DCDM-INTRA: DYNAMICALLY CONFIGURABLE 3D-HEVC DEPTH MAPS INTRA-FRAME PREDICTION ALGORITHM

[1,2]*Gustavo Sanchez*, [1]*Ramon Fernandes*, [3]*Luciano Agostini*, [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, ramon.fernandes}@acad.pucrs.br, agostini@inf.ufpel.edu.br, cesar.marcon@pucrs.br

ABSTRACT

This work proposes the Dynamically Configurable 3D-HEVC Depth Maps Intra-Frame Prediction (DCDM-Intra), which explores the fact that Rough Mode Decision (RMD) was inherited from texture coding without taking advantages of the depth maps simplicity. DCDM-Intra classifies the HEVC Intra-frame Prediction Modes (IPMs) according to their BD-rate impact when encoding depth maps. Then, the application or the user can dynamically define the number of IPMs supported by the depth maps intra-prediction, according to the system status or application requirements, removing the IPMs that less affect the encoding efficiency. DCDM-Intra allows 35 distinct operation points with different encoding efficiency impacts. Experimental results demonstrate that the exclusion of 20% of IPMs causes a BD-rate increase of 0.03%, and the removal of almost 51% of IPMs rises 0.11% the BD-rate.

Index Terms— 3D-HEVC, Intra-Frame Prediction, Depth Maps Coding, IPMs, RMD.

1. INTRODUCTION

Several reasons allowed the 3D High Efficiency Video Coding (3D-HEVC) [1] to reduce the bitrate significantly for an encoded video with a similar quality when compared to its 3D standard predecessor [2]. Among these reasons, it is worth to mention the use of (i) advanced tools proposed in High Efficiency Video Coding (HEVC) [3], (ii) new encoding tools to explore 3D video characteristics, and (iii) Multiview Video plus Depth (MVD) data format [4].

MVD associates to each texture frame a depth map informing the distance between the objects and the camera. Although these depth maps are not directly displayed for the viewers, they play an essential role when generating synthesized views at the decoder side. Techniques, such as Depth-Image-Based Rendering (DIBR) [5], allow the interpolation of texture views based on these depth maps, generating a dense set of high-quality synthesized views. However, the encoding efficiency and quality provided by this new data format significantly increase the encoding computational effort because the depth maps are composed of large regions of homogeneous values and well-defined edges. Therefore, the Joint Collaborative Team on 3D Video Coding Extension Development (JCT-3V) experts introduced a new set of encoding tools to explore the depth maps characteristics. These encoding tools can be used together with the traditional texture coding tools to achieve a high encoding efficiency.

Regarding 3D-HEVC intra-frame prediction, depth maps require 5.8 times the encoding effort spent in texture coding [6] because new tools were inserted, like Depth Intra Skip (DIS) [7], Depth Modeling Modes (DMMs) [8] and Segment-wise Direct Component Coding (SDC) [9]. The HEVC Transform-Quantization (TQ) tools and intra-frame prediction are also supported by the 3D-HEVC, contributing to the high encoding effort required by this standard. On one hand, the new tools were specifically developed for depth maps coding, enabling the efficient exploration of the depth maps characteristics. On the other hand, the HEVC intra-frame prediction and the TQ tools were inherited from the texture coding without considering that the depth maps content is much simpler than texture. Therefore, there is a significant space for innovative solutions intending to reduce the effort spent in these inherited tools, exploring their behavior when processing depth maps.

This work presents the Dynamically Configurable 3D-HEVC Depth Maps Intra-Frame Prediction (DCDM-Intra) to dynamically define the depth maps encoding effort and, contributing to accelerate the depth maps encoding process and also to reduce energy consumption. DCDM-Intra ranks the HEVC Intra-frame Prediction Modes (IPMs) according to their BD-rate impact. Subsequently, the application or the user can select different operation points to accelerate the process and/or to reduce the required energy consumption, where a variable number of IPMs with lower impacts in BD-rate can be removed from the encoder evaluation.

2. RELATED WORK

Most of the works that speed up the depth map intra-frame prediction focus on the new encoding tools; however, only a few works focus on HEVC intra-frame prediction. Zhang et al. [10] classify the encoding block by visualizing its pixels and the neighborhood in three categories: homogeneous, edge and normal. They accelerate the DMMs evaluation skipping homogeneous and edges blocks from HEVC intra-frame prediction assessment. Zhang et al. [11] normally evaluate the Rate-Distortion (RD) list until it obtains an RD-value below a fixed threshold. Silva et al. [12] use a Sobel filter to speed up the HEVC intra-frame prediction. Our previous work [13] reduces the RD-list size and therefore accelerate the TQ and SDC flows for the HEVC intra-frame prediction modes.

Notice that among these related works, the ones that evaluate the intra-frame prediction mode (i.e., [10] and [12]) require a preprocessing to perform their decisions, which increases the encoding computational effort and rises the latency when these algorithms are applied into a real-time encoding system. Furthermore, when these algorithms are implemented into a dedicated hardware design, they require additional hardware resources to compute these decisions.

The works [11] and [13] focus on further evaluating the modes inside RD-list and do not require extra encoding effort to perform a preprocessing. However, the work [11] requires a variable number of operations according to the obtained RD-cost values, which can also be a problem when designing a dedicated hardware. Moreover, the works [11]-[13] require the implementation of all HEVC IPMs, which, in a hardware design, requires the insertion of additional hardware and consequently, an energy consumption increase.

3. 3D-HEVC INTRA-FRAME PREDICTION BACKGROUND

Fig. 1 displays the dataflow model used on the 3D-HEVC depth maps intra-frame prediction. Several tools are used in the decision process, including HEVC intra-frame prediction, DMM-1, DMM-4, SDC, TQ, and entropy coding. All available combinations of these tools are evaluated, searching for the best Rate Distortion Cost (RD-Cost), which is calculated according to the encoded video quality and the required bits to encode that block.

The HEVC intra-frame prediction has 35 IPMs as described in [14]. The 3D-HEVC Test Model (3D-HTM) locally evaluates these modes intending to reduce the number of modes evaluated by the full RD-Cost (see Fig. 1), which is much more complex. Two algorithms are used to define which modes must be evaluated by the full RDO: Rough Mode Decision (RMD) [15] and Most Probable Modes (MPM). These algorithms were inherited from texture coding without any modification.

RMD evaluates all intra-prediction encoding modes locally, aiming to define a few modes among all IPMs to insert into RD-list (see Fig. 1). These modes are further evaluated in the TQ and SDC flows and then a full RD-Cost evaluation is done to define the best IPM to encode each input block. The RMD evaluation applies the Sum of Absolute Transformed

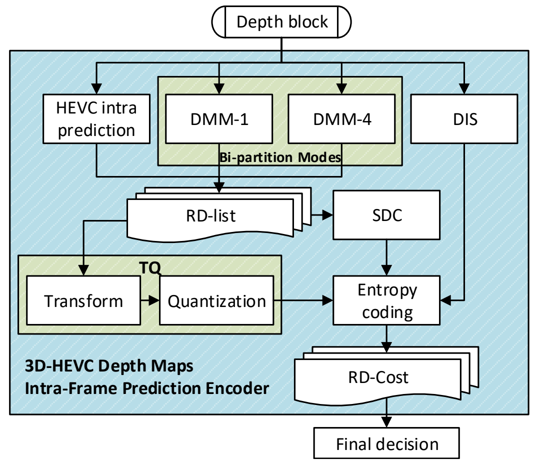


Fig. 1: Dataflow model employed on the 3D-HEVC depth maps intra-frame prediction [6].

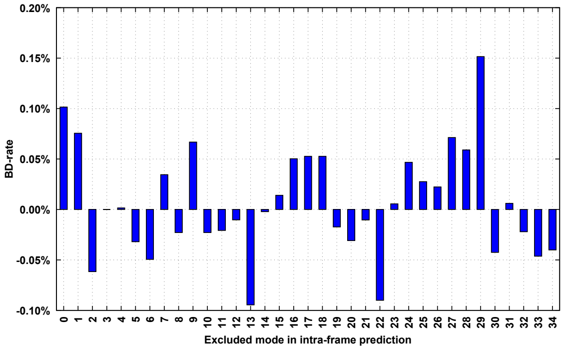
Differences (SATD) comparing the prediction generated by each IPM with the original encoding block. One of the IPMs that result in lower SATD values probably will be the one selected as the best IPM after the full RD-Cost evaluation, then, only these IPMs are selected to be evaluated through the full RD-Cost. Consequently, 35 IPMs are locally evaluated, requiring significant computational effort, but only a few modes are evaluated through the full RD-Cost, reducing a lot the global encoder computational effort.

MPM is a light-weight algorithm that inserts into RD-list the modes selected by neighbor blocks because they tend to have a high correlation with the current block decision. Moreover, the MPM usage can provide a gain in bitrate because it has a special signalization that uses a lower quantity of bits.

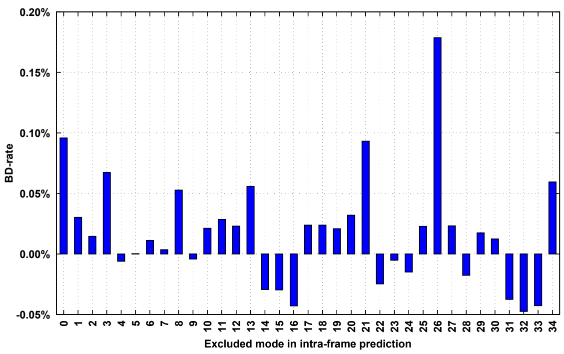
4. STATISTICAL ANALYSIS AND THE FH-INTRA ALGORITHM

Since depth maps contain simpler characteristics than texture and the HEVC intra-frame prediction was inherited from texture coding without any modification to fit the depth maps characteristics, we performed an experiment for assessing the impact of removing the evaluation of each IPM on a set of video sequences. The experiment covers the encoding of 10 frames in three views and two video sequences using the four quantization parameters defined in Common Test Conditions (CTC) [16]. All evaluations performed in this paper considered the all intra scenario, where all frames are encoded using only the I-frames. The Balloons and Newspaper_CC test sequences were randomly selected among the available sequences with their associated depth maps.

Fig. 2 presents the Bjontegaard Delta-rate (BD-rate) [17] effects of removing each IPM. Fig. 2a and 2b show the results for the Balloons and Newspaper_CC video sequences, respectively. The BD-rate was measured at the synthesized views, where the depth maps quality losses affect the user vi-



(a) Balloons



(b) Newspaper_CC

Fig. 2: Analysis of BD-rate according to the exclusion of one IPM in the HEVC intra-frame prediction of depth maps coding for (a) Balloons and (b) Newspaper_CC video sequences.

sual quality experience [18]. One can notice, observing 2, that there is a significant variation of the impacts considering each considered IPM, showing that the removal of some specific mode could not degrade the depth maps encoding performance, while the removal of other encoding modes significantly reduces the encoding performance. This behavior was explored to define the DCDM-Intra algorithm.

DCDM-Intra is a dynamically configurable algorithm, which allows the definition of multiple operation points according to the application and/or the user requirements. Each operation point considers N best-ranked HEVC intra-frame prediction modes with a respective encoding effort reduction and a consequent BD-rate degradation. Considering the experiment presented in Section 4, the most relevant IPMs to encode depth maps were classified and this classification is presented in Table 1.

In this work, the encoding effort is considered as the required number of IPM evaluations. N defines the set of best-ranked IPMs used to encode the depth maps, according to the results presented in Table 1. For example, an N = 7 will consider the seven best-ranked IPMs (26, 0, 29, 1, 27, 21 and, 18).

The application or the user choose the reduction in IPMs evaluation and the BD-rate degradation changing the N

Table 1: Relevance of each IPM to encode depth maps.

Pos	IPM	Pos	IPM	Pos	IPM	Pos	IPM	Pos	IPM
1st	26	8th	17	15th	8	22nd	23	29th	5
2nd	0	9th	3	16th	34	23rd	10	30th	6
3rd	29	10th	9	17th	12	24th	4	31st	13
4th	1	11th	25	18th	11	25th	15	32nd	2
5th	27	12th	28	19th	16	26th	30	33rd	32
6th	21	13th	7	20th	19	27th	31	34th	33
7th	18	14th	24	21st	20	28th	14	35th	22

value according to their requirements. DCDM-Intra running with N=35 did not insert any restriction in the original intra-prediction algorithm, them, this operation point did not present any impact on BD-rate since all available IPMs are evaluated. All other operation points will present some BD-rate impact. As higher is the N value, as lower is the coding effort reduction and the encoding efficiency degradation. High values of N tend to remove only the evaluations of IPMs with less importance for depth maps coding, while low values of N tend to discard the evaluation of more modes, including some with significant BD-rate impact.

The DCDM-Intra can run supporting a static (design time) or a dynamic (runtime) definition of N value. In the static version, the N value is defined in the beginning of the encoder process. Therefore, a dedicated hardware, for example, can be designed supporting only a limited set of IPMs, reducing the required area and contributing to increase the throughput and to reduce the energy consumption. The dynamic definition of N allows runtime changes in the number of IPMs that are evaluated by the encoder, then the intra-prediction effort can be adjusted accordingly with the system status, like the battery level, for example.

5. RESULTS

The DCDM-Intra was evaluated with an experiment varying N from 1 to 35. Each group of N best-ranked modes (1 to 35) was evaluated using ten additional frames of three views of the same two video sequences used in the previous experiment. Fig. 3 displays the average BD-rate results reached with this experiment. Fig. 3 also presents the reduction of the evaluated IPMs, which can be interpreted as the encoder effort reduction. Note that the use of only one IPM causes a significant loss of more than 23% in BD-rate but reducing more than 97% the evaluated IPMs. The encoding loss is still significant until N=9 when the BD-rate drops below 1% with a reduction of almost 75% of the evaluated IPMs.

The DCDM-Intra supports N values from 1 to 35, but N values lower than nine must be carefully evaluated since the BD-rate impacts are very high. Increasing the N value allows higher BD-rate reduction with a linear impact on the number of evaluated IPMs. Two operation points where defined to allow a depth discussion of DCDM-Intra results. These operation points are detached in Fig. 3: (i) N=17 and (ii) N=28. In

Table 2: BD-rate in synthesized views results.

Sequence type	Video	FH-intra 51%	FH-intra 20%
Training Sequences	Balloons	0.22%	0.06%
	Newspaper_CC	0.24%	0.06%
	Average	**0.21%**	**0.06%**
Evaluating Sequences	Kendo	0.26%	0.08%
	GT_Fly	0.04%	0.01%
	Poznan_Hall2	0.07%	0.00%
	Poznan_Street	0.12%	0.02%
	Undo_Dancer	0.04%	0.00%
	Shark	0.13%	0.05%
	Average	**0.11%**	**0.03%**

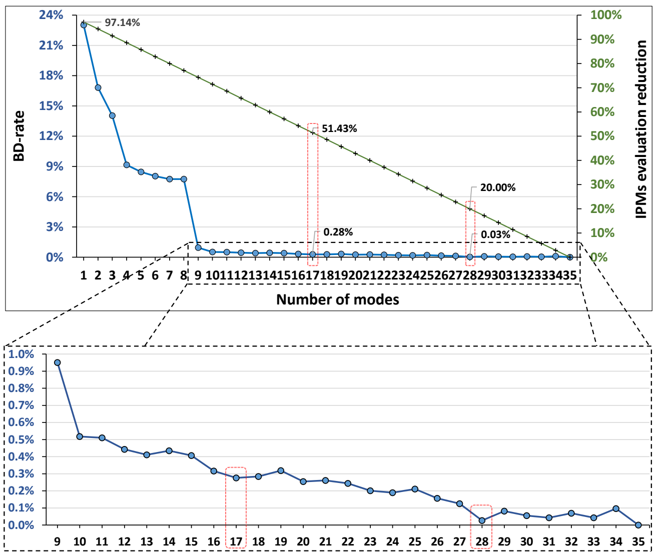


Fig. 3: Analysis of the best N intra-prediction modes usage according to the BD-rate.

(i) the IPMs evaluation is reduced more than 51% with a BD-rate increase of only 0.28%, while in (ii) a reduction of 20% at the IPMs evaluations caused a BD-rate increase of 0.03%.

These two operation points were further evaluated considering the complete CTC [16] scenario to better evaluate the reached results. These results are displayed in Table 2. The reductions of evaluated IPMs are 51% for N=17 and 20% for N=28, independently of the considered video sequence.

The results of the two videos used in the analysis section (i.e., Balloons and Newspaper_CC) were separated from the remaining video sequences to show that DCDM-Intra is able to reach good results independently of the processed sequence. One can notice that the worst BD-rate results in Table 2 were obtained in lower resolution videos (Balloons, Newspaper_CC, and Kendo video sequences). This occurs because these videos are more susceptible to BD-rate variations when the prediction process introduces a higher residual error.

Discarding the training sequences, DCDM-Intra algorithm reached an average BD-rate increase of 0.11% for N=17, and 0.03% for N=28. Considering N=17, the BD-rate increase varies from 0.04% (GT_Fly and Undo_Dancer sequences) to 0.26% (Kendo sequence). For N=28, the BD-rate varies from 0% to 0.08%.

Comparing with related works, DCDM-Intra presents competitive results and allow a dynamic selection of the operation point, which is not allowed by the related works. The work of Zhang et al. [10] reduces the IPMs evaluation to 1 for smooth blocks, between 2 and 18 for edge blocks, and preserve the original evaluation without simplification for the remaining blocks. Their approach implies a BD-rate increase of 1.03%, on average. The same authors in [11] speed up the RD-list evaluation in 81% of cases (only RMD and one MPM are evaluated by SDC and TQ flows), remaining 19% of the cases with the traditional evaluation. This speedup is obtained increasing 0.94% the BD-rate. The work of Silva et al. [12] limited in nine the maximum number of evaluated IPMs, dynamically selecting the modes through a Sobel filter application. Their work implies a BD-rate increase of 0.53%, on average. Our previous work [13] increases the BD-rate to 0.098%, on average, limiting the RD-list size to five for small block sizes and to two for higher block sizes, always inserting only one MPM.

DCDM-Intra can be integrated with solutions that reduce the effort spent in the RD-list evaluation, such as the works [11] and [13], to speed up, even more, the encoding process.

6. CONCLUSION

This paper presented the DCDM-Intra algorithm. DCDM-Intra is a dynamically configurable algorithm, which allows the reduction of intra-prediction the effort when encoding the depth maps, with reduced impacts in BD-rate. DCDM-Intra allows the definition of distinct operation points according to the application and/or the user requirements. By analyzing the impact of removing each IPM and ranking them, it was possible to determine the best candidate modes to be eliminated from the encoder evaluation. The N value determines the number of evaluated IPMs and as lower is the N value, as lower is the encoder effort but also as higher is the BD-rate impact. For N lower than nine, the BD-rate increases are higher than 1% with a reduction of more than 75% of the evaluated IPMs. When N=17 and N=28 DCDM-Intra reduces 51% and 20% of the evaluated IPMs, with a BD-rate increase of only 0.11% and 0.03%, respectively. Compared to the related works, our algorithm provides competitive results and is the unique solution that allows a dynamic operation point selection.

7. ACKNOWLEDGMENTS

This paper was achieved in cooperation with Hewlett-Packard Brazil Ltda. using incentives of Brazilian Informatics Law (Law n 8.248 of 1991). Authors would like to thanks CNPq, FAPERGS and CAPES Brazilian research agencies to support the development of this work.

8. REFERENCES

[1] G. Tech, Y. Chen, K. Müller, J. R. Ohm, A. Vetro, and Y. K. Wang, "Overview of the multiview and 3D extensions of high efficiency video coding," *IEEE Transactions on Circuits and Systems for Video Technology*, vol. 26, no. 1, pp. 35–49, Jan 2016.

[2] G. Balota, M. Saldanha, G. Sanchez, B. Zatt, M. Porto, and L. Agostini, "Overview and quality analysis in 3D-HEVC emergent video coding standard," in *2014 IEEE 5th Latin American Symposium on Circuits and Systems*, Feb 2014, pp. 1–4.

[3] G. J. Sullivan, J. R. Ohm, W. J. Han, 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, pp. 1649–1668, Dec 2012.

[4] 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," *Image Commun.*, vol. 22, no. 2, pp. 217–234, Feb. 2007.

[5] C. Bal and T. Q. Nguyen, "Multiview video plus depth coding with depth-based prediction mode," *IEEE Transactions on Circuits and Systems for Video Technology*, vol. 24, no. 6, pp. 995–1005, June 2014.

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

[7] J. Y. Lee, M. W. Park, and C. Kim, "3D-CE1: Depth intra skip (DIS) mode," *JCT-3V Doc., JCT3V-K0033, Geneva, CH*, 2015.

[8] P. Merkle, K. Müller, D. Marpe, and T. Wiegand, "Depth intra coding for 3D video based on geometric primitives," *IEEE Transactions on Circuits and Systems for Video Technology*, vol. 26, no. 3, pp. 570–582, March 2016.

[9] H. Liu and Y. Chen, "Generic segment-wise DC for 3D-HEVC depth intra coding," in *2014 IEEE International Conference on Image Processing (ICIP)*, Oct 2014, pp. 3219–3222.

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

[11] H. B. Zhang, C. H. Fu, W. M. Su, S. H. Tsang, and Y. L. Chan, "Adaptive fast intra mode decision of depth map coding by low complexity RD-Cost in 3D-HEVC," in *2015 IEEE International Conference on Digital Signal Processing (DSP)*, July 2015, pp. 487–491.

[12] T. L. da Silva, L. V. Agostini, and L. A. da Silva Cruz, "Fast mode selection algorithm based on texture analysis for 3D-HEVC intra prediction," in *2015 IEEE International Conference on Multimedia and Expo (ICME)*, June 2015, pp. 1–6.

[13] G. Sanchez, L. Agostini, and C. Marcon, "Complexity reduction by modes reduction in RD-list for intra-frame prediction in 3D-HEVC depth maps," in *2017 IEEE International Symposium on Circuits and Systems (ISCAS)*, May 2017, pp. 1–4.

[14] J. Lainema, F. Bossen, W. J. Han, J. Min, and K. Ugur, "Intra coding of the HEVC standard," *IEEE Transactions on Circuits and Systems for Video Technology*, vol. 22, no. 12, pp. 1792–1801, Dec 2012.

[15] L. Zhao, L. Zhang, S. Ma, and D. Zhao, "Fast mode decision algorithm for intra prediction in HEVC," in *Visual Communications and Image Processing (VCIP), 2011 IEEE*. IEEE, 2011, pp. 1–4.

[16] K. Müller and A. Vetro, "Common test conditions of 3DV core experiments, joint collaborative team on 3D video coding extensions (JCT-3V) document JCT3V-G1100," in *7th Meeting: San José, US*, 2014.

[17] G. Bjontegaard, "Calculation of average PSNR differences between rd-curves," in *ITU-T Q. 6/SG16 VCEG, 15th Meeting, Austin, Texas, USA, April, 2001*, 2001.

[18] A. Dricot, J. Jung, M. Cagnazzo, B. Pesquet, F. Dufaux, P. T. Kovcs, and V. K. Adhikarla, "Subjective evaluation of super multi-view compressed contents on high-end light-field 3D displays," *Signal Processing: Image Communication*, vol. 39, pp. 369 – 385, 2015, Recent Advances in Vision Modeling for Image and Video Processing.