3D-HEVC Depth Maps Intra Prediction Complexity Analysis

[1,2]Gustavo Sanchez, [1]Rodrigo Cataldo, [1]Ramon Fernandes, [3]Luciano Agostini, [1]César Marcon
[1]Pontifical Catholic University of Rio Grande do Sul – Porto Alegre, Brazil
[2]Instituto Federal Farroupilha– Alegrete, Brazil
[3]Federal University of Pelotas – Pelotas, Brazil
{gustavo.sanchez, rodrigo.cataldo, ramon.fernandes}@acad.pucrs.br, agostini@inf.ufpel.edu.br, cesar.marcon@pucrs.br

***Abstract*— This paper presents a complexity analysis of 3D High Efficiency Video Coding (3D-HEVC) depth maps intra prediction. The 3D-HEVC inserts new coding tools in depth maps intra prediction such as Depth Intra Skip (DIS), Depth Modeling Modes (DMMs) and Segment-wise DC (SDC). Therefore, it is important to understand the complexity of each module to allow the design of new complexity reduction techniques to encode the depth maps. This paper aims to guide other works to the most time-consuming tools that could be simplified to achieve a real-time design according to the encoding context.**

Keywords—3D-HEVC; Intra-prediction; Computational Complexity; Depth Maps

I. INTRODUCTION

Video coding techniques have obtained relevant advances in the last few years. These advances lead to the creation of the High Efficiency Video Coding (HEVC) standard [1], providing much better encoding efficiency than previous encoding standards. However, inserting new techniques to achieve bitrate gain increases the algorithmic complexity, making 2D video coding a challenging process.

Further advances in video coding provide extensions of HEVC to 3D video coding generating the 3D-HEVC standard [2]. In 3D videos, multiple cameras capture multiple views at the same instant. The use of many cameras increases a lot the encoding complexity when compared to 2D coding. Then, complexity reduction techniques are desirable to allow real-time processing in 3D HEVC encoders.

The 3D-HEVC improves the encoding efficiency adopting the Multiview plus Depth (MVD) data format [3], which associates a depth map to each texture view, providing the geometrical information according to the objects distance from the camera. Fig. 1 presents a (a) texture view extracted from *Balloons* video sequence and its (b) depth map. A 3D screen does not display the depth maps during a 3D video presentation. However, when applying rendering techniques at the decoder, using both texture and depth maps, it is possible to generate high-quality virtual synthesized views located between originally encoded views [3].

A depth map contains some characteristics, such as vast areas of nearly constant depth values at background or object bodies and sharp edges at objects borders. Traditional 2D-HEVC algorithms were designed focusing on characteristics of texture with smooth transitions between pixels. Thus, using these algorithms to encode 3D videos produce a low efficient compression, with decreasing the depth maps quality and consequently decreasing also the synthesized views quality. Therefore, 3D-HEVC inserted new tools for depth maps coding. Many of these tools are related to depth maps intra prediction such as Depth Intra Skip (DIS), Depth Modeling Modes (DMMs) and Segment-wise DC (SDC).

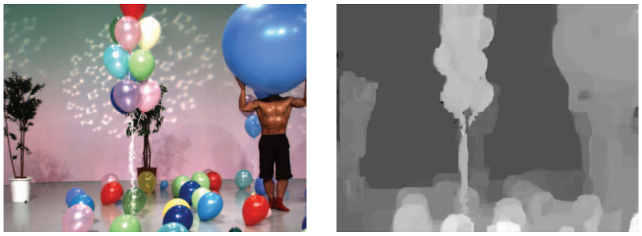

(a) (b)

Fig. 1. First frame of *Balloons* video sequence (a) texture (b) depth map.

Many works already proposed complexity reduction techniques for these new tools, such as [4]-[9], considering the high complexity associated with depth maps intra-frame prediction. Moreover, the 3D-HEVC Test Model (3D-HTM) 16.0 already applies complexity reduction techniques in the intra prediction. However, its computational complexity is still prohibitive for real-time applications.

Ahmad et al. [10] presented a time profiling of 3D-HEVC encoding at random access mode. This profiling was performed per function in the 3D-HTM C++ code, which can not be used to identify how each encoding mode affects the 3D-HEVC complexity since multiple modes can use the same function. No study under all intra case is performed in [10].

Our paper presents a complexity analysis of the tools used on 3D-HEVC depth maps intra prediction. In this paper, computational complexity is measured regarding execution time. By identifying the complexity of each tool, this paper should serve as a guide to other works that focuses on reducing 3D-HEVC intra prediction computational complexity.

II. DEPTH MAPS INTRA-PREDICTION

For a given depth block, Fig. 2 presents the dataflow model of 3D-HEVC depth maps intra prediction. The 3D-HEVC encoding process is based on the Rate-Distortion (RD) cost

978-1-5090-6113-6/16/$31.00 ©2016 IEEE

computation, where many possible encoding modules are evaluated by their RD-cost and the cost that obtained the lower RD-cost is selected as the encoding mode. The RD-cost is a value that groups the visual quality and the required bits of the block encoded by that mode.

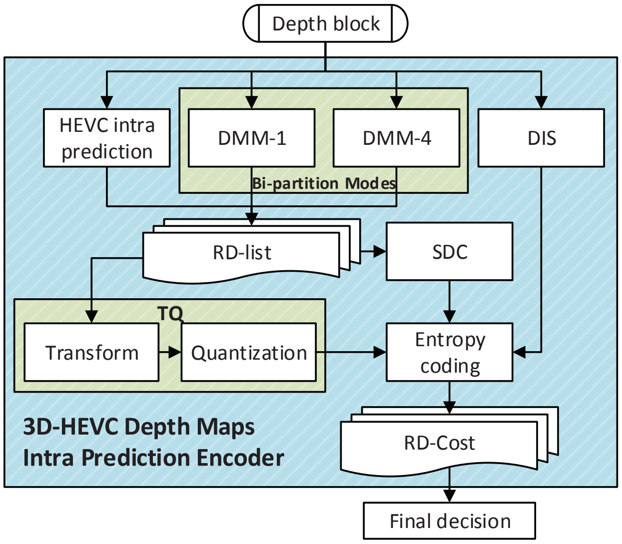


Fig. 2. Dataflow model used for the 3D-HEVC depth maps intra prediction.

The Depth Intra Skip (DIS) mode can be evaluated in parallel with HEVC intra prediction, DMM-1, and DMM-4. This mode does not require encoding residual information, and then it only encodes entropy to obtain its RD-cost.

The intra-frame prediction encoder creates an RD-list, where modes selected by HEVC intra prediction, a DMM-1 pattern, and a DMM-4 pattern can be added to be later evaluated by their RD-cost. A set of HEVC intra prediction modes, among the possible ones, are always inserted in this list (see Section II.B). The 3D-HTM 16.0 applies a complexity reduction technique [4], inserting the DMM modes into RD-list only if a condition is achieved. This condition requires that the first mode selected in HEVC intra prediction is not planar, and the encoding block variance is higher than a threshold, which is adapted according to the Quantization Parameter (QP).

Each mode inserted into the list RD-list is evaluated according to the RD-cost in the Transform-Quantization (TQ) and Segment-wise DC (SDC) flows. The TQ flow transforms and quantifies the residual information of the predicted block. The SDC flow predicts the input block by (i) a single value if a HEVC intra mode is under evaluation; or (ii) two values if any DMM is under evaluation, being one for each region. After accomplishing SDC or TQ flows, the algorithm performs the Entropy coding step to obtain the RD-cost.

The 3D-HTM 16.0 applies the technique proposed in [5]; consequently, SDC evaluation is skipped for some modes inside RD-list. This skipping is achieved if the encoding block is nearly homogeneous.

After obtaining the RD-cost of all encoding modes, the encoding mode with the lowest RD-cost value is selected for that block. The following sections describe some relevant details of DIS, HEVC intra prediction, DMM-1 and, DMM-4.

A. Depth Intra Skip (DIS)

The DIS mode considers that most of the depth maps information is smooth areas. In general, this smooth are not relevant for synthesized views quality. Therefore, DIS skips the residual coding for the smooth areas.

The DIS mode includes four different prediction modes: (1) vertical intra prediction mode; (2) horizontal intra prediction mode; (3) vertical single depth mode and; (4) horizontal single depth mode. The HEVC vertical and horizontal intra directions are used to generate the first two predictions. The vertical and horizontal intra single depth modes fill the predicted block with a single depth value that is derived from the neighbor upper and neighbor left blocks, respectively.

B. HEVC Intra Prediction

The HEVC intra prediction defines the planar mode, the DC mode and, 33 directional modes, whose directions are presented in Fig. 3(a).

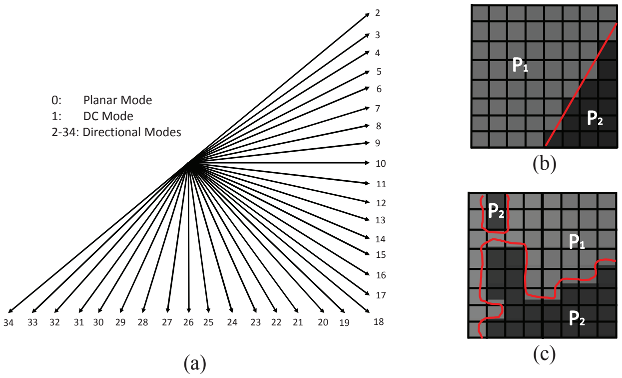


Fig. 3. (a) HEVC Intra Prediction Directions (b) wedgelet partition and (c) contour partition.

The HEVC intra prediction employs Rough Mode Decision (RMD) and Most Probable Modes (MPM) [11]. In RMD algorithm, the Sum of Absolute Transformed Differences (SATD) enables to evaluate all HEVC intra modes locally without the complete RD-cost evaluation. This algorithm inserts a few modes with the lowest SATDs ordered into the RD-list. The MPM gets the modes used in the encoded neighbor blocks and inserts them into the RD-list. Table I presents the total number of directions defined by HEVC intra prediction and the number of selected modes by RMD.

TABLE I. QUANTITY OF HEVC INTRA MODES, RMD SELECTED MODES, DMM-1 WEDGELETS AND EVALUATED WEDGELETS PER BLOCK SIZE

Block size	HEVC Intra Prediction		DMM-1	
	Total Modes	RMD Selection	Total Wedgelets	Wedgelets before Refinement
4x4	18	8	86	58
8x8	35	8	802	314
16x16	35	3	510	384
32x32	35	3	510	384
64x64	4	3	-	-

C. Depth Modeling Mode 1

The DMM-1 algorithm is based on wedgelets, where a wedgelet is a straight line that segments a block into two regions. Fig. 3(b) presents an example of a wedgelet partition.

There are many possible wedgelets as Table I describes, which varies according to the encoding block size. However, only a subset of them should be evaluated before the refinement. After the DMM-1 refinement, that evaluates at most eight wedgelets, the wedgelet that obtained the best prediction result, according to a similarity criterion, is inserted into the RD-list.

D. Depth Modeling Mode 4

The DMM-4 algorithm uses a technique called inter-component prediction to find the best contour partition. This inter-component prediction uses previously encoded information from the texture view during the depth map prediction.

The DMM-4 algorithm creates a partition dynamically from the texture information, which consists of arbitrary shapes or even disconnected regions. Fig. 3(c) exemplifies a DMM-4 partition, which is inserted into the RD-list.

III. EXPERIMENTS AND ANALYSIS

All videos in Common Test Conditions (CTC) have been evaluated under all intra case [12]. During the evaluation, the execution times spent in texture and depth encoding were stored. In this work, we will use the execution time as a metric to evaluate complexity. Table II presents the percentage of the complexity for each component per video and per QP pair (texture/depth maps) related to the total execution time. The QP defines the compression rate and influences the image quality directly. Lower QP values tend to introduce fewer image distortions, also achieving lower compression rates. The QP pairs presented in our evaluation are defined in CTC [12].

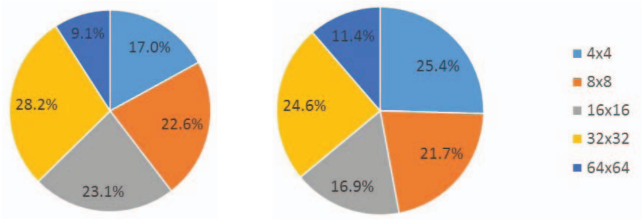
TABLE II. PERCENTUAL COMPLEXITY FOR TEXTURE AND DEPTH CODING IN THE 3D-HEVC ALL INTRA CASE.

Videos	QP = (25/34)		QP = (30/39)		QP = (35/42)		QP = (40/45)	
	texture	depth	texture	depth	texture	depth	texture	depth
Balloons	15.7	84.3	14.4	85.6	14.9	85.1	14.5	85.5
Kendo	15.9	84.1	15.2	84.8	15.4	84.6	13.4	86.6
Newspaper	13.2	86.8	11.0	89.0	11.7	88.3	11.8	88.2
GT_Fly	14.8	85.2	13.3	86.7	14.5	85.5	14.8	85.2
PoznanHall	18.2	81.8	16.2	83.8	16.8	83.2	15.6	84.4
PoznanStreet	13.6	86.4	12.2	87.8	13.7	86.3	13.5	86.5
UndoDancer	18.0	82.0	16.7	83.3	16.0	84.0	14.8	85.2
Shark	16.1	83.9	14.7	85.3	15.2	84.8	14.5	85.5

*QP = (QPtexture/QPdepth)

The complexity spent during texture coding ranges from 11.0% to 18.2%, which is much lower than depth coding complexity. It occurs in the all intra configuration because, in this case, texture coding only applies traditional HEVC intra prediction, while depth coding still uses DMMs, DIS and SDC evaluations. On average, depth maps coding is 5.8 times the texture coding complexity.

When considering the above statements, the information of the complexity of each depth maps intra prediction encoding steps and block sizes are important when designing new complexity reduction algorithm to achieve better results. In this direction, Fig. 4 presents the average complexity spent during depth maps coding on each block size for the corners cases QPs (i.e., QPs equal to 34 and 45 for depth maps). One can notice that when the QP varies, a different complexity distribution occurs.



(a) QP 34 (b) QP 45

Fig. 4. Complexity distribution according to the block size (a) QP = 34 (b) QP = 45.

Fig. 5 presents the average complexity distribution among the intra prediction encoding steps for each block size and the corners QPs of the encoding modes, complementing the previous experiment. First, notice that in Fig. 5(a) and (b) DIS and SDC evaluation are not presented because these modes are only applied for block sizes from 8×8 to 64×64. Moreover, DMM-1 and DMM-4 are not presented in Fig. 5(i) and (j) because DMMs are only applied from 4×4 to 32×32 blocks.

For 4×4 blocks, the evaluation in TQ flow is the most complex operation. Considering 8×8 to 32×32 blocks and lower QPs, DMM-1 is the most complex step, due to the high effort spent to search the initial wedgelet set. The SDC evaluation is a complex operation for all available block sizes (i.e., 8×8 to 64×64). Even being a complex operation, its complexity still arises for 64×64 blocks since SDC represents almost 50% of complexity. It occurs because the simplification applied in SDC [5] skips this evaluation if the encoding block is almost homogeneous. Since a larger block contains more samples inside it, this block tends to be heterogeneous, then, SDC should be less skipped here.

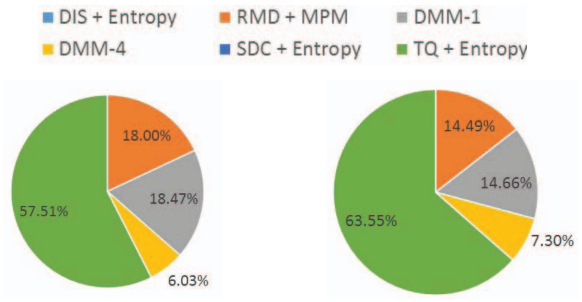
The complexity of DMM-4 and DIS predictions with entropy coding is low for almost all encoding block sizes. The DIS complexity is only higher than 10% at 64×64 encoding blocks. However, Fig. 4 demonstrates that 64×64 is the case that lower complexity is spent during the depth maps encoding. Then, reducing DMM-4 and DIS complexity tends to achieve low complexity reduction in all scenarios.

Considering the complexity profiling explained in previous paragraphs, for 4×4 blocks a complexity reduction technique should consider reducing the amount of modes inserted into RD-list. All modes inside this list should be evaluated by TQ flow, and then reducing its quantity of modes tends to achieve high complexity reduction for 4×4 blocks.

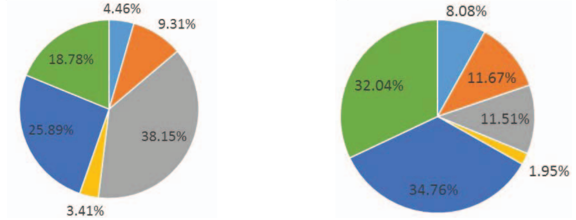
When considering 8×8 to 32×32 blocks and lower QPs, the complexity reduction techniques should target the DMM-1 to reach most impressive results. When using higher QP values, the algorithm proposed in [4] uses a higher threshold that results in DMMs more cut. Consequently, DMM-1 complexity is already reduced, then reducing DMM-1 complexity should not achieve impressive results in this case and other tools such as TQ and SDC flows evaluation should be the focus of simplification. As SDC evaluation represents almost 50% of complexity in 64×64 blocks, it should be focused on its complexity reduction techniques.

The RMD and MPM processes are responsible for a considerable amount of complexity for selecting the candidates that should be inserted into the RD-list in all available block sizes (4×4 to 64×64). These algorithms were developed for texture coding and just replicated for depth coding without modifications. Then, reducing RMD and MPM complexity by introducing depth maps characteristics in the encoding process should achieve good complexity reduction results, and few works proposed solutions in this direction.

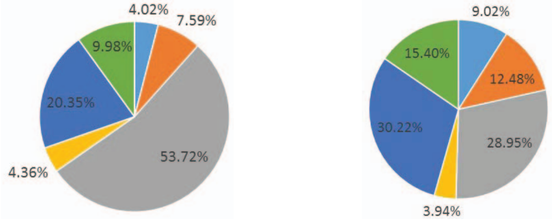
Moreover, if RMD and MPM selection reduces the quantity of modes inserted into RD-list, consequently the TQ and SDC flows complexity will also decrease since a low amount of modes should be evaluated by them.



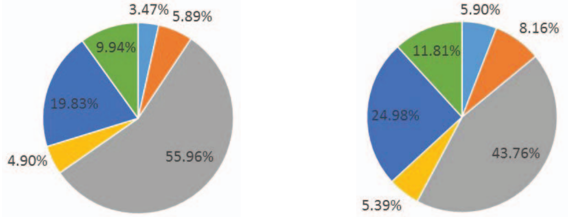
(a) block size 4×4 and QP 34 (b) block size 4×4 and QP 45



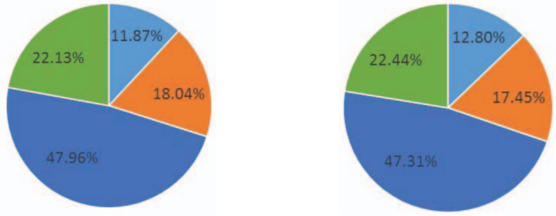
(c) block size 8×8 and QP 34 (d) block size 8×8 and QP 45



(e) block size 16×16 and QP 34 (f) block size 16×16 and QP 45



(g) block size 32×32 and QP 34 (h) block size 32×32 and QP 45



(i) block size 64×64 and QP 34 (j) block size 64×64 and QP 45

Fig. 5. Complexity distribution of each intra prediction step for different block sizes and two QPs.

Summarizing this section, one can conclude that each block size has a different complexity profiling. When designing new complexity reduction techniques for depth intra coding, it is important to notice the impact that the target tool has in the global system.

IV. CONCLUSIONS

This paper presented 3D-HEVC depth maps intra-frame prediction complexity analysis. Many new tools were specially developed for depth maps during 3D-HEVC standardization, and also demonstrated the complexity of those tools when running the 3D-HTM 16.0 under the CTCs and using the all intra configuration.

This paper showed that the depth maps encoding behavior is highly dependent on the encoding context (block size and QP). We also presented the complexity distribution among the encoding tools for the different contexts. This information allows us to conclude that an efficient complexity reduction technique developed aiming real-time systems must be context-adaptive to explore better this depth map encoding proprieties. To the best of our knowledge, this is the first work in literature that provides a complexity analysis of depth maps intra prediction in 3D-HEVC.

REFERENCES

[1] G. Sullivan, J. Ohm, W. Han, T. Wiegand. “Overview of the high efficiency video coding (HEVC) standard,” *Transactions on circuits and systems for video technology*, v. 22, n. 12, pp. 1649-1668, Dec. 2012.

[2] G. Tech et al. “Overview of the Multiview and 3D extensions of High Efficiency Video Coding,” *IEEE Transactions on Circuits and Systems for Video Technology (TCSVT)*, v. 26, n. 1, pp. 35-49, Jan. 2016.

[3] P. Kauff et al. “Depth map creation and image-based rendering for advanced 3DTV services providing interoperability and scalability,” *Image Communication*, v. 22, n. 2, pp. 217-234, Feb. 2007.

[4] Z. Gu et al. “Fast Depth Modeling Mode selection for 3D HEVC depth intra coding,” *IEEE International Conference on Multimedia and Expo Workshops (ICMEW)*, pp. 1-4, 2013.

[5] Z. Gu, J. Zheng, N. Ling, P. Zhang. “Fast Intra SDC Coding for 3D-HEVC Intra Coding,” ITU-T SG 16 WP 3 and ISO/IEC JCT 1/SC 29/WG 11, JCT3V-I0123, 9th meeting: Sapporo, JP, Jul. 2014.

[6] G. Sanchez, et al. “Complexity reduction for 3D-HEVC depth maps intra-frame prediction using simplified edge detector algorithm,” *International Conference on Image Processing (ICIP)*, pp. 3209-3213, 2014.

[7] Q. Zhang et al. **“**Fast intra mode decision for depth coding in 3D-HEVC,” *Multidimensional Systems and Signal Processing*, pp. 1-24, 2016.

[8] H. Zhang, et al. “Early determination of intra mode and segment-wise DC coding for depth map based on hierarchical coding structure in 3D-HEVC,” Asia-Pacific Signal and Information Processing Association Annual Summit and Conference (APSIPA), 2015.

[9] G. Sanchez et al. “A Complexity reduction algorithm for depth maps intra prediction on the 3D-HEVC,” *Visual Communications and Image Processing (VCIP)*, pp. 137-140, 2014.

[10] W. Ahmad, et al. “Complexity and Implementation Analysis of synthesized view distortion estimation architecture in 3D High Efficiency Video Coding”, International Conference on 3D Imaging, 2015.

[11] L. Zhao, L. Zhang, S. Ma, D. Zhao. “Fast Mode Decision Algorithm for Intra Prediction in HEVC,” *IEEE Visual Communications and Image Processing (VCIP)*, pp. 1-4, Nov. 2011.

[12] D. Rusanovskyy, K. Muller, A. Vetro. “Common Test Conditions of 3DV Core Experiments,” *ISO/IEC JTC1/SC29/WG11 MPEG2011/N12745*, Geneva, Jan. 2013.