Bitrate Allocation for Multiview Video Plus Depth Simulcast Coding

Yasir Al-Obaidi, Tomasz Grajek, Olgierd Stankiewicz, Marek Domański
Poznań University of Technology, Poznań, Poland
{yalobaidi, tgrajek, ostank}@multimedia.edu.pl, marek.domanski@put.poznan.pl

Abstract - **For such applications as virtual navigation, free-viewpoint television or virtual reality, the cameras are sparsely distributed around a scene. For such camera arrangements, the simulcast HEVC coding is nearly as efficient as 3D-HEVC. Therefore, the problem of control of the multiple video and depth map HEVC encoders is considered in the paper. In particular, the optimum choice of quantization steps for the views and the depth maps is considered. Similarly, the paper deals with the bitrate allocation between the abovementioned views and depth maps. With the use of the results of the experiments, the respective mathematical models are derived for the choice of the quantization parameters and bitrate allocations.**

Keywords - **HEVC; Simulcast Coding; Multiview plus depth; FVC.**

I. INTRODUCTION

Multiview plus depth (MVD) [1] is widely used video representation for such applications like virtual navigation [2], free-viewpoint television [3] and virtual reality. The compression of MVD video is called 3D video compression, in particular, if the technique uses special coding tools related to depth. In that respect, 3D video coding differs for multiview (MV) video coding that is aimed at compression of several views that exploit the similarities between views, and does not deal with depth.

3D compression is an important research area and several important results have been published in the last years, e.g. [4,5]. The development of new 3D coding techniques led even to development of the international standards on 3D video coding that extend the video coding standards such as AVC [6,7] and HEVC [8,9].

In this paper, we focus on the state-of-the-art HEVC video coding technology. For the multiview plus depth video coding we could consider two variants:

- Simulcast coding of views and depth maps using HEVC individually applied the video sequences corresponding to the individual views and dynamic depth maps.
- Joint coding of all these views and depth maps using 3D HEVC technology.

The latter approach has proved to be more efficient as the total bitrate may be reduced even up to 50% as compared to simulcast [9]. Unfortunately, this well-known result holds only for rectified MVD video acquired using cameras with the parallel optical axes and densely distributed on a line. Such an application scenario is important for video feeding autostereoscopic displays. For the already mentioned applications in virtual navigation, free-viewpoint television and virtual reality systems, quite different scenarios are relevant. These applications imply that cameras are sparsely located around a scene and their optical axes are convergent with the parallax of 10-20 degrees of arc even. The experimental results, obtained by the authors for MVD video acquired with such camera arrangements, demonstrate that the gain due to application the 3D-HEVC coding technology is significantly lower and varies mostly between 0% and 15% [10,11]. Therefore, a question arises, if for such a small gain it is worth to use the special 3D-HEVC technology that needs a specific codec architecture. Therefore, the simulcast coding appears to be an interesting option for the abovementioned applications as long a more efficient 3D-HEVC extension is not developed and standardized for arbitrary camera locations. It is worth to mention that MPEG has already started exploratory activities towards an efficient HEVC extension for MVD video acquired from cameras circularly located around a scene [12]. In response to this requirement, some new techniques have been proposed [5,10,11,13]. Nevertheless, the respective standardization activity is not launched yet. Thus, the simulcast HEVC coding is still a reasonable compression solution for MVD video acquired from cameras with arbitrary locations around a scene, in particular, for a circular camera arrangement around a scene.

The goal of the paper is to study the problem of the bitrate allocation between the views and depth maps in simulcast coding of MVD video.

II. BITRATE ALLOCATION TO VIEWS AND DEPTHS

It is quite obvious that the bitrate allocation between views and depth maps influences the compression efficiency, that is measured as the quality of the synthesized virtual views versus the total bitrate for the real views and the corresponding depth maps transmitted. This problem has been already considered in several papers in the context of AVC and its multiview extensions [14-17]. In these papers, useful relations between quantization parameters, virtual view quality and bitrates have been described and even close-form formulas have been proposed for the control of the encoders.

In this paper, similarly as in [14-18], we consider the problem of the optimum bitrate allocation between views and depth maps. In contrary to the papers [14-18] we consider this problem in the context of the cameras sparsely distributed around a scene. As consequence, we consider the simulcast

HEVC compression that is nearly as efficient as 3D-HEVC [10,11] but simpler and faster. In particular, we are going to find the roles for the optimum choice of the quantization parameters for views QP and depth maps QD.

The problem of the optimum bitrate allocation was already considered in the context of 3D-HEVC used for MVD video coding for video sequences obtained from cameras being sparsely distributed along a line [19-21]. Nevertheless, it is hard to find such results for circular camera arrangements and simulcast HEVC compression.

III. METHODOLOGY

In the experiments, the relation between bitrates allocated to the views and depth maps and the quality of the virtual views is investigated. The view bitrate is calculated jointly for 3 views, and the depth bitrate is calculated for the corresponding 3 depth maps. The choice of the views and depth maps is defined according to the Common Test Conditions (CTC) used by MPEG for testing 3D video coding [22] (of Table I). The virtual views are synthesized using the extended version of the VSRS software [23] that is used by MPEG also for rendering views from the views acquired by convergent cameras.

For the experiments, the version HM16.18 of the HEVC reference software [24] is used. The encoder is configured according to the MPEG common test conditions for 3D video [22]. For the sake of simplicity it is assumed that the quantization parameter QP is constant for all views, and the quantization parameter QD is constant for all depth maps. The basic version of the software for 4:2:0 videos is used, therefore, the depth maps are encoded with all-zero chrome components. This results in negligible bitrate overhead in depth coding, but corresponds to practical straightforward approach.

The quality of the virtual views is measured as luminance PSNR with the reference to the collocated real view available from the relevant multiview test sequence. The sequences exploited in the experiments are summarized in Table I. The coding efficiency is assessed using the Bjøntegaards metric [28].

TABLE I. TEST SEQUENCES USED IN THE EXPERIMENTS

Sequence Name	Resolution	Used views	Synthesized view
Ballet [25]	1024×768	3, 5	4
Breakdancers [25]	1024×768	2, 4	3
BBB Butterfly [26]	1280×768	49, 51	50
BBB Flowers [26]	1280×768	39, 41	40
Poznan_Block2 [27]	1920x1080	2, 6	4
Poznan_Fencing [27]	1920x1080	2, 6	4

IV. OPTIMUM QP-QD SETTING

In order to find the optimum QP-QD settings, all possible QP-QD pairs were tested (QP and QD values both from 15 to 50). Figure 1 shows the quality of virtual view synthesized with the use of encoded views and depth maps with all QP-QD pairs for Breakdancers sequence. Figure 2 shows optimum QP-QD pairs for Ballet and Flower sequences. Optimum QP-QD pairs belong to the peak envelope over cloud of PSNR-bitrate points (Fig.1) that form the best R-D (rate-distortion) curve.

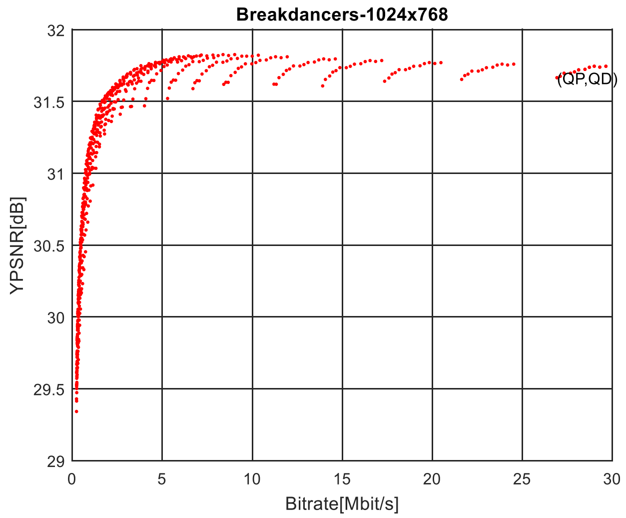

Figure 1. Quality of the synthesized view for all QP-QD pairs.

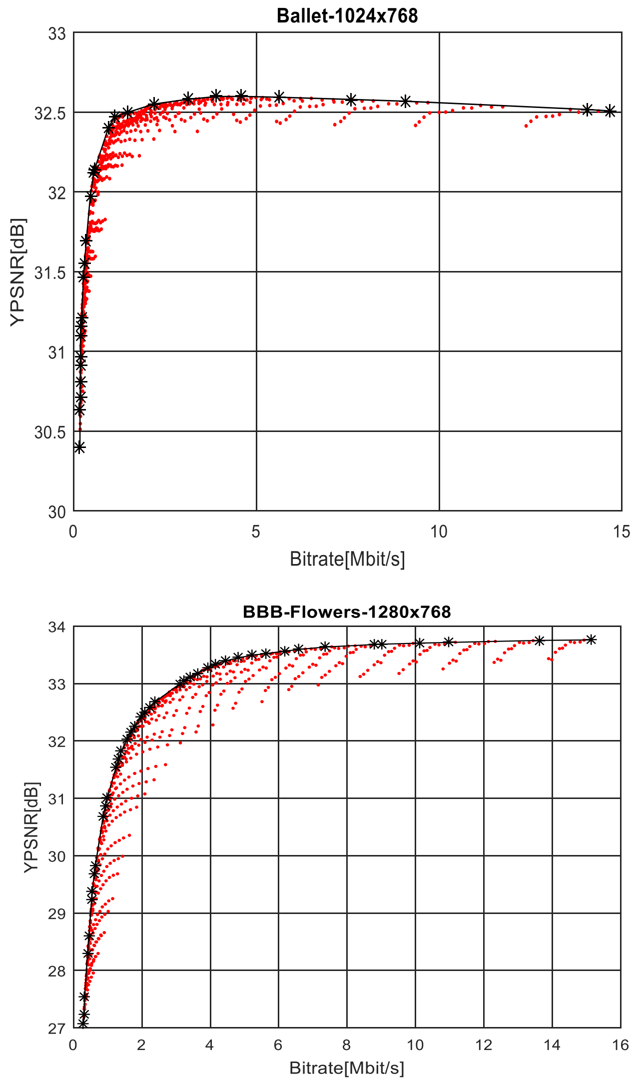

Figure 2. The best R-D curve with optimum QP-QD pairs.

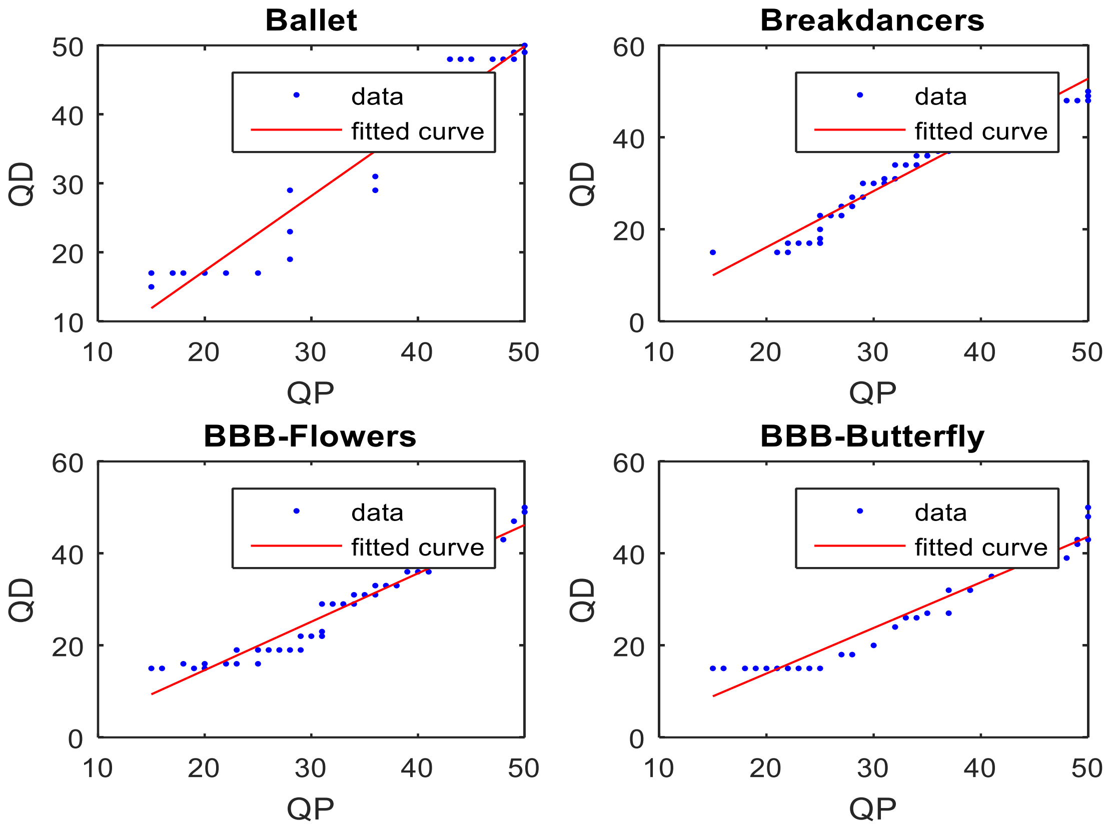

Figure 3. The approximated relationship $QD(QP)$ for the optimum pairs for exemplary sequences with the use of linear regression.

V. General Model for QD as a Function of QP

Based on the results obtained in previous section (in the form of optimum QP-QD pairs), linear regression has been applied in order to develop formula for QD value derivation based on QP settings:

$$QD(QP) = \alpha \cdot QP + \beta \quad . (1)$$

The pairs of parameters α and β have been evaluated with the usage of the least squares fitting to the optimum QP-QD pairs. The results obtained individually for four test sequences are collected in Table II. Figure 4 shows optimum pairs of QP-QD curves for four considered test sequences.

TABLE II. Parameters α and β for linear regression model approximation (equation 1) for optimum QP-QD pairs

Sequence	α	β
Ballet	1.0845	-4.3553
Breakdancers	1.2215	-8.3082
BBB.Butterfly	0.9911	-5.934
BBB.Flowers	1.0523	-6.4204
Average	1.0874	-6.2545

In order to verify developed formula, we have compared quality of the synthesized virtual views achieved using the proposed method and with the QP=QD approach for two different test sequences: Poznan_Block2 (see fig. 5) and Poznan Fencing.

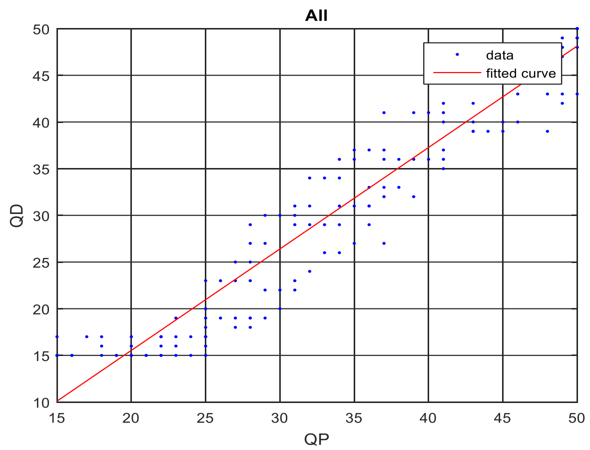

Figure 4. Optimum pairs of QP-QD curves for four considered sequences.

Table III gathers bitrate reductions calculated by Bjøntegaards rates between the proposed algorithm and:

- QP=QD approach,
- Optimum QP-QD pairs

Moreover, Figure 6 shows comparison between the proposed method and optimum QP-QD pairs for Poznan_Block2 sequence.

TABLE III. BITRATE REDUCTIONS CALCULATED BY BJØNTEGAARDS RATES BETWEEN THE PROPOSED ALGORITHMIC (QP-QD) AGAINST CODING WITH REFERENCE (QP=QD) AND OPTIMUM QP-QD PAIRS.

Sequence	Bjøntegaards rates	Proposed vs Reference	Proposed vs Optimum
Poznan_Block2	$\Delta PSNR$	-8.71%	15.82%
	$\Delta Bitrate$	0.18%	-0.27%
Poznan_Fencing	$\Delta PSNR$	-2.58%	8.74%
	$\Delta Bitrate$	0.17%	-0.41%

The performed experiments showed that proposed formula led to decrease of total bitrate and improve of the virtual view quality for sequences compared to the reference (QD=QP) while the comparison between proposed formula and optimum QD-QP pairs for sequences led to increase of total bitrate and decrease of virtual view quality.

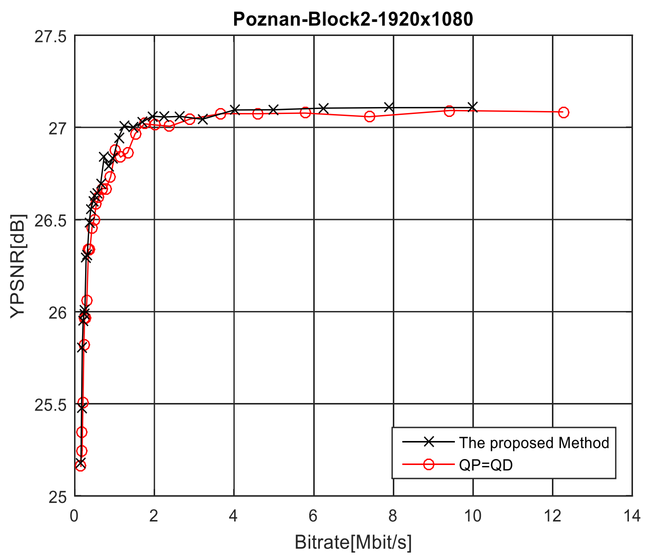

Figure 5. Comparison between the proposed method and QP=QD approach for Poznan-Block2 sequence.

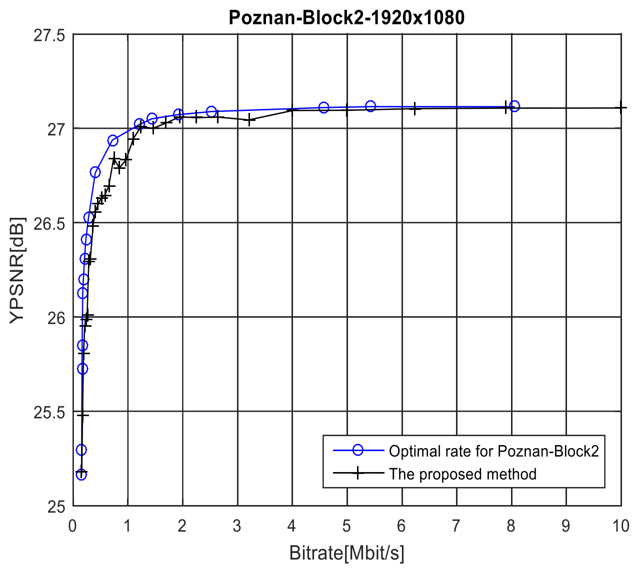

Figure 6. Compare between proposed method and optimum QP-QD pairs for Poznan-Block2 sequence.

VI. GENERAL MODEL FOR VIEW BITRATE

The model is derived for the bitrate analysis between videos and depth maps as a function of the quantization parameter (QP) applied for the videos. In the previous section we proposed the formula how to choose appropriate QD value for the given QP value in order to maximize the PSNR of the virtual view. Right now we would like to check what part of the bitstream represents the views and what depth maps. Based on the data gathered for the optimum QP-QD pairs we plotted (View_bitrate/Total bitrate)-QP points (see fig. 7) and applied the polynomial regression in order to find formula describing mentioned relationship:

$$\frac{View_{Bitrate}}{Total_{Bitrate}} = f(QP) = \gamma \cdot QP^2 + \delta \cdot QP + \theta \quad . (2)$$

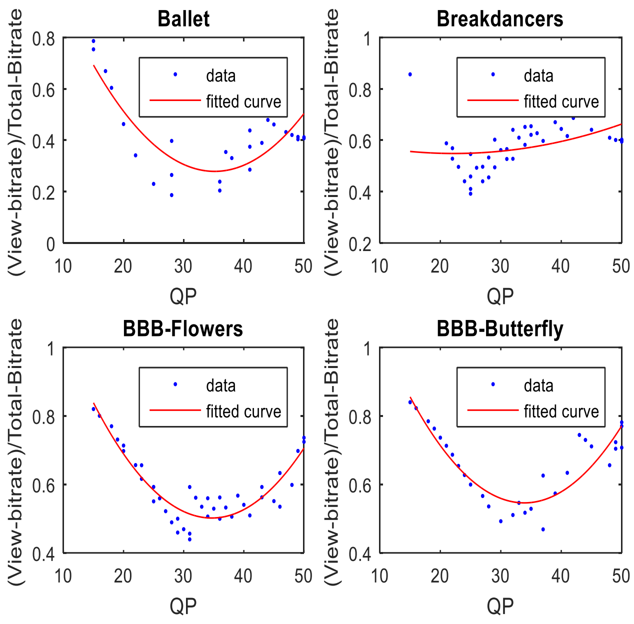

Figure 7. The approximate relationship $\frac{View_{Bitrate}}{Total_{Bitrate}} = f(QP)$ for the optimum pair in some sequences with the use of polynomial regression.

Parameters γ, δ and θ have been evaluated with the usage of least squares fitting to the optimum (View_bitrate/Total bitrate)-QP pairs produced by the proposed algorithm. The obtained results are collected in Table IV. Figure 8 shows optimum pairs of (View_bitrate/Total bitrate)-QP curves for four considered sequences.

TABLE IV. PARAMETERS γ, δ AND θ FOR POLYNOMIAL REGRESSION MODEL APPROXIMATION (EQUATION (2)) OF OPTIMUM VIEW_BITRATE-QP CURVE ALGORITHM.

Sequence	γ	δ	θ
Ballet	0.001	-0.0717	1.5381
Breakdancers	0.0001	-0.0066	0.6223
BBB.Butterfly	0.0009	-0.0587	1.5415
BBB.Flowers	0.0009	-0.0602	1.5462
Average	0.0007	-0.0493	1.3120

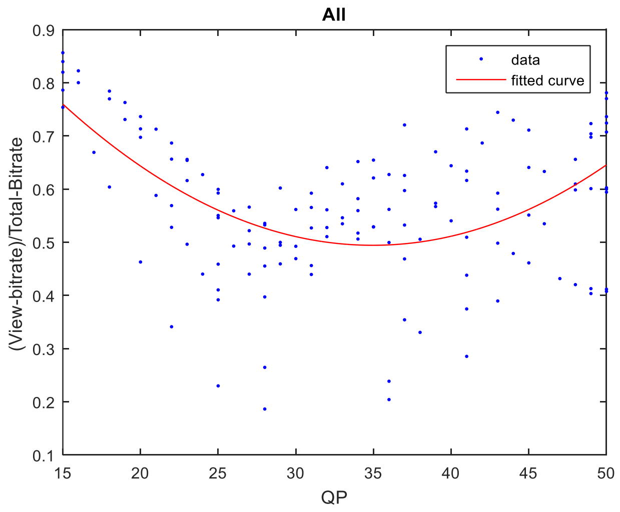

Figure 8. Optimum pairs of (View_bitrate/Total bitrate)-QP curves for four considered Sequences.

VII. CONCLUSIONS

In the paper a closed–form formula is derived for QD as a function of QP for the simulcast HEVC coding of multiview plus depth (MVD) video acquired from cameras located around a scene. The proposed choice of QD maximizes the PSNR of the virtual views. The complementary model is derived for the bitrate breakdown between views and depth maps as a function of the quantization parameter QP used for the views.

It if shown that the application of the proposed formulas results in 2-10% of bitrate reduction as compared the straightforward setting QP=QD.

ACKNOWLEDGMENT

The presented work has been funded by the Polish Ministry of Science and Higher Education within the status activity task "Theory and algorithms of multidimensional signal processing" in 2018.

REFERENCES

[1] K. Muller, P. Merkle and T. Wiegand, "3-D video representation using depth maps," Proc. IEEE, vol. 99, pp. 643-656, April 2011.

[2] O. Stankiewicz, M. Domanski, A. Dziembowski, A. Grzelka, D. Mieloch, J. Samelak, " A free-viewpoint television system for horizontal virtual navigation", IEEE Trans. Multimedia, 2018, early access.

[3] M. Tanimoto, M. Panahpour, T. Fujii, T. Yendo, "FTV for 3 D spatial communication," Proc. IEEE, vol. 100, pp. 905-917, Feb. 2012.

[4] K. Muller, H. Schwarz, D. Marpe, C. Bartnik, S. Bosse, H. Brust, T. Hinz, H. Lakshman, P. Merkle, F. Rhee, G. Tech, M. Winken, T. Wiegand, "3D high-efficiency video coding for multi-view video and depth data", IEEE Trans. Image Proc.,vol. 22, pp. 3366-3378, Sep. 2013.

[5] M. Domański, O. Stankiewicz, K. Wegner, M. Kurc, J. Konieczny, J. Siast, J. Stankowski, R. Ratajczak, T. Grajek," High efficiency 3D video coding using new tools based on view synthesis", IEEE Trans. Image Proc., vol. 22, pp.3517-3527, Sep. 2013.

[6] "Coding of audio-visual objects, Part 10: Advanced Video Coding," ISO/IEC IS 14496-10, 2014.

[7] M. Hannuksela, D. Rusanovskyy, W. Su, L. Chen, R. Li, P. Aflaki, D. Lan, M. Joachimiak, H. Li, M. Gabbouj, "Multiview-video-plus-depth coding based on the advanced video coding standard", IEEE Trans. Image Proc., vol. 22, pp. 3449 - 3458, Sep. 2013.

[8] "High efficiency coding and media delivery in heterogeneous environments - Part 2: High efficiency video coding," ISO/IEC IS 23008-2, ITU-T Rec. H.265, 2015.

[9] G. Tech, Y. Chen, K. Müller, J. Ohm, A. Vetro, Y.Wang," Overview of the multiview and 3D extensions of high efficiency video coding", IEEE Trans. Circuits Syst. Video Techn., vol. 26, pp. 35-49, January 2016.

[10] J. Samelak, J. Stankowski, M. Domański, "Adaptation of the 3D-HEVC coding tools to arbitrary locations of cameras", Int. Conf. Signals and Electronic Systems, ICSES 2016, Kraków, Poland, Sep. 2016.

[11] M. Domański, A. Dziembowski, A. Grzelka, Ł. Kowalski, D. Mieloch, J. Samelak, O. Stankiewicz, J. Stankowski, K. Wegner, "[FTV AHG] Extended results of Poznan University of Technology proposal for call for evidence on free-viewpoint television", ISO/IEC JTC1/SC29/WG11 MPEG2016, M38246, Geneva, Switzerland, 30 May - 03 June 2016.

[12] G. Lafruit, K. Wegner, M. Tanimoto, "Call for evidence on free-viewpoint television: Super-multiview and free navigation", ISO/IEC JTC1/SC29/WG11, MPEG2015/N15348, June 2015, Warsaw.

[13] L. Yu, Q. Wang, A. Lu, Y. Sun, "Response to call for evidence on free-viewpoint television: Zhejiang university", ISO/IEC JTC1/SC29/WG11, MPEG2016/m37608, San Diego, US, February 2016.

[14] F. Shao, G. Jiang, M. Yu, K. Chen, Y. Ho, "Asymmetric coding of multi-view video plus depth based 3-D video for view rendering", IEEE Trans. Multimedia, vol. 14, pp. 157 - 167, Feb. 2012.

[15] F. Shao, G. Jiang, W. Lin, M. Yu, Qi. Dai, "Joint bit allocation and rate control for coding multi-view video plus depth based 3D video", IEEE Trans. Multimedia, vol. 15, pp. 1843 - 1854, December 2013.

[16] K. Klimaszewski, O. Stankiewicz, K. Wegner, M. Domański, "Quantization optimization in multiview plus depth video coding", IEEE Int. Conf. Image Proc. (ICIP), Paris, France, October 2014.

[17] J. Stankowski, D. Karwowski, T. Grajek, K. Wegner, J. Siast, K. Klimaszewski, O. Stankiewicz, M. Domański, "Bitrate distribution of syntax elements in the HEVC encoded video", Int. Conf. Signals and Electronic Systems (ICSES), Poznań, Poland, September 2014.

[18] E. Bosc, F. Racape, V.Jantet, P. Riou, M. Pressigout, L. Morin," A study of depth/texture bit-rate allocation in multi-view video plus depth compression", annals of telecommunications, Vol. 68, pp. 615–625, 2013.

[19] S. Tan, J. Si, S. Ma, S. Wang, W. Gao, "Adaptive frame level rate control in 3D-HEVC", IEEE Visual Communications and Image Proc. Conf., pp. 382 - 385, 2014.

[20] Y. Song, K. Jia, Z. Wei, "Improved LCU level rate control for 3D-HEVC", Visual Communications and Image Proc. (VCIP), pp. 1-4, 2016.

[21] M. Zhong, X. Lan, K. Wang, "Saliency-based joint distortion model for 3D video coding", Chinese Automation Congress, pp. 720-725, 2017.

[22] K. Müller, A. Vetro, "Common test conditions of 3DV core experiments," ISO/IEC JTC1 SC29/WG11 and ITU-T SG 16 WP 3, Doc. JCT3V G1100, San José, Jan. 2014.

[23] O. Stankiewicz, K. Wegner, M. Tanimoto, M. Domański, "Enhanced View Synthesis Reference Software (VSRS) for Free-viewpoint Television", ISO/IEC JTC 1/SC 29/WG11, Doc. M31520, Geneva, 2013.

[24] 2D HEVC reference codec available online https://hevc.hhi.fraunhofer.de/svn/svn_HEVCSoftware/tags/HM-16.18.

[25] L. Zitnick, S.B. Kang, M. Uyttendaele, S. Winder, R. Szeliski, "High-quality video view interpolation using a layered representation," ACM SIGGRAPH, pp. 600-608, 2004.

[26] P. Kovacs, "[FTV AHG] Big Buck Bunny light-field test sequences". ISO/IEC JTC1/SC29/WG11, Doc. MPEG M35721, Geneva, 2015.

[27] M. Domański, A. Dziembowski, A. Grzelka, D. Mieloch, O. Stankiewicz, K. Wegner, "Multiview test video sequences for free navigation exploration obtained using pairs of cameras", ISO/IEC JTC1/SC29/WG11, Doc. MPEG M38247, Geneva, 2016.

[28] G. Bjontegaard, "Calculation of Average PSNR Differences between RD-curves," ITU-T SG16,Doc. VCEG-M33, USA, 2001.