Influence of Depth Map Fidelity on Virtual View Quality

Yasir Al-Obaidi
Chair of Multimedia Telecommunications and Microelectronics
Poznan University of Technology
Poznań, Poland
yalobaidi@multimedia.edu.pl

Tomasz Grajek
Chair of Multimedia Telecommunications and Microelectronics
Poznan University of Technology
Poznań, Poland
tgrajek@multimedia.edu.pl

***Abstract*—The paper deals with optimal bitrate distribution between video and depth maps in multiview compression. Influence of noise in depth data on the optimal bitrate allocation is studied. For simplicity simulcast approach is used. The results show that for high bitrates, virtual view quality depends on the depth map fidelity. However for low bitrates, this influence is negligibly small.**

Keywords*—High Efficiency Video Coding (HEVC), view synthesis, depth map, multiview video plus depth (MVD), free viewpoint video (FTV)

I. Introduction

Free Viewpoint Television (FTV) [1] and 3-dimensional Television (3DTV) [2-3] are the most important applications of multiview video [4]. These systems employ many cameras to capture a scene. Typically, cameras in such systems are located around a scene on an arc or circle. Cameras in a system produce synchronized video streams that have to be converted into a representation that allows an user to freely select or alter the viewpoint he or she wants to watch. The most commonly used representation for the mentioned applications is multiview video and depth (MVD) format [5]. The multiview systems offer viewers very realistic depth impression through 3D content, that contains much more presented views of the scene than in fact have been acquired. These additional views are called virtual views. Virtual views are created by means of view synthesis from the video and depth maps acquired by the camera system.

In all those applications the produced representation needs to be transmitted to the viewer. Currently there are many works focused on efficient multiview and depth compression. Several techniques included in international standards have been developed in recent years, namely MVC [6], MV-HEVC and 3D-HEVC [7] to tackle the problem. Many more have been proposed within literature, e.g. [8-10].

One of the particular important problems is the bitrate allocation between texture and depth data. The problem have been studied quite extensively already [11-17] but there are still many questions opened.

In our previous paper [18] we have analyzed optimal bitrate allocation between video and depth data in case of HEVC simulcast compression for MVD video acquired from cameras with arbitrary locations around a scene. We have shown how to divide available bitrate between those two components to assure the best quality of the virtual views presented to the viewer in FTV or 3DTV systems.

During this study, we have observed that not all MVD video material behaved the same way. Some of test sequences exhibits strange, counterintuitive behavior. Namely, reducing the bit allocation for the depth component (effectively lowering total bit allocation) results in the increasing quality of the virtual views in some bitrate range. This strange behavior have been observed only for some, but not all test sequences.

Our assumption is that the quality of the depth data and thus its fidelity for view synthesis is not equal in all test data. Thus we will try to investigate the influence of the noise, or in general lowered quality of the depth data on the optimal bitrate distribution between texture and depth data.

The goal of the research is to investigate the influence of fidelity of depth maps on the quality of virtual view synthesis.

II. Methodology of Experiments

In order to assess how fidelity of depth maps influences optimal bitrate allocation between depth and video data under the highest possible quality of virtual views assumption, experiments have been conducted according to block diagram presented in Fig. 1. At first, two views (*i+1 -th* and *i-1 -th*) with two associated depth maps have been independently encoded and decoded using HEVC [19]. Then, decoded views together with associated depth maps have been used to create a virtual view at position of *i-th* camera. This synthesized view is compared with the view acquired by the real camera exactly at the same position in 3D space as virtual one. Finally, the PSNR of virtual view created and total bitrate of all the data necessary for creating it, have been gathered together.

The experiments have been conducted on four test multiview video sequences with depth maps [20-21]. Summary of the test multiview sequences are given in Table 1. Camera positions for experiments have been chosen according to the Common Test Conditions (CTC) used by MPEG for testing 3D video compression [22]. To compress views and depth maps the reference test model of HEVC, namely HM v.16.18 [23] has been used. Since basic version of the software is prepared mainly for the video compression in 4:2:0 chroma format, depth maps have been encoded with all-zero chrome components. This results in negligible bitrate overhead for depth map coding, but corresponds to practical straightforward approach. For view synthesis the reference model software VSRS v.3.5 [24] has been used.

The presented work has been funded by the Polish Ministry of Science and Higher Education for the status activity consisting of research and development and associated tasks supporting development of young scientists and doctoral students in 2018 in Chair of Multimedia Telecommunications and Microelectronics.

DOI: 10.1109/ICSES.2018.8507272

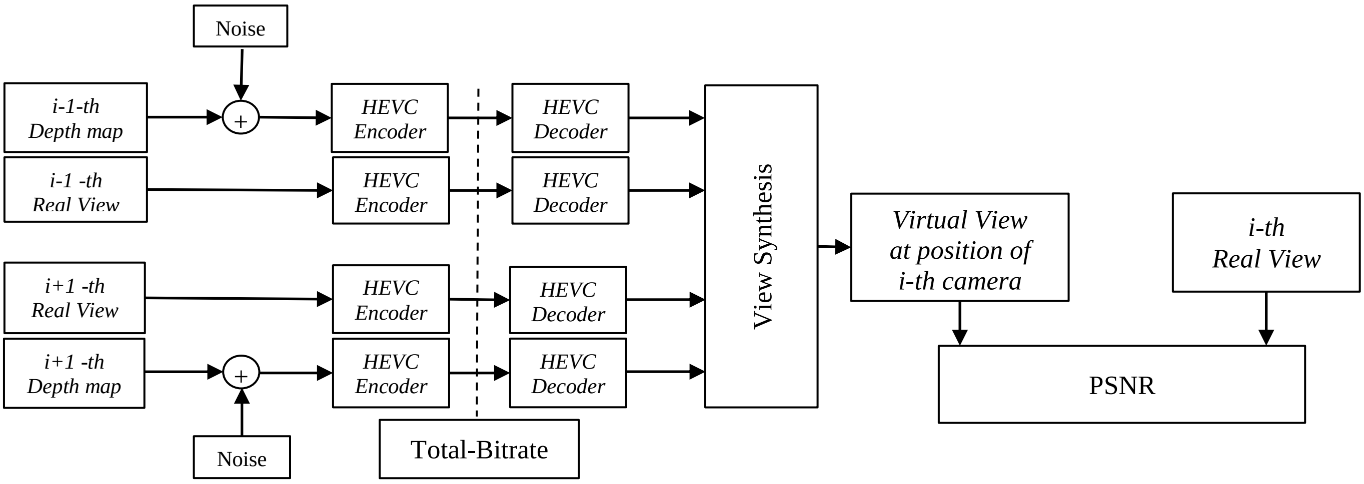

Fig. 1. Structure of the performed experiments.

TABLE I. TEST SEQUENCES USED IN EXPERIMENTS

Sequence Name	***Resolution***	***Used views***	***Synthesized view***
Ballet [20]	1024×768	3, 5	4
Breakdancers [20]	1024×768	2, 4	3
BBB Butterfly [21]	1280×768	49, 51	50
BBB Flowers [21]	1280×768	39, 41	40

The simulcast compression of MVD data is controlled by four quantization parameters: two for video (*QP*) and two for depth (*QD*). Commonly two quantization parameters for video (*QP*) are set equal and two quantization parameters for depth (*QD*) also are set equal, resulting in only two free parameters one for both videos and one for both depth maps.

In [18] we propose a method for appropriate choosing *QP* and *QP* values (so called optimum *QP-QD* pairs) in order to reach highest possible virtual view quality.

Having optimum *QP-QD* pairs for unaltered (original) depth maps from [18], we have repeated experiments, this time with noise added depth maps. We have try several noise amplitudes. Adding noise *N* (see Fig. 1) is understood as adding random uniformly distributed values between (-*N* and +*N*) to each of the values stored in both depth maps. We have considered only that *N* is integer number (such as 1,2,3, etc.). For each *N* value, optimal *QP-QD* pairs have been extracted from data gathered during experiments.

III. EXPERIMENTAL RESULTS

As in [18], to find the optimum *QP-QD* settings, all *QP-QD* pairs were tested (*QP* and *QD* values both from 15 to 50). In Fig. 2, red points represent results of compression with all possible combinations of quantization parameters for views and depth maps for four considered sequences, while black lines represent optimum *QP-QD* pairs for each test sequences. The optimum *QP-QD* pairs belong to envelope over cloud of PSNR-bitrate points that form the best R-D (rate-distortion) curve. For natural test sequences (*Ballet* and *Breakdancers*), when decreasing bitrate (applying stronger compression, i.e. selecting higher quantization parameters values) quality of virtual (synthesized) view is first increasing and after passing some maximum point, starts decreasing. This "increasing" part of the curve is very surprising and interesting, because as a result of the reduced number of bits needed to represent MVD sequence we get better quality of virtual view (of course up to some point). However, this observation does not hold for artificial (computer generated) sequences, in our case *BBB Butterfly* and *BBB Flowers*.

Fig. 3 shows the relationship between the quality of synthesized view and total bitrate for exemplary test sequences with different quality of depth maps. The lines represent the optimum R-D curves. The blue lines represent quality of virtual view produced from video and depth maps without any noise. The magenta, green, and red lines represent quality of virtual view synthesized from video and depth maps with various noise (*N*) added. Black squares indicate *QP-QD* pair with the highest quality of virtual view for given depth map quality. Despite the amount of noise added, the best quality of virtual view is achieved for the same *QP* value (for our test sequences *QP=22*). However, the more noise we add the greater the *QD* value should be.

Fig. 4 presents relations between fidelity of depth maps and quality of the synthesized views for different amount of noise added. Fidelity of depth maps (*Depth PSNR*) is calculated as average depth maps quality associated with i*+1* –th and *i-1 -th* views.

From Fig. 3 and Fig. 4 we clearly see the influence of added noise on the quality of virtual view, but only for high bitrates, when compression is not very strong. If we apply stronger compression (lower bitrates) the influence of added noise becomes negligible. The most straightforward explanation is: from some point, quantization error (quantization noise) introduced by compression becomes stronger than the noise added to depth maps before compression. If we recall, generally quantization removes high frequency components from images, so it will remove noise (noise in fact is high frequency signal). Adding stronger noise to the depth maps requires greater *QD* value (stronger quantization) to remove this added noise during compression of MVD sequences.

Abovementioned observation is also clearly visible in Fig. 5, where optimum *QP-QD* pairs have been plotted. Again for values of quantization parameters (*QP*) below 25÷30, the stronger noise we add, the higher *QD* should be set to get the highest possible quality of virtual views.

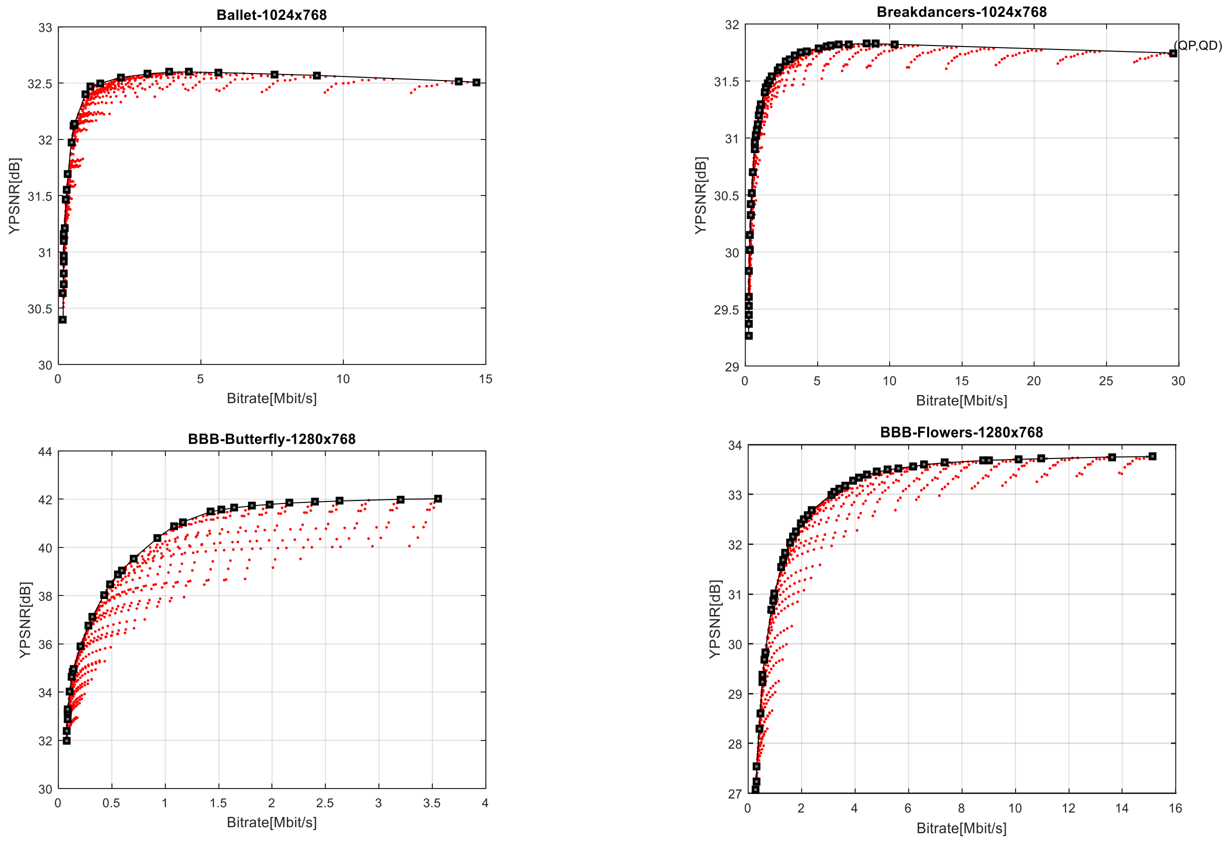

Fig. 2. Optimum R-D curves for unmodified depth maps with optimum *QP-QD* pairs.

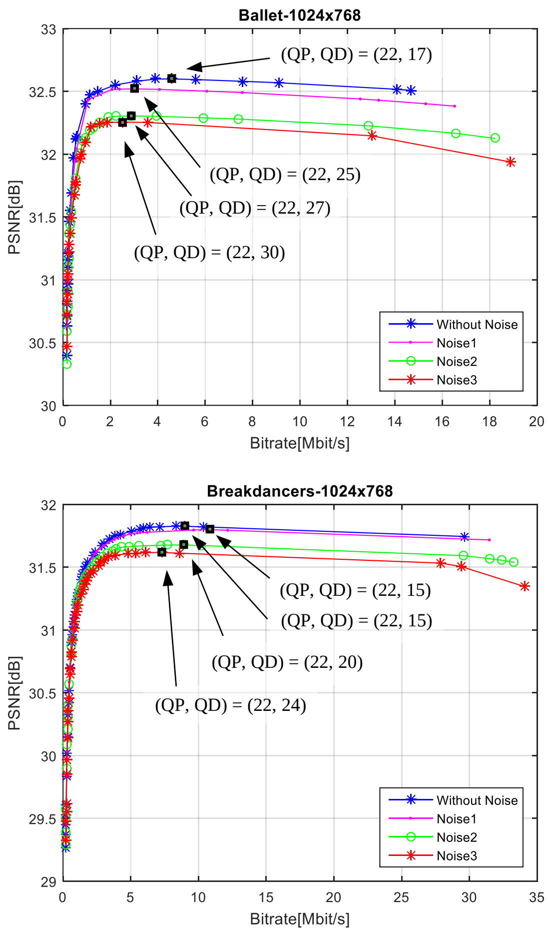

Fig. 3. Optimum R-D curves for different quality of depth maps.

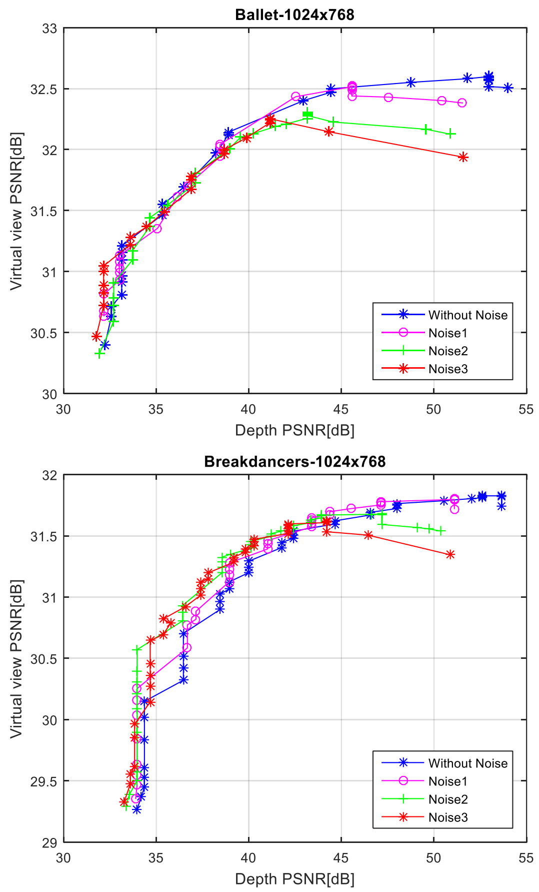

Fig. 4. Influence of fidelity of depth maps on virtual view quality. *Depth PSNR* is calculated as average depth map quality associated with *i+1* and *i-1* views.

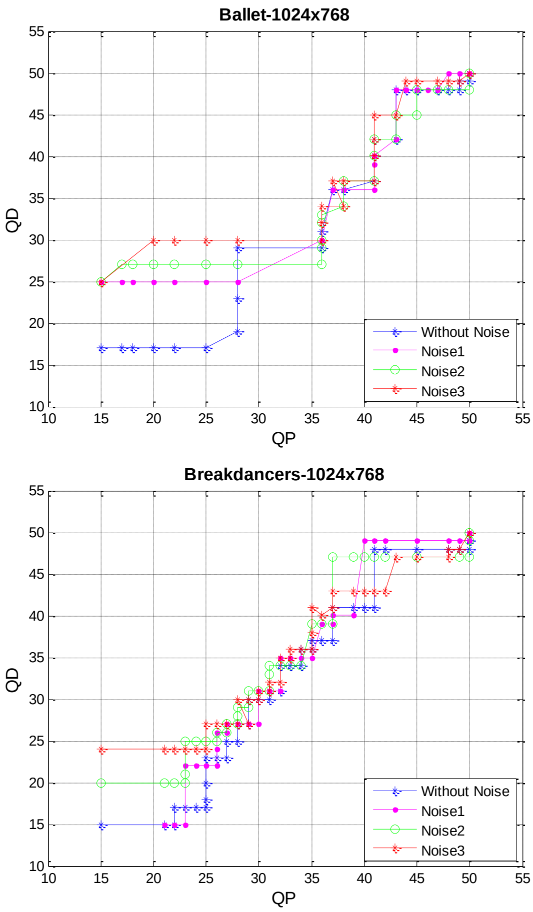

Fig. 5. Optimum *QP-QD* pairs for exemplary test sequences for different fidelity of depth maps.

IV. CONCLUSIONS

In this paper we studied the performance of the optimum *QP-QD* pairs and showed the influence of the fidelity of depth maps on the quality of virtual views.

As can be expected, increasing the amount of noise added to depth maps leads to decrease of virtual views quality. However this observation holds only to some point, i.e. to the point in which quantization error (noise) introduced by compression is lower than the added noise. For stronger compression, when quantization noise is greater than added one, difference between quality of virtual views achieved for various amount of added noise is negligible.

For rather high bitrates, the stronger noise we add to depth maps, the higher *QD* should be applied to get the highest possible quality of virtual views.

REFERENCES

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

[2] A. Kubota, A. Smolic, M. Magnor, M. Tanimoto, T. Chen, C. Zhang, "Multiview imaging and 3DTV," IEEE Signal Proc. Magazine, vol. 24, pp. 10-21, 2007.

[3] H. Ozaktas, L. Onural, Three-dimensional television: capture, transmission, and display, Springer, Heidelberg, 2007.

[4] G. Lafruit, M. Domański, K. Wegner, T. Grajek, T. Senoh, J. Jung, P. Kovács, P. Goorts, L. Jorissen, A. Munteanu, B. Ceulemans, P. Carballeira, S. García, M. Tanimoto, "New visual coding exploration in MPEG: Super-MultiView and Free Navigation in Free viewpoint TV," IST Electronic Imaging, Stereoscopic Displays and Applications XXVII, San Francisco, pp. 1-9, 2016.

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

[6] A. Vetro, T. Wiegand, and G. J. Sullivan, "Overview of the stereo and multiview video coding extensions of the H.264/AVC standard," Proc. IEEE, vol. 99, no. 4, pp. 626–642, Apr. 2011.

[7] 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 Sys. Video Techn., vol. 26, pp. 35-49, Jan. 2016.

[8] 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.

[9] J. Stankowski, Ł. Kowalski, J. Samelak, M. Domański, T. Grajek, K. Wegner, "3D-HEVC extension for circular camera arrangements," 3DTV-Conference: The True Vision - Capture, Transmission and Display of 3D Video (3DTV-CON), Lisbon, pp. 1-4, 2015.

[10] 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.

[11] A. Banitalebi-Dehkordi, M. Pourazad, P. Nasiopoulos, "A study on the relationship between depth map quality and the overall 3D video quality of experience," 3DTV-Conference: The True Vision-Capture, Transmission and Display of 3D Video (3DTV-CON), Aberdeen, Scotland, pp. 1-4, 2013.

[12] C. Hewage, S. Worrall, S. Dogan, S. Villette, A. Kondoz, "Quality evaluation of color plus depth map-based stereoscopic video," IEEE Journal of Selected Topics in Signal Proc., vol.3, no.2, pp.304-318, Apr. 2009.

[13] G. Nur, S. Dogan, H.K. Arachchi, A.M. Kondoz, "Impact of depth map spatial resolution on 3D video quality and depth perception," 3DTVConference: The True Vision - Capture, Transmission and Display of 3D Video (3DTV-CON), Tampere, Finland, pp. 1-4, 2010.

[14] C. Hewage, S. Worrall, S. Dogan, A. Kondoz, "Prediction of stereoscopic video quality using objective quality models of 2-D video," Electronics Letters, vol.8, pp.963-965, 2008.

[15] Q. Zhang, P. An, Y. Zhang, Z. Zhang, "Efficient rendering distortion estimation for depth map compression," IEEE Int. Conf. Image Proc., Brussels, Belgium, pp.1105-1108, 2011.

[16] K. Klimaszewski, K. Wegner, M. Domanski, "Distortions of synthesized views caused by compression of views and depth maps," 3DTV-Conf.: The True Vision Capture, Transmission and Display of 3D Video, Potsdam, Germany, 2009.

[17] P. Merkle, Y. Morvan, A. Smolic, D. Farin, K. Mueller, P. H. N de With, and T. Wiegand, "The effects of multiview depth video compression on multiview rendering," Singal Proc.: Image Communication, vol. 24, pp. 73–88, January 2009.

[18] Y.Al-Obaidi, T. Grajek, O. Stankiewicz, M. Domanski, "Bitrate allocation for multiview video plus depth simulcast coding," Int. Conf. on Systems, Signals and Image Proc. (IWSSIP), Maribor, Slovenia, Jun. 2018.

[19] G. Sullivan, J. Ohm, W. Han, T. Wiegand, "Overview of the high efficiency video coding (HEVC) Standard," IEEE Trans. Circuit and System for Video Techn., vol.22, pp. 1649-1668, 2012

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

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

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

[24] 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, Switzerland, 2013.