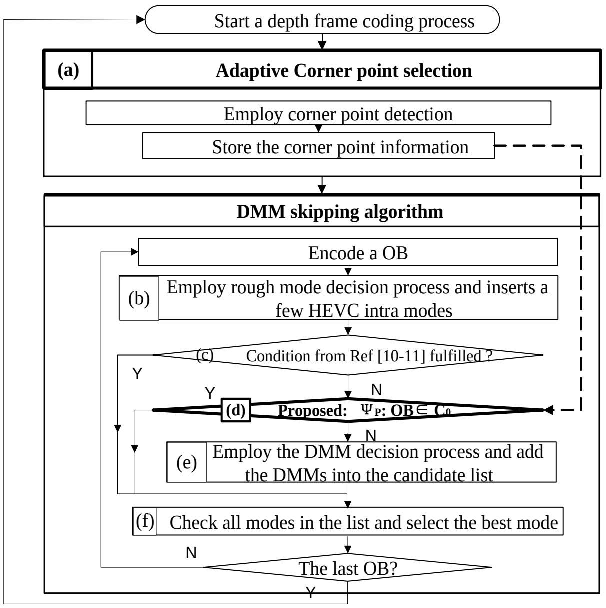

Fig. 3 Flowchart of the proposed DMM skipping algorithm

owed the common test condition (CTC) [15]. Since the proposed algorithm focuses on depth intra coding, all frames were encoded as I-frames. Seven test sequences recommend by CTC were tested, including *Kendo*, *Balloons*, *Newspaper1*, *Undo_Dacner*, *GT_Fly*, *Poznan_Hall2* and *Poznan_Street*. The encoding time percentage difference was used to evaluate the coding efficiency.

The BDBR was used to evaluate the coding performance, which comes from the Bjontegaard metrics[16] and calculated by the overall transmitted bitrate(Texture+Depth) and the Peak signal-to-noise ratio (PSNR) of synthesis views. For each recommended sequence, three views were encoded and 6 virtual views were rendered by the decoded textures and the corresponding depth maps. Meanwhile, another 6 views should be rendered by the un-encoded textures and depth maps in advance as an anchor for calculating the Peak signal-to-noise ratio (PSNR) of synthesis views.

The experimental results are presented in Table 3. It can be found that the propose method provides about 17% time reduction, averagely, compared with the original 3D-HEVC encoder. In particular, the algorithm performs better in the sequences *Undo_Dancer*, *GT_Fly* and *Poznan_Hall2* since they contained more smooth and simple PUs. Besides, the proposed algorithm introduce only about 0.19% BDBR increase averagely, which means that the bit rate increase or quality loss is extremely small compared to the HTM-13.0 including [10]-[11]. However, the BD performance degradation is a litter bit higher for sequences *Undo_Dancer* and *Poznan_Hall2*. In these sequences, the depth values changes slightly and gradually, which leads to C_0 classification and DMM decision skipping. But the color video is more complicated in the corresponding region, which results in the degradation of quality in the synthesized

Table 3. Comparison of coding performance between Park's method and proposed algorithm.

Test sequence	Park's [12]		Proposed	
	ΔBDBR	ΔT	ΔBDBR	ΔT
Kendo	0.08	-10.2	0.12	-17.45
Balloons	0.03	-8.42	0.16	-15.31
Newspaper1	0.16	-4.26	0.12	-11.74
1024x768	**0.09**	**-7.62**	**0.13**	**-14.83**
Undo_Dancer	0.14	-5.73	0.37	-18.83
GT_Fly	0.02	-6.10	0.03	-20.65
Poznan_Hall2	0.47	-16.15	0.52	-23.01
Poznan_Street	0.00	+6.51	0.00	-9.14
1920x1088	**0.16**	**-5.37**	**0.23**	**-17.91**
Average	**0.11**	**-5.14**	**0.19**	**-16.59**

views. At the same time, the sequence *GT_Fly* also has regions with gradually changed pixel values. But the BD increasing is only 0.03%. It is because the corresponding region in color video is a plain background (desert). The observation is also consistent with our statistics in Table 2, where *Undo_Dancer* and *Poznan_Hall2* has relatively large value of p($DMM|C_0$). This mismatch case between depth map and color video will be further investigated in our future work.

Table 3 also shows the comparison between the proposed algorithm and the state-of-the-art method [12]. As shown in Table 3, our method has similar BD performance with the Park's method. While it is about 11% better than Park's algorithm in terms of time saving. It is because not only the PUs with horizontal or vertical edges but also those with any strictly straight orientation of edges can benefit from the proposed algorithm. This result, more or less, can be explained by Fig. 1, where the OB_4 and OB_5 could also be classified into C_0 by the proposed Ψ_P. Besides, compared with Park's edges detection, the threshold used in the proposed corner point detection is adaptive based on QP. This is another advantage of the proposed algorithm.

5. CONCLUSION

In this paper, we formulate the DMM decision as a classification problem and analyze the features of previous works. After that, a feature corner point is raised to evaluate the orientation of edges. The proposed feature is used to skip the cases, which could be predicted well by both conventional HEVC modes and DMM. DMM is skipped in more of PUs by the proposed classifier compared with the existing methods. Experimental results show that the proposed algorithm can provide about 17% time reductions while maintaining the coding performance.

ACNOWLEGEMENT

This work is supported in part by the Fundamental Research Funds for the Central Universities (No. 30917011319), and a grant from the Research Grants Council of the Hong Kong Special Administrative Region, China (Grant No. PolyU 152052/15E).

REFERENCES

[1] C. H. Xu, Y. H. Wang, T. N. Tan, and L. Quan, "Depth vs. intensity: which is more important for face recognition," in the proceedings of ICPR, pp. 342-345, Aug. 2004.

[2] W. Q. Li, Z. Y. Zhang, and Z. C. Liu, "Action recognition based on a bag of 3D points," in the proceeding of Computer Vision and Pattern Recognition Workshops (CVPRW), San Francisco, CA, pp. 9-14, June 2010.

[3] X. Bai, Z. Zhu, and L. Latecki, "3D Shape Matching via Two Layer Coding," IEEE Transaction on Pattern Analysis and Matching Intelligence, Vol. pp, no. 99, pp.1, April 2015.

[4] Helten, T., M. Muller, H.-P. Seidel, and C. Theobalt, "Real-time Body Tracking with one Depth Camera and Inertial Sensors," in the proceedings of international conference on Computer vision (ICCV), Sydney, NSW, pp. 1105-1112, Dec. 2013.

[5] Muller, K.; Merkle, P.; Wiegand, T., "3-D Video Representation Using Depth Maps," in Proceedings of the IEEE , vol.99, no.4, pp.643-656, April 2011.

[6] Muller, K.; Schwarz, H.; Marpe, and more authors, "3D High-Efficiency Video Coding for Multi-View Video and Depth Data," IEEE Transactions on Image Processing, vol.22, no.9, pp.3366-3378, Sept. 2013.

[7] Chan, S.C.; Heung-Yeung Shum; King-To Ng, "Image-Based Rendering and Synthesis," IEEE Signal Processing Magazine, vol.24, no.6, pp.22-33, Nov. 2007.

[8] Lainema, J.; Bossen, F.; Woo-Jin Han; Junghye Min; Ugur, K., "Intra Coding of the HEVC Standard," IEEE Transactions on Circuits and Systems for Video Technology, vol.22, no.12, pp.1792-1801, Dec. 2012.

[9] Muller, K.; Merkle, P.; Tech, G.; Wiegand, T., "3D video coding with depth modeling modes and view synthesis optimization," IEEE international conference APSIPA ASC, pp.1-4, 3-6 Dec. 2012.

[10] Z. Gu, J. Zheng, N. Ling, and P. Zhang, "Fast Intra Prediction Mode Selection for Intra Depth Map Coding," document JCT3V-E0238, ITU-T SG 16 WP 3 and ISO/IEC JTC 1/SC 29/WG 11, Aug. 2013

[11] Z. Y. Gu; J. H. Zheng; N. Ling; Zhang, P., "Fast Depth Modeling Mode selection for 3D HEVC depth intra coding," IEEE International Conference on Multimedia and Expo Workshops (ICMEW), pp.1-4, 15-19 July 2013.

[12] Chun-Su Park, "Edge-Based Intra mode Selection for Depth-Map Coding in 3D-HEVC," IEEE Transactions on Image Processing, vol.24, no.1, pp.155-162, Jan. 2015

[13] Zhang, Hong Bin, et al. "Quadtree Decision for Depth Intra Coding in 3D-HEVC by Good Feature." IEEE International Conference on Acoustics, Speech and Signal Processing IEEE, 2016.

[14] "MPEG 3D-HTM Software V13.0." [Online]. Available: https://hevc.hhi.fraunhofer.de/svn/svn 3DVCSoftware/tags/13.0

[15] F. Jäger, "Description of Core Experiment 5 (CE5) on Depth Intra Modes," ITU-T SG 16 WP 3 and ISO/IECJTC 1/SC 29/WG 11, JCT3V-E1105, 5th Meeting: Vienna, AT, 27 Jul. –2 Aug. 2013.

[16] G. Bjontegard, "Calculation of Average PSNR Differences between RD curves," ITU-T Q.6/SG16, doc. VCEG-M33, Austin, TX, USA, April 2001.

[17] Zhang H B, Fu C H, Chan Y L, et al. Probability-based Depth Intra Mode Skipping Strategy and Novel VSO Metric for DMM Decision in 3D-HEVC[J]. IEEE Transactions on Circuits & Systems for Video Technology, 2016, PP(99):1-1.

[18] H.-B. Zhang, C.-H. Fu, W.-M. Su, S.-H. Tsang, W.-C. Siu, "Efficient Depth Intra Mode Decision by Reference Pixels Classification in 3D-HEVC", IEEE international conference on image processing, pp. 961-965, Sept. 2015.

[19] Hong-Bin Zhang, Chang-Hong Fu, Yui-Lam Chan, Sik-Ho Tsang, Wan-Chi Siu, Wei-Min Su, "Efficient wedgelet pattern decision for depth modelling modes in three-dimensional high-efficiency video coding", Journal of Electronic Imaging, 25(3), 033023, May 2016.