

# A Fast Depth Intra Mode Selection Algorithm

Jieling Fan\*, Qiang Li and Jianlin Song

(Chongqing Key Laboratory of Signal and Information Processing (CQKLS&IP), Chongqing University of Posts and Telecommunications (CQUPT), Chongqing, 400065, China)

\*Corresponding author

**Abstract**—For the purpose of reducing the complexity of depth intra coding, a fast intra prediction mode decision algorithm is proposed. First, the fast decision of depth modeling modes (DMMs) is obtained by using Laplace Operator detection method. Second by using the correlation between the wedgelet pattern and texture features, only the wedgelet patterns, which are related to the intra angle mode, is traversed. As a consequence, the number of wedgelet patterns is reduced through this algorithm. Experiment results indicate that the proposed algorithm can reach the same goal with the reduction of computational complexity by 24.7%, while the quality of the video is almost unchanged.

**Keywords**—depth intra coding; depth modeling mode; wedgelet pattern

## I. INTRODUCTION

3D (three dimensional) video can provide the tridimensional expression of the natural scene, and bring people visual experience. 3D movies, 3D television, tablet PCs and gaming terminals have entered the lives of the people. Due to the popularization of 3D equipment, 3D video coding technology [1] faces higher requirements. In order to develop and popularize 3D video coding technology, the ITU Video Working Group VCEG (video coding experts group) and MPEG (moving picture experts group) co-sponsored 3D video joint group JCT-3V, and established 3D video coding standards. 3D video coding based on H.265 / HEVC (High Efficiency Video Coding) [2] is extended to 3D-HEVC [3], which is a new generation of 3D video coding standard.

Multi-view video plus depth (MVD) is a standard coded data format of 3D-HEVC, which can provide users with free choices of user's viewing point without transmitting all viewpoint information. In the MVD, the depth map is used to synthesize the virtual viewpoint rather than being directly visible to the viewer. The quality of the encoded edges at the sharp edges determines the accuracy of the synthetic virtual viewpoint. The block-based prediction algorithm, which is the same as the texture image, has significant distortion at the edges. Therefore, on the one hand, the MVD inherits the HEVC quad-tree coding structure and the traversal process of the prediction mode; On the other hand, according to the characteristics of the depth map, a new intra prediction mode, namely depth modeling modes (DMMs), is proposed, which can be used to predict the depth of the mode. [4]. The adoption of DMM enhances the coding quality of edge contours in depth map, but brings sharp increase of coding complexity. The proposed algorithm reduces the coding complexity effectively with the ensurance of the quality of the depth map.

In this paper, a fast decision algorithm is proposed for the depth intra prediction mode. The Laplacian detection method is used to determine whether the DMMs is traversed or not. To achieve fast decision-making, the correlation between the intra prediction mode and the DMMs is used to reduce the number of wedgelet patterns. Experimental results show that the average coding time is greatly reduced and the video quality is almost the same when the average coding bit rate is little increased.

## II. FAST DEPTH INTRA PREDICTION MODE DECISION

### A. Depth Intra Prediction Mode Decision Process

3D-HEVC Depth Map Intra prediction includes the traditional 35 intra prediction modes and two DMMs. The intra prediction consists of 33 angle prediction modes (Figure. 1), one planar mode and one DC mode. The DMMs include wedgelet pattern mode DMM1 and contour division mode DMM4.

The intra prediction mode decision of the depth map is based on the prediction unit (PU). The optimal mode decision of the PU block is obtained through the rough mode decision, the most likely mode decision and the best mode decision. The concrete process is as follows:

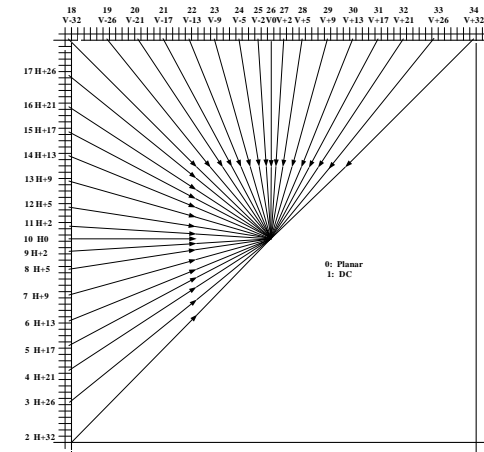


FIGURE 1. INTRA PREDICTION MODES IN HEVC (0: PLANAR 1: DC)

a) A subset of all intra prediction modes is obtained by calculating the SATD in Rough Mode Decision (RMD). The number of the subset is pre-determined to 8 for 4\*4 and 8\*8 PUs, 3 for 16\*16, 32\*32 and 64\*64 PUs.

$$Cost = SATD + \lambda \times Bits \quad (1)$$

In the calculation of the *Cost* value (1), *SATD* is the sum of the absolute values of the residuals cost after the Hadamard transform of the prediction residuals in PU block;  $\lambda$  is the Lagrangian multiplier factor; *Bits* is bits of a prediction mode Number.

b) 0~2 modes are selected from the final intra prediction modes of the left and upper blocks of the currently coded PU blocks as the most likely mode(MPMs).

c) The Rough modes, MPMs and DMMs are added to the full search list RDModeList. Then, calculate the full RD values from all the prediction modes in the list. The prediction mode with the minimum full RD value is the optimal prediction mode of the PU block. The total RD value is calculated as shown in (2).

$$J = D + l_s \times \lambda \times R \quad (2)$$

Where  $J$  is the rate distortion value,  $\lambda$  is the Lagrangian operator,  $D$  is the weighted average of the depth image and the synthetic viewpoint distortion,  $l_s$  is the scaling factor, and  $R$  is the bit rate to be consumed in each decision mode.

The DMM decision takes nearly 40% of the total intra coding time in the intra-prediction mode during the decision-making process of the depth map [5]. If the algorithm can decide whether traverse the DMMs or not and reduce the number of the wedgelet patterns then the fast decision of the intra prediction mode of depth map can be achieved with the time of 3D video coding being reduced.

#### B. Fast DMMs Decision

DMMs are used to encode steep boundaries of depth maps. Since the Dancer depth map in Figure. II has a flat or slowly varying area over a large area in the depth map, it is unlikely to choose DMMs as the optimal mode in the coding process.



FIGURE II. DEPTH MAP DANCER (1920×1088)

In this paper, the 3D-HEVC test sequence is selected as the coding target. At Table I, the content is the coding parameters of the seven test sequences. Table II shows the probability of selecting the DMMs as the optimal mode 1 of the test sequences. The probability of choosing DMMs as the optimal

mode is 0.09% and the maximum is only 5.65%. If the DMMs are always added to the RDModeList during depth intra coding, it will lead to an increase in the computational complexity of the prediction mode decision.

TABLE I. TEST SEQUENCE PARAMETERS

Test Sequences	fps	Resolution	original viewpoint
Dancer	25	1920×1088	1-5-9
Poznan_Hall	25	1920×1088	7-6-5
Poznan_street	25	1920×1088	5-4-3
GT_Fly	25	1920×1088	9-5-1
Kendo	30	1024×768	1-3-5
Ballons	30	1024×768	1-3-5
Newspaper	30	1024×768	2-4-6

TABLE II. DMMS STACTICAL RATE AS BEST MODE

Test Sequence	QP=45	QP=42	QP=39	QP=34
Dancer	0.88%	0.83%	0.93%	1.41%
Poznan_Hall	0.10%	0.09%	0.23%	0.37%
Poznan_Street	0.52%	0.38%	0.82%	2.69%
GT_Fly	0.63%	1.27%	2.19%	2.03%
Kendo	0.47%	0.46%	0.94%	2.16%
Ballons	0.38%	0.63%	1.40%	3.12%
Newspaper	0.51%	0.93%	2.40%	5.65%

In the candidate list of the PU block, if the first pattern is Planar, the PU block is a flat region. Therefore, in this case, DMMs do not need to be added to the RDModeList. Hence, the number of traversal DMMs can be reduced, and the process of the intra prediction mode decision is accelerated.

When choosing DMMs as the best mode of the PU block, its variance is greater than the variances of other intra prediction modes. Hence, the standard algorithm often uses variance to determine whether traversing DMMs. The first step is to compute the variance of the brightness of the PU block and then compare it with a threshold. If the variance is greater than the threshold, the DMMs will be added to the RDModeList. The first step is to compute the variance of the brightness of the PU block. Then compare the variance with threshold. If the variance is greater than the threshold, DMMs will be added to the RDModeList. Threshold and variance are calculated as (3) ~ (5).

$$V_{th} = \max\left(\frac{QP}{8} - 1, 3\right) \quad (3)$$

$$T_{th} = V_{th} \times V_{th} - 8 \quad (4)$$

$$Var = \frac{\sum_{i=1}^n (f_i - M)^2}{n} \quad (5)$$

In (3), QP is the depth quantization step of the PU block, max represents the maximum value of the two values;  $T_{th}$  represents threshold, and represents variance.

In (5),  $f_i$ ,  $M$  and  $n$  are the brightness value, the brightness average value and the number of PU block pixels, respectively.

0	1	0
1	-4	1
0	1	0

FIGURE III. LAPLACE OPERATOR

Variance is a rough estimate of the brightness of a PU block. The larger the value is, the higher the brightness contrast is. Therefore, according to the size of the variance value, this algorithm can determine whether there is an edge of the PU block or not. But this method has a large amount of computation. In this paper, we use the Laplace operator in Figure.III to detect the edge of PU block. The proposed algorithm uses (6) to calculate  $\nabla^2 f$  (two order difference value of the PU block). If  $\nabla^2 f$  is not zero, then DDMs will be added to the RDModeList.

$$\begin{aligned} \nabla^2 f = & f(i+1, j) + f(i-1, j) + f(i, j+1) \\ & + f(i, j-1) - 4f(i, j) \end{aligned} \quad (6)$$

$f(i+1, j)$ ,  $f(i-1, j)$ ,  $f(i, j+1)$ ,  $f(i, j-1)$  and  $f(i, j)$  are brightness values.

Table. III shows the comparison of the computational complexity of both the Laplace detection method and the variance method. When the PU block is  $32 \times 32$  and  $16 \times 16$ , the Laplace detection method requires more add operations than the variance method, but the Laplace detection method does not require multiplication. For  $32 \times 32$ ,  $16 \times 16$  and  $8 \times 8$  of the PU block, we use Laplace detection method, and for  $4 \times 4$  PU block we use the variance method for detection.

### C. Fast Wedgelet Pattern

DMMs divide the PU blocks in the depth map into two non-rectangular regions. When the final segmentation pattern is determined, the block constants for each region will be calculated. The wedgelet pattern divides the PU block into two non-rectangular regions by a straight line. After traversing all the possible start points and end points of the PU block, the wedgelet mode with the least distortion will be selected as the optimal mode. When the algorithm is initialized, wedgelet list will be established containing all the wedgelet pattern.

TABLE III. COMPARISON OF OPERATION FOR VARIANCE AND LAPLACE DETECTION METHOD

Size	Variance		Laplace	
	Number of additions	Number of multiplications	Number of additions	Number of multiplications
$32 \times 32$	3073	1025	4-3600	0
$16 \times 16$	768	257	4-786	0
$8 \times 8$	192	65	4-144	0

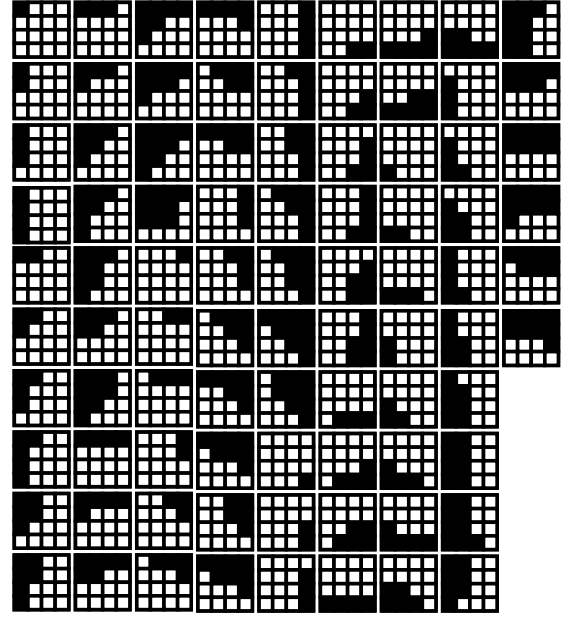


FIGURE IV.  $4 \times 4$  PU 86 WEDGELET PATTERN

In the decision process of the PU mode, each pattern in the PU wedgelet list is traversed. Therefore, the computational complexity of this algorithm is very large.

Figure IV shows 86 kinds of full search wedgelet pattern for the  $4 \times 4$  PU block. Black area and white area refer to both sides of the wedge partition line area. When the PU block becomes large, such as  $16 \times 16$ , the decision-making process of wedgelet pattern will be very time-consuming.

3D-HEVC improves the algorithm of wedgelet full-search segmentation, and uses double-layer search algorithm, which is a combination of coarse search and fine search, to reduce the computational complexity. The process of the two-layer search algorithm decision is as follow:

a) *The algorithm establishes a rough search list set.* In Figure V(a), both horizontal and vertical coordinates at every other point will be chosen as the start point and the end point. In Figure V(b), the dashed line represents a fine search containing eight reference directions around coarse segmentation pattern;

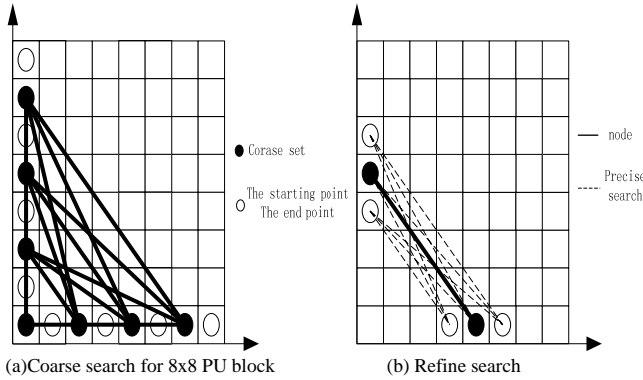


FIGURE V. STANDARD SEARCH

TABLE IV. COMPARISON OF FULL SEARCH AND DOUBLE SEARCH

PU Size	Full search	Double-Layer Search	
		Coarse search	Refined search
4x4	86	58	8
8x8	766	310	8
16x16	1349	384	8

b) Traverse all the patterns in the wedgelet coarse search set of the PU block, then the algorithm obtains the minimum distortion wedgelet pattern index;

c) By traversing the corresponding 8 kinds of precise searches, we obtain the minimum rate-distortion wedgelet segmentation pattern index. Eventually, the minimum distortion wedge segmentation mode is the final segmentation mode of the PU block.

Comparing with the full search algorithm, the double-layer search algorithm reduces both the number of wedgelet patterns and the computational complexity. Table III shows a comparison of the number of these two methods traversing the split pattern.

As we can see from Table IV, the double-layer search algorithm still needs to traverse wedgelet segmentation patterns. So the algorithm complexity is still large. Considering the wedgelet patterns have a strong correlation with texture features, correlation can be used to reduce the number of wedgelet segmentation patterns. If the best mode of the intra prediction is a certain angle mode, only the wedge division pattern related to the angle pattern can be searched. Based on the correlation, this paper proposes an improved algorithm for fast wedgelet segmentation decision. The decision-making process is as follows:

a) There is the set of full wedgelet search and wedgelet patterns are related to intra angle mode. In Figure.VI, the intra prediction mode 7 corresponds to four types of wedgelet patterns, and the intra prediction mode 20 corresponds to three kinds of wedgelet patterns.

b) After a rough search and the most likely pattern search, RdModelList is generated. According to the angle mode in the list, the corresponding wedgelet pattern is traversed. Then, the least rate-distortion wedgelet segmentation mode becomes the final wedgelet pattern mode of the PU block.

After above procedures, the complexity of the calculation is greatly reduced as shown in Table V.

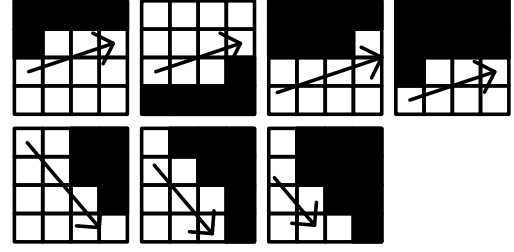


FIGURE VI. ANGLE 7 AND 20 MODE CORRESPONDING TO THE WEDGELET PATTERNS

TABLE V. THE NUMBER OF WEDGELET PATTERNS IN DMM1 FOR STANDARD SEARCH AND FAST ALGORITHM

PU Size	Double-Layer	Fast search
4x4	58+8	13-21
8x8	310+8	113-166
16x16	384+8	8-54

Double-Layer search is the standard search method in HTM-13, and the fast search is the proposed algorithm.

#### D. Fast depth Intra Mode Decision

Based on above analysis, this paper proposed a fast depth intra mode decision algorithm. The flow chart is shown in Figure. VII.

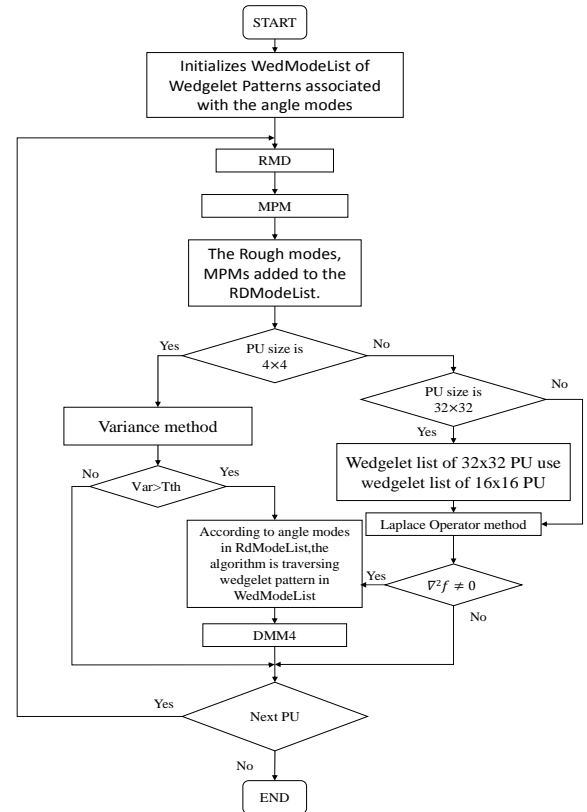


FIGURE VII. FLOW CHART OF THE PROPOSED FAST DMM SELECTION ALGORITHM

### III. EXPERIMENTAL RESULTS

In this paper, the improved algorithm is implemented in HTM-13.0 [6], and the standard test sequence is used as the coding target. Test conditions [7] are based on Tektronix image quality analyzer PQA600A server. The evaluation criteria were measured using BD-PSNR, BDBR [8] and DT (%) [9] values, where BD-PSNR represents the difference in PSNR-Y between the two methods at a given code rate, BDBR (%) indicates the bit rate savings of the two methods under the same objective quality and DT (%) refers to the decreased time of the algorithm. Positive and negative values represent the time increment and decrement respectively compared with the algorithm in HTM-13.0.

TABLE VI. COMPARISON OF TIMING SAVING FOR DEPTH MAP

Sequence	QP=48	QP=45	QP=42	QP=39
Balloons	21.7%	28.0%	29.4%	32.1%
Dancer	17.8%	18.9%	22.5%	28.5%
Newspaper	25.7%	28.1%	29.7%	33.9%
PoznanHall	14.9%	14.4%	21.0%	29.8%
PoznanStreet	16.6%	20.4%	25.1%	33.3%
Kendo	23.9%	26.2%	29.1%	33.0%
GT_Fly	14.7%	15.8%	23.1%	34.4%

TABLE VII. BDBR AND BD-PSNR RESULTS

Sequence	BDBR	BD-PSNR
Balloons	1.31%	0.07
Dancer	1.51%	0.29
Newspaper	1.02%	0.04
PoznanHall	0.63%	0.21
PoznanStreet	0.32%	0.32
Kendo	0.84%	0.18
GT_Fly	0.27%	0.05

Table VI and Table VII show the reduction ratio of the depth image encoding time and the ratio of the reduction of BDBR and BD-PSNR, respectively. In this paper, when the bit rate increases by 0.84%, the average encoding time is reduced by 24.8%. The test data show that the proposed algorithm can effectively reduce the complexity of the depth model. According to the reduction of time, time for large complexity sequences, such as "Newspaper", "Ballons" and "Kendo", is generally longer than the sequences with low complexity such as "Poznan\_Hall" and "Poznan\_street", while the loss of its coding efficiency is negligible. Experimental results show that, for large complexity sequences, the depth of DMM decision process can be skipped and the coding complexity is greatly reduced.

### IV. CONCLUSION

In this paper, an improved intra-prediction mode decision algorithm is proposed, which makes full use of the correlation

among the prediction modes and reduces the number of traversals in the prediction mode. By classifying the number of wedgelet pattern in DMM1, the number of wedgelet pattern is reduced. Compared with HTM-13.0, the proposed algorithm reduces the complexity of the 3D-HEVC encoder with little increase in the coding bit rate and low video quality loss.

### ACKNOWLEDGMENT

This work is supported by the National Natural Science Foundation of China (No.61102131).

### REFERENCES

- [1] Stankowski, J., et al. "3D-HEVC extension for circular camera arrangements." 3dtv-Conference: the True Vision - Capture, Transmission and Display of 3d Video IEEE, 2015.
- [2] Sullivan, G. J., Ohm, J., Han, W. J., & Wiegand, T. (2012). Overview of the high efficiency video coding (hevc) standard. IEEE Transactions on Circuits & Systems for Video Technology, 22(12), pp .1649-1668..
- [3] Muller, K., Schwarz, H., Marpe, D., Bartnik, C., Bosse, S., & Brust, H., et al. (2013). 3d high-efficiency video coding for multi-view video and depth data. IEEE Transactions on Image Processing, 22(9), 3366-3378.
- [4] LIU H,JIA J. Depth modeling mode coding and decoding method and video codec: WO,2014114168 A1. 2014.
- [5] TAKESHI T. JCT3V-I0110 Lookup table size reduction in DMM1 . Japan: ISO/EC,2014.
- [6] JCT3V. HTM-13. [EB/OL]. (2015-02-11)[2016-07-11]. [https://hevc.hhi.fraunhofer.de/svn/svn\\_3DVCSoftware/tags/HTM-13.0/](https://hevc.hhi.fraunhofer.de/svn/svn_3DVCSoftware/tags/HTM-13.0/).
- [7] KARSTEN M,ANTHONY V. JCT3V-G1100 Common test conditions of 3DV core experiment,US: ISO/EC,2014.
- [8] BJONTEGAARD G. Calculation of average PSNR difference between RD-curves. ITU-Telecommunications Standardization Sector Study Group 16 Question6 Video Coding Experts Group(VCEG) 13th Meeting. Austin, US,ITU,2001: 1-4.
- [9] XIE Hong,WEI Lisha,XIE Wu. A fast coding unit size decision algorithm for the depth map based on texture in 3D-HEVC. Applied Science and Technology,2016,43(2): 14-18.