Complexity Reduction by Modes Reduction in RD-List for Intra-frame Prediction in 3D-HEVC Depth Maps

[1,2]Gustavo Sanchez, [3]Luciano Agostini and [1]César Marcon
[1]Pontifical Catholic University of Rio Grande do Sul – Porto Alegre, Brazil
[2]IF Farroupilha – Alegrete, Brazil
[3]Federal University of Pelotas – Pelotas, Brazil
gustavo.sanchez@acad.pucrs.br, agostini@inf.ufpel.edu.br, cesar.marcon@pucrs.br

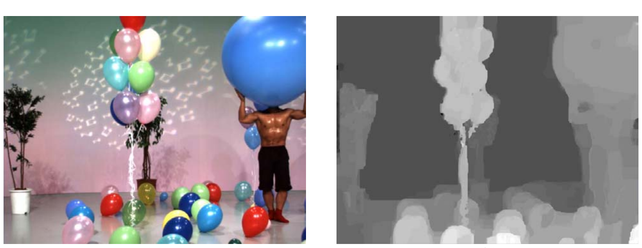
***Abstract*— This paper presents a complexity reduction technique for 3D High Efficiency Video Coding (3D-HEVC) using Rate-Distortion (RD) list reduction in the depth maps intra-frame prediction. In the 3D-HEVC standardization, the HEVC intra-frame prediction was inherited from texture to depth maps. However, considering depth maps simpler behavior than texture, the intra-frame prediction contains an unnecessary complexity since there is a high number of modes evaluated by their RD-cost. This paper evaluates the usage of fewer best-ranked modes in RD-list. As a case study, 23.9% of complexity reduction was achieved by reducing the Rough Mode Decision (RMD) selection and the Most Probable Modes (MPM) algorithm to the selected configuration with small impact in BD-rate.**

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

I. Introduction

The 3D High Efficiency Video Coding (3D-HEVC) [1] has provided a significant advance in comparison against previous state-of-the-art 3D encoders [2]. One of the key factors regarding its efficiency is related to the adoption of the Multiview Video plus Depth (MVD) data representation. In this representation, the cameras that capture the texture videos are equipped with infrared sensors, enabling to capture depth maps with the distance between the camera and the objects.

Fig. 1 shows a (a) texture view extracted from Balloons video sequence and its (b) depth map. A 3D scene shows only texture views; however, depth maps allows generating a dense set of high-quality virtual views between the originally encoded views [3], using techniques such as Depth-Image-Based Rendering (DIBR) [4]. The 3D scene presentation uses these virtual views, resulting in a 3D video with better quality.



(a) (b)

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

One can notice that depth maps have significant differences that contrast with texture views, because texture views contain complex behavior, while depth maps contain homogeneous regions (objects bodies and background) and sharp changes (objects borders).

Traditional 2D-HEVC algorithms were designed focusing only on texture characteristics without considering depth maps proprieties; therefore, requiring more computation time and producing low-quality 3D videos. 3D-HEVC inserted new tools for depth maps coding. Many of these tools are related to the intra-frame prediction of depth maps such as Depth Intra Skip (DIS), Depth Modeling Modes (DMMs) and Segment-wise Direct Component Coding (SDC) [5]. These new 3D-HEVC algorithms are alternatives to the traditional HEVC algorithms. However, HEVC algorithms still provide significate compression rates; thus, completely removing their evaluation is undesired. During standardization process, many HEVC tools were just inherited to depth maps coding, such as HEVC intra-frame prediction. If these tools consider depth maps simpler behavior than texture, complexity minimization can be achieved without reducing the encoding performance. In this paper complexity refers to the encoding time.

Many solutions have already addressed depth maps complexity; however, they are mainly focused on the new tools inserted in depth maps coding, such the works [6]-[9]. We found only the works [10]-[12] that focus HEVC intra-frame prediction and require a variable number of encoding modes according to its decision or the RD-cost calculations. Zhang et al. [10] propose to classify the encoding block regarding the block content and neighbor reference samples. This classification allows reducing the evaluation of the HEVC intra-frame prediction and DMMs. Zhang et al. [11] propose an intra-frame prediction algorithm that starts evaluating the Rate-Distortion (RD) list and stops its evaluation when an RD-value is below a compared threshold value. Silva et al. [12] apply a Sobel filter in the encoding process. Both works, [11] and [12], are capable of reducing the evaluations of HEVC intra-frame prediction and DMMs. The works [10] and [12] require a preprocessing algorithm for obtaining their final decisions that can increase latency when designing a real-time solution.

This paper proposes a new technique for complexity reduction of depth maps coding in HEVC intra-frame prediction. This technique reduces the number of elements evaluated in the RD-list provided by RMD and Most Probable Mode (MPM) algorithms. It is performed using only the best-

978-1-4673-6853-7/17/$31.00 ©2017 IEEE

ranked modes in RMD, instead of evaluating the entire RD-list, since last ranked modes in RD-list rarely are selected as the best-encoded mode. This technique does not increase the encoding latency, guaranteeing a constant number of HEVC intra-frame prediction encoding modes per block size, according to the block size.

The remaining of this paper is organized as follows: Section II presents the dataflow model of the intra-frame prediction in 3D-HEVC depth maps. Section III describes our statistical analysis and the proposed RD-list reduction algorithm. Section IV describes and discusses the experiments. Section V shows the results of a case study. Finally, Section VI renders the conclusions of this paper.

II. Depth Maps Intra-Frame Prediction

Fig. 2 illustrates the dataflow model of 3D-HEVC depth maps intra-frame prediction. The 3D-HEVC decision is based on the RD Optimization (RDO), where all available encoding modes are evaluated by their RD-cost and the mode with the lower RD-cost is selected as the encoding one. In this evaluation, HEVC intra-frame prediction, DMM-1 and DMM-4 are evaluated in both Transform-Quantization and SDC flows, while DIS algorithm is directly entropy encoded.

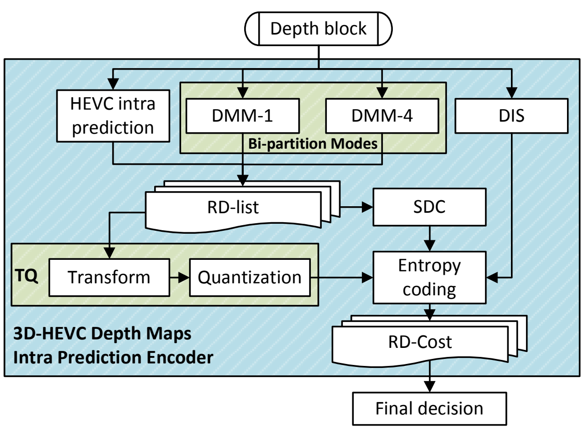


Fig. 2. Dataflow model for depth maps intra-frame prediction of 3D-HEVC.

The HEVC intra-frame prediction employs RMD and MPM algorithms to select the directional modes that should be evaluated by their RD-cost [13]. The selected ones are inserted in the RD-list to be later evaluated. Fig. 3 depicts all available directional modes of HEVC intra-frame prediction.

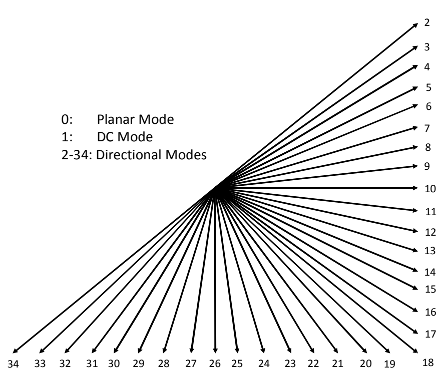


Fig. 3. Directional modes of HEVC intra-frame prediction.

Table I presents the quantity of modes direction modes and the number of modes selected by RMD, according to the encoding block size.

Table I. Number of directional modes of HEVC intra-frame prediction and the RMD selection

Block Size	Total of directional modes	RMD selection
4×4	18	8
8×8	35	8
16×16	35	3
32×32	35	3
64×64	4	3

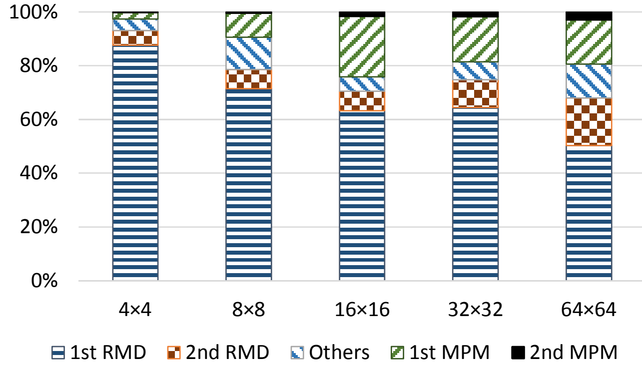
The RMD algorithm performs the Sum of Absolute Transform Differences (SATD) between the original and predicted blocks with the intra-frame directional modes, instead of evaluating all available block sizes by their RD-cost. Lower SATD values tend to produce a better prediction intra-frame direction for the current encoding block. Using the SATD rank of the directional modes allows inserting into the RD-list only a few modes. The MPM algorithm is responsible for reusing data from neighbor left and top encoded blocks in the current process. The MPM can insert at most two encoding modes into the RD-list according to the modes used during the encoding process of the neighbor blocks. The RMD and MPM algorithms were developed focusing on texture coding, where there is a complex behavior. Since depth maps coding is simpler than texture, it is not necessary to evaluate all those modes selected by RMD and MPM process.

III. Statistical Analysis and the Proposed Algorithm

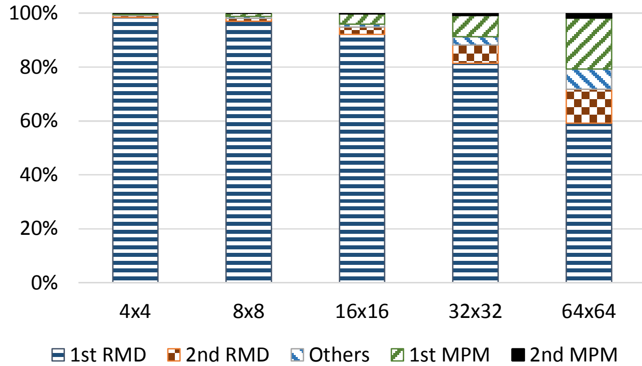
We encoded all videos of Common Test Conditions (CTC) for 3D experiments [14] using two Quantization Parameter (QP) for depth coding in the all-intra-test case using the 3D-HEVC Test Model (3D-HTM) version 16.0 [15]. The best directional mode selected after computing the RD-cost during each block computation was stored and compared against its position in RMD rank or MPM decision.

The statistical usages of each rank in RMD selection and MPM usage are presented in Fig. 4(a) and (b) for QPs equal to 34 and 45, respectively. The others in Fig. 4 represent the selection of modes ranked in the positions 3 to 8 in RMD for 4×4 and 8×8 blocks, and the mode ranked in the position 3 in RMD for blocks ranging from 16×16 to 64×64. Notice that the first RMD possesses the most representative selection, representing more than 50% in all encoding cases. Moreover, when increasing the encoding QP, the selection of the first RMD mode is still higher. Furthermore, there are still positions in RD-list that contains a significant selection, i.e., the second RMD mode and the first MPM.

One can conclude from our statistical analysis that depth maps coding does not require the complete RD evaluation for all modes inside RD-list. It happens because the RMD and MPM pre-analysis were designed considering texture coding that contains a behavior that is more complex. Additionally, due to many modes inserted into depth maps intra-frame prediction, the removal of some RD evaluation of some HEVC intra-frame prediction modes should reduce the complexity significantly without a large impact on the coding efficiency.



(a) QP = 34.



(b) QP = 45.

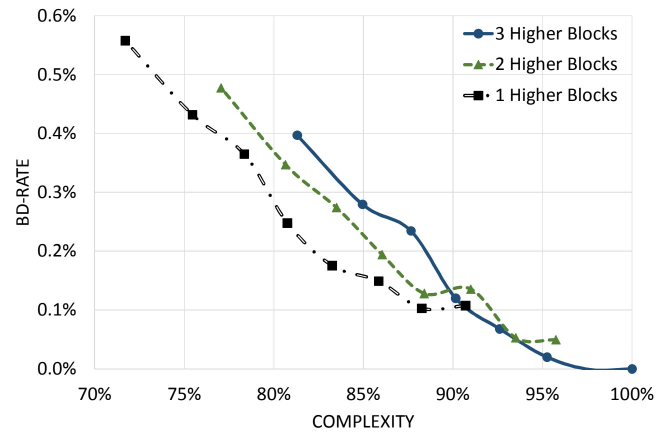
Fig. 4. Statistical analysis of RMD and MPMs selection.

IV. RD-Complexity Impact Analysis

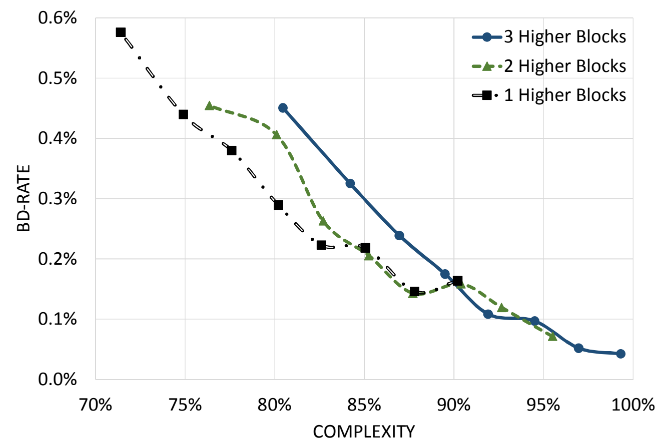
Considering that many encoding modes selected by RMD and MPM algorithms have almost no impact in the encoding selection, 48 experiments have been performed to identify the encoding efficiency and complexity impact of the proposed RD-list reduction. In these experiments, ten frames of *Balloons* and *Undo_Dancer* sequences were encoded following the CTCs under all-intra-case.

Fig. 5 (a) presents the results of these experiments when using two MPMs and (b) using only the first MPM. Each curve in Fig. 5 shows the results when varying the maximum number of modes selected in RMD analysis in 4×4 and 8×8 blocks. In each curve, the left marks, i.e., with more complexity reduction, use only the best-ranked mode by RMD, while the right marks use all modes (i.e., 8). Moreover, each curve uses a different number of modes selected by RMD for higher block sizes. By higher blocks, we refer to the blocks ranging from 16×16 to 64×64.

First, notice that removing one intra mode from larger blocks sizes represents higher complexity reduction than removing one mode from lower block sizes. Moreover, removing one evaluation of MPM represents only small complexity reduction and a negligible impact on BD-rate. Consequently, in real-time systems, reducing the MPM evaluation to only one mode results in a sound tradeoff. The removal of RMD modes from lower block sizes reduces its complexity almost linearly, due to its small rise in the BD-rate.



(a) Using two MPMs.



(b) Using the first MPM.

Fig. 5. BD-rate and Complexity for some number of selected RMD modes.

From these experiments, one can conclude that different kind of complexity reduction could be achieved according to the system configuration. Notice that there is no optimum configuration for our proposed algorithm and the system should be configured according to its application constraints such as the required rate and the maximum power dissipation.

V. Case Study

As a case study, we selected the operation point that reached the best complexity reduction for a BD-rate under 0.3%, supposing that this is our application constraint. The selected condition uses the four best-ranked modes in RMD selection and one MPM for 4×4 and 8×8 blocks. For 16×16 to 64×64 blocks, the proposed algorithm uses only the best-ranked mode in RMD selection and the first MPM since this configuration was capable of reaching a complexity reduction of almost 20% in our previous experiment.

The evaluation performed in Section IV was applied only to ten frames of two videos. Then, in this case study, we evaluated the selected scenario to all videos and frames in all-intra-test case with the original 3D-HTM 16.0 fast algorithms being applied. The results of this evaluation are presented in Table II. On average, it was possible to reduce in 23.9% depth maps complexity with a drawback of 0.167% in BD-rate of synthesized views. When analyzing all encoded views BD-rate, it drops to 0.098% because our algorithm does not affect texture views.

Completing this analysis, Fig. 6 presents the RD-plot of *Balloons* video sequence when encoded by 3D-HTM and our solution of RD-list reduction. From Fig. 6, one can conclude that our proposed solution produces similar rate and quality; however, it provides an impressive complexity reduction that should be considered when designing a real-time system. Similar results were found for other video sequences (omitted due to the lack of space).

TABLE II. BD-RATE AND COMPLEXITY REDUCTION OF RD-LIST REDUCTION EVALUATION UNDER ALL-INTRA

Video	BD-rate Synthesized views	BD-rate All views	Depth Complexity reduction
Balloons	0.266%	0.198%	23.6%
Kendo	0.284%	0.212%	22.6%
Newspaper_CC	0.306%	0.226%	23.0%
GT_Fly	0.096%	0.066%	24.1%
Poznan_Hall2	0.205%	0.148%	26.2%
Poznan_Street	0.148%	0.107%	25.5%
Undo_Dancer	0.097%	0.070%	22.4%
Shark	0.083%	0.060%	23.6%
Average	**0.167%**	**0.098%**	**23.9%**

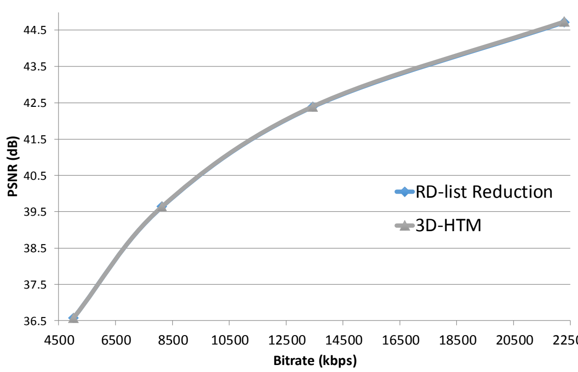


Fig. 6. RD-plot of Ballons video sequence.

The works [10] and [11] were evaluated under 3D-HTM-8.1 and were capable of obtaining an average time saving of 27.9% and 34%, with a BD-rate increase of 1.03% and 0.94%. The work [12] was evaluated under 3D-HTM 12.1 and was capable of achieving a complexity reduction of 26.2% with a BD-rate increase of 0.53% in synthesized views. These results cannot be compared directly to our solution because it was developed in a different 3D-HTM version. However, our solution increases a few the BD-rate in synthesized views with a significant complexity reduction when compared indirectly to those works.

VI. CONCLUSIONS

This paper presented an intra-frame prediction complexity reduction for 3D-HEVC depth maps. The 3D-HEVC inherited many tools from texture coding, and since depth maps coding presents a simpler behavior than texture coding, these tools should be simplified. The HEVC intra-frame prediction selects a set of modes to be entirely evaluated by its RD-cost by applying the RMD and MPM algorithms. This paper proposes the reduction of RD-list when coding depth maps. As a case study, we selected a configuration that was capable of obtaining a complexity reduction of 23.9% with a drawback of 0.167% BD-rate increase in synthesized views. Other applications could achieve higher complexity reduction or lower BD-rate increase by using a different configuration.

REFERENCES

[1] G. Tech, Y. Chen,K. Müller, J.-R. Ohm, A. Vetro, Y.-K. Wang. “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.

[2] G. Sullivan, J. Boyce, Y. Chen, J.-R. Ohm, C. Segall, A. Vetro. “Standardized Extensions of High Efficiency Video Coding (HEVC)”, *IEEE Journal of Selected Topics in Signal Processing*, v. 7, n. 6, pp. 1001-1016, Dec. 2013.

[3] P. Kauff, N. Atzpadin, C. Fehn, M. Müller, O. Schreer, A. Smolic, R. Tanger. “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] C. Fehn. “Depth-image-based rendering (DIBR), compression, and transmission for a new approach on 3D-TV”, *SPIE, Stereo-scopic Displays and Virtual Reality Systems*, v. 5291, pp. 93-104, May 2004.

[5] G. Sanchez, R. Cataldo, R. Fernandes, L. Agostini, C. Marcon. “3D-HEVC depth maps intra prediction complexity analysis”, IEEE International Conference on Electronics, Circuits and Systems (ICECS), pp. 348-351, 2016.

[6] G. Sanchez, M. Saldanha, G. Balota, B. Zatt, M. Porto, L. Agostini. “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, Y. Yang, H. Chang, W. Zhang, Y. Gan. “Fast intra mode decision for depth coding in 3D-HEVC”, *Multidimensional Systems and Signal Processing*, pp. 1-24, 2016.

[8] H.-B. Zhang, S.-H. Tsang, Y.-L. Chan, C.-H. Fu, W.-M. Su. “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)*, pp. 374-378, 2015.

[9] G. Sanchez, M. Saldanha, G. Balota, B. Zatt, M. Porto, L. Agostini. “A Complexity reduction algorithm for depth maps intra prediction on the 3D-HEVC”, *Visual Communications and Image Processing (VCIP)*, pp. 137-140, 2014.

[10] H.-B. Zhang, C.-H. Fu, Y.-L. Chan, S.-H. Tsang, W.-C. Siu. “Efficient depth intra mode decision by reference pixels classification in 3D-HEVC”, *International Conference on Image Processing (ICIP)*, pp. 961-965, 2015.

[11] H. Zhang et al. “Adaptive fast intra mode decision of depth map coding by low complexity RD-cost in 3D-HEVC”, *International Conference on Digital Signal Processing (DSP)*, pp. 487-491, 2015.

[12] T. Silva, L. Agostini, L. Cruz. “Fast Mode Selection algorithm based on texture analysis for 3D-HEVC intra prediction”, *International Conference on Multimedia and Expo (ICME)*, pp. 1-6, 2015.

[13] 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.

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

[15] Y. Chen, G. Tech, K. Wegner, S. Yea. ”Test Model 11 of 3D-HEVC and MV-HEVC”, *document JCT3VK1003 of JCT-3V*, Geneva, CH, Feb. 2015.