COMPLEXITY REDUCTION FOR 3D-HEVC DEPTH MAPS INTRA-FRAME PREDICTION USING SIMPLIFIED EDGE DETECTOR ALGORITHM

Gustavo Sanchez, Mário Saldanha, Gabriel Balota, Bruno Zatt, Marcelo Porto, Luciano Agostini

Federal University of Pelotas - UFPEL
Group of Architecture and Integrated Circuits – GACI
{gfsanchez, mrdfsaldanha, gmbalota, zatt, porto, agostini}@inf.ufpel.edu.br

ABSTRACT

This paper presents a new mode decision for the depth maps intra-frame prediction in 3D-HEVC. The proposed technique decides if the traditional High Efficiency Video Coding-based (HEVC) intra-frame prediction should be performed or skipped. This technique is inspired by the fact that traditional intra-frame prediction may generate artifacts in the synthesized views when an edge is encoded. The Simplified Edge Detector (SED) algorithm has been proposed to classify if a block contains an edge or a nearly constant region demanding a minimum processing overhead. Through software evaluations, SED algorithm was capable to obtain an average complexity reduction of 23.8% for depth maps coding with no quality losses.

***Index Terms*—** 3D-HEVC, Complexity reduction, Depth maps, Intra-frame prediction

1. INTRODUCTION

With the growing three-dimensional (3D) encoding devices market, new 3D coding standards have been proposed and studied in the last few years. The 3D-High Efficiency Video Coding (3D-HEVC) is the most advanced emergent standard [1-2], which is based on the well-known High Efficiency Video Coding (HEVC) [3].

To reduce the necessary bandwidth for a 3D video transmission, the 3D-HEVC adopts a model based on the Multi-View plus Depth (MVD) [4][5] concept where every frame of each view is associated to a depth map. Instead of encoding all cameras presented in a scenario, MVD encodes only a subset of those cameras along with their depth maps. The 3D-HEVC decoder is capable to adaptively synthesize intermediary virtual views [5], located between the transmitted texture views using techniques such as Depth Image Based Rendering (DIBR) [6-7].

Depth maps contains characteristics highly different from texture. Depth maps are characterized by large areas of constant values (within objects) and sharp edges (objects borders) [2]. The traditional HEVC algorithms presents limited efficiency when encoding the sharp edges presented in depth maps and, if applied wrongly, can generate undesired artifacts in the synthesized views. During the coding process it is important to preserve the edges information for view synthesis since the distortion on edges may lead to mistaken interpretation of background/foreground pixels causing significant synthesis distortion and perceptional artifacts [8]. Thus, Depth Modeling Modes (DMM) [1-2] have been designed to better encode depth maps in 3D-HEVC.

In this new context, the depth coding represents 23.4% of the whole 3D coder complexity. The extra effort inserted in the 3D-HEVC encoder to deal with depth maps is undesired, since it contributes to the increase in power consumption and goes on the opposite direction of embedded-systems desired characteristics. For this reason, reducing the depth maps encoding complexity is an interesting research topic.

The contribution of this work is an optimization in 3D-HEVC depth maps mode decision by classifying the encoding block as an edge or a nearly constant region. With this classification it is possible to remove coding operations that have a low probability to be chosen as the best block match, such as DMM when encoding a nearly constant region. The proposed algorithm is named Simplified Edge Detector (SED). The rest of this paper is divided as follows: in Section 2 the 3D-HEVC intra-prediction algorithm is presented. In Section 3 the proposed algorithm to reduce the intra-prediction complexity is presented along with the statistical analysis supporting the algorithm decisions. Section 4 shows the results compared to the standard algorithms for 3D-HEVC and presents a comparison to related works. Finally, in Section 5 the conclusions of this paper are presented.

2. 3D-HEVC INTRA PREDICTION

In the 3D-HEVC emergent standard, DMM algorithms have been used in the scope of depth maps intra-frame prediction. By using DMM, a depth block is approximated by a model that partitions the block into two regions, where each region is represented by a constant value, commonly referred to as Constant Partition Value (CPV). The information required for the model are the partitioning information, specifying how the block is partitioned, and the CPV, specifying a constant value for the samples of the corresponding region [1-2]. The DMMs are integrated as an alternative to the conventional intra-prediction modes defined in HEVC standard.

In 3D-HEVC, the DMMs allows two partitioning strategies: Wedgelet and Contour partitioning. In a Wedgelet partition the two regions, labeled by P_1 and P_2, are separated by a straight line that can be described by a linear equation. However, in digital image processing discrete signal space is used. The samples of the predicted Wedgelet are mapped to a binary value which indicates the region each pixel belongs. An example of Wedgelet partition is presented in Fig. 1 (a).

In Contour partition, two arbitrarily shaped regions, namely P_1 and P_2, are used to predict the block. Each region in a Contour partition can consist of multiple parts. The separation line of a Contour partition of a block cannot be easily modeled by a geometrical function, as the Wedgelet could. The partition pattern is derived individually for each block from the reference block. An example of Contour partition is presented in Fig. 1 (b).

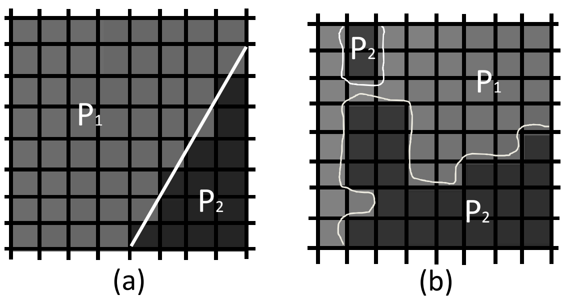

Fig. 1. Example of (a) Wedgelet Partition and (b) Contour Partition.

After obtaining the partition information, either in form of a Wedgelet or a Contour, the next information required for modeling the depth block is the CPV of the two regions P_1 and P_2. The better approximation is reached using the mean value of the original depth signal in each partition.

The traditional intra predictions algorithms are also evaluated for depth maps intra prediction (mainly to encode homogeneous regions). In fact, a mode decision algorithm is responsible to choose which mode fits best the encoded block. The best matching considers two criteria in the search process: (1) Distortion: the distortion between original and the predicted block; and (2) Rate: the number of bits necessary to represent the predicted partition. In the mode decision process, a rate-distortion (RD) search list is created where all modes in this list are evaluated for RD cost calculation. In this list, they are inserted the Most Probable Modes (MPM) for traditional HEVC intra-frame prediction [12], which are chosen from 36 HEVC intra-prediction possibilities. Also, all DMMs are added in this list.

The main problem of this process it that the RD search process does not consider the information of the distortion of the synthesized views, which is the view that will be visualized by the end-user. The RD search process takes a local decision only considering the distortion of the encoded depth map.

Based on the high complexity for evaluating this RD search list, the next section presents a statistics-based mode decision algorithm, the Simplified Edge Detector (SED), capable to classify the encoded block in edge or constant region. With this classification, the SED is capable to reduce the RD search list and, thus, reduce the complexity of the whole coding process.

3. STATISTICAL ANALYSIS & SED ALGORITHM

By analyzing the 3D-HEVC intra-prediction for depth maps, one can observe that many modes are tested with a low probability of being chosen as the best block match. These operations could be simplified by determining the class of encoding block (edge or constant region).

Fig. 2 presents a depth map for *Undo_Dancer* sequence, where edges are detached in blue blocks (1-4) and nearly constant regions are detached in red blocks (5-8). A zoom of these detached blocks are also presented in the bottom of Fig. 2.

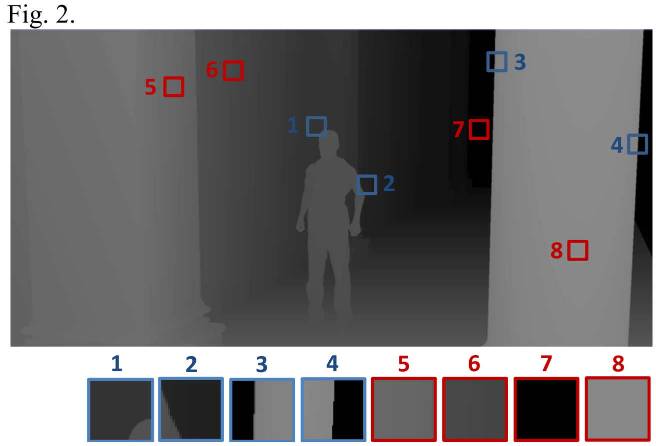

Fig. 2. Depth map with detached edges and constant area blocks.

As can be seen on the highlighted blocks on the Fig. 2, edges are presented in a block only when the corners of a block contain a significant difference among them. Motivated by this context we propose a technique to classify a block as edge or nearly constant region based on the highest difference among the corners of the encoded block (D_{max}). If the difference is significant, the block is classified as an edge, otherwise it is classified as a nearly constant region.

Subsection 3.1 presents a statistical analysis to support the blocks classification and subsection 3.2 presents the proposed Simplified Edge Detector (SED) algorithm.

3.1. Statistical Analysis

To classify the blocks, a statistical analysis has been performed using two videos of the Common Test Condition

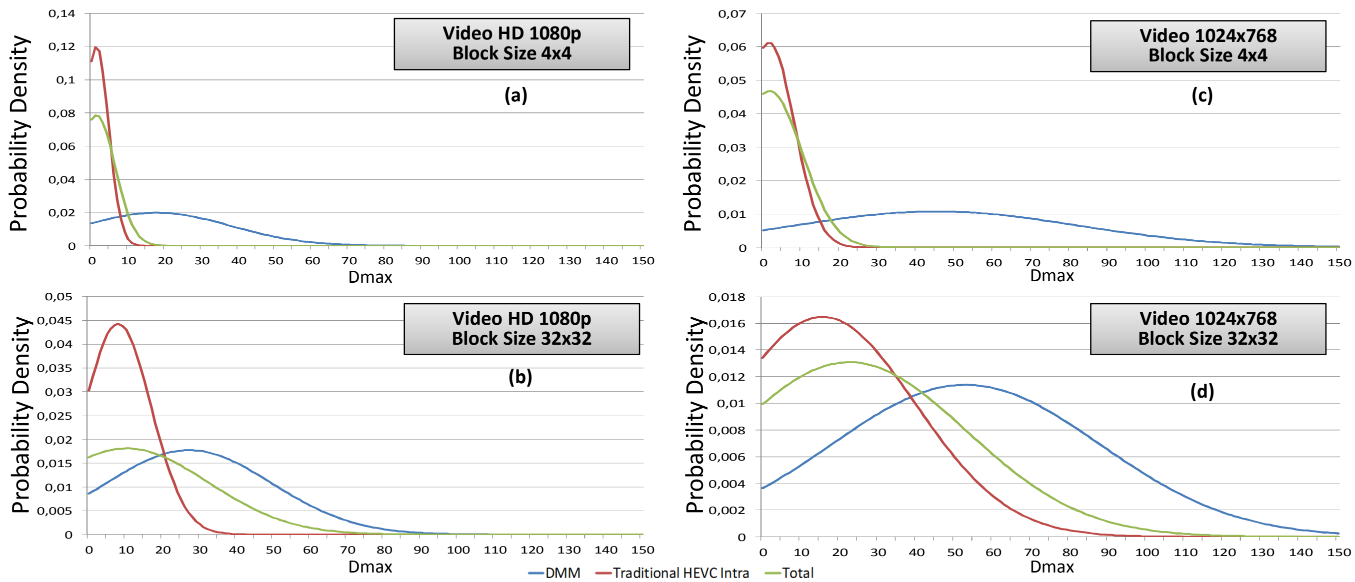

Fig. 3. Statistical Analysis.

(CCT) *Kendo* and *Poznan_Street*. These videos were simulated for the following Quantization Parameters (QP): 25, 30, 35, and 40. In this analysis, the encoded block was separated according to video resolution and block size. This separation is important because different block sizes and video resolutions have a different representation scale.

For each depth encoded block, it was stored the information of the best mode selected, DMM or traditional HEVC intra prediction. The D_{max} was also stored. The probability distribution was considered to follow a Gaussian pattern. Fig. 3 presents the probability density obtained in the proposed analysis for the following conditions: (a) block size 4x4 and video 1080p, (b) block size 32x32 and video 1080p, (c) block size 4x4 and video 1024x768, (d) block size 32x32 and video 1024x768. This study has also been performed considering the block size 8x8 and 16x16, but only the corner cases are presented. By analyzing Fig. 3 (a) and (c) (i.e. the 4x4 blocks), the HEVC intra prediction has been chosen as the best block match for low values of D_{max}. For values of D_{max} over 18, the HEVC intra prediction presents an insignificant impact. At this point, the choice of DMM coding is responsible to obtain better results. For Fig. 3 (b) and (d) (i.e. for blocks 32x32) the probability for HEVC intra prediction be chosen is only negligible for higher values of D_{max} because a 32x32 block contains more information than a 4x4 block. Moreover, the difference between the resolutions is notable because a block *N*x*N* in a given resolution has a smaller representation factor in a higher resolution. Also, through our simulation analysis, the DMMs are only chosen in 1.22% of cases for depth intra coding whereas traditional HEVC intra prediction is chosen in 98.78% of cases, when DMM could be discarded.

Considering this scenario, the DMMs can be discarded in the points that it presents a low probability to be chosen compared to traditional HEVC intra prediction without significant impact on the video quality. This statistical analysis was used to define the Simplified Edge Detector (SED) algorithm, which is presented in the next subsection.

3.2. Simplified Edge Detector (SED)

As shown in the previous subsection, an edge can be detected by comparing the D_{max}, considering a threshold that can be defined according to the block size and video resolution.

As DMMs have only a low probability to be chosen, it can be discarded when traditional HEVC intra prediction has a dominant probability without impacting the resulting intra prediction algorithms efficiency.

Through experimental analysis, the thresholds used in this algorithm were selected as the D_{max} average added to 1.5 standard deviation of the traditional intra prediction obtained in the statistical analysis for 1024x768 videos. For HD 1080p videos, the thresholds were selected as the D_{max} average added to one standard deviation. The selected thresholds are presented in Table I.

Table I. SED thresholds

Resolution \ Block Size	4x4	8x8	16x16	32x32
1024x768	12	20	34	55
1080p	8	11	16	25

The block diagram of SED algorithm is presented in Fig. 4. The SED algorithm compares the D_{max} with a threshold (according to Table I) to classify the block. When a block is classified as an edge, no simplification is performed and the HEVC intra prediction using MPM algorithm and DMMs are inserted in RD-list, which is the conventional 3D-HEVC flow. However, when the encoded block is classified as a constant region, the SED algorithm does add the DMM to the RD-list but only the traditional HEVC intra prediction, resulting in a discard of over 84% in DMM evaluation.

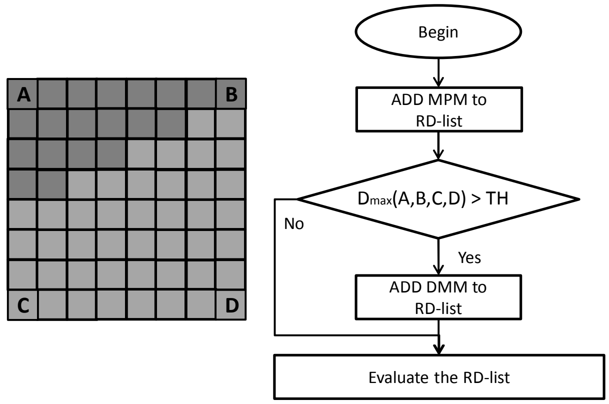

Fig. 4. Simplified Edge Detector Flowchart.

4. SIMULATION RESULTS AND COMPARISON

The proposed algorithm has been evaluated in HTM-7.0 using the Common Test Conditions (CTC) [9] for Random Access mode. The results are presented in Table II, using the well-accepted BD-rate measurement [10]. As video 0 (texture) does not have any data dependencies with depth maps, the video 0 results has not been changed. Videos 1 and 2 present a few variations because the texture coding of dependent views (i.e. all texture views except the 0) present data dependencies with depth coding. However, in average case the SED presents only a 0.027% in bit rate increase.

The most important results of this work are obtained calculating the BD-rate with the accumulated bit rate for all texture and depth views and the quality for the synthesized views. In average, the SED presented an average BD-rate saving of 0.064%. This gain is mainly related to *Newspaper_CC* sequence, where it was obtained a BD-rate saving of 2.062%. This gain occurred because the 3D-HEVC depth coding takes a local decision to encode the depth map considering only the predicted depth map distortion and the number of bits necessary to encode that depth map. In fact, the depth map is not visualized but only used for view synthesis. As the 3D-HEVC does not consider the distortion of synthesized views in the decision mode, it is possible to reduce the complexity of the decision process and even obtain a bit rate saving. The video *Newspaper_CC* has many regions containing lots of edges with a small difference in D_{max}. When applying the SED algorithm, those blocks containing small D_{max} have been chosen to be encoded using traditional HEVC intra prediction, while previously, many of these edge blocks have been encoded using DMM. On the others videos, in average, there was a 0.269% in BD-rate increase.

Analyzing the complexity of the proposed algorithm it was possible to obtain a complexity reduction of 5.9%, in average, considering the whole 3D-HEVC encoder. In fact, considering only the depth channel it was possible to obtain a complexity reduction of 23.8%.

The work [11] was the only work found in the literature proposing algorithms to reduce the complexity of depth maps intra-frame prediction in the 3D-HEVC. This work [11] presents an algorithm to reduce the complexity of DMM coding and was adopted in the HTM-8.0 version. It discards the DMM coding when the first mode in RD search list is planar mode. In [11] it was possible to reduce the complexity of the whole coder by 1.34% with a bit rate increase of 0.11% in synthesized views. The SED showed to be faster than [11] and capable to obtain a better coding efficiency. Both solutions may be combined in 3D-HEVC coder to obtain a better complexity reduction.

5. CONCLUSIONS

This work presented an approach to reduce the complexity on the depth intra-frame prediction of the 3D-HEVC emergent standard. The Simplified Edge Detector (SED) algorithm is a statistic-based algorithm capable to classify the blocks as edges or constant regions. By classifying the block in one of these two possibilities, the SED algorithm is able to reduce the complexity of the depth intra-frame by excluding the DMM evaluations on blocks classified as nearly constant region. The presented results show that SED achieves a reduction of 5.9% considering the whole encoder and 23.8% considering only the depth maps without coding efficiency losses.

Table II. BD-rate results for 3-view case under CTC, Random Access.

Videos	Video 0 (BD-rate)	Video 1 (BD-rate)	Video 2 (BD-rate)	Video Only (BD-rate)	Synthesis Only (BD-rate)	Encoding Time	Time Reduction	Depth Time Reduction
Balloons	0.000%	0.107%	0.427%	0.095%	0.282%	94.6%	5.4%	20.8%
Kendo	0.000%	-0.025%	-0.169%	-0.020%	0.153%	94.9%	5.1%	20.1%
Newspaper_CC	0.000%	-0.162%	0.222%	-0.001%	-2.062%	94.0%	6.0%	22.9%
GT_Fly	0.000%	0.172%	-0.216%	0.029%	0.173%	94.0%	6.0%	24.1%
Poznan_Hall2	0.000%	0.036%	0.260%	0.084%	0.487%	94.0%	6.0%	26.7%
Poznan_Street	0.000%	0.008%	-0.153%	-0.016%	0.217%	92.9%	7.1%	26.6%
Undo_Dancer	0.000%	0.220%	-0.123%	0.015%	0.303%	94.3%	5.7%	25.3%
Avg. 1024x768	0.000%	-0.027%	0.160%	0.025%	-0.542%	94.5%	5.5%	21.3%
Avg. 1920x1088	0.000%	0.109%	-0.058%	0.028%	0.295%	93.8%	6.2%	25.7%
Average	**0.000%**	**0.051%**	**0.036%**	**0.027%**	**-0.064%**	**94.1%**	**5.9%**	**23.8%**

6. REFERENCES

[1] G. Tech; K. Wegner; Y. Chen; S. Yea; “3D HEVC Test Model 3”. Document:JCT3V-C1005. Draft 3 of 3D-HEVC Test Model Description. Geneva, 2013.

[2] Muller, K. ; Schwarz, H. ; Marpe, D. ; Bartnik, C. ; Bosse, S. ; Brust, H. ; Hinz, T. ; Lakshman, H. ; Merkle, P. ; Rhee, F.H. ; Tech, G. ; Winken, M. ; Wiegand, T., “3D High-Efficiency Video Coding for Multi-View Video and Depth Data” IEEE Transactions on Image Processing, , 2013.

[3] H.265: High Efficiency Video Coding. Available at: <http://www.itu.int/rec/T-REC-H.265-201304-I>. Access in 2013.

[4] Merkle, P.; Smolic, A.; Muller, K.; Wiegand, “T. Efficienct Compression of Multi-view Depth Data Based on MVC”, 3DTV Conference, 2007.

[5] K. Muller, P.; Merkle, T.; Wiegand; “3-D Video Representation Using Depth Maps”, Proceedings of IEEE, 2011.

[6] A. Smolic, K. Müller, K. Dix, P. Merkle, P. Kauff, and T. Wiegand, “Intermediate view interpolation based on multiview video plus depth for advanced 3D video systems,” Proc .IEEE Int. Conf. on Image Proc., San Diego, USA, pp. 2448-2451, Oct. 2008.

[7] P. Kauff, N. Atzpadin, C. Fehn, M. Müller, O. Schreer, A. Smolic and R. Tanger, “Depth map creation and image based rendering for advanced 3DTV services providing interoperability and scalability,” Signal Process., Image Commun., Special Issue 3DTV, vol. 22, no.2, pp. 217–234, Feb. 2007.

[8] X. Zhao, Y. Chen, L. Zhang, and M. Karczewicz, "Texture Mode Dependent Coding in 3D-HEVC" ICIP, 2013.

[9] “Common test conditions for 3DV experimentation”, ISO/IEC JTC1/SC29/WG11 MPEG2011/N12745, Geneva, Switzerland, Apr. 2012.

[10] G. Bjontegaard, “Calculation of Average PSNR Differences Between RD Curves,” VCEG-M33, ITU-T SG16/Q6 VCEG, 13th VCEG Meeting: Austin, USA, Apr. 2001.

[11] Z. Gu; J. Zheng; N. Ling; P. Zhang; “Fast Depth Modeling Mode Selection for 3D HEVC Depth Intra Coding”, International Conference on Multimedia and Expo Workshops (ICMEW), 2013.

[12] L. Zhao, L. Zhang, S. Ma and D. Zhao, “Fast Mode Decision Algorithm for Intra Prediction in HEVC” Visual Communications and Image Processing (VCIP), November, 2011 IEEE.