Journal of Electronic Imaging

JElectronicImaging.org

DFPS: a fast pattern selector for depth modeling mode 1 in three-dimensional high-efficiency video coding standard

Gustavo Sanchez
Lucas Jordani
César Marcon
Luciano Agostini

Gustavo Sanchez, Lucas Jordani, César Marcon, Luciano Agostini, "DFPS: a fast pattern selector for depth modeling mode 1 in three-dimensional high-efficiency video coding standard," *J. Electron. Imaging* **25**(6), 063011 (2016), doi: 10.1117/1.JEI.25.6.063011.

DFPS: a fast pattern selector for depth modeling mode 1 in three-dimensional high-efficiency video coding standard

Gustavo Sanchez,[a,b,*] Lucas Jordani,[a] César Marcon,[a] and Luciano Agostini[c]
[a]Pontifícia Universidade Católica do Rio Grande do Sul, PPGCC, Avenida Ipiranga n. 6681, Porto Alegre CEP 90619-900, Brazil
[b]Instituto Federal Farroupilha, RS-377, Alegrete CEP 97555-000, Brazil
[c]Universidade Federal de Pelotas, PPGC–GACI, Rua Gomes Carneiro n. 1, Pelotas CEP 96160-000, Brazil

Abstract. Depth modeling mode 1 (DMM-1) processes its prediction unit (PU) without taking into account its neighborhood. However, an extensive analysis of neighbor pattern selection showed that the processing PU could use the same pattern of its neighbors or an adaptation of these patterns. Therefore, this work proposes the DMM-1 fast pattern selector (DFPS) algorithm that includes lightweight and medium-weight DMM-1 pattern predictors. DFPS starts using the lightweight predictor, whose output is compared with a threshold and then the algorithm employs a DMM-1 refinement or the medium-weight predictor. The result of this last predictor is compared against a new threshold, where the encoder decides if the remaining patterns of the DMM-1 algorithm or only its refinement should be evaluated. Experiments evaluating DFPS encoding efficiency in the all-intra mode demonstrate that by reducing 71% and 84% of the DMM-1 patterns' evaluation, the complexity is reduced 11.7% and 13.4%, respectively, without significantly affecting the quality of the synthesized views.
[DOI: 10.1117/1.JEI.25.6.063011]

Keywords: three-dimensional high-efficiency video coding; depth modeling mode 1; complexity reduction; pattern predictor.

Paper 16682 received Aug. 9, 2016; accepted for publication Nov. 7, 2016; published online Nov. 30, 2016.

1 Introduction

The popularization of three-dimensional (3-D) video coding applications comes from the visual experience that allows users to enjoy the depth perception of the scene, bringing an experience that goes beyond two-dimensional (2-D) videos. Since 3-D videos present high bandwidth and storage requirements, many researchers were motivated to investigate and develop new solutions for the upcoming 3-D video applications that include free viewpoint video and 3-D television. According to Merkle et al.,[1] higher compression rates can be achieved using the 3-D video coding standard that explores the redundancies between views instead of encoding each view separately.

In response to the need for higher compression rates for 3-D video applications, the 3D high-efficiency video coding (3D-HEVC)[2,3] was developed by the Joint Collaborative Team on 3-D Video Coding Extension Development (JCT-3V)[4] as an extension of HEVC.[5] The 3D-HEVC standardization process was finished in 2015, and today is the most advanced 3-D video coding standard.

Real-time encoding for high-definition videos is a challenging process due to advances in HEVC for providing a better bitrate with similar encoding quality. Consequently, the 3D-HEVC brings a worse scenario because it requires encoding multiple cameras with more encoding techniques.

A multiview video sequence consists of many views captured from multiple cameras, placed very close to each other, to cover different portions of a 3-D scene.[6] Many coding features have already been included in the previous 3-D standard[7] to encode videos from multiple cameras efficiently. Nevertheless, 3D-HEVC uses even more advanced texture coding tools such as neighboring block based on disparity vector derivation,[8,9] interview motion prediction, interview residual prediction,[10] and illumination compensation.[11]

Multiview video plus depth (MVD)[12] representation is a key factor for 3D-HEVC efficiency. In MVD, each texture view is associated with a depth map which provides the geometrical information according to the object's distance from the camera.[13] Figure 1 shows the (a) texture view of the first frame of the Newspaper_CC video and its associated (b) depth map.

The depth maps should not be displayed on a screen during a 3-D video presentation; however, applying techniques such as depth image-based rendering[12] at the decoder allows to synthesize virtual views between the original cameras that compose the 3-D scene.[14,15]

Depth maps have distinct characteristics from texture views, as exemplified in Fig. 1. While the texture is a multicolor view containing regions with smooth transitions between pixels, the depth maps are gray shades views that contain large areas with the same depth value, representing background or a body of objects, and sharp edges, representing borders of objects.[16] The traditional 2-D video coding techniques apply low-pass filters to smooth the transitions between pixels in texture views. When a low-pass filter is used in depth maps, the objects' borders (high-frequency components) are smooth, which should result in low-quality synthesized views. Therefore, using only traditional 2-D video coding techniques in depth maps compression introduces blocking artifacts, mosquito noise, and edge blurring[17] in

*Address all correspondence to: Gustavo Sanchez, E-mail: gustavo.sanchez@acad.pucrs.br

1017-9909/2016/$25.00

(a) (b)

Fig. 1 (a) Texture view of the first frame of Newspaper_CC video and its associated (b) depth map.

the synthesized views. Therefore, it is important to encode the depth maps as precisely as possible, preserving the edges (i.e., without smoothing them) and avoiding errors in the video synthesis process. To overcome this problem, 3D-HEVC inserted new tools during depth maps coding as an alternative to the traditional 2-D encoding techniques. These new tools include the depth modeling modes (DMMs) (also called bipartition modes) that are innovative intraprediction methods for depth maps coding.[18] DMMs are edge-aware algorithms capable of providing a better depth map coding and consequently high-quality synthesized views.

These new modes increase the complexity of 3D-HEVC and, according to Refs. 19 and 20, DMM-1 has a high-computational complexity in intraprediction, which is a problem in the design of a real-time application and consequently, requires aggressive complexity reduction algorithms when running in an embedded system.

This article proposes the DMM-1 fast pattern selector (DFPS) algorithm, which accelerates DMM-1 processing using neighbor prediction units (PU) information in current PU DMM-1 prediction. Our previous works in Refs. 19 and 21 use the information inside the current PU for decision-making, and our new solution is capable of using neighbor PU decisions to accelerate current PU decision. Therefore, the main contributions of this article are highlighted next:

- Analysis of similarity between neighbor blocks DMM-1 pattern selection.
- Proposition of inserting two predictors and an early termination technique for DMM-1 complexity reduction.
- Objective and subjective quality analysis of the proposed solution.

This remaining of this article is divided as follows. Section 2 presents the related work about complexity reduction schemes and encoding techniques related to 3D-HEVC. The DMM-1 algorithm is described in Sec. 3 along with a statistical analysis motivating the use of pattern predictors. Section 4 describes the proposed DFPS algorithm along with its threshold's definition. Section 5 evaluates the DFPS algorithm under common test conditions, and Sec. 6 renders the conclusions of this article.

2 Related Work

Most of the complexity reduction techniques are based on early termination techniques by using encoding PU, encoding frames, or past encoded frames characteristics. Several works have been proposed for H.265/HEVC (state-of-the-art in 2-D video coding) and H.264/MVC (previous 3-D encoding standard).

Techniques recommended for H.265/HEVC and H.264/MVC tend to maintain good results regarding encoding quality and low complexity when applied to 3D-HEVC. However, as they are mainly designed for texture coding and as they do not consider depth maps' characteristics, they lead to an inefficient prediction concerning timesaving and encoding quality.

There are already works designed especially for 3D-HEVC, focused on depth maps coding, such as in Refs. 20, 18, 21–26, where Refs. 23 and 18 are already available by default in 3D-HEVC test model (3D-HTM) version 16.0. The work in Ref. 22 stops splitting the encoding quadtree in dependent-views quickly by analyzing the disoccluded information provided by a hole-filing map. The recursive quadtree expansion has been limited by Ref. 18, resulting in meaningful complexity reduction for both intra and interprediction modes.

The works in Refs. 20, 21, 23, and 24 proposed simplifications in the intraframe prediction of depth maps by skipping the entire DMMs evaluation when the encoding PU tends to be well-encoded by traditional HEVC intraprediction, i.e., in smooth regions. These works are based on a high-level decision of the 3D-HEVC encoder. However, they do not simplify the DMM-1 algorithm directly. These decisions focus on removing the DMMs' evaluation entirely; however, if the DMMs are selected to be evaluated, the traditional DMM-1 computation is performed without any simplification.

The work in Ref. 25 proposes a simplification of the process to obtain the distortion cost of the DMM-1 without changing the number of wedgelets evaluated in the process. This article focuses on the reduction of the number of wedgelet evaluations, which is the major concern of the works in Refs. 19, 20, and 26. Sanchez et al.[19] propose a gradient filter in the border of the PU to detect the most promising positions for evaluating DMM-1 patterns. Saldanha et al.[26] show advances in its gradient filter providing aggressive wedgelets evaluation skips. These filters require a complexity overhead for computing the gradient and insert significant impact on DMM-1 processing. Zhang et al.[20] present a high-level decision that is capable of removing the entire DMMs' evaluation when the PU is smooth. Moreover, in the case where

the PU is not smooth, a DMM-1 complexity reduction is applied. They reduce the DMM-1 complexity by evaluating only wedgelets with six predefined HEVC intraprediction directions (notice that there can be many wedgelets following an HEVC intradirection).

None of these DMM-1 complexity reduction techniques uses neighbor PU encoding information in the current PU DMM-1 encoding. This work shows that if this approach is considered during the DMM-1 encoding process, there is still space to provide aggressive solutions for DMM-1 complexity reduction; in addition, it can be integrated with other solutions.

3 Depth Modeling Mode 1 Algorithm

The DMMs are only evaluated for blocks from 4×4 to 32×32 pixels. DMMs model the encoding depth block, approximating its original signal into two rectangular regions, where each region is represented by a constant value called constant partition value (CPV). It is necessary to pack the pattern containing the segmentation information in the bitstream, specifying the region to which each sample belongs, the residues, and the CPV of each region.[27] Figure 2 depicts the block diagram of DMM-1 encoding algorithm, which segments blocks based on a search over a predefined wedgelet set.

The DMM-1 encoding algorithm encompasses three stages: main stage, refinement stage, and residue stage. The main stage evaluates an initial wedgelet set (i.e., wedgelets that should be assessed before the refinement) searching for the best match and is performed using a distortion criterion that compares each wedgelet pattern against the original block. Finding the best match requires a high-computational complexity, where the following steps are necessary for each wedgelet. First, the encoding block is mapped into the wedgelet pattern. According to this mapping, the average values of all samples mapped into regions 0 and 1 are computed, and the predicted block is defined as the average value of each region (prediction step). A distortion criterion, such as sum of absolute differences (SAD), is applied to obtain the distortion of the predicted block

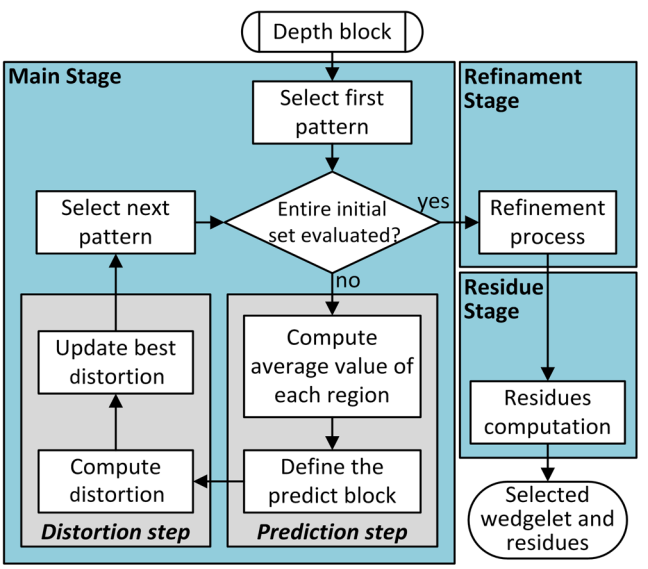

Fig. 2 Main blocks of the DMM-1 encoding algorithm.

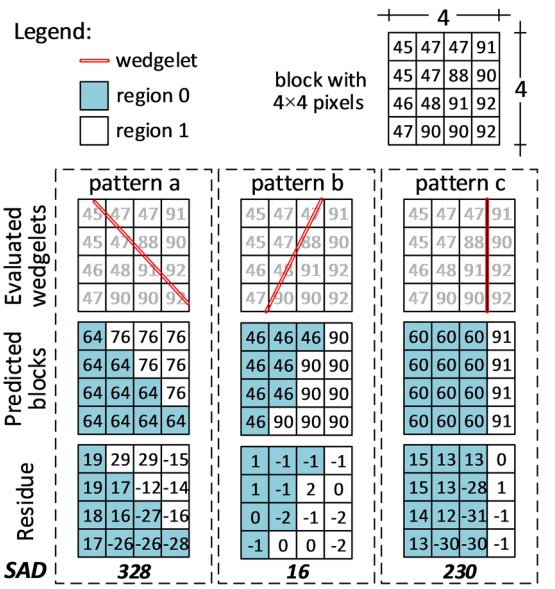

Fig. 3 Example of encoding a block with the DMM-1 algorithm.

compared against the encoding block. The wedgelet pattern with the lowest distortion is defined as the best wedgelet (distortion step). Refinement stage evaluates up to eight wedgelets around the wedgelet selected in the main stage. Again, the wedgelet with the minimum distortion is chosen as the best wedgelet. Finally, the residue stage subtracts the predicted block of the elected wedgelet from the original one.

Figure 3 exemplifies the DMM-1 encoding of a 4×4 depth block along with the evaluation of three different wedgelet patterns. In addition, Fig. 3 shows that the prediction process of DMM-1 encodes the depth block sample according to the evaluated wedgelet (i.e., patterns a, b, and c), using SAD as the distortion criterion.

This procedure maps the pixels of the block sample in one of the two regions. Subsequently, the predicted block step computes the average value of all pixels in the region (e.g., the average value of regions 0 and 1 of pattern a are 64 and 76, respectively). The residue step annotates the position corresponding to each pixel with the difference between the predicted and original depth sample. The SAD of all patterns is attained, and finally, pattern b is selected since it has the lowest SAD.

Table 1 exhibits the number of possible wedgelets regarding the block size, which was defined by 3D-HEVC. Clearly, the main stage, where the search over the initial wedgelet set is performed, is the most time-consuming stage of DMM-1, and there are considerable possibilities to save time here.

A wedgelet will have one of the six possible orientations as presented in Fig. 4. The wedgelet orientation is based on

Table 1 Number of evaluated wedgelets in DMM-1.

Block size	Total possible wedgelets	Wedgelets in initial set
4×4	86	58
8×8	802	314
$>= 16 \times 16$	510	384

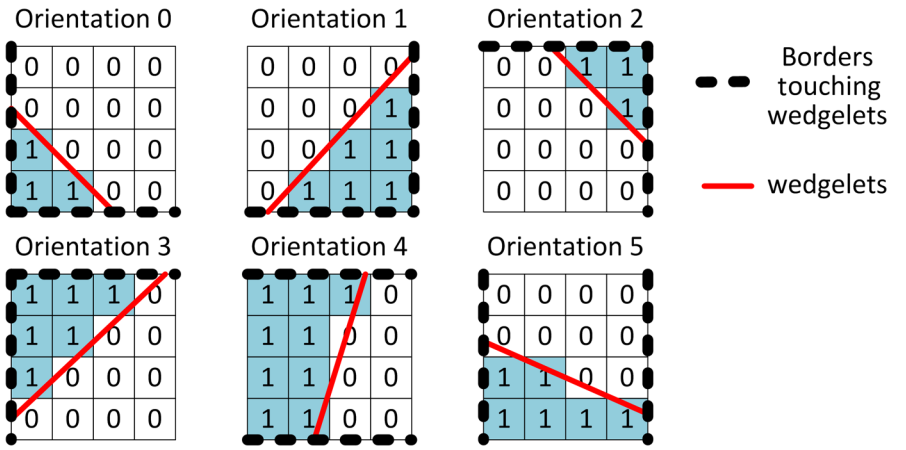

Fig. 4 Examples of possible wedgelets orientation.

the position that the wedgelet touches the border; e.g., orientation 0 means that the wedgelet touches left border and bottom border. This information is not used during the DMM-1 encoding process and can be explored to obtain significant complexity reduction.

3.1 *Statistical Analysis and Discussion*

We propose the following experiment to explain the relevance of DMM-1 pattern selection based on the neighbor PUs. Each block size (i.e., 4×4, 8×8, 16×16, and 32×32 pixels) is associated with a vector that is filled with zeros. At the end of DMM-1 main stage, the selected pattern number is inserted in the corresponding vector position. Moreover, four integer variables are responsible for storing the last encoded pattern information of each block size.

Figure 5 shows the current PU and the neighbor upper and left PUs. Let $\text{Pattern}_{\text{up}}$, $\text{Pattern}_{\text{left}}$, and $\text{Pattern}_{\text{curr}}$ be the pattern selected by the upper, left, and current PU in the DMM-1 main stage, respectively. Two predictors P_{copy} and P_{extend} are defined considering this information according to Eqs. (1) and (2), respectively. The first predictor P_{copy} searches for the pattern with the smallest distortion between copying $\text{Pattern}_{\text{left}}$ and $\text{Pattern}_{\text{up}}$ to the current PU. The second predictor P_{extend} searches for the pattern with the smallest distortion among extended orientation of $\text{Pattern}_{\text{left}}$ and $\text{Pattern}_{\text{up}}$.

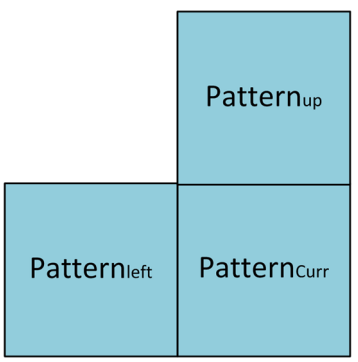

Fig. 5 An example of neighbor PUs and selected patterns in DMM-1 main stage.

$$P_{\text{copy}}:\ \text{Pattern}_{\text{curr}} = P\text{MinDist}(\text{Pattern}_{\text{up}}, \text{Pattern}_{\text{left}}), \tag{1}$$

$$P_{\text{extend}}:\text{Pattern}_{\text{curr}} = P\text{MinDist}(\text{Extend}(\text{Pattern}_{\text{up}}, \text{Pattern}_{\text{left}})), \tag{2}$$

where PMinDist searches for the pattern with the smallest distortion and Extend returns all patterns with extended orientation.

Table 2 shows the extended orientation for P_{extend} according to their reference orientation. Notice that if the wedgelet does not touch the right or bottom border according to its position, its orientations cannot be extended.

Every video sequence in the common test conditions (CTC)[28] were evaluated using the conditions described in Table 3.

During the evaluations, the success rate of each part of each predictor was analyzed according to the block size. Equation (3) describes the success rate, which is the number of cases for which the proposed predictors had success divided by the total number of evaluations. The P_{extend} predictor is only evaluated when P_{copy} fails.

$$\text{Success rate} = \frac{\text{Number of cases that the predictor had success}}{\text{Total number of evaluations}}(\%). \tag{3}$$

Table 2 Orientation types for P_{extend}.

Left PU DMM-1 orientation	Extended orientation	Upper PU DMM-1 orientation	Extended orientation
0	–	0	2
1	3	1	3
2	0	2	–
3	–	3	–
4	–	4	4
5	5	5	–

Table 3 Important test configurations.

Views	3
Bits per pixel	8
DMM evaluation	On
Depth intraskip	On
QP (texture)	25, 30, 35, 40
QP (depth)	34, 39, 42, 45
VSO with depth fidelity term	On
1024×768 sequences	Balloons, Kendo, Newspaper_CC
1920×1088 sequence	GT_Fly, Poznan_Hall2, Poznan_Street, Undo_Dancer, and Shark
Rate control	Off
Encoding mode	All intra
3D-HTM version	16.0

Tables 4–7 display the statistical results, for 4×4, 8×8, 16×16, and 32×32 pixel blocks, respectively. Since the DMMs are not applied in 64×64 PUs, then it is not presented in this analysis.

The total success rate (Total column of each table) is the union of all predictors' success rates (i.e., considers that any of the predictors had success), which is relevant information when combining these techniques into a single solution. Analyzing these results, one can observe that smaller PUs obtain a higher success in P_{copy}, while larger PUs obtain better results in P_{extend}. Moreover, the combination of both predictors allows obtaining a high-success rate independent of the encoding block size.

Table 4 Predictors success rate for 4×4 pixel blocks.

Predictor	P_{copy}		P_{extend}		
Video	Left	Up	Left	Up	Total
Balloons	71.67	70.06	3.31	4.85	84.87
Kendo	78.03	73.93	2.86	4.22	88.87
Newspaper_CC	55.95	53.84	4.88	6.66	73.81
Gt_Fly	77.7	36.68	16.25	2.00	83.74
Poznan_Hall2	87.82	87.31	1.09	2.65	96.30
Poznan_Street	43.96	42.82	6.66	4.77	60.77
Undo_Dancer	71.64	74.26	3.14	2.19	85.64
Shark	68.53	62.37	4.01	4.25	79.91

Table 5 Predictors success rate for 8×8 pixel blocks.

Predictor	P_{copy}		P_{extend}		
Video	Left	Up	Left	Up	Total
Balloons	19.78	25.78	11.84	17.76	53.89
Kendo	26.32	30.26	13.31	16.77	60.09
Newspaper_CC	14.18	17.79	10.94	17.77	45.11
Gt_Fly	35.44	17.23	42.97	6.45	56.85
Poznan_Hall2	33.45	64.09	9.71	36.99	84.85
Poznan_Street	9.23	8.29	16.94	11.54	36.94
Undo_Dancer	36.35	40.97	15.9	17.43	67.91
Shark	10.44	10.24	13.60	10.51	35.18

Table 6 Predictors success rate for 16×16 pixel blocks.

Predictor	P_{copy}		P_{extend}		
Video	Left	Up	Left	Up	Total
Balloons	11.72	16.46	12.10	21.75	48.54
Kendo	17.42	20.86	14.07	19.99	54.54
Newspaper_CC	8.15	9.92	11.98	19.7	40.47
Gt_Fly	32.33	5.42	57.84	5.12	63.79
Poznan_Hall2	27.65	52.92	12.52	41.24	82.75
Poznan_Street	5.42	5.05	26.94	13.02	43.67
Undo_Dancer	20.47	19.91	26.4	33.31	69.23
Shark	8.68	7.06	16.51	11.74	35.63

Table 7 Predictors success rate for 32×32 pixel blocks.

Predictor	P_{copy}		P_{extend}		
Video	Left	Up	Left	Up	Total
Balloons	7.83	11.14	12.07	17.58	38.94
Kendo	9.76	13.25	14.19	19.37	45.63
Newspaper_CC	4.90	5.57	12.26	18.74	34.96
Gt_Fly	39.03	11.35	71.90	2.80	74.16
Poznan_Hall2	23.12	45.45	13.13	39.42	77.41
Poznan_Street	5.50	5.07	39.95	11.62	52.67
Undo_Dancer	20.88	11.21	24.98	39.52	68.16
Shark	8.43	4.01	20.13	9.89	34.40

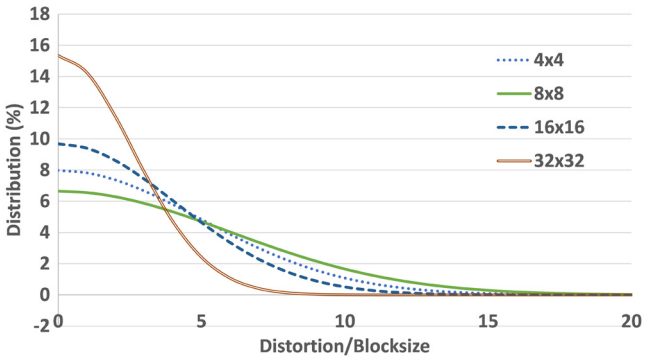

Fig. 6 Probability density function of P_{copy} having success according to the distortion divided by the block size. This function was achieved executing the *GT_Fly* video sequence under all-intra mode.

Along with this analysis, Fig. 6 presents the probability density function of having success in P_{copy} according to the obtained distortion divided by the block size, which was used as an evaluation criterion since larger block sizes tend to have larger distortion values.

This function, which follows a Gaussian distribution, was achieved executing the GT Fly video sequence under all-intra mode. The results show a high probability of these predictors in having success for small values of distortion, having almost no chance to have success for values of distortion divided by the block size over 15 since larger distortion values mean that the PU should be encoded by a different pattern. Similar analysis has been performed for P_{extend} resulting in similar conclusions. Thus, the value of distortion divided by the block size obtained by the predictors is directly related to the success/fail rate of the predictor. Hence, it can be used to perform an early termination decision according to a threshold criterion. Moreover, these experiments enable us to conclude that there is a high probability of our predictors having success by performing low evaluations (at most two evaluations should be performed) of DMM-1 in P_{copy} or performing medium assessments (a subset of patterns should be evaluated before requiring the refinement) of DMM-1 in P_{extend}, for all available block sizes.

4 Depth Modeling Mode 1 Fast Pattern Selector

Figure 7 shows the dataflow model of DFPS inside the DMM-1 encoding algorithm, which was designed based on the analysis presented in Sec. 3.1. Looking for a light-weight solution, capable of skipping many DMM-1 evaluations, DFPS starts finding the minimum distortion in P_{copy}. If the distortion divided by the block size is lower than the THreshold 1 (TH_1) defined by the offline analysis described in Sec. 4.1, then the DMM-1 main process is finalized, and the refinement is performed followed by the residue computation.

When TH_1 criterion is not met, further evaluations are required to obtain a good prediction on DMM-1 Main Stage. Therefore, instead of evaluating the entire DMM-1 initial set, we propose a medium-weight solution where P_{extend} is evaluated extending left and upper PUs wedgelets' orientation. Once more, the minimum distortion divided by the block size is compared with THreshold 2 (TH_2). If it is smaller than TH_2, then the main stage is finalized, and the refinement stage can be performed. Otherwise, further

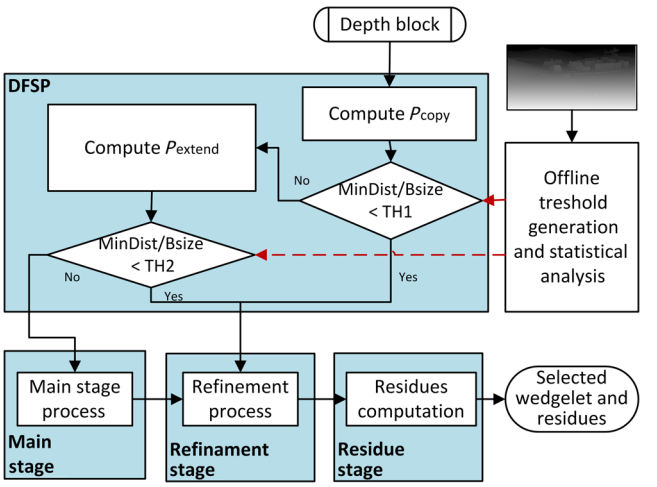

Fig. 7 DFPS dataflow model.

evaluations are required, and the remaining wedgelets are evaluated without any simplification.

In the worst case (i.e., no simplification is performed) the rate-distortion cost (RD-cost) calculated by our solution is the same as the traditional approach. Notice that our solution always stores the information of the DMM-1 Main Stage pattern in a vector and in an integer variable. Consequently, even if other modes have been selected in the entire RD-cost evaluation, the DMM-1 information can be used to accelerate neighbor encoding PUs.

4.1 *Definition of Thresholds*

The threshold values lead to light or aggressive solutions regarding complexity reduction. Thus, we employed an extensive experimental analysis that evaluates 16 scenarios to explain the impact of the threshold variation. In the analysis presented in Sec. 3.1, the average and standard deviation of the distortion divided by the block size has been saved from GT_Fly video sequence, and it was used to determine the evaluations scenarios.

The evaluation scenarios used for each threshold (called here TH_n), the Eq. (4), where k is empirically selected ranging from 0 to 3, and u_n and std_n represent the average and standard deviation value for the TH_n computation, respectively.

$$TH_n = u_n + k \times std_n.$$

Table 8 presents all evaluated thresholds, with $u_1 = 1$, $u_2 = 2$, $std_1 = 5$, and $std_2 = 25$.

Figure 8 depicts the results of these evaluations highlighting the percentage of DMM-1 patterns evaluation in the main stage according to the threshold criterion.

Figure 9 displays the percentage of wedgelets' reduction according to the well-accepted Bjontegaard Delta rate (BD-rate) criterion.[29] The proposed solution allows different operation points that are capable of providing better coding effectiveness or higher pattern skips. We detached two points in Figs. 8 and 9 (case-2 and sase-7 in Table 8) that we considered as the best operation points for the study case evaluated in Sec. 5.

Table 8 Thresholds for the evaluated scenarios.

Case	TH_1	TH_2	Case	TH_1	TH_2
Case-1	1	2	Case-9	1	52
Case-2	6	2	Case-10	6	52
Case-3	11	2	Case-11	11	52
Case-4	16	2	Case-12	16	52
Case-5	1	27	Case-13	1	77
Case-6	6	27	Case-14	6	77
Case-7	11	27	Case-15	11	77
Case-8	16	27	Case-16	16	77

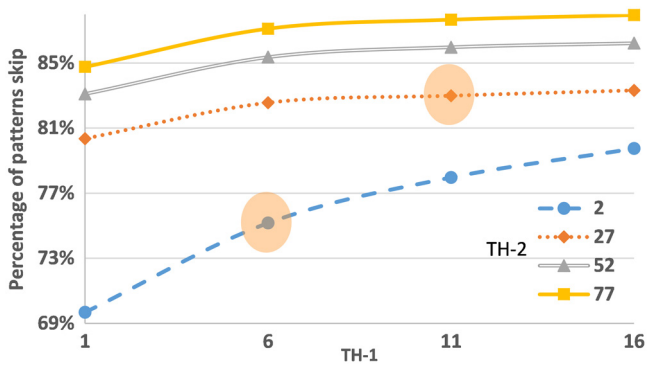

Fig. 8 Percentage of skips on DMM-1 pattern evaluation according to some threshold values.

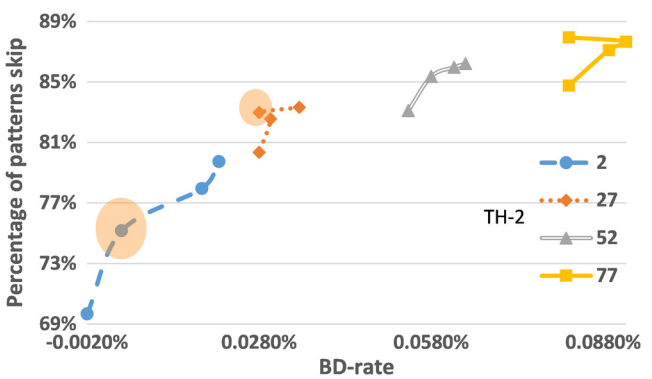

Fig. 9 Percentage of skips on DMM-1 patterns evaluation according to the BD-rate impact.

5 Case Study: Results and Discussion for Case-2 and Case-7

This section describes the results of DFPS complexity reduction technique under all-intra and random access modes for both case-2 and case-7 (as presented in Sec. 4.1). All experiments conducted in this article were evaluated using 3D-HTM 16.0 and using CTC for 3-D core experiments, where the encoding effectiveness of the videos encoded by the proposed solution are compared with the 3D-HTM 16.0 software without modifications. The same conditions detailed in Table 3 were applied in this evaluation. The CTC 3-views case is used, where three texture cameras and three depth maps are encoded. Since depth maps are not presented to TV users, six views are synthesized between the encoded texture views using View Synthesis Reference Software (VSRS) provided by JCT-3V.[30] We employed the BD-rate criterion to obtain the compression rate for the same objective quality compared to 3D-HTM 16.0. A comparison against related work is provided in Sec. 5.3. In addition, a subjective quality assessment is presented in Sec. 5.4, where absolute category rating with hidden reference (ACR-HR)[31] is applied comparing the proposed solutions against the original 3D-HTM solution, corroborating the objective analysis.

5.1 *All-Intra Evaluation*

The all-intra evaluation performs only the intracoding flow; consequently, DMM-1 has a high impact on this assessment. As our solution only targets intraprediction, this is the fairest evaluation scenario for the proposed algorithm.

Table 9 illustrates the results of all-intra evaluation for case-2 and case-7. In Table 9, the DMM-1 skipped column means a reduction in DMM-1 evaluations when DFPS is applied, the synthesis only BD-rate column represents an increase in the encoded video bitrate for a similar video quality and, the encoding time savings column represents the execution time reduction considering the whole encoder (texture and depth coding).

In average, the designed solutions are capable of achieving 11.7% and 13.4% of the encoder complexity reduction when the all-intra mode is considered, with a drawback of 0.0467% and 0.1714% in BD-rate, respectively. The BD-rate impact is negligible if compared with the achieved time-saving, showing the quality of our approach. These results are obtained because the proposed approach is capable of achieving more than 70% and 83% skips on DMM-1 evaluations under case-2 and case-7, respectively.

Figures 10 and 11 depict detailed results of DMM-1 evaluation per block size and per 3-D video sequence for case-2 and case-7, respectively. For both cases, higher skips are obtained in 4×4 pixel blocks where, as presented in Sec. 3.1, there is a higher probability that the proposed predictors have success. On the other hand, larger block sizes require the evaluation of more than 300 wedgelets making it hard to achieve success, which corroborates the obtained results.

Figure 12 illustrates the RD-plot for balloons' video sequence under all-intra mode. One can notice that the RD-plot does not present a visible variation between the 3D-HEVC, case-2, and case-7 curves, which happened to all other videos (omitted due to the limit of space). Applying a high-level of zoom, one can notice that case-2 and case-7 represent the minor impact of peak signal-to-noise ratio (PSNR) in synthesized views for the same bitrate. This impact is acceptable considering that many wedgelets patterns are not being evaluated.

5.2 *Random Access Evaluation*

The proposed algorithm has also been evaluated in random access mode, which considers the I–P–B frame structure,

Table 9 Results of all-intra evaluation for case-2 and case-7.

		Case-2			Case-7		
Resolution	Videos	DMM-1 skipped	Synthesis only BD-rate	Encoding time savings	DMM-1 skipped	Synthesis only BD-rate	Encoding time savings
1024 × 768	Balloons	58.75%	0.0152%	8.1%	77.60%	0.1249%	12.6%
	Kendo	58.75%	0.0189%	6.5%	77.60%	0.1366%	12.1%
	Newspaper_cc	57.60%	0.0733%	10.8%	77.34%	0.3547%	12.9%
	Average	58.37%	0.0358%	8.5%	77.51%	0.2054%	12.5%
1920 × 1088	Gt_fly	75.17%	0.0045%	15.3%	83.00%	0.0281%	10.5%
	Poznan_hall2	89.77%	0.1847%	13.3%	95.92%	0.4898%	14.3%
	Poznan_street	79.07%	0.0380%	14.6%	91.14%	0.1540%	17.0%
	Undo_dancer	82.87%	0.0298%	13.7%	87.74%	0.0346%	14.0%
	Shark	65.67%	0.0152%	11.0%	80.41%	0.0488%	13.8%
	Average	78.51%	0.0533%	13.6%	87.64%	0.1511%	13.9%
Average		70.96%	0.0467%	11.7%	83.84%	0.1714%	13.4%

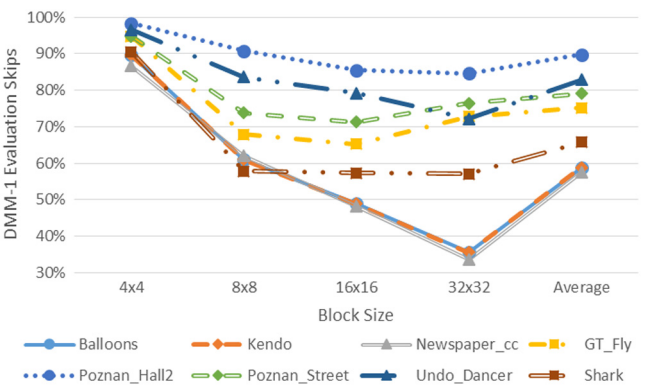

Fig. 10 Percentage of DMM-1 evaluation skips when applying DFPS case-2 under all-intra mode.

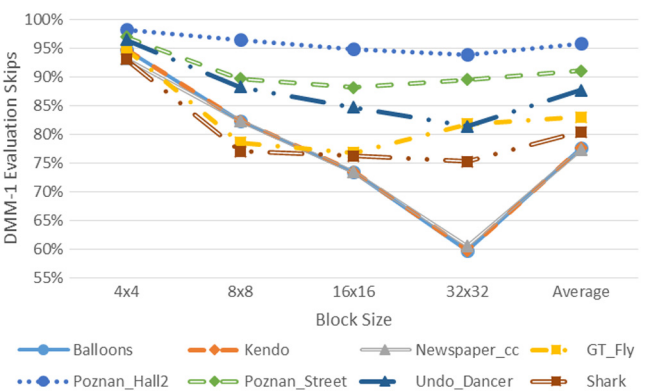

Fig. 11 Percentage of DMM-1 evaluation skips when applying DFPS case-7 under all-intra mode.

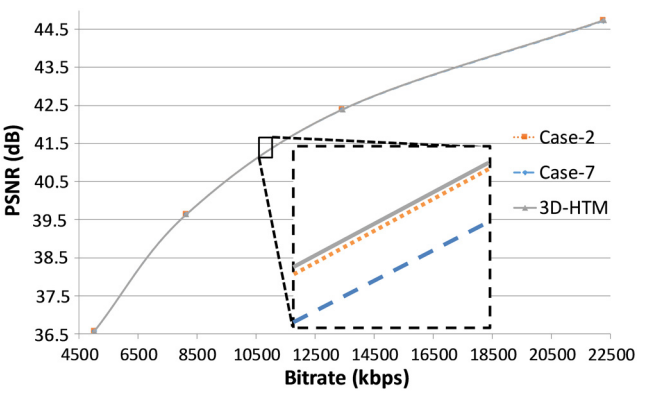

Fig. 12 RD-curve for balloons video sequence under all-intra mode.

where only intra-algorithms predict the first encoded frame. The frames in the same layer of one frame are predicted using disparity compensated prediction and intra-algorithms, and the remaining frames are predicted using all available encoding techniques. This scenario presents the realistic impact of the proposed technique inside a 3-D video encoder; however, due to the high complexity of P and B frames, the intratechniques present a small relative complexity.

Table 10 shows the results of random access evaluation for case-2 and case-7. In random access, our solution is capable of reducing the encoding time by 1.2% and 1.3% with a minor BD-rate impact for case-2 and case-7, respectively. In addition, the BD-rate is reduced by 0.0048% in case-2, showing that the proposed solution is capable of achieving gains in both complexity and quality axis.

Table 10 Results of random access evaluation for case-2 and case-7.

Resolution	Videos	Case-2			Case-7		
		DMM-1 skipped	Synthesis only BD-rate	Encoding time savings	DMM-1 skipped	Synthesis only BD-rate	Encoding time savings
1024 × 768	Balloons	52.48%	−0.0951%	0.8%	71.36%	0.0493%	1.0%
	Kendo	57.63%	−0.0247%	1.1%	74.29%	0.0135%	1.4%
	Newspaper_cc	52.39%	0.0230%	1.1%	71.52%	0.2447%	1.4%
	Average	54.16%	−0.0323%	1.0%	72.39%	0.1025%	1.3%
1920 × 1088	Gt_fly	64.62%	−0.0054%	0.8%	74.18%	−0.0436%	0.9%
	Poznan_hall2	85.92%	−0.0432%	1.3%	93.80%	0.1620%	0.7%
	Poznan_street	71.62%	0.0795%	1.7%	86.12%	0.1927%	1.8%
	Undo_dancer	74.01%	0.0218%	1.5%	79.97%	−0.0161%	1.7%
	Shark	52.84%	0.0053%	0.9%	68.60%	−0.0223%	1.3%
	Average	69.80%	0.0132%	1.2%	80.53%	0.0737%	1.3%
Average		63.94%	−0.0048%	1.2%	77.48%	0.0725%	1.3%

The timesaving gains are related to the high number of DMM-1 evaluations skips that are detailed according to the block sizes in Figs. 13 and 14 for both cases.

Figure 15 presents the RD-plot for balloons video sequence under random access mode for 3D-HTM, case-2, and case-7, corroborating our analysis. Again, observing the curve without zoom demonstrates that the proposed algorithm does not represent any impact in the encoding effectiveness. Additionally, when looking at the curve with zoom, one can see that case-2 is capable of providing minor gains in synthesized views for the same bitrates than 3D-HTM, while case-7 provides a smaller level of PSNR.

5.3 *Wedgelets Evaluation Comparisons*

Table 11 summarizes our comparison against the related works that also focus on reducing DMMs wedgelets evaluations.[19,20,26] Since these works were developed in

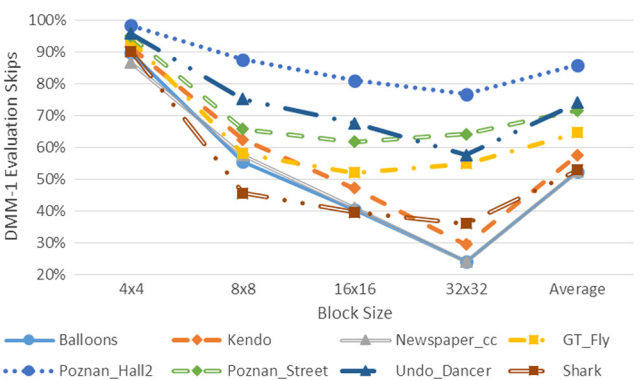

Fig. 13 Percentage of DMM-1 evaluation skips when applying DFPS case-2 under random access mode.

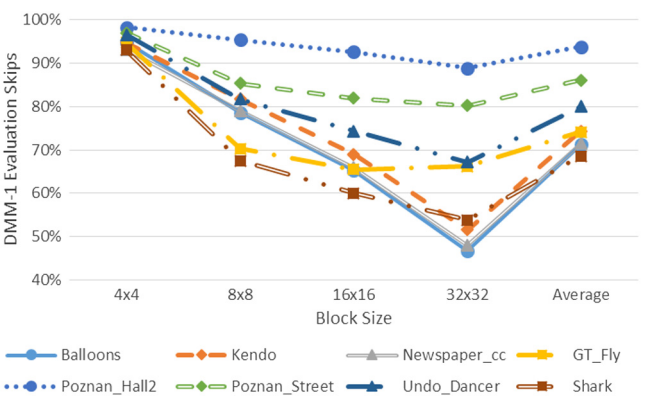

Fig. 14 Percentage of DMM-1 evaluation skips when applying DFPS case-7 under random access mode.

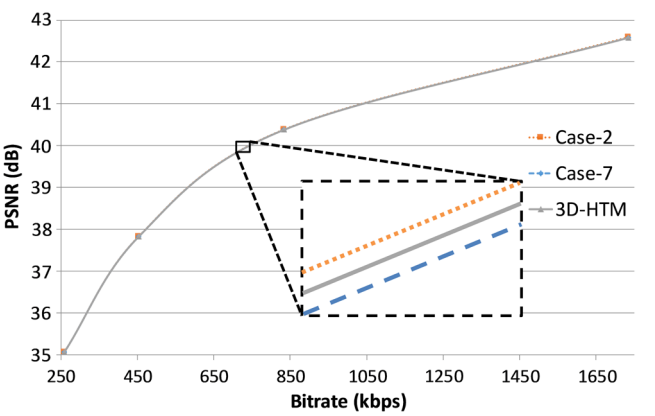

Fig. 15 RD-curve for balloons video sequence under random access mode.

Table 11 Comparison with related works.

Work	3D-HTM version	Preprocessing	Wedgelets evaluation (reduction)	Evaluation reduction	BD-rate
DFPS-2	16.0-AI	—	Two wedgelets or two wedgelet + one orientation set or traditional process	70.96%	0.0467%
DFPS-7	16.0-AI			83.94%	0.1714%
DFPS-2	16.0-RA			63.84%	−0.0048%
DFPS-7	16.0-RA			77.48%	0.0725%
Ref. 19	7.0-RA	Apply a gradient filter in the borders	Wedgelets touching borders in 8 positions + refinement	73.95%	−0.047%
Ref. 20	11.0-AI	—	Wedgelets following horizontal mode, vertical mode, angular mode 5, angular mode 14, angular mode 21, and angular mode 30	—	—
Ref. 20	11.0-RA			—	—
Ref. 26	10.2-AI	Apply a gradient filter in the borders	1 wedgelet + refinement	—	1.06%

different 3D-HTM versions, they cannot be directly compared with this work. The main important aspect in the proposed comparison is verifying the proposed technique and the reduction of the DMMs wedgelet evaluation that was achieved when applying the proposed technique.

The work in Refs. 19 and 26 applied a preprocessor in the borders of the encoding PU. This preprocessor computes the gradients of the entire borders and sorts these gradients according to their values. In Ref. 19, only the wedgelets that touch in the eight highest gradients were evaluated, followed by the original refinement, resulting in a reduction of 73.95%. In Ref. 26, only the nearest wedgelet of the two highest gradients from two distinct borders is evaluated, followed by a refinement. The reduction in DMM-1 evaluation is not presented in Ref. 26. The solution proposed in Ref. 20, focusing on DMM-1 complexity reduction, evaluates wedgelets following horizontal mode, vertical mode, and angular modes with 5, 14, 21, and 30 directions. The only detail of reducing DMM-1 complexity in Ref. 20 (since it was evaluated with high-level decision algorithms) is that this technique is capable of achieving the same results of traditional DMM-1 algorithm 80% of the time.

Considering the reductions obtained by related works that focus on DMM-1 wedgelet search, one can conclude that our solution is capable of reducing this DMM-1 search complexity significantly. When analyzing the BD-rate obtained by the other solutions, it is clear that the proposed solution presents a negligible impact on the BD-rate performance. Moreover, our solution is the only one that presents a subjective quality analysis, as presented in Sec. 5.4.

The techniques proposed in Refs. 19, 20, and 26 could also be integrated with DFPS for building a third-level decision without requiring the processing of the entire wedgelet list in the worst case. Moreover, the proposed technique could be easily integrated with a high-level decision algorithm such as the algorithms provided by Refs. 20, 21, 23, and 24 to obtain a high reduction in the DMMs evaluation and a DMM-1 complexity reduction when the DMMs are required to be evaluated.

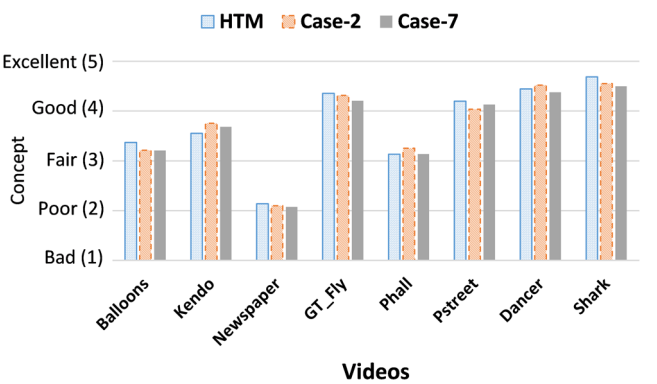

Fig. 16 Subjective assessment results.

5.4 *Subjective Quality Analysis*

A subjective quality analysis using ACR-HR [31] has also been performed to assess the quality of the synthesized views better when applying DFPS, complementing the objective analysis presented in the previous sections. This experiment shows synthesized views from all videos in CTCs using all-intra mode and the quantization parameter pair (30, 39) to 13 observers with ages ranging from 18 to 43 years. Figure 16 demonstrates the average results obtained with this analysis for 3D-HTM, case-2, and case-7.

Notice that a small variation in quality is achieved according to the video sequence. Case-2 and case-7 lose 0.015 and 0.069 points in comparison with 3D-HTM, on average. This difference is insignificant compared to the high gains provided by the complexity reduction of our solutions. Moreover, by applying student's t-test [32] with 95% of confidence per video, no statistical difference in quality results has been noticed when comparing these two versions of DFPS with 3D-HTM-16.0.

6 Conclusions

This article presents the DFPS algorithm that uses some information of neighbor blocks encoding to increase the

encoding efficiency of the current block, reducing the DMM-1 complexity. DFPS employs two predictors, a lightweight and a medium-weight. If the lightweight predictor ends its analysis by a threshold criterion, DFPS skips an aggressive DMM-1 pattern evaluation. Otherwise, the medium-weight predictors are applied. If the medium-weight predictor has success, the first DMM-1 search process is finalized. Otherwise, the traditional DMM-1 approach is performed. The thresholds have been defined under extensive experimentation. The proposed algorithm was evaluated under CTC[28] for two operation points, obtaining a complexity reduction of 11.7% and 13.4% with a drawback of only increasing the BD-rate by 0.047% and 0.171%, respectively, under all-intra mode. These results are achieved by skipping almost 71% and 84% of the DMM-1 pattern evaluation. A subjective quality assessment corroborates the quality of our solution, demonstrating that DFPS is capable of obtaining a high complexity reduction without impacting the view synthesis process.

References

1. P. Merkle et al., "Coding efficiency and complexity analysis of MVC prediction structures," in *15th European Signal Processing Conf. (EUSIPCO '07)*, pp. 5–9 (2007).
2. K. Muller et al., "3D high-efficiency video coding for multi-view video and depth data," *IEEE Trans. Image Process.* **22**(9), 3366–3378 (2013).
3. G. Tech et al., "Overview of the multiview and 3D extensions of high efficiency video coding," *IEEE Trans. Circuits Syst. Video Technol.* **26**(1), 35–49 (2016).
4. ITU-T/ISO/IEC, "Joint collaborative team on 3D video coding extension development, JCT-3V," 2016, http://phenix.int-evry.fr/jct2/ (23 May 2016).
5. G. Sullivan et al., "Overview of the high efficiency video coding (HEVC) standard," *IEEE Trans. Circuits Syst. Video Technol.* **22**(12), 1649–1668 (2012).
6. Z. Liu et al., "H.264/MVC interleaving for real-time multiview video streaming," *J. Real-Time Image Process.* **10**(3), 501–511 (2015).
7. A. Vetro, T. Wiegand, and G. Sullivan, "Overview of the stereo and multiview video coding extensions of the H.264/MPEG-4 AVC standard," *Proc. IEEE* **99**(4), 626–642 (2011).
8. L. Zhang, Y. Chen, and M. Karczewicz, "Disparity vector based advanced inter-view prediction in 3D-HEVC," in *IEEE Int. Symp. on Circuits and Systems (ISCAS '13)*, pp. 1632–1635 (2013).
9. J. Kang et al., "Low complexity neighboring block based disparity vector derivation in 3D-HEVC," in *IEEE Int. Symp. on Circuits and Systems (ISCAS '14)*, pp. 1921–1924 (2014).
10. X. Li, L. Zhang, and Y. Chen, "Advanced residual prediction in 3D-HEVC," in *IEEE Int. Conf. on Image Processing (ICIP '13)*, pp. 1–5 (2013).
11. H. Liu et al., "3D-CE2.h: results of illumination compensation for inter-view prediction," in *Joint Collaborative Team on 3D Video Coding Extension Development (JCT-3V '12)*, Doc. JCT3V-B0045, Shanghai (2012).
12. P. Kauff et al., "Depth map creation and image-based rendering for advanced 3DTV services providing interoperability and scalability," *Image Commun.* **22**(2), 217–234 (2007).
13. Z. Pan, Y. Zhang, and S. Kwong, "Fast mode decision based on texture-depth correlation and motion prediction for multiview depth video coding," *J. Real-Time Image Process.* **11**(1), 27–36 (2016).
14. Y. Zhao et al., "Boundary artifact reduction in view synthesis of 3D video: from perspective of texture-depth alignment," *IEEE Trans. Broadcast.* **57**(2), 510–522 (2011).
15. C. Fehn, "Depth-image-based rendering (DIBR), compression, and transmission for a new approach on 3D–TV," *Proc. SPIE* **5291**, 93–104 (2004).
16. A. Smolic et al., "Intermediate view interpolation based on multiview video plus depth for advanced 3D video systems," in *IEEE Int. Conf. on Image Processing (ICIP '08)*, pp. 2448–2451 (2008).
17. L. Vostersm, C. Varekamp, and G. Haan, "Overview of efficient high-quality state-of-the-art depth enhancement methods by thorough design space exploration," *J. Real-Time Image Process.* 1–21 (2015).
18. E. Mora et al., "Initialization, limitation and predictive coding of the depth and texture quadtree in 3D-HEVC," *IEEE Trans. Circuits Syst. Video Technol.* **24**(9), 1554–1565 (2014).
19. G. Sanchez et al., "A complexity reduction algorithm for depth maps intra prediction on the 3D-HEVC," in *IEEE Visual Communications and Image Processing Conf.*, pp. 137–140 (2014).
20. Q. Zhang et al., "Fast intra mode decision for depth coding in 3D-HEVC," *Multidimens. Syst. Signal Process.* 1–24 (2016).
21. G. Sanchez et al., "Complexity reduction for 3D-HEVC depth maps intra-frame prediction using simplified edge detector algorithm," in *Int. Conf. on Image Processing (ICIP '14)*, pp. 3209–3213 (2014).
22. L. Guo et al., "Hole-filling map-based coding unit size decision for dependent views in three-dimensional high-efficiency video coding," *J. Electron. Imaging* **25**(3), 033020 (2016).
23. Z. Gu et al., "Fast depth modeling mode selection for 3D HEVC depth intra coding," in *IEEE Int. Conf. on Multimedia and Expo Workshops (ICMEW)*, pp. 1–4 (2013).
24. C. Park, "Edge-based intramode selection for depth-map coding in 3D-HEVC," *IEEE Trans. Image Process.* **24**(1), 155–162 (2015).
25. H. Lei et al., "Early termination algorithm for the depth modeling mode in three-dimensional extension of high-efficiency video coding," *J. Electron. Imaging* **25**(5), 050503 (2016).
26. M. Saldanha et al., "Solutions for DMM-1 complexity reduction in 3D-HEVC based on gradient calculation," in *Latin American Symp. on Circuits and Systems (LASCAS '16)*, pp. 211–214 (2016).
27. P. Merkle et al., "3D video: depth coding based on inter-component prediction of block partitions," in *Picture Coding Symp. (PCS '12)*, pp. 149–152 (2012).
28. K. Muller and A. Vetro, "Common test conditions of 3DV core experiments," ISO/IEC JTC1/SC29/WG11, document JCT3V-G1100, San José (2014).
29. G. Bjontegaard, "Calculation of average PSNR differences between RD curves," 13th VCEG Meeting, VCEG-M33, ITU-T SG16/Q6 VCEG, Austin, Texas (2001).
30. K. Muller et al., "View synthesis for advanced 3D video systems," *EURASIP J. Image Video Process.* **2008**, 438148 (2009).
31. ITU—International Telecommunication Union, "P. 910: subjective video quality assessment methods for multimedia applications," Geneva, Switzerland (2009).
32. D. W. Zimmerman, "Teacher's corner: a note on interpretation of the paired-samples t test," *J. Educ. Behav. Stat.* **22**(3), 349–360 (1997).

Gustavo Sanchez has been a professor at the Federal Institute Farroupilha, Brazil, since 2014. He received his master's degree in computer science from the Federal University of Pelotas (2014). He is currently pursuing his PhD in computer science at the Pontifical Catholic University of Rio Grande do Sul. He has 7+ years of experience on designing algorithms and hardware for video coding. His interests include complexity reduction algorithms and dedicated hardware design for video coding.

Lucas Jordani has been an undergraduate in computer engineering at Pontifical Catholic University of Rio Grande do Sul since 2012. His research interests include complexity reduction and control algorithms for 3-D video coding.

César Marcon has been a professor in the graduate program in Computer Science of Pontifical Catholic University of Rio Grande do Sul, Brazil, since 1995. He received his PhD in computer science from Federal University of Rio Grande do Sul, Brazil, in 2005. He is a member of IEEE. He has more than 100 papers published in prestigious journals and conference proceedings. Since 2005, he has coordinated 9 research projects in areas of telecom and telemedicine.

Luciano Agostini has been a professor since 2002 at Federal University of Pelotas, Brazil, where he leads the group of architectures and integrated circuits and since 2013 he has been the executive vice president for research and graduate studies. He received his PhD from UFRGS, Brazil, in 2007. He has more than 200 published papers in journals and conference proceedings. He is a senior member of IEEE and a member of ACM, SBC, SBMicro.