International Journal of Multimedia and Ubiquitous Engineering
Vol.11, No.12 (2016), pp.171-180
http://dx.doi.org/10.14257/ijmue.2016.11.12.16

A Fast Intra Prediction Algorithm for DMM Mode in Depth Map Coding

Zhenyan Sun, Jianjun Lei, Xuguang Mei*, Jinhui Duan and Lele Li

Tianjin University, School of Electronic Information Engineering
meixg@tju.edu.cn

Abstract

As the extension of High Efficiency Video Coding (HEVC) for 3D video coding, 3D-HEVC achieves high efficiency for the compression of the multi-view videos plus depth (MVD) format. In order to ensure the performance of depth map coding, a new depth intra coding tool called Depth Modeling Mode (DMM) is introduced. However, the process of DMM significantly increases the computational complexity of depth map coding due to the blind traversal of all wedgelet partitions. In this paper, a fast intra prediction algorithm for DMM mode in depth map coding is proposed. In the first step, the unnecessary DMM mode is skipped by judging whether the best prediction mode in Rough Mode Decision (RMD) is DC mode. In the second step, the direction information represented by a permitted angle range is acquired based on the best prediction mode achieved in RMD. In the third step, a searching subset is obtained based on the direction information and the position information represented by the coordinates of the pixel with biggest depth value change in each boundary of PU. Then the patterns within the searching subset are tested by view synthesis optimization (VSO) to find the minimum distortion partition. Compared with the coarse-refinement algorithm, the proposed algorithm shows significant time saving with acceptable performance loss.

***Keywords**: 3D-HEVC, Depth map coding, DMM, Fast intra prediction, Wedgelet*

1. Introduction

Nowadays, 3D video technology has been used more and more widely. How to further improve video coding efficiency has drawn more attention. A lot of research on 3D video coding are underway.

3D video coding methods can be classified into two categories based on the format of video representation. In the first category, the coding methods are based on multi-view video (MVV) format. In order to compress MVV effectively, the extension of H.264/AVC [1] named multi-view video coding (MVC) [2] has been adopted, which introduces inter-view prediction and results in considerable coding performance. As the next generation video coding standard, High Efficiency Video Coding (HEVC) [3-4] achieves better coding efficiency compared with the conventional H.264/AVC, which benefits its multi-view extension (MV-HEVC). MV-HEVC encodes different views with the HEVC standard and also uses inter-view prediction, which significantly improves the coding performance. Although the coding methods for MVV can offer good 3D perception, it has limitations on 3D video acquisition and coding efficiency.

The second category of 3D video coding is proposed depending on multi-view video plus depth (MVD) [5] format. Compared with MVV, MVD introduces the corresponding depth video sequences. At the receiver side, the MVD based video system can generate arbitrary intermediate views by means of depth image based rendering (DIBR). Depth videos in the MVD format have unique characteristics which are different from texture videos. Therefore, the 3D video coding standard 3D-HEVC [6] is developed, which is the state-of-the-art video coding standard for the compression of MVD format.

ISSN: 1975-0080 IJMUE

The coding performance of depth map deeply affects the quality of synthesized view rendered on the decoding side, which is important in 3D video coding [7]. In actual applications, depth map can be regarded as texture video which only contains the luminance component, thus the encoding mechanism for depth map can be inherited from texture video [8]. However, the characteristic of depth map is quite different from that of texture video. The traditional intra prediction modes do not suit edge regions, and significant coding artifacts at sharp edges may be observed in synthesized intermediate views. As a result, many new tools are proposed for depth coding in 3D-HEVC, one of them is DMM [9]. However, the process of DMM is very time-consuming because of the blind traversal of all wedgelet partitions. It's meaningful to research fast algorithms for depth video coding.

Many fast algorithms for depth video coding have been proposed in recent years. Shen *et al.* proposed a low complexity depth coding algorithm, which took advantage of correlations between depth maps and color videos, science there are strong correlations between depth maps and texture videos [10]. In [11], a fast mode decision algorithm for depth video coding was proposed. The algorithm showed significant encode time saving by utilizing inter-view and inter-component correlations. Chung *et al.* [12] proposed a fast quadtree structure determination scheme, which could terminate the quadtree structure determination for coding tree units as early as possible and achieve significant encoding-time saving. In [13], all PUs were classified into two classes, parent PUs and children PUs. Based on the connections between parent PUs and children PUs, the search process of the children PUs could be skipped to achieve a time-saving effect. In [14], Sanchez *et al*. saved coding time by skipping the blocks with homogeneous regions in DMM. Blocks were classified as edges or homogeneous regions by using a texture descriptor. The results reported that the solution was able to provide time reduction with negligible BD-rate increasing. In [15], Zhang *et al*. classified wedgelet patterns into the subsets according to the angle of prediction modes. The algorithm utilized the relation between traditional prediction modes and DMM to reduce the searching subset. Although the algorithm achieved considerable time saving, the BD-rate increasing was non-negligible. In [16], Chun-Su analyzed the value of RD cost to determine whether DMMs can be skipped. Experimental results showed that the proposed algorithm could reduce encoding complexity significantly without coding performance loss. In this paper, a fast intra prediction algorithm is proposed for DMM mode. The algorithm skips unnecessary DMM mode for homogeneous PU firstly. Then the position information and the direction information are combined to simplify the searching subset in the process of selecting the optimal wedgelet pattern. Experimental results demonstrate that the proposed algorithm can save significant time for DMM prediction.

The rest of the paper is organized as follows. The background is presented in Section 2. Section 3 discusses the proposed algorithm in detail. Experimental results of the proposed algorithm are shown in Section 4. Section 5 provides conclusions.

2. Background

As shown in Figure 1, there are 35 traditional intra prediction modes used for depth intra coding in HEVC, which include a DC mode, a planar mode, and 33 angular intra prediction modes [17]. The new prediction mode DMM for PU with sharp edge can be divided into two categories according to the different partition: DMM Mode 1 (wedgelet partition) and DMM Mode 4 (contour partition) [18]. DMM Mode 1 adapts some search algorithms which are based on wedgelet patterns to select the optimal matching pattern.

Wedgelet partition approximately divides a depth PU into two non-rectangular regions, which are separated by a straight line consisting of a start point and an end point on different PU borders. Region value information in each segment is represented by a constant value called constant partition value (CPV). CPV utilizes the mean value of the original depth samples in each region. The whole partition information is stored as a partition pattern represented by an array of binary values. As illustrated in Figure 2, for a depth PU segmented by a separation straight line, the two regions are represented by P_1 and P_2 [18]. Two extreme points called start point S and end point E are on different borders. Depending on its location relative to the separation line, each complete sample is assigned to only one segment. Then, a binary partition pattern of the block is derived.

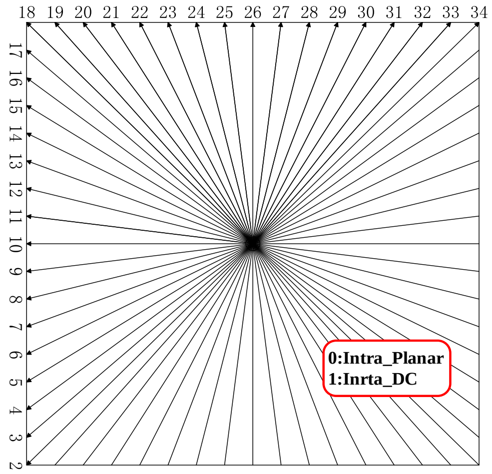


Figure 1. Conventional HEVC Intra Prediction Modes

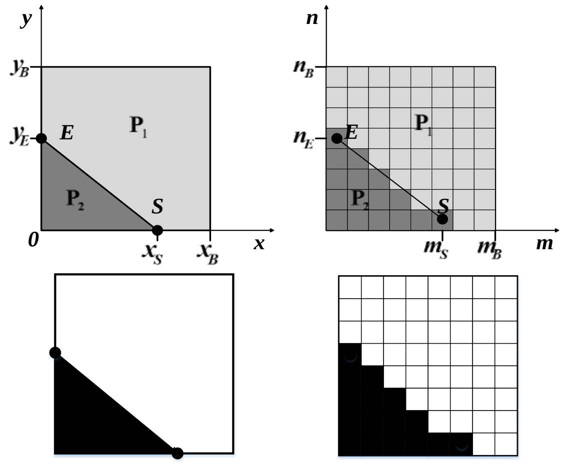


Figure 2. Wedgelet Partition of a Depth PU

When making the full search, a lookup list of wedgelet partition patterns should be initialized at the beginning of encoding. The resolution for the start and end positions are different for PUs with different sizes (for 32x32 PU with an accuracy of two samples, 16x16 PU with full-sample accuracy, and for 8x8 and 4x4 PUs with half-sample accuracy). The best matching wedgelet partition is obtained by traversing all the patterns. During that search process, the main principle is whether the wedgelet partition yields the minimum distortion between the original depth signal and the estimated approximation. Once DMM Mode 1 is finally selected, only the partition information is necessary to be transmitted in the form of bits. In order to save more encoding time, a coarse-refinement algorithm is proposed. The difference between full search and coarse-refinement is that the coarse-refinement algorithm introduces a coarse set and its corresponding node set.

Although the search process of coarse-refinement algorithm is simplified, the number of patterns need to be tested is still large. There are also some unnecessary wedgelet patterns included in the search process. Consequently, a fast intra prediction algorithm for DMM mode in depth map coding is proposed to simplify the search process.

3. Proposed method

A fast intra prediction algorithm based on a three-step process for DMM mode in depth map coding is proposed. First, the DMM mode is skipped selectively by judging whether the best prediction mode is DC mode. The best prediction mode refers to the mode with minimum cost in RMD. Second, the direction information represented by a permitted angle range is achieved based on the best prediction mode. Third, a searching subset is obtained by utilizing edge detection and the permitted angle range. Finally, the patterns within the searching set are tested by VSO to find the minimum distortion partition.

3.1. Unnecessary DMM Mode Skip

As introduced in Section 2, 35 intra prediction modes are used for depth intra coding in HEVC, which includes a DC mode, a planar mode, and 33 angular intra prediction modes. Before DMM, the process of RMD should be conducted in 3D-HEVC, and 3 or 8 modes are selected by calculating the Sum of Absolute Transformed Difference (SATD). It is proved that the mode with minimal SATD in the RMD process has strong relevance with the texture of PU [19]. The proposed algorithm selects the mode with minimal SATD in RMD as the best prediction mode. The DC mode is mainly suitable for the homogeneous regions. If the best prediction mode is DC mode, the PU may be homogeneous and the probability of adopting DMM at last is very small. As a result, the DMM mode can be skipped.

3.2. Acquirement of Permitted Angle Range

In this work, the direction information can be obtained by inheriting the best prediction mode achieved in RMD process and extending the best angular mode to a permitted angle range. As illustrated in Figure 2, the separation line of wedgelet partition also has strong relevance with the sharp edge of the PU. It can be concluded that the direction of the final optimal wedgelet partition is close to the direction of the best prediction mode. In order to improve the accuracy of prediction, a permitted angle range is obtained by calculating the corresponding permitted mode set Ω_p as follows:

$$\Omega_p = \left\{ n_i \,\middle|\, |n_i - n_b| \le n_t \right\}. \tag{1}$$

The symbol $\boldsymbol{n}$ represents the index of the mode, the subscript $\boldsymbol{i}$ represents the permitted prediction mode, the subscript $\boldsymbol{b}$ represents the best prediction mode, the subscript $\boldsymbol{t}$ indicates the threshold, the value of $\boldsymbol{n_t}$ is confirmed by experimental results in Part 4. Calculated by (1), the $\boldsymbol{n_t}$ in Ω_p may be out of the range from 2 to 34. In this case, it is further defined as follow:

$$n_i = \begin{cases} n_i + 32, & if \quad n_i \le 1 \\ n_i - 32, & if \quad n_i \ge 35 \end{cases}. \tag{2}$$

The value of $\boldsymbol{n_t}$ in the set Ω_p is updated according Eq. (2). Then, the permitted angle range is obtained based on Ω_p. For example, if $\boldsymbol{n_b}$ = 10 and $\boldsymbol{n_t}$ = 4, the permitted mode set will be [6, 14], which can be illustrated as Figure 3. The permitted angle range obtained as above will be used in the next step.

It should be note that in 33 angular modes, one mode corresponds to one angle. If the best prediction mode is planar mode without the direction information, the process of acquiring permitted angle range will be skipped.

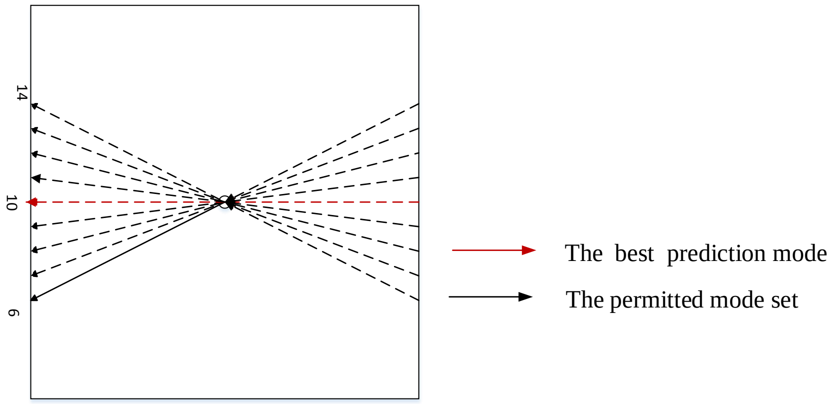


Figure 3. The Permitted Mode Set

3.3. Obtainment of Optimal Wedgelet Pattern

In this part, the position information is first achieved, which is represented by the coordinates of the pixel with biggest depth value change in the boundary of PU. Then, the position information and the direction information obtained in Section 3.2 will be combined to acquire the searching subset.

First, the pixel coordinates with the maximal gradient change in each boundary of PU are acquired by edge detection. As shown in Figure 2, the points with the maximal gradient change can be seen as the start or end point in the corresponding partition pattern. Thus, the possible separation lines can be represented by these points as shown in Figure 4. In the figure, one possible separation line corresponds to one wedgelet pattern.

Then, the angles of the possible separation lines are calculated respectively. If the achieved angle belongs to the permitted angle range, the corresponding wedgelet pattern is added to the searching subset. The search subset is derived by traversing all the possible wedgelet patterns.

Finally, the patterns in the searching subset are tested by VSO to get the one with minimum distortion and its corresponding node subset can also be got by extending the range of [-1, 1] for the start and end point. All patterns in the node subset are tested to select the final optimal one.

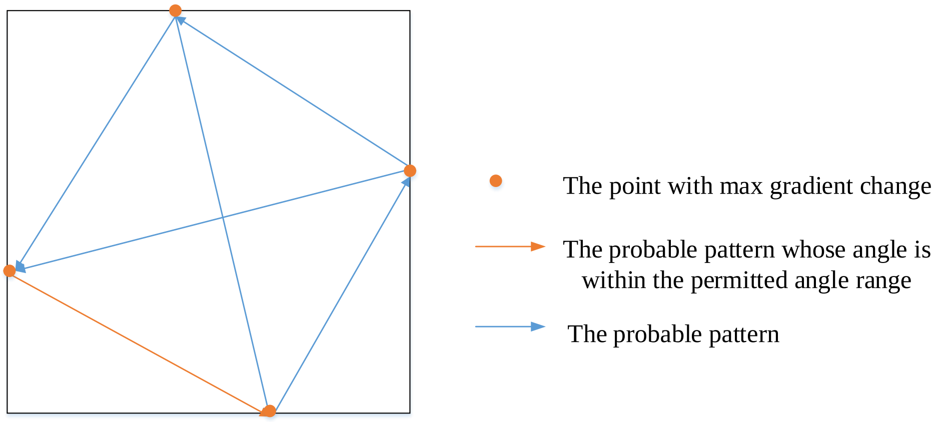


Figure 4. The Possible Separation Lines Based on Edge Detection

3.4. The Overall Flow

The overall flow chart of the proposed method is illustrated in Figure 5. First, the given PU is judged by whether it meet the condition of skipping. The skipping condition includes two cases: 1) the best prediction mode is DC mode as described in Section 3.1; 2) there are no points with biggest depth value change in edge detection as described in Section 3.3. Second, a permitted angle range is obtained based on the best prediction mode achieved in RMD. Finally, a searching subset is obtained by combining the direction information and the position information. All the patterns within the searching subset and its corresponding node set are tested by VSO to find the final optimal wedgelet pattern.

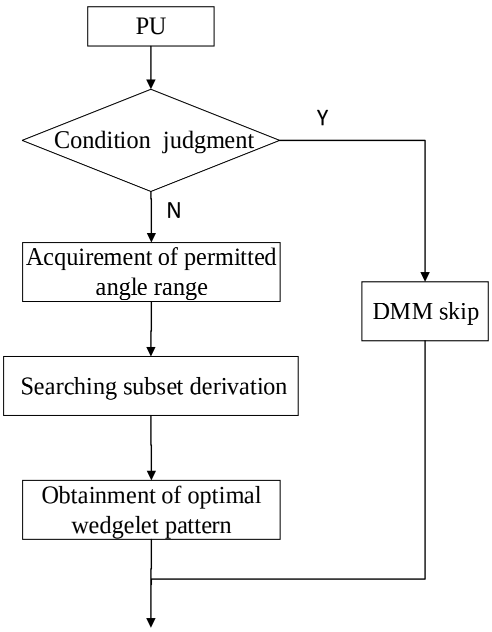


Figure 5. The Overall Flow Chart of the Proposed Method

4. Experimental Results

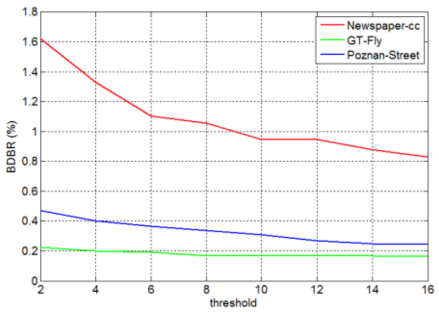
We implemented the proposed algorithm on the 3D-HEVC reference software HTM 10.0 [20]. Three standard test video sequences are tested including Newspaper (1024×768), Poznan_Street and GhostTownFly (1088×1920) [21]. 50 frames are tested in each sequence based on common test conditions (CTC) and all-intra encoder configuration. The settings of QP pairs (for texture and depth) for the experiments are (25, 34), (30, 39), (35, 42), and (40, 45). The proposed algorithm is compared with coarse-refinement algorithm.

The coding performance is evaluated in terms of Bjontegaard delta bitrate (BDBR) and Bjontegaard delta Peak Signal-to-Noise Rate (BDPSNR) [22-24]. The total encoding time saving $\triangle \boldsymbol{T_s}$ is computed as:

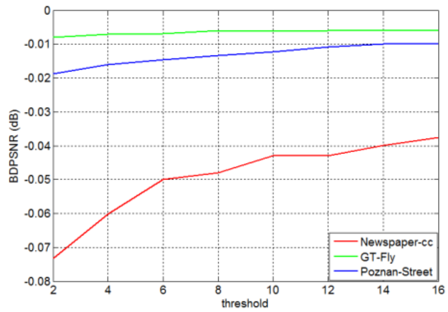
$$\Delta T_s = \frac{T_r - T_p}{T_r} \times 100\% \quad (3)$$

Where the subscript ***r*** represents the reference algorithm; ***p*** denotes the proposed algorithm. To confirm the value of the threshold $\boldsymbol{n_t}$, a statistical analysis has been performed. 17 frames of each sequence are tested. In order to show the performance intuitively, the curves of coding performance for Newspaper, Poznan_Street and

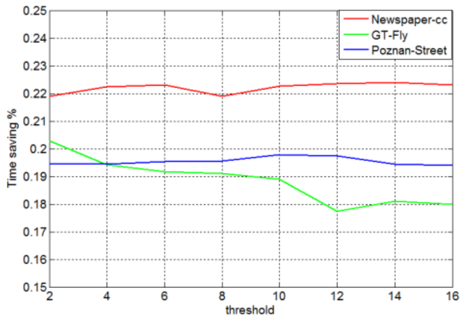
GhostTownFly are plotted in Figure 6.The BDBR and BDPSNR for synthesized view against the coarse-refinement algorithm are shown in Figure 6 (a) and Figure 6 (b), respectively. Figure 6 (c) shows the change of time saving with the vary of $\boldsymbol{n_t}$. As shown in the figure, the quality of synthesized views improves with the increase of $\boldsymbol{n_t}$, and the total encoding time saving fluctuates slightly with the increase of $\boldsymbol{n_t}$.



(a)



(b)



(c)

Figure 6. Encoding Performance Vary with N_t. (a) BDBR Vary with N_t. (b) BDPSNR Vary with N_t. (c) Total Encoding Time Saving Varies with N_t.

Considering the trade-off between the encoding time and the quality of synthesized views, we define $\boldsymbol{n_t}$ as 10 and the final experimental results are shown in TABLE 1. As shown in the table, compared with the coarse-refinement algorithm, the proposed algorithm reduces the total encoding time about 21.05% on average. Meanwhile, the BDBR for synthesized views increases 0.497% on average, and the BDPSNR for synthesized views decreases 0.0212 dB on average.

Table 1. Experimental Result Compared with Coarse-Refinement

	Coarse-Refinement Vs. Proposed		Time saving
Sequence	BDBR	BDPSNR	
Newspaper	0.986%	-0.0446dB	24.15%
GhostTownFly	0.212%	-0.0076dB	22.21%
Poznan_Street	0.286%	-0.0114dB	16.80%
Average	0.497%	-0.0212dB	21.05%

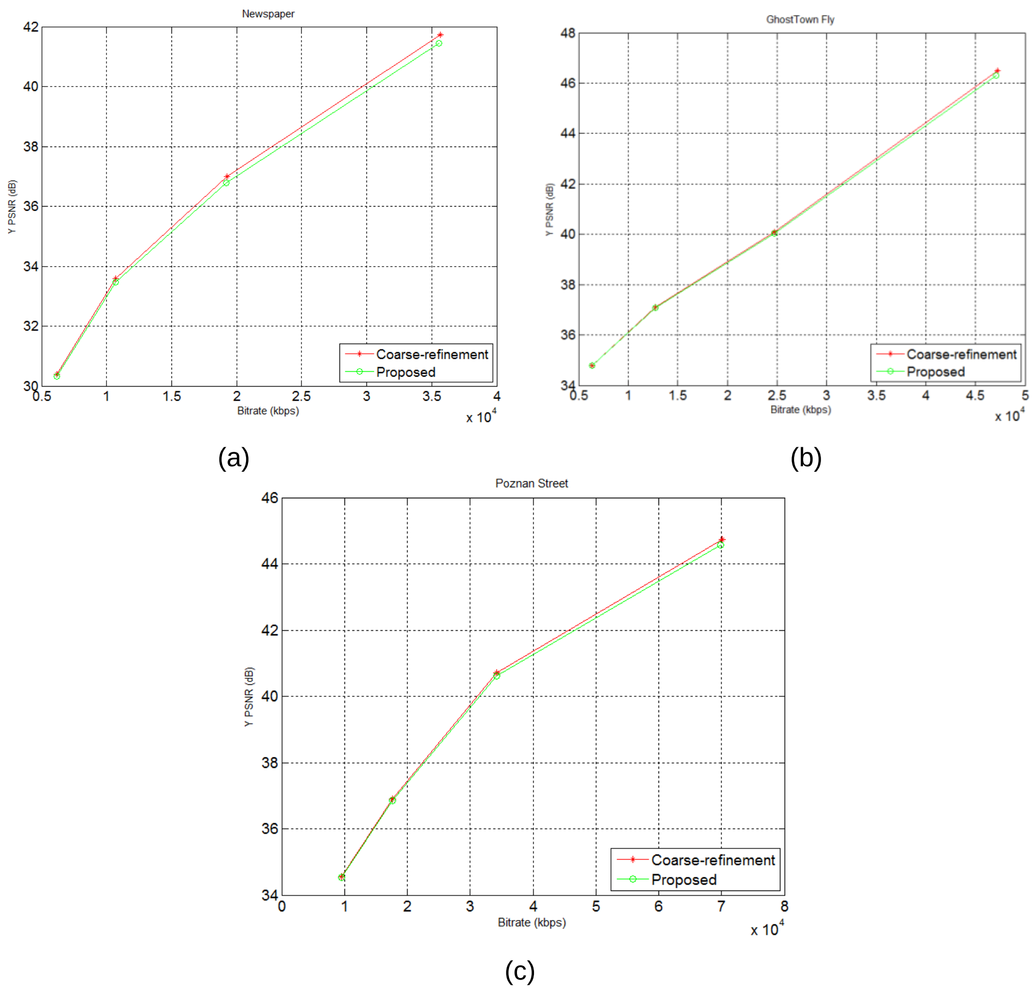


Figure 7. RD Curves for Synthesized Views

In order to show the performance intuitively, the RD curves as synthesized view PSNR against the total bit rate for coarse-refinement and the proposed algorithm are plotted in Figure 7. It can also be concluded from the RD curves that the proposed algorithm almost has the same RD performance with the coarse-refinement algorithm. According to the results, the proposed algorithm achieves significant encoding time saving than coarse-refinement algorithm with acceptable performance loss for synthesized views.

4. Conclusion

The process of DMM in 3D-HEVC is time-consuming because of the blind traversal. In this paper, a fast intra prediction algorithm for DMM mode is proposed. The algorithm first skips the unnecessary DMM mode by judging whether the best prediction mode is DC mode. Then, the direction information which represented by a permitted angle range is acquired. Finally, the direction information and the position information of a given PU are combined to simplify the process of searching the final optimal wedgelet pattern. Experimental results show that the proposed algorithm can achieve significant encoding time saving with acceptable performance loss for synthesized views in all intra case.

References

[1] ITU-T and ISO/IEC JCT 1, “Advanced video coding for generic audiovisual services”, ITU-T Recommendation H.264 and ISO/IEC 14496-10 (MPEG-4 AVC), Version 10, **(2010)**.
[2] V. T. Wiegand and G. J. Sullivan, “Overview of the stereo and multiview video coding extensions of the H. 264/MPEG-4 AVC standard”, Proceedings of the IEEE, vol. 99, no. 4, **(2011)**, pp. 626-642.
[3] G. J. Sullivan, J. Ohm, H. W. Jin and T. Wiegand, “Overview of the High Efficiency Video Coding (HEVC) standard”, IEEE Transactions on Circuits and Systems for Video Technology, vol. 22, no. 12, **(2012)**, pp. 1649-1668.
[4] F. Bossen, B. Bross, K. Sühring and D. Flynn, “HEVC complexity and implementation analysis”, IEEE Transactions on Circuits and Systems for Video Technology, vol. 22, no. 12, **(2012)**, pp. 1685-1696.
[5] P. Merkle, A. Smolic and T. Wiegand, “Multi-View Video Plus Depth Representation and Coding”, 2007 IEEE International Conference on Image Processing, vol. 1, **(2007)**, pp. 201-204.
[6] S. Yunseok and H. Yo-Sung, “Unified depth intra coding for 3D video extension of HEVC”, Signal, image and video processing, Springer, vol. 8, no. 6, **(2014)**, pp. 1031-1037.
[7] H. Schwarz, C. Bartnik, S. Bosse, H. Brust, T. Hinz, H. Lakshman, D. Marpe, P. Merkle, K. Müller, H. Rhee, G. Tech, M. Winken and T. Wiegand, “3D video coding using advanced prediction, depth modeling, and encoder control methods”, Picture Coding Symposium (PCS), **(2012)**, pp. 1-4.
[8] J. Lei, S. Li, C. Zhu, M.-T. Sun and C. Hou, “Depth coding based on depth-texture motion and structure similarities”, IEEE Transactions on Circuits and Systems for Video Technology, vol. 25, no. 2, **(2015)**, pp. 275-286.
[9] P. Merkle, C. Bartnik, K. Muller, D. Marpe and T. Wiegand, “3D video: Depth coding based on inter-component prediction of block partitions”, Picture Coding Symposium (PCS), **(2012)**, pp. 149-152.
[10] L. Shen, P. An, Z. Liu and Z. Zhang, “Low complexity depth coding assisted by coding information from color video”, IEEE Transactions on Broadcasting, vol. 60, no. 1, **(2014)**, pp. 128-133.
[11] J. Lei, J. Sun, Z. Pan, S. Kwong, J. Duan and C. Hou, “Fast mode decision using inter-view and inter-component correlations for multiview depth video coding”, IEEE Transactions on Industrial Informatics, vol. 11, no. 4, **(2015)**, pp. 978-986.
[12] K. Chung, Y. Huang, C. Lin and J. Fang, “Novel bitrate-saving and fast coding for depth videos in 3D-HEVC”, IEEE Transactions on Circuits and Systems for Video Technology, DOI: 10.1109/TCSVT.2015.2473296, **(2015)**, pp. 1-11.
[13] Z. Pan, J. Lei, Y. Zhang, X. Sun and S. Kwong, “Fast motion estimation based on content property for low-complexity H.265/HEVC encoder”, IEEE Transactions on Broadcasting, DOI: 10.1109/TBC.2016.2580920, **(2016)**, pp. 1-10.
[14] G. Sanchez, M. Saldanha, G. Balota, B. Zatt, M. Porto and L. Agostini, “Complexity reduction for 3D-HEVC depth maps intra-frame prediction using simplified edge detector algorithm”, IEEE International Conference on Image Processing (ICIP), **(2014)**, pp. 3209-3013.
[15] M. Zhang, C. Zhao, J. Xu and H. Bai, “A fast depth-map wedgelet partitioning scheme for intra prediction in 3D video coding”, 2013 IEEE International Symposium on Circuits and Systems (ISCAS), **(2013)**, pp. 2852-2855.
[16] P. Chun-Su, “Efficient intra-mode decision algorithm skipping unnecessary depth-modelling modes in 3D-HEVC”, Electronics Letters, vol. 51, **(2015)**, pp. 756-758.
[17] J. Lainema, F. Bossen, H. Woo-Jin, M. Junghye and K. Ugur, “Intra Coding of the HEVC Standard”, IEEE Transactions on Circuits and Systems for Video Technology, vol. 22, no. 12, **(2012)**, pp. 1792-1801.
[18] Y. Chen, G. Tech and K. Wegner, “Test Model 9 of 3D-HEVC and MV-HEVC. ITU-T SG16 WP3 & ISO/IEC JTC1/SC29/WG11 JCT3V-I1003”, **(2014)**.
[19] M. Zhang, C. Zhao and J. Z. Xu, “An adaptive fast intra mode decision in HEVC”, IEEE International Conference on Image Processing (ICIP), **(2012)**, pp. 221-224.
[20] https://hevc.hhi.fraunhofer.de/svn/svn_3DVCSoftware/tags/HTM-10.0/.
[21] D. Rusanovskyy, K. Müller and A Vetro, “Joint Collaborative Team on 3D Video Coding Extension Development of ITU-T SG 16 WP 3 and ISO/IEC JTC 1/SC 29/WG 11 JCT3V-E1100”, **(2013)**.
[22] Z. Pan, S. Kwong, M.-T. Sun and J. Lei, “Early MERGE mode decision based on motion estimation and hierarchical depth correlation for HEVC”, IEEE Transactions on Broadcasting, vol. 60, no. 2, **(2014)**, pp. 405-412.
[23] L. Shen, Z. Zhang and Z. Liu, “Effective CU Size Decision for HEVC Intra coding”, IEEE Transactions on Image Processing, vol. 23, no. 10, **(2014)**, pp. 4232-4241.
[24] G. Bjontegaard, “Calculation of average PSNR difference between RD curves”, document ITU-T VCEG-M33, **(2001)**.