Hardware-Oriented Wedgelet Evaluation Skip for DMM-1 in 3D-HEVC

[1,2]Gustavo Sanchez, [3]Mário Saldanha, [3]Luciano Agostini, [1]César Marcon
[1]*Pontifical Catholic University of Rio Grande do Sul – Porto Alegre, Brazil*
[2]*IF Farroupilha – Alegrete, Brazil*
[3]*Video Technology Research Group (ViTech) – Federal University of Pelotas (UFPel) – Pelotas, Brazil*
gfsanchez@acad.pucrs.br, mrdfsaldanha@inf.ufpel.edu.br, agostini@inf.ufpel.edu.br, cesar.marcon@pucrs.br

***Abstract*—This work presents a pattern-based gradient hardware-friendly simplification for Depth Modeling Mode 1 (DMM-1) encoding effort reduction in 3D High Efficiency Video Coding (3D-HEVC). The proposed algorithm is an advance of a related work that uses a line-based gradient algorithm. This novel algorithm computes the gradient of the borders of the encoding depth block and verifies the best patterns that should be evaluated by DMM-1. When designing this algorithm, hardware features were analyzed to allow simple and efficient implementation. The pattern-based algorithm can achieve an encoding effort reduction of 7.6% with a negligible BD-rate loss of 0.01%. These results represent almost the same encoding effort reduction of related works and nearly 10 times less impact on the encoder performance regarding compression rate and video quality.**

Keywords—3D-HEVC, Depth Maps, Intra-frame Prediction, DMM, Hardware-Oriented.

I. Introduction

The 3D-High Efficiency Video Coding (3D-HEVC) [1] standard was developed as an extension of the High Efficiency Video Coding (HEVC) [2] to deliver higher compression rates for 3D videos. One of the key features to achieve its efficiency is the new data format to represent 3D videos, known as Multi-View plus Depth (MVD) [3], where each original video information (texture frames) is associated with the distance between the objects and the camera (depth maps frames).

MVD format is capable of achieving a considerable reduction in the bandwidth for a 3D video transmission, since only a subset of views, along with their depth maps, is required to be transmitted or stored. In a 3D-HEVC decoder, virtual views are generated using view synthesis techniques such as Depth Image Based Rendering (DIBR) [4], therefore, without needing to encode or to transmit intermediary views.

In contrast with texture frames, which contains complex behavior, depth maps contain homogeneous regions (background and objects bodies) and sharp edges (object borders). The edges information should be preserved during the encoding process, since distortion on edges information may lead to incorrect interpretation between pixels of foreground and background in view synthesis. Incorrect information in edges may change the 3D video structure and consequently affects the video quality negatively.

Depth maps use some algorithms inherited from HEVC texture coding (e.g., HEVC intra-frame prediction). However, these algorithms do not consider depth maps properties, leading to low-quality encoded 3D videos. Based on this fact, efficient tools for coding sharp edges and homogeneous regions of depth maps were adopted by the 3D-HEVC standard.

Depth Intra Skip (DIS) [5] and Segment-wise Direct Component Coding (SDC) [6] tools were designed taking into account that homogeneous regions of depth maps share the same or similar depth values. A block with these characteristics can be represented in the coded data with less information by applying a low-complexity technique to predict it.

Depth Modeling Modes (DMMs) [7], composed of DMM-1 (wedgelet) and DMM-4 (contour), were developed as an alternative to HEVC intra-frame prediction to coding edge regions. DMM modes tend to divide edge regions in depth maps into two well-behaved regions, and then, to code each one independently. However, DMM-1 requires costly computational algorithms to find suitable partitions to predict edge regions efficiently [8]. As a drawback of high quality in synthesized views, DMM modes increases the 3D-HEVC encoder complexity significantly. Thus, novel techniques have been proposed to reduce the encoding effort of depth maps intra-frame prediction.

Sanchez et al. [9] and Saldanha et al. [10] propose some variations of techniques based on the Gradient Mode One Filter (GMOF) to reduce the DMM-1 encoding effort. These techniques were based on a study in which borders with high gradient values tend to be good candidates to be evaluated in the DMM-1 process. In fact, high-grade results in the encoding effort reduction with low impact on video quality were achieved. However, those works use line-based approach to find the best wedgelet and without taking account that the DMM-1 evaluations are pattern-based, thereby, missing an important aspect during the decision process. Moreover, these algorithms have a high difficult to be implemented in hardware because the line-based approach requires floating point number, increasing the complexity when designing a hardware. Allied to it, the hardware designs found in the literature [11]-[13] that allows encoding DMM-1 in real-time, requires a considerable high area, power dissipation, and/or Bjontegaard Delta-Rate impact (BD-rate) [14].

Therefore, this work describes a detailed analysis of DMM-1 pattern-based process and proposes a novel pattern-based gradient wedgelet evaluation skip for DMM-1, which is capable of achieving a better approximation to the wedgelet selected when compared to line-based solutions. The algorithm was designed considering hardware implementation characteristics, avoiding float point operations as defined by [9] and [10]. Additionally, the algorithm proposed here reduces in 10 times

978-1-5386-7431-4/18/$31.00 ©2018 IEEE

the impact on video quality and maintains similar results regarding encoding effort, when compared with [10].

The remainder of this paper is divided as follows. Section II describes the depth maps intra-frame prediction encoding flow. In Section III, our motivational analysis, related works, and the proposed algorithm are presented. Experimental results and comparisons are shown in Section IV. The conclusions are discussed in Section V.

II. Depth Maps Intra-Frame Prediction

The 3D-HEVC depth maps intra-frame prediction includes four main prediction modes: HEVC intra-frame prediction, DMM-1, DMM-4, and DIS, as demonstrated in Fig. 1. Rate-Distortion Optimization (RDO) performs the mode decision in 3D-HEVC similarly to HEVC. In a full-RDO evaluation, all encoding modes are evaluated by their Rate Distortion-Cost (RD-Cost), and the mode with the lower encoding RD-cost is selected to encoding the current depth block.

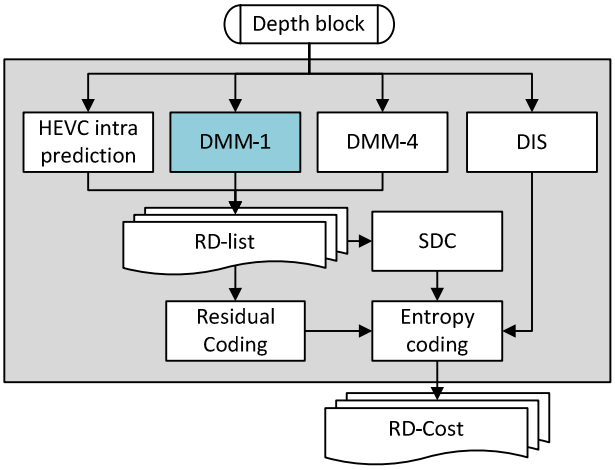

Fig. 1. 3D-HEVC depth maps intra-frame prediction.

HEVC intra-frame prediction and DIS are selected when samples in depth block share the same or similar values; i.e., in case of homogeneous blocks. In DIS mode, four predictions are evaluated: vertical and horizontal (the same of HEVC intra-frame prediction), and two single depth value considering only one sample. Similar to SKIP of HEVC inter-frame prediction, DIS does not transmit any residue values, only an index for the best-predicted mode for decoder side.

Whereas, HEVC intra-frame prediction is preferred for depth blocks which contain smooth transitions in their samples, such as, a diagonal gradient. In this case, one among 35 valid modes [15] of HEVC intra-frame prediction can achieve good results by encoding the current depth block. This flow does not characterize sharp edges regions well, and distortion in the depth map effectively may result in blurring artifacts at the boundary of objects in the synthesized views. As an alternative, 3D-HEVC adds DMM-1 and DMM-4. These modes are essential to represent edges and can achieve high-grade quality results at the cost of a high encoding effort.

The DMM-1 (which is the focus of this work) flowchart is presented Fig. 2. It divides the current encoding block into two distinct regions separated by a straight line called a wedgelet. This straight line is discretized resulting in a binary pattern that should be used during the DMM-1 search process. The traditional process requires evaluating several patterns iteratively, searching for the best distortion as shown in Fig. 2. Table I illustrates the number of wedgelets searched in DMM-1 execution, where one can notice that there are up to 384 evaluations for larger blocks sizes. These wedgelets were defined in the standardization of 3D-HEVC. The difference between the number of total possible and the evaluated wedgelets in the main stage is because there is a high number of wedgelets that can be only reachable by the refinement process. Each DMM-1 evaluation has a high encoding effort associated with it since 3D-HEVC Test Model (3D-HTM) uses Synthesized View Distortion Change (SVDC) [16] as the distortion criterion. Therefore, DMM-1 evaluation is an expensive computational task. Thus, techniques to reduce the search by the DMM-1 patterns are desired.

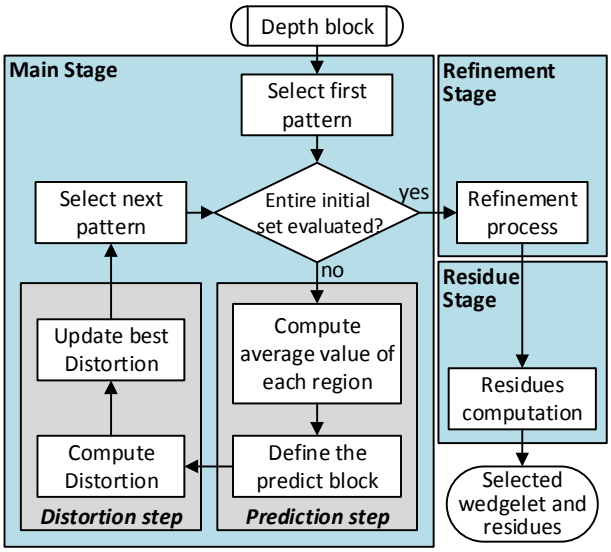

Fig. 2. DMM-1 Flowchart [11].

Table I. Number of Total and Evaluated Wedgelets

Block size	Total possible wedgelets	Evaluated wedgelets in Main Stage	Refinement wedgelets
4×4	86	58	1-8
8×8	802	314	1-8
>= 16×16	510	384	1-8

The DMM-4 represents an inter-component tool that performs a contour partition based on texture information. Two regions that can be shaped arbitrarily and may consist of multiples disconnected parts divide the current depth block.

III. Depth Block Analysis and the Proposed Solution

Fig. 3 presents an example of an 8×8 depth block with the selected wedgelet when the DMM-1 original algorithm is applied. The gradient values of the block borders are also presented in this figure. The gradient of the upper row, the left column, the bottom row and the right column are obtained by applying the equations (1), (2), (3), and (4), respectively. The x in these equations ranges from 1 to *size*-1, where *size* represents the block width. Notice that in the example of Fig. 3, the best

encoding wedgelet is found near the positions with the highest gradient. Therefore, one can conclude that the borders positions with high gradient values tend to be good candidates for the DMM-1 wedgelet decision.

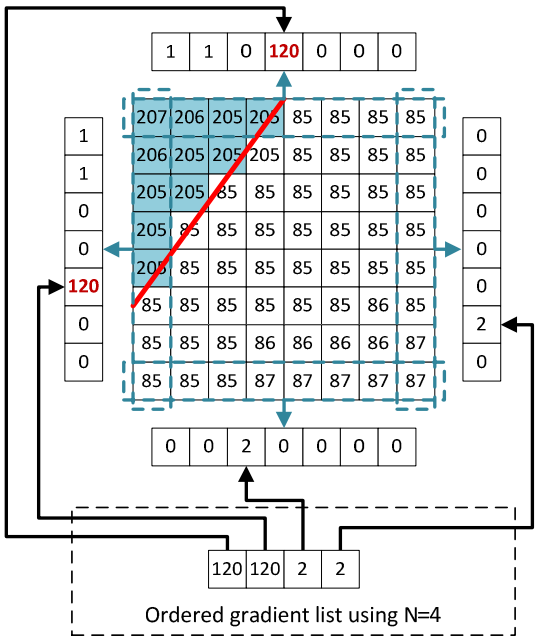

Fig. 3. Example of an encoding depth block wedgelet selection with an analysis of its borders gradient.

$$Grad_{Upper}(x) = |P(1,x) - P(1,x+1)| \quad (1)$$

$$Grad_{Left}(x) = |P(x,1) - P(x+1,1)| \quad (2)$$

$$Grad_{Bottom}(x) = |P(size,x) - P(size,x+1)| \quad (3)$$

$$Grad_{Right}(x) = |P(x,size) - P(x+1,size)| \quad (4)$$

Since the DMM-1 algorithm is one of the most time-consuming operations in depth maps intra-frame prediction, Sanchez et al. [9] have proposed the Gradient-based Mode One Filter (GMOF) algorithm to reduce the number of evaluated wedgelets considerably in DMM-1, and Saldanha et al. [10] presented some optimizations of this algorithm. All these algorithms reduce the number of wedgelets by searching for wedgelets whose straight line starts in a position that there is a considerable high change of gradient value. The solution proposed by these works are line-based because it considers the position of the stored straight line. The details of the GMOF solutions are presented in Subsection III.A, and in Subsection III.B we present the proposed modification to create the new Pattern-Based GMOF, which is capable of achieving a better approximation to the wedgelet selected by the original DMM-1 algorithm when compared to these line-based solutions. Also, the new solution was designed being easier to be implemented in hardware than those works.

A. Line-based GMOF variations

The original line-based GMOF algorithm creates a gradient list with *N* positions ordered by the highest gradients of the borders. The positions in this list point to the position change that obtained that gradient, as presented in the bottom of Fig. 3. The algorithm evaluates only wedgelets whose straight line is located in those positions. By applying intense experimentation, the best *N* selected in [9] was 8.

The other GMOF variations [10] are based on the original GMOF algorithm. The Strong GMOF (S-GMOF) algorithm selects the two highest gradients from two distinct borders. The Single Degree of Freedom GMOF (SDF-GMOF) selects the highest gradient among all possible positions and set a wedgelet line start into it. The ending of the wedgelet is selected searching in the wedgelet list for the nearest end position compared against the second gradient position. The Double Degree of Freedom GMOF (DDF-GMOF) searches the entire wedgelet list for the wedgelet that has the lowest distance compared to the two highest gradients positions. The algorithms S-GMOF, SDF-GMOF, and DDF-GMOF are capable of reducing the wedgelet list evaluation to a single wedgelet evaluation. After this initial evaluation, the original DMM-1 refinement is applied in all these algorithms. However, often when using these line-based algorithms (including original GMOF), it is not possible to reach the best wedgelet approximation.

Fig. 4 shows an example of encoding a 4×4 depth block with the best pattern selected by the original DMM-1 algorithm. In this example, one can notice that the best wedgelet pattern is obtained starting the wedgelet in an intermediary pixel and there is a gradient value only a half pixel distance from there. It occurs because the DMM-1 wedgelets evaluations are defined in starting with a resolution of half-pixel and the original DMM-1 process is pattern-based and not line-based as explored in [9] and [10].

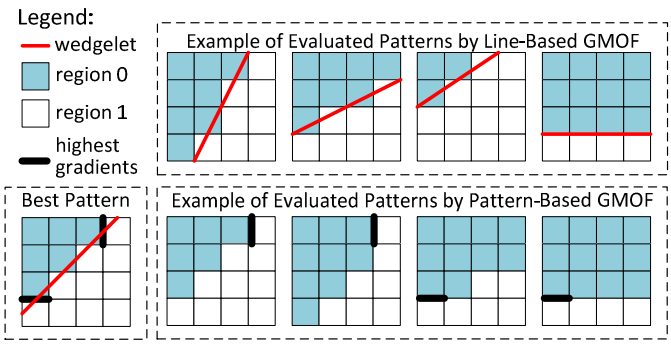

Fig. 4. Best pattern and the evaluated patterns using line-based GMOF and pattern-based GMOF of a 4×4 encoding depth block.

The line-based approach requires working with float point number since it needs to compute a pixel with real coordinates where the wedgelets should start/end. When designing dedicated hardware, it gets harder to achieve a high processing rate, requiring a higher area and power dissipation.

B. The New Pattern-based GMOF

Considering that the DMM-1 evaluation does not use the wedgelet line position, but uses only the binary pattern, a new pattern-based GMOF algorithm is proposed. The dataflow model of pattern-based GMOF is presented in Fig. 5.

The gradient list constructed by original line-based GMOF is maintained in this algorithm by applying the equations (1) to (4). After constructing this list, the pattern of each wedgelet in the initial wedgelet pattern set is analyzed with the positions in the gradient list. If there is a division in the pattern in any of those

positions (i.e., it changes from one region to another in that position), the wedgelet is evaluated by its RD-cost. Otherwise, the next wedgelet is applied to the same process. Furthermore, when the entire initial wedgelet set is analyzed, the original refinement is applied and the wedgelet that obtained the lowest RD-cost is selected as the best-encoded wedgelet.

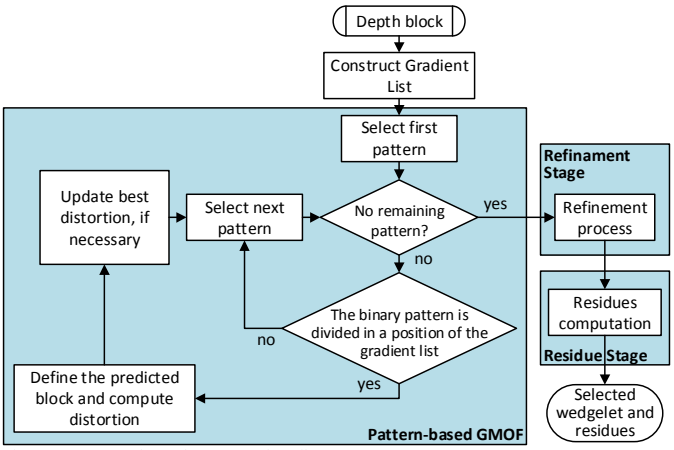

Fig. 5. Pattern-based GMOF dataflow.

The same example presented in Fig. 4 for line-based GMOF is also presented for pattern-based GMOF. In this case, the pattern-based version is capable of achieving the same wedgelet that would be selected by the original DMM-1 algorithm.

When designing a hardware containing the pattern-based GMOF solution, it is required to compute the gradient of the borders and ranking it, as solutions designed by Saldanha et al. [10] require. However, when evaluating if the encoding wedgelet requires the computation of its distortion, or if the encoding process of that wedgelet can be skipped, the process is easier when implemented in hardware because only the binary pattern of the wedgelet needs to be compared with these gradients position since these solutions considers the wedgelet pattern and not the wedgelet line. Therefore, a hardware design for DMM-1 with our solution can be implemented using fewer hardware resources and consequently dissipate a smaller power.

IV. Results and Comparisons

The proposed pattern-based GMOF has been evaluated in 3D-HTM 16.0 [17] under Common Test Conditions (CTC) for 3D experiments [18]. Results of the proposed algorithm and a comparison against GMOFs implementation [10] are presented in Table II. For a fair comparison, we evaluated our pattern-based GMOF solution using N=8, the same value of line-based GMOF provided in [9] and [10]. In those works, Shark video sequence was not evaluated since that work used an older version of CTC.

The pattern-based solution is capable of achieving an average timesaving of 7.6%, with an increase in BD-rate of 0.01%. The highest BD-rate impact is noticed in Newspaper_CC video (but still negligible), where 0.03% BD-rate increase is noticed. This BD-rate is negligible compared against [9] and [10] algorithms that are capable of obtaining an average timesaving ranging from 4.9% to 8.2% with an increase in BD-rate of 0.33% to 1.47% (according to the selected solution).

Our new solution is capable of reaching almost the same encoding effort reduction of those solutions, with a BD-rate increase at least more than 10 times lower. Fig. 6 shows the wedgelets evaluation skips obtained by pattern-based GMOF solution, explaining the depth timesavings results. Notice that more than 50% of wedgelets are skipped in all evaluated videos, with only a small variation between the encoded video, due to their characteristics. In this evaluation, on average, 58% of wedgelets were skipped.

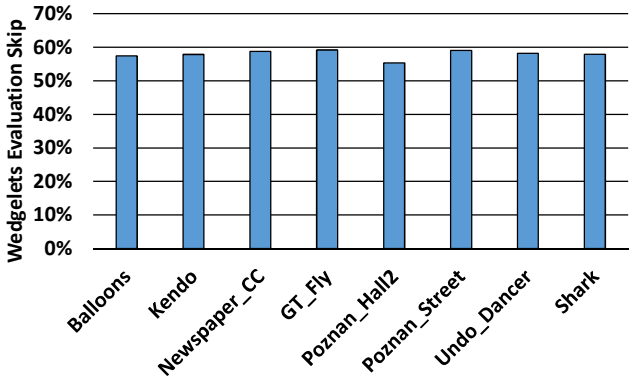

Fig. 6. Wedgelets evaluation skip by pattern-based GMOF.

Different values of N can be used when applying the proposed algorithm according to the system specification; e.g.,

Table II. Experimental results of pattern-based GMOF and comparisons with line-based GMOF

Solution and Reference Software	**Line-based GMOF variations [10] 3D-HTM 10.2**		**Pattern-based GMOF 3D-HTM 16.0**	
Videos	Synthesis only (BD-rate)	Depth Timesaving	Synthesis only (BD-rate)	Depth Timesaving
Balloons	0.13% - 1.20%	5.3% - 10.0%	0.02%	7.8%
Kendo	0.24% - 1.10%	4.9% - 10.0%	0.01%	6.1%
Newspaper_CC	0.76% - 2.81%	5.7% - 11.7%	0.03%	7.6%
GT_Fly	0.25% - 0.73%	4.7% - 5.5%	0.00%	9.8%
Poznan_Hall2	0.17% - 1.62%	4.9% - 8.9%	0.00%	7.4%
Poznan_Street	0.22% - 0.92%	4.4% - 6.5%	0.01%	8.4%
Undo_Dancer	0.53% - 1.91%	4.1% - 5.0%	0.01%	6.8%
Shark	Not available	Not available	0.01%	7.1%
Average	**0.33% - 1.47%**	**4.9% - 8.2%**	**0.01%**	**7.6%**

for a system that requires lower encoding effort, *N* can be reduced and for a system that requires better BD-rate *N* can be increased. Moreover, considering that the hardware designed for DMM-1 in the literature are pattern-based [11]-[13], the proposed solution in this paper is more feasible for a real-time hardware implementation compared against the line-based GMOFs variations.

Therefore, our main contributions are reducing the BD-rate increase required by [9] and [10] when designing a DMM-1 encoding effort reduction solution and providing a hardware-oriented solution that is easier to be implemented when designing a dedicated hardware design.

V. CONCLUSIONS

This work presented a pattern-based gradient approach for reducing the DMM-1 encoding effort. The proposed algorithm is an improvement of the line-based GMOF algorithm proposed in related works. The improvements in this work introduced hardware characteristics in the process and allowed the achieving a better approximation in the prediction process. The pattern-based solution was evaluated under 3D-HTM 16.0 using CTC. Simulation results show that the proposed solution is capable of achieving an encoding effort reduction of 7.6%, with only 0.01% of the BD-rate increase, which is negligible. Comparing against line-based GMOFs, our novel algorithm is capable of reducing 10 times the BD-rate with almost the same encoding effort reduction achievement. Moreover, the proposed solution is easier to be integrated into a hardware design, since the DMM-1 hardware designs found in the literature store the wedgelets binary patterns.

ACKNOWLEDGMENT

This paper was achieved in cooperation with Hewlett-Packard Brazil Ltda. using incentives of Brazilian Informatics Law (Law nº 8.248 of 1991). Authors also would like to thank CAPES, CNPq and FAPERGS Brazilian research agencies to support the development of this work.

REFERENCES

[1] G. Tech, Y. Chen, K. Muller, J. Ohm, A. Vetro, Y. 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. J. Sullivan, J. Ohm, W.-J. Han, T. Wiegand. "Overview of the high efficiency video coding (HEVC) standard." *IEEE Transactions on circuits and systems for video technology (TCSVT)*, v. 22, n. 12, pp. 1649-1668, Dec. 2012.

[3] P. Kauff, N. Atzpadin, C. Fehn, M. Muller, 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 Syst*, v. 5291, pp. 93-104, May 2004.

[5] J. Y. Lee, M. W. Park, and C. Kim, *3D-CE1: Depth Intra Skip (DIS) Mode*, document JCT3V-K0033, Geneva, Switzerland, Feb. 2015.

[6] H. Liu and Y. Chen, “Generic segment-wise DC for 3D-HEVC depth intra coding,” in *Proc. IEEE Int. Conf. Image Process. (ICIP)*, Paris, France, pp. 3219–3222, Oct. 2014.

[7] P. Merkle, K. Muller, D. Marpe, T. Wiegand. “Depth Intra Coding for 3D Video based on Geometric Primitives,” *IEEE Transactions on Circuits and Systems for Video Technology*, v. 26, n. 3, pp. 570-582, Feb. 2015.

[8] G. Sanchez, L. Agostini, C. Marcon. “3D-HEVC depth maps intra prediction complexity analysis”, *in Proceedings of the IEEE International Conference on Electronics, Circuits, & Systems (ICECS)*, pp. 1-4, 2016.

[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 Conference (VCIP)*, pp. 137-140, 2014.

[10] M. Saldanha, B. Zatt, M. Porto, L. Agostini, G. Sanchez, “Solutions for DMM-1 complexity reduction in 3D-HEVC based on gradient calculation”, *Latin American Symposium on Circuits & Systems (LASCAS)*, pp. 211-214, 2016.

[11] G. Sanchez, C. Marcon, L. Agostini. “Real-Time Scalable Architecture for 3D-HEVC Bipartition Modes,” *Journal of Real-Time Image Processing (JRTIP)*, v. 13, n. 1, pp 71-83, Mar. 2017.

[12] F. Amish, E. Bourennane. “An Efficient Hardware Solution for 3D-HEVC Intra-Prediction,” *Journal of Real-Time Image Processing (JRTIP)*, pp. 1-13, Jan. 2017.

[13] G. Sanchez, L. Agostini, F. Mór, C. Marcon, “Low-area scalable hardware architecture for DMM-1 encoder of 3D-HEVC video coding standard," 2017 30th Symposium on Integrated Circuits and Systems Design (SBCCI), Fortaleza, pp. 36-40, 2017.

[14] G. Bjontegaard. “Calculation of average PSNR differences between RD curves. ITUT SC 16/SG16VCEG-m33, Apr. 2001.

[15] J. Lainema, F. Bossen, W. Han, J. Min, K. Ugur, “Intra Coding of the HEVC Standard,” *IEEE Transactions on Circuits and Systems for Video Technology (TCSVT)*, v. 22, n. 12, pp. 1792-1801, Dec. 2012.

[16] G. Tech, H. Schwarz, K. Müller, T. Wiegand, “3D video coding using the synthesized view distortion change,” in *Proc. Picture Coding Symposium (PCS)*, pp. 25-28, May 2012.

[17] 3D-HEVC Test Model. Available at: https://hevc.hhi.fraunhofer.de/svn/svn_3DVCSoftware/tags/HTM-16.0/, access in Sep. 2017.

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