Low Complexity Intra Mode Decision Algorithm for 3D-HEVC

Hamza HAMOUT
Computer Systems and Vision Laboratory
Faculty of Sciences Ibn-Zohr University
Agadir Morocco
Email: hamzahamout@gmail.com

Abderrahmane ELYOUSFI
Department of Computer Science
National Engineering School of Applied
Sciences Ibn-Zohr University Agadir Morocco
Email: elyousfiabdo@yahoo.fr

***Abstract*—The 3D High Efficiency Video Coding (3D-HEVC) has recently been approved by the Joint Collaborative Team on 3D Video coding (JCT-3V) and standardized as an extension of HEVC. To improve the intra coding efficiency of Multi-view Video plus Depth maps (MVD), the 3D-HEVC brings a good intra coding solution, but the computational complexity load increase significantly, which restricts the use of 3D-HEVC encoders in real world application. Therefore, an efficient intra mode decision algorithm is needed. To overcome the aforementioned problem, this paper presents a new mode prediction method enabling high complexity reduction of 3D-HEVC intra coding. The experimental results show that the low complexity intra mode decision algorithm increases the speed of intra coding significantly with negligible loss of encoding efficiency.**

***Index Terms*—3D-HEVC, Intra coding, JCT-3V, Tensor structure, isotropic Gaussian filter.**

I. Introduction

The Joint Collaborative Team on 3D Video coding (JCT-3V) finalizing a new standard for 3D video coding. The new standard is known as 3D video extension of High Efficiency Video Coding (3D-HEVC) [1-3]. To improve the multimedia techniques, the Multi-view Video plus Depth maps (MVD) is opted as the best new format for 3D video representation [4]. In the Multi-view Video plus Depth maps format, a video signal is composed of a depth and a texture component. The both components are encoded and multiplexed into a 3D bitstream. After encoding, the decoder uses the Depth Image Based Rendering technique (DIBR) to synthesize the intermediate rendered views spatially positioned between each view [5].

3D-HEVC intra coding is one of the most important technique used in 3D video compression. Depth intra coding employs both conventional HEVC intra prediction mode and new intra prediction modes to select the best coding mode [6], whilst texture intra coding procedure uses only the conventional HEVC intra coding modes [3] (DC ,planar and 33 directional intra prediction modes). To reach the highest coding efficiency, 3D-HEVC encoder examines all intra coding modes corresponding to each coding unit (CU) using a Rate Distortion cost (RD-cost) [3], which involves a high computational complexity knowing the crucial role that it plays. Thus, it is very needed to develop a low complexity intra mode decision algorithm for 3D-HEVC that decreases the computational complexity with negligible loss of coding efficiency.

Recently, a number of fast algorithms have been proposed to reduce the computational complexity of 3D-HEVC intra coding [7-12]. An adaptive mode selection method is presented in [7] to improve the 3D-HEVC depth intra coding based on the correlation between the best intra prediction and the smallest rough-RD cost value. In [8], a fast technique based on Simplified Edge Detector (SED) to reduce the complexity of the depth intra coding. The algorithm proposed in [9] opted to skipped unnecessary Depth Modeling Modes (DMMs) by investigating the edge classification in Hadamard transform domain. The authors of [10] propose a fast DMM selection by using the information of the Most Probable Modes (MPMs) in the Rough Mode Decision (RMD). The work presented in [11] proposes a threshold method to skip DMMs when the variance of a PU is bigger than that threshold. A fast method is presented in [12] that optimize the 3D-HEVC intra prediction process by analyzing each prediction block using a Sobel filter to estimate the edge direction.

To further reduce the computational complexity of the 3D-HEVC intra coding, we propose an efficient edge detection algorithm based on structure tensors for conventional HEVC intra prediction mode. The rest of this paper is structured as follows. Section II presents an overview of the 3D-HEVC intra coding process. Followed by Section III which gives an introduction on structure tensor and a description of our intra prediction algorithm. Section IV presents the simulation results, whilst Section V concludes this paper.

II. OVERVIEW OF 3D-HEVC INTRA CODING

3D-HEVC introduces a hierarchical structure subdivision using three units: coding unit (CU), prediction unit (PU) and transform unit (TU) [1]. A picture is divided into largest coding units (LCU) and 2N×2N CU. Each 2N×2N CU can be recursively splitted into four N×N CUs up to maximum coding unit depth. PU is utilized as an inter and intra prediction unit from 64×64 down to 4×4. TU is the third unit used for transform and quantization, where both the size and the PU are allowed to be uncorrelated. The maximum transform size is 32×32 and the minimum size is 4×4 [3].

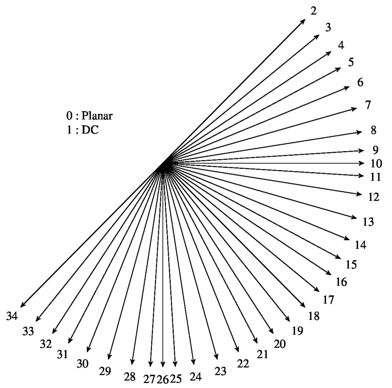

Fig. 1. Conventional HEVC intra prediction.

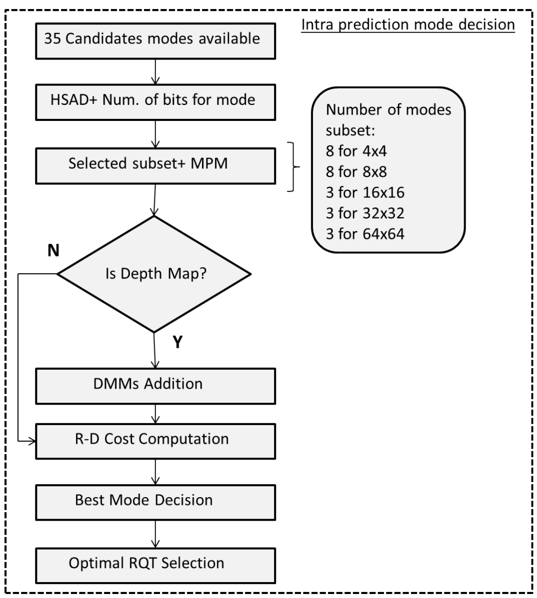

Fig. 2. 3D-HEVC intra mode decision process.

The 3D-HEVC intra prediction performs each PU using the conventional HEVC intra prediction, DC and planar [3], as well as 33 directional intra prediction modes as shown in Fig. 1. This perform is evaluated using the Most Probable Modes in the Rough Mode Decision (Fig. 2) [13]. Note that the number of available intra modes is related to the PU size: 4 modes for 64×64 PUs; 35 modes for 8×8, 16×16 and 32×32 PUs; and 18 modes for 4×4 PUs [14]. In the next step (but only for depth map), 3D-HEVC encoder uses DMM (Depth Modeling Modes) as a new additional intra coding tools [6].

The DMMs performs two splitting strategies: Wedgelet and Contour splitting [1]. Each splitting strategy subdivides the prediction block into two areas and each area will be predicted using a Constant Partition Value (CPV) of the two areas P_1 and P_2 [15]. For a Wedgelet splitting, the two areas are separated by a straight line L_{SE} as demonstrated in Fig. 3 top, where S and E are respectively the start and end points of the straight line L_{SE}. As can be shown in Fig. 3 top, the splitting information is stored in form of a binary N×N matrix.

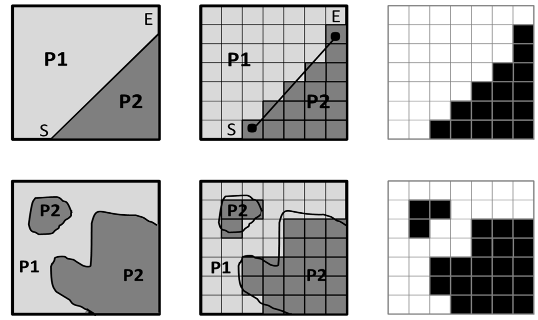

Fig. 3. Wedgelet partition (top) and contour partition (bottom) of a depth block: continuous (left) and discrete signal space (middle) with corresponding partition pattern (right).

Each sample in the matrix describes whether the pixel belongs to the segment P_1 or P_2 , which are represented by the black and white pixels. All possible wedgelet pattern are evaluated by calculating the Rate Distortion cost (RD-Cost) and the optimal wedgelet pattern and their corresponding indices are stored in advance at both the encoder and decoder sides [1]. In the Contour splitting mode, the pattern splitting differs from Wedgelet splitting in the way that it is not a geometry guided block division, but texture guided block segmentation. This is called an inter-component predicted mode [16], which uses co-located texture block to generate block splitting. The thresholding method which is based on the mean value of the luminance of the texture block, is used to define two areas as showing in Fig. 3 bottom [1].

III. PROPOSED ALGORITHM

This section gives an introduction to the theory of structure tensors for edge detection, followed by the description of the proposed method.

A. Theory of Structure Tensors

The structure tensor has been utilized extensively in image processing to resolve so many problems such as motion detection, anisotropic filtering, local structure estimation and local orientation [17]. The structure tensor is defined in terms of the first derivative of the image. This method is based on the information of the gradient to estimate the orientation of the edges [18].

Let I be a 2D image and I(x, y) a point of this image. At this point, the tensor matrix and the structure tensor are respectively defined as follows:

$$J_0 = \begin{bmatrix} g_x^2 & g_x g_y \\ g_x g_y & g_y^2 \end{bmatrix} \tag{1}$$

$$J_\sigma = \begin{bmatrix} g_x^2 * G_\sigma & g_x g_y * G_\sigma \\ g_x g_y * G_\sigma & g_y^2 * G_\sigma \end{bmatrix} = \begin{bmatrix} J_{xx} & J_{xy} \\ J_{xy} & J_{yy} \end{bmatrix} \tag{2}$$

where $G_\sigma = \frac{1}{2\pi\sigma^2} e^{-\frac{x^2+y^2}{2\sigma^2}}$ is a two-dimensional Gaussian function with 0 mean and standard deviation σ , g_x and g_y are respectively the horizontal and vertical gradient vector

$$Gradient\_X = \begin{bmatrix} 1 & 0 & -1 \\ 1 & 0 & -1 \\ 1 & 0 & -1 \end{bmatrix} \quad Gradient\_Y = \begin{bmatrix} 1 & 1 & 1 \\ 0 & 0 & 0 \\ -1 & -1 & -1 \end{bmatrix}$$

Fig. 4. Gradient masks (left:g_x, right:g_y).

components at each pixel, the symbol "*" is convolution operator. The structure tensor J_σ is positive semi-definite and symmetric. So it has two orthogonal eigenvectors $\hat{e}_{\lambda_1}$ and $\hat{e}_{\lambda_2}$ corresponding to two positive eigenvalues λ_1 and λ_2, where $\lambda_1 \geq \lambda_2 \geq 0$. The characteristic equation of structure tensor matrix J_σ is:

$$\lambda^2 - (J_{xx} + J_{yy})\lambda + J_{xx}J_{yy} - J_{xy}^2 = 0. \tag{3}$$

The solutions of this characteristic equation are the eigenvalues of structure tensor J_σ:

$$\lambda_1 = \frac{1}{2}\left(J_{yy} + J_{xx} + \sqrt{(J_{yy} - J_{xx})^2 + 4J_{xy}^2}\right) \tag{4}$$

$$\lambda_2 = \frac{1}{2}\left(J_{yy} + J_{xx} - \sqrt{(J_{yy} - J_{xx})^2 + 4J_{xy}^2}\right) \tag{5}$$

The corresponding eigenvector for each eigenvalue is defined as follows:

$$\hat{e}_{\lambda_1} = \begin{bmatrix} 2J_{xy} \\ J_{yy} - J_{xx} + \sqrt{(J_{yy} - J_{xx})^2 + 4J_{xy}^2} \end{bmatrix} \tag{6}$$

$$\hat{e}_{\lambda_2} = \begin{bmatrix} J_{yy} - J_{xx} + \sqrt{(J_{yy} - J_{xx})^2 + 4J_{xy}^2} \\ -2J_{xy} \end{bmatrix} \tag{7}$$

The eigenvector $\hat{e}_{\lambda_1}$ corresponds to the direction in which the intensity varies the most and it is collinear to the gradient (normal direction). Then $\hat{e}_{\lambda_2}$ corresponds to the direction orthogonal to the gradient, so if there is a principal orientation (tangent direction), it corresponds to $\hat{e}_{\lambda_2}$ [17].

The orientation degree ϕ of an eigenvector is given by:

$$\phi = \tan^{-1} \frac{\lambda - J_{xx}}{J_{xy}} \tag{8}$$

B. Overall algorithm

For each pixel I(x, y) of the PU, the corresponding gradient direction is computed using a pair of gradient filters as showing in Fig. 4. Based on these filters, we can compute for each pixel the degree of variation in horizontal (g_x) and vertical (g_y) direction. It becomes very easy to obtain the structure tensor matrix J_σ by smoothing each component of J_0 (Eq. 1) with a Gaussian kernel G_σ (Eq. 2). Finally the orientation ϕ is computed (Eq. 8) using eigenvalue λ_2 (Eq. 5).

To decide if a block of prediction unit contains an edge, and how solid this edge is, a directional edge histogram is created by adding the largest eigenvalue λ_1 (Eq. 4) of each direction from all pixels of the prediction unit. For depth map, we select every direction that has the histogram value different to 0. For Texture, the direction with the largest value is selected as the best intra prediction mode. To reduce the error during the calculation, we add eight other modes, four are selected from left and another four is chosen of the right side in respect with the best intra predicted mode. Note that DC and Planar mode does not present an angular direction, so they will not be included in the directional edge histogram and they will be always selected among the intra modes candidates. Fig. 5 illustrates the overall proposed algorithm.

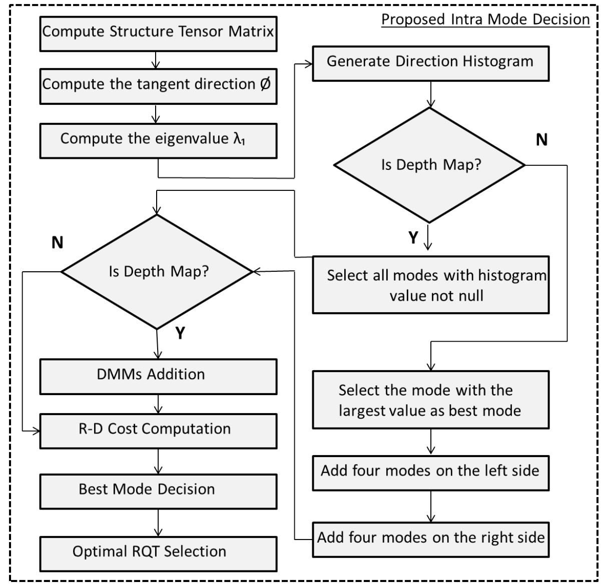

Fig. 5. The proposed 3D-HEVC intra mode decision algorithm.

IV. EXPERIMENTAL RESULTS

In order to perform our proposed algorithm, the experiments were carried out under the Common Testing Conditions (CTC) [19] required by JCT-3V using the recent reference software HTM-16.0 [20]. The video sequences consist in two resolution classes, 1024×768 and 1920×1088, which are listed as flow, Balloons, Kendo, and Newspaper1 with resolution of 1024×768 pixels and GTFly, PoznanHall2, PoznanStreet, UndoDancer and Shark with resolution of 1920×1088 pixels. For each test sequence, three texture views and their corresponding three depth views are encoded. The simulations were performed using all-intra configuration and four QP combinations for texture and depth are tested: (25, 34), (30, 39), (35, 42), (40, 45). The test platform is Intel(R) Xeon(R) CPU E3-1225 v5 @ 3.30GHz, 8GB RAM with Microsoft VS C++ 2010 compiler.

The coding performance is evaluated based on Bjontegaard metrics (BD-BR, BD-PSNR) [21][22], and encoding time saving. Table 1 summarizes the coding performance of both our proposed method and the one presented in [12], in comparison with 3D-HEVC, where in this table, the "only video" columns present the encoding efficiency results corresponding

TABLE I
EXPERIMENTAL RESULTS OF THE PROPOSED ALGORITHM IN ALL-INTRA CASE FOR 3-VIEW UNDER CTC

Sequences	BD-rate only video (%)		BD-PSNR only video (dB)		BD-rate only synthesis (%)		BD-PSNR only synthesis (dB)		Time encoding reduction (%)	
	Proposed	[12]	Proposed	[12]	Proposed	[12]	Proposed	[12]	Proposed	[12]
Balloons	0.08	0.34	-0.004	-0.019	-0.08	0.52	0.012	-0.028	30.8	25.9
Kendo	0.17	0.38	-0.003	-0.018	-0.06	0.62	0.013	-0.010	34.5	26.3
Newspaper1	0.06	0.31	-0.004	-0.014	-0.45	0.44	0.057	-0.020	23.6	24.7
GTFly	0.26	0.39	-0.011	-0.014	-1.50	0.62	0.062	-0.022	35.0	27.8
PonznanHall2	0.69	0.20	-0.004	-0.005	-2.27	1.26	0.029	-0.029	40.4	25.9
PonznanStreet	-0.01	0.23	-0.006	-0.009	-1.44	0.29	0.074	-0.012	24.2	24.0
UndoDancer	-0.01	0.27	-0.012	-0.010	-0.58	0.18	0.023	-0.000	35.4	27.3
Shark	0.14	0.31	-0.016	-0.016	-1.56	0.21	0.024	-0.009	31.1	28.1
1024x768	0.11	0.34	-0.004	-0.017	-0.20	0.53	0.027	-0.019	29.6	25.6
1920x1088	0.21	0.28	-0.010	-0.011	-1.47	0.54	0.042	-0.014	33.2	26.6
Average	0.17	0.30	-0.007	-0.013	-0.99	0.53	0.037	-0.016	31.9	26.2

to the texture views compression, the "only synthesis" columns present the performance results of the synthesized views, and the "Time encoding reduction" column presents the complexity reduction for texture and depth map compression.

The experimental results presented in Table I indicate that the proposed approach can greatly reduce the coding time with negligible coding efficiency loss. It can save about 31.9% in coding time over all sequences in comparing with the entire 3D-HEVC encoder. Concerning the coding efficiency, it can observe that the coding efficiency loss is very negligible with 0.007dB PSNR drop and the bitrate increases about 0.17% on average for the only video. Regarding the synthesized views, the BD-rate has been reduced with 0.99% on average, whereas the average PSNR increases for all the test sequences with a value equals to 0.037 dB on average.

The proposed algorithm achieves a larger complexity reduction compared with the resulting values in [12], 26.2% on average. The former work improved the encoding efficiency with a bit rate increase of 0.30% for only video and 0.53% for the synthesized views when compared to HTM-12.1

Figures 6 and 7 show the more detailed experiment results under different QPs compared to 3D-HEVC when encoding the Kendo sequence (1024768). As present in figure 6, the proposed algorithm can efficiently reduce coding time while keeping nearly the same RD performance as 3D-HEVC, whilst figure 7 shown almost the same coding efficiency from low to high bitrate compared to 3D-HEVC.

V. CONCLUSION

This paper presents a new 3D-HEVC intra mode decision to reduce the intra coding computational complexity. The proposed algorithm exploits two important characteristics at each pixel, the eigenvalue and the direction using the structure tensor matrix. Based on these characteristics, a local edge direction histogram is established. The mode with the highest eigenvalue and its adjacent modes are selected to be evaluated in RDO process. The experimental results prove the efficiency of the proposed algorithm in terms of obtaining a reduced complexity value by 31.9% in average, with small impact in quality and bitrate.

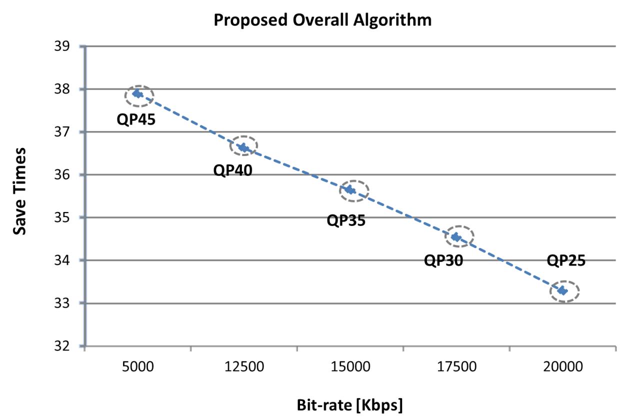

Fig. 6. Encoder time of the proposed algorithm in comparison with HTM-16.0 (Kendo 1024x768).

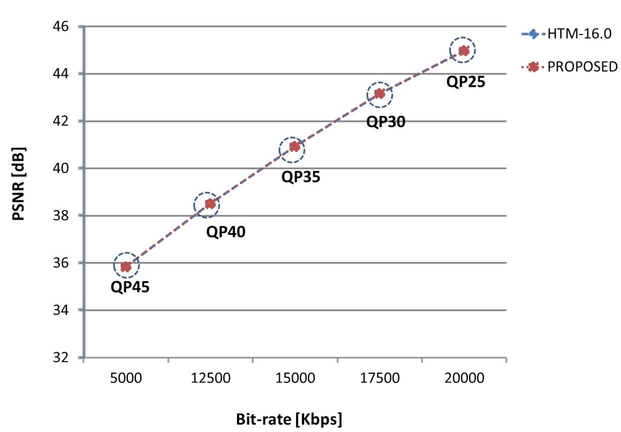

Fig. 7. RD curves of the proposed algorithm (Kendo 1024x768).

References

[1] K. Müller, H. Schwarz, D. Marpe, C. Bartnik, S. Bosse, H. Brust, T. Hinz, H. Lakshman, P. Merkle, F. H. Rhee, G. Tech, M. Winken and T. Wiegand, 3D High-Efficiency Video Coding for Multi-View Video and Depth Data, IEEE Transactions on Image Processing, vol. 22, no. 9, pp. 3366-3378, Sept. 2013.

[2] G. J. Sullivan, J. M. Boyce, Y. Chen, J. R. Ohm, C. A. Segall and A. Vetro, Standardized Extensions of High Efficiency Video Coding (HEVC), IEEE Journal of Selected Topics in Signal Processing, vol. 7, no. 6, pp. 1001-1016, Dec. 2013.

[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] M. Chen,Y. Yang,Q. Zhang, X. Zhao, X. Huangb,Y. Gan, Low complexity depth mode decision for HEVC-based 3D video coding, Optik -International Journal for Light and Electron Optics, vol. 127, no. 11, pp. 4758-4767, June. 2016.

[5] C. Fehn, Depth-image-based rendering (DIBR), compression, and transmission for a new approach on 3D-TV, in Proc. Stereoscopic Displays and Virtual Reality Systems XI, Jan. 2004, pp. 93-104.

[6] Q. Zhang, Y. Yang, H. Chang, W. Zhang, Y. Gan, Fast intra mode decision for depth coding in 3D-HEVC, Multidimensional Systems and Signal Processing, doi. 10.1007/s11045-016-0388-1, pp. 1-24, Feb. 2016.

[7] Z. Gu, J. Zheng, N. Ling and P. Zhang, Fast bi-partition mode selection for 3D HEVC depth intra coding, in Proc. IEEE International Conference on Multimedia and Expo (ICME), Jul. 2014, pp. 1-6.

[8] G. Sanchez, M. Saldanha, G. Balota, B. Zatt, M. Porto and L. Agostini, Complexity reduction for 3D-HEVC depth maps intra-frame prediction using simplified edge detector algorithm, in Proc. IEEE International Conference on Image Processing (ICIP), Jan. 2014, pp. 3209-3213.

[9] C. S. Park, Edge-Based Intramode Selection for Depth-Map Coding in 3D-HEVC, IEEE Transactions on Image Processing, vol. 24, no. 1, pp. 155-162, Jan. 2015.

[10] Zhouye Gu, Jianhua Zheng, Nam Ling and P. Zhang, Fast Depth Modeling Mode selection for 3D HEVC depth intra coding, in Proc. IEEE International Conference Multimedia and Expo Workshops (ICMEW), Oct. 2013, pp. 1-4.

[11] Z. Y. Gu, J. H. Zheng, N. Ling, and P. Zhang, Fast Intra Prediction Mode Selection for Intra Depth Map Coding, ITU-T SG16 WP3 and ISO/IEC JTC 1/SC29/WG 11, JCT3V- E0238, 5th Meeting: Vienna, Austria, Aug. 2013.

[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 Proc. IEEE International Conference on Multimedia and Expo (ICME), Jul. 2015, pp. 1-6.

[13] Y. Piao, J. Min and J. Chen, Encoder improvement of unified intra prediction, JCT-VC Document JCTVC-C207, Oct. 2010.

[14] T. Wiegand, W.-J. Han, B. Bross, J.-R. Ohm, andG.J. Sullivan, High efficiency video coding (HEVC) text specification draft 7, JCTVC-I1003, JCT-VC Meeting-Geneva. May. 2012.

[15] C. H. Fu, H. B. Zhang, W. M. Su, S. H. Tsang and Y. L. Chan, Fast wedgelet pattern decision for DMM in 3D-HEVC, in Proc. IEEE International Conference on Digital Signal Processing (DSP), Jul. 2015, pp. 477-481.

[16] Y. Chen, G. Tech, K. Wegner, S. Yea: Test Model 10 of 3D-HEVC and MV-HEVC. In ITU-T/ISO/IEC JCT3V-J1003, pp. 1-29, Oct. 2014.

[17] A. Baghaie, Z. Yu, Structure Tensor Based Image Interpolation Method, AEU - International Journal of Electronics and Communications, vol. 69, no. 2, pp. 515-522, Feb. 2015.

[18] M. Faraklioti, M, Petrou, The Use of Structure Tensors in the Analysis of Seismic Data, in Proc. Mathematical Methods and Modelling in Hydrocarbon Exploration and Production, Springer Berlin Heidelberg, 2005, pp. 47-88

[19] K. Müller, A. Vetro, Common test conditions of 3DV core experiments, in: Joint Collaborative Team on 3D Video Coding Extensions (JCT-3V) document JCT3V-G1100, 7th Meeting, San Jos, US, January, 2014.

[20] Joint Collaborative Team on 3D video coding (JCT-3V) of ITU-T SG 16 WP 3 and ISO/IEC JTC 1/SC 29/WG 11. (2015, Nov.). HTM 16.0 Reference Software [Online]. Available: https://hevc.hhi.fraunhofer.de/trac/3dhevc/browser/3DVCSoftware/tags/HTM-16.0.

[21] K. Müller, A. Vetro, Common test conditions of 3DV core experiments, JCT3V-G1100, 2014.

[22] G. Bjontegaard, Calculation of Average PSNR Differences Between RD Curves,2001: ITU-T Video Coding Experts Group (VCEG).