DISTORTIONS OF SYNTHESIZED VIEWS CAUSED BY COMPRESSION OF VIEWS AND DEPTH MAPS

Krzysztof Klimaszewski, Krzysztof Wegner, Marek Domański

Chair of Multimedia Telecommunications and Microelectronics,
Poznań University of Technology, Poland

ABSTRACT

The paper deals with prospective 3D video transmission systems that would use compression of both multiview video and depth maps. The paper addresses the problem of quality of views synthesized from other views transmitted together with depth information. For the state-of-the-art depth map estimation and view synthesize techniques, the paper proves that AVC/SVC-based Multiview Video Coding technique can be used for compression of both view pictures and depth maps. The paper reports extensive experiments where synthesized video quality has been estimated by use of both PSNR index and subjective assessment. Defined is the critical value of depth quantization parameter as a function of the reference view quantization parameter. For smaller depth map quantization parameters, depth map compression has negligible influence on fidelity of synthesized views.

Index Terms— multiview, depth map compression, free viewpoint television, 3D television, MVC.

1. INTRODUCTION

Recently, a lot of attention is paid to multiview video systems [10]. Prospective 3D video applications include free-viewpoint television as well as various stereoscopic displays [1]. These facts stimulate research on compression of 3D video. In particular, Multiview Video Coding (MVC) technology has been developed on the basis of the well-known AVC/H.264 video compression standard together with its scalable extension (SVC) [11,12]. Interesting are systems that transmit very limited set of views but together with respective depth maps [10,13]. In such systems, videos corresponding to other viewpoints are synthesized at decoders. Both video sequences as well as depth maps have to be compressed prior to transmission.

Virtual view synthesis requires reference views from real cameras and the corresponding depth maps. In this paper we assume that depth maps are created at the transmitter side, and transmitted to the receiver (see Fig. 1).

This approach requires efficient depth map compression algorithms. There are attempts to design special depth map compression algorithms, like in [3]. Other approaches use standard compression techniques [2,10,15]. In [2], a special depth compression technique is compared to AVC/H.264 intraframe coding, and their influence on synthesized view quality is considered, similarly as in this paper.

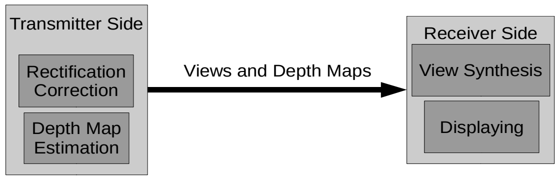

Figure 1. Schematic of an envisaged multiview system.

Here, considered is compression of views and disparity maps using standard Multiview Video Coding (MVC) with all its modes including interframe coding [11,12]. The assumption is that the state-of-the-art depth map estimation and view synthesize techniques are used as described in [3-5]. Obtained original depth maps are not ideal and have some errors. Depth map compression adds additional distortions that have obvious influence on synthesized view quality. Here, we study subjective and objective quality of synthesized views derived from compressed reference views and compressed depth maps. Estimated are the values of quantization steps and bitrates needed for views and depth maps in order to obtain good quality of synthesized views.

2. DEPTH MAP COMPRESSION

Features of depth maps are quite different from those for regular views (see Table 1). Nevertheless the compression results obtained with a special technique from [3] are not much better than those obtained using simple standard techniques[2]. Therefore, we have tested standard MVC codec [6] used to compress the depth maps. The main advantage of this approach is that the technology is readily available and standard reference software can be used.

Table 1. Views and depth maps features.

Depth Maps	Views
only one color component	three color components
limited texture	abundance of texture
prominent contours	less prominent contours
accurate depth value important	accurate value not necessary

3. TEST SETUP

In order to evaluate compression influence on view synthesis performance, we performed many experiments. Two views together with the corresponding disparity maps were compressed and decompressed. The disparity maps represent depths and therefore are called as such. A synthesized viewpoint video is compared to the original video registered from this viewpoint (Figure 2).

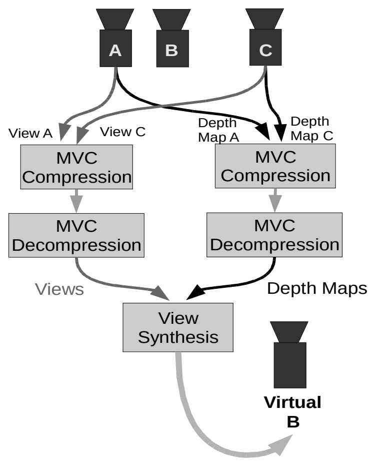

Figure 2. Data flow in experiments described in the paper.

For the experiments, depth maps were created using depth estimation reference software (DERS) [4], kindly provided by Nagoya University and improved [5] in order to provide good temporal consistency. In the compression experiments we used reference JMVC software [6]. We used standard prediction structure. For two views, as in our case, this leads to limited inter-view prediction, only for anchor frames.

Quantization parameter indices QP and QD covered wide range of values. The reconstructed depth maps and views were used as input to view synthesis that was performed with VSRS reference software [4] version 2.0. Synthesized views were compared to the original views from the actual cameras, both in terms of PSNR measure and subjective quality assessment. We present results for three different multiview test sequences: "Book Arrival" [7], "Newspaper" [8] and "Pantomime"[9]. For other sequences similar results have been obtained.

5. RESULTS

The results are presented as functions of total bitrate (for 2 views and 2 depth maps) as well as functions of bitrate ratio (Figs. 3 and 4)

$$bitrate\,ratio = \frac{depth\,map\,bitrate}{view\,bitrate + depth\,map\,bitrate} \quad . \qquad (1)$$

The value of quantization parameter index QP in view compression is given next to each line on each figure. Moving along each line corresponds to changing depth quantization parameter index QD. The dashed horizontal line marks PSNR obtained by using uncompressed depth maps and uncompressed reference views.

Depth quantization parameter index QD has negligible (less than 0.2 dB decrease) impact on synthesized view quality as long its value is smaller than the following threshold estimated from the experiments

$$Threshold\ QD = 49 - 0.005(76 - QP)^2 \quad . \qquad (2)$$

The points corresponding to the above mentioned threshold are denoted by the dashed curves in Fig. 4. From this figure, we conclude that for high- and medium-quality depth map compression, synthesized view quality is independent from depth map bitrate. For higher reference view quality ($QP \leq 21$), the quality of synthesized views is approximately the same as that for uncompressed views, PSNR differences being less than 0.2 dB and sometimes even outperform results for uncompressed data.

The above mentioned results have been verified by subjective quality assessment for synthesized views. MOS is expressed in a continuous scale (0 means "very bad with annoying artifacts", 5 – "imperceptible artifacts"). Three sets of randomly ordered synthesized sequences have been presented to 15 viewers (Fig. 5). The original video from actual real camera was used as hidden reference. In general, the results confirm conclusions drawn from PSNR results.

6. CONCLUSIONS

Our most important conclusion is that the quality of views is much more important than the quality of the depth maps for good view synthesis. Our experiments also show that MVC codec can be successfully used as a depth map compression tool. Although not optimized for this purpose it is good and readily available solution.

From Figure 4 it can be observed that the threshold above which increase of depth bitrate does not increase the quality of synthesized view can be approximately calculated from Formula (2). For useful quality of views, this point corresponds to bitrate ratios below 0.3. Depth bitrate for depth map should be between 10% and 30% of total bitrate.

For $QP < 21$ and $QD <$ Threshold QD, significant loss of quality in synthesized views is caused rather by imperfect depth estimation and view synthesis. Below the abovementioned values, further increase of view and depth bitrate does not improve synthesized view quality.

These results have been obtained for the state-of-the-art depth estimation and view synthesis techniques [4-6] but may be modified when more precise techniques for depth estimation and view synthesis would be available.

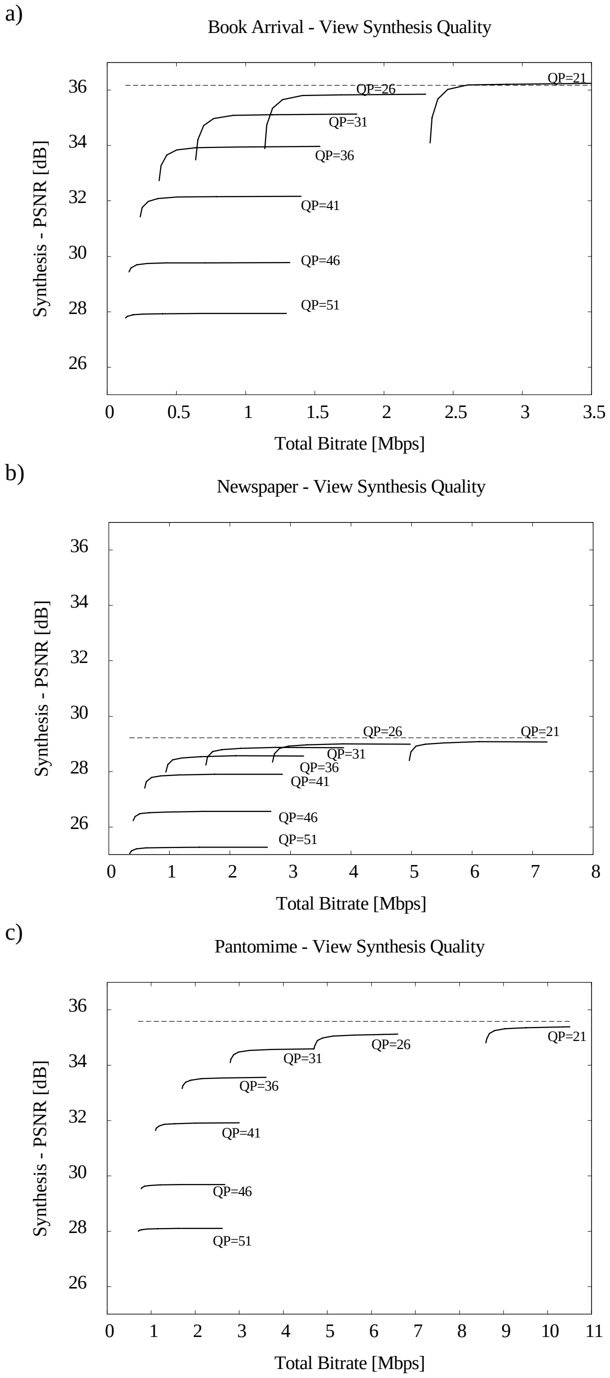

Figure 3. Synthesis quality as a function of total bitrate of two views and depth maps for three test sequences. Seven lines correspond to different QP setting for reference view compression. Dashed horizontal line marks level of quality of synthesis using uncompressed input data.

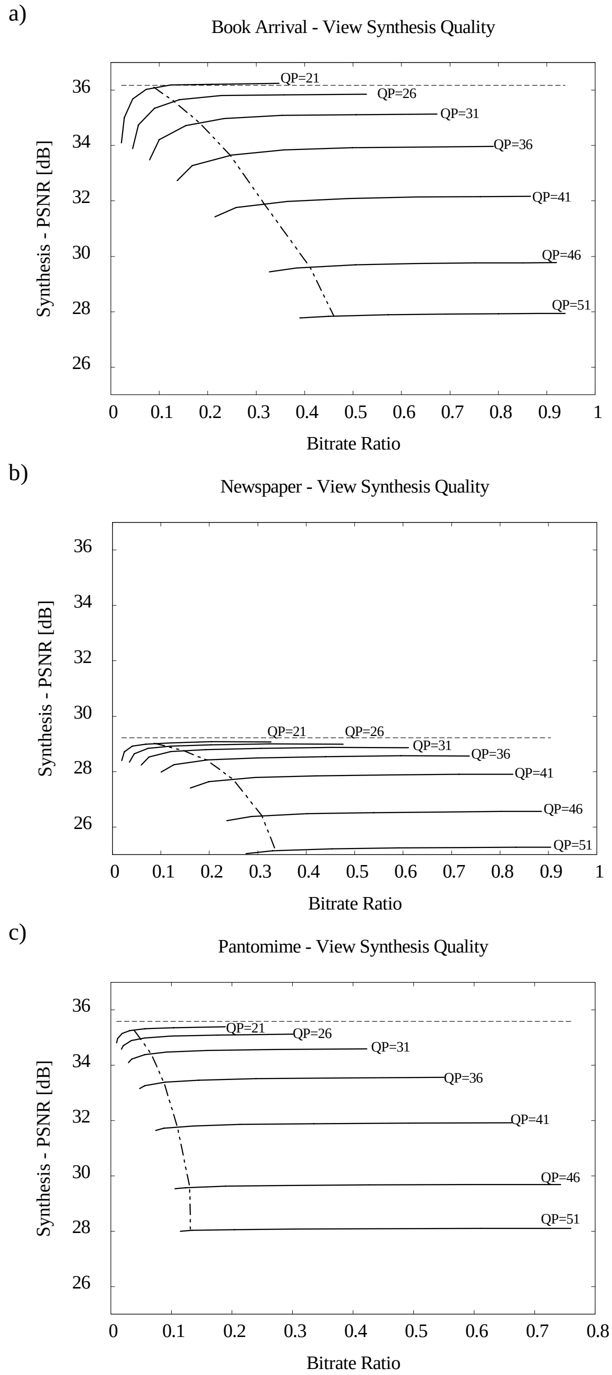

Figure 4. Synthesis quality as a function of bitrate ratio for three test sequences. Seven lines correspond to different QP setting for reference view compression. Dashed horizontal line marks level of quality of synthesis using uncompressed input data. Dashed curved line connects threshold QD points calculated from equation (2).

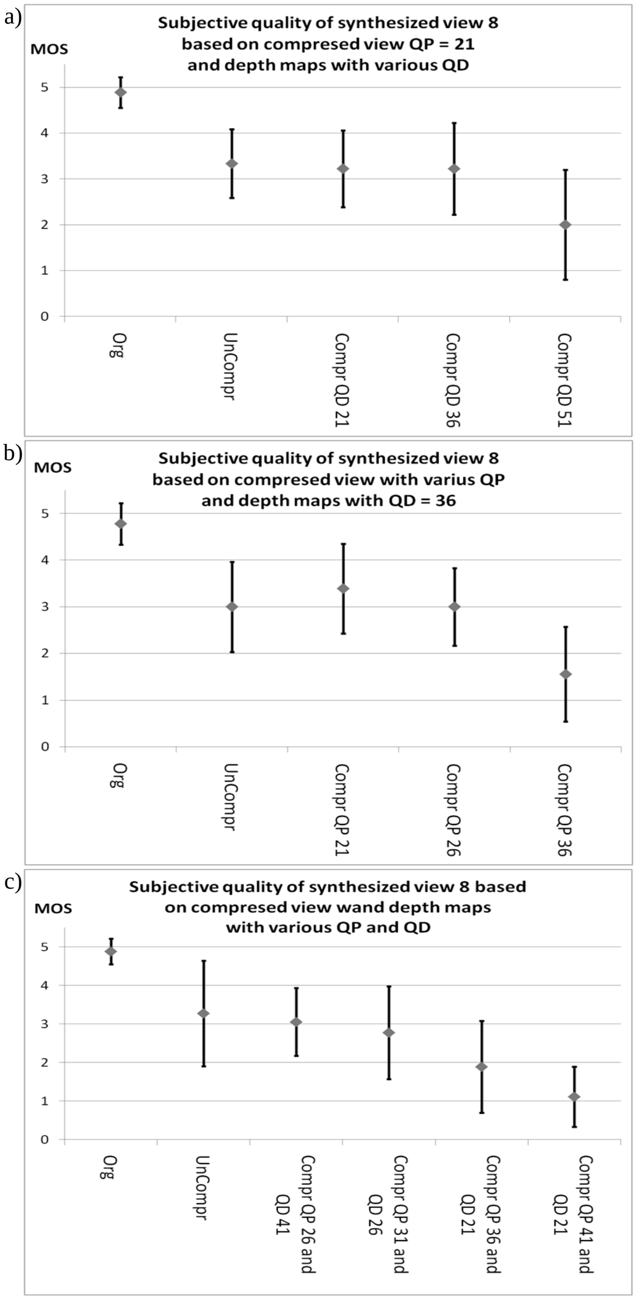

Figure 5. Subjective quality for synthesis for Book Arrival sequence. From left to right – original view, view synthesized from uncompressed input data, views synthesized from compressed data. In Figure 5a – variable QD, 5b – variable QP, 5c – with total bitrate around 1.5 Mbps.

REFERENCES

[1] Jung-Young Son, Bahram Javidi, Kae-Dal Kwack, "*Methods for Displaying Three-Dimensional Images*", Proceedings of the IEEE, vol. 94, no. 3, March 2006.

[2] P. Merkle, Y. Morvan, A. Smolic, D. Farin, K. Muller, P.H.N. de With, T. Wiegand, "*The Effect of Depth Compression on Multiview Rendering Quality*", Proceedings of 3DTV Conference: The True Vision - Capture, Transmission and Display of 3D Video, pp. 245-248, May 2008.

[3] Y. Morvana, P.H.N. de Witha, D. Farina, "*Platelet-Based Coding of Depth Maps for the Transmission of Multiview Images*", Proceedings of SPIE, Stereoscopic Displays and Applications, vol. 6055 p. 93-100, January 2006, San Jose (CA), USA

[4] M. Tanimoto, T. Fujii, K. Suzuki, N. Fukushima, Y. Mori, "*Reference Softwares for Depth Estimation and View Synthesis*", ISO/IEC JTC1/SC29/WG11 MPEG2008/M15377, Archamps, France, April 2008.

[5] Sang-Beom Lee, Yo-Sung Ho, "*Enhancement of Temporal Consistency for Multi-view Depth Map Estimation*", ISO/IEC JTC1/SC29/WG11 MPEG2008/M15594, Hannover, Germany, July 2008.

[6] P. Pandit, Y. Chen, S. Yea, "*Text of ISO/IEC 14496-5:2001/PDAM 15 Reference Software for Multiview Video Coding*", ISO/IEC JTC1/SC29/WG11 MPEG2008/W9974 Hannover, Germany, July 2008.

[7] I. Feldmann, M. Mueller, F. Zilly, R. Tanger, K. Mueller, A. Smolic, P. Kauff, T. Wiegand, "*HHI Test Material for 3D Video*", ISO/IEC JTC1/SC29/WG11 MPEG 2008/M15413, Archamps, France, April 2008.

[8] Yo-Sung Ho, Eun-Kyung Lee, Cheon Lee, "*Multiview Video Test Sequence and Camera Parameters*", ISO/IEC JTC1/SC29/WG11 MPEG2008/M15419, Archamps, France, April 2008.

[9] M. Tanimoto, T. Fujii, N. Fukushima, "*1D Parallel Test Sequences for MPEG-FTV*", ISO/IEC JTC1/SC29/WG11 MPEG2008/M15378, Archamps, France, April 2008.

[10] G.B. Akar, A.M. Tekalp, C. Fehn,; M.R. Civanlar, **"***Transport Methods in 3DTV—A Survey*", IEEE Trans. Circuits Systems for Video Techn., Vol. 17, Nov. 2007, pp.1622 – 1630.

[11] ISO/IEC 14496-10:2008/FDAM 1:2008(E), Information technology – Coding of audio-visual objects – part 10: Advanced Video Coding, Amendment 1: Multiview Video Coding.

[12] M. Flierl, B. Girod, "*Multiview Video Compression*", IEEE Signal Processing magazine, Vol. 24, Nov. 2007, pp. 66 – 76.

[13] Text of ISO/IEC FDIS 23002-3 Representation of Auxiliary Video and Supplemental Information, ISO/IEC JTC1/SC29/WG11, Jan. 2007, Doc. N8768, Marrakesh, Morocco.

[14] P. Merkle, A. Smolic, K. Müller, and T. Wiegand, "*Multi-view Video Plus Depth Representation and Coding*", ICIP 2007, IEEE International Conference on Image Processing, San Antonio, TX, USA, September 2007.