INTERNATIONAL ORGANISATION FOR STANDARDISATION
ORGANISATION INTERNATIONALE DE NORMALISATION
ISO/IEC JTC1/SC29/WG11
CODING OF MOVING PICTURES AND AUDIO

ISO/IEC JTC1/SC29/WG11 MPEG2014/m35070
October 2014, Strasbourg, France

Source Poznan University of Technology
Status Input
Title Super multiview image compression: results for Bee sequence (FTV EE3).
Author Marek Domański,
Dawid Mieloch,
Adrian Dziembowski,
Olgierd Stankiewicz,
Krzysztof Wegner (kwegner@multimedia.edu.pl)

1 Introduction

Super multiview is a hot research topic recently. Acquisition of super-multiview video is typically performed with a large number of cameras and thus large number of views is produced. As direct transmission of large number of views (necessary for super-multiview display) is not feasible, efficient representation of such material is needed. Therefore, development of efficient representation of super-multiview video is currently an important research topic.

Compression of super multiview should be standardized. Due to that, the MPEG experts group recently has started exploration activities. The most recent ones, established during 109th MPEG meeting in Sapporo by Free Viewpoint Television (FTV) ad hoc group, target at developing the know-how that will enable MPEG to develop the said super-multiview video format.

One of the areas of the above-mentioned exploration is related to the question about efficiency of the recently developed multiview and 3DV compression technologies – MV-HEVC [3] and 3D-HEVC [4], respectively. This question is considered in so called Exploration Experiment 3 (EE3) to which this document is related to.

2 Experiment description

Exploration Experiment 3 (EE3) defines two parts of experiments:

- EE3 A1, in which all of the views of super-multiview video are coded in order to compare performance gains of MV-HEVC and 3D-HEVC related to simulcast coding with HEVC. Also, distribution of quality among the coded views is assessed.
- EE3 A2, in which only a subset of views of super-multiview video is coded and the remaining views are synthesized with use of depth. This experiment is performed in context of 3D-HEVC and its goal is to asses distribution of quality among coded and synthesized views.

This document presents experimental results for both EE3 parts (A1 and A2) performed on Bee sequence [5]. This sequence is provided by NICT for MPEG FTV standardization.

3 Software and hardware platform

For coding we have used the latest version of MV-HEVC and 3D-HEVC coding techniques implemented in HTM software, version 11.2, with bug fixes provided over the FTV reflector.

For view synthesis we have used the latest version of VSRS 6.0.

All of the simulations results were generated on an about 80-core cluster system. The cluster platform's processing units have the following specifications:

- processor: Intel Xeon X5675 ,
- clock Speed: 3.06 GHz,
- memory: approx. 4 GB per Core,
- OS: 64-bit Windows Server 2008,
- compiler: Microsoft Visual Studio 2008 (64 bit).

4 Exploration Experiment 3 A1: Evaluation of the current multiview and 3DV coding technologies

The goal of EE3 A1 is to provide a reference for advanced experiments related to super-multiview coding. Therefore, depth data is not considered (video only) and all of the views are coded (without view subsetting). This means that evaluation of the quality for all of the views is possible without performing view synthesis.

The coding was performed with use of three techniques: HEVC in simulcast, MV-HEVC and 3D-HEVC. Various constant QP values has been used for coding, same for all used codecs. In coding with MV-HEVC and 3D-HEVC, one of the views (number 16) is the base view. Other views are side views which are coded predictively basing on the base view.

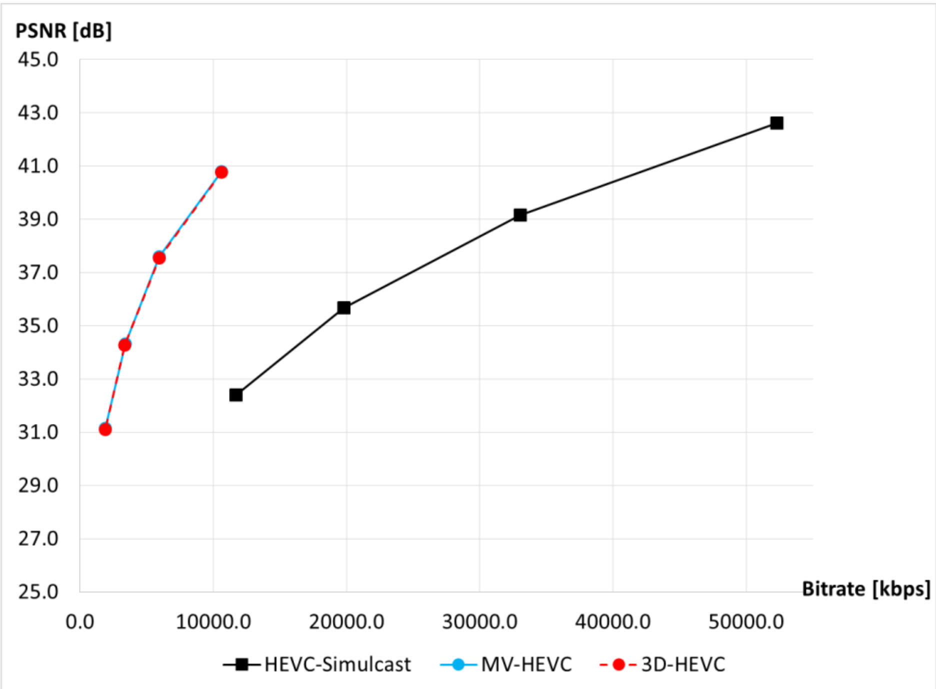

Figure 1. Rate-distortion (luminance PSNR) curves for coding of Bee sequence with use of HEVC-simulcast, MV-HEVC and 3D-HEVC.

The coded views have been compared with the original uncompressed views and luminance PSNR values have been calculated. Such results are presented in Fig. 1. Basing on them, Bjøntegaard BD-rates were calculated and resulted in the following values:

- MV-HEVC vs HEVC-simulcast, BD-Rate is -77.628%,
- 3D-HEVC vs HEVC-simulcast, BD-Rate is -77.632%.

Therefore it can be seen, that HEVC-simulcast has significantly worse performance than both MV-HEVC and 3D-HEVC (by about 77%). Moreover, it can be noticed that practically MV-HEVC and 3D-HEVC have the same coding performance. Such results is clear, because the coding was performed without depth data, and therefore 3D-HEVC is comparable to MV-HEVC.

Next, distribution of quality (luminance PSNR) over views was assessed. In Figure 2 it can be seen that in case of HEVC simulcast, such distribution is uniform. When inter-view coding is used (MV-HEVC, 3D-HEVC), it can be noticed that the base view (view 16) has higher quality (by about 1÷2 dB) than in the side views.

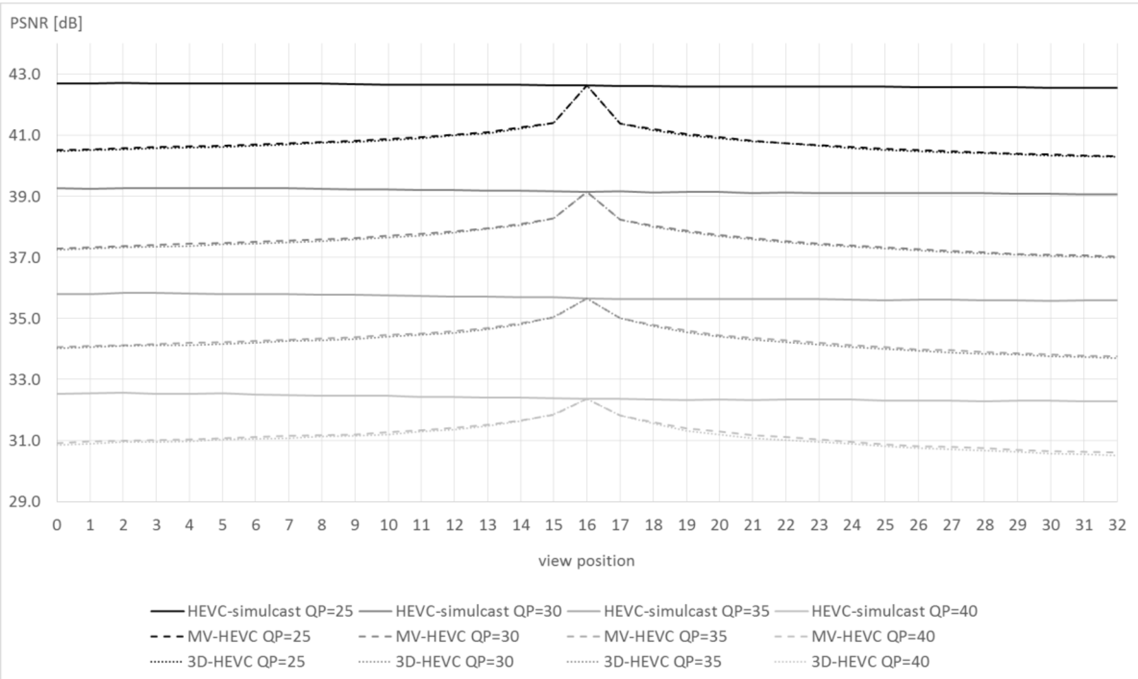

Figure 2. Quality of views coded in EE3 A1, measured as luminance PSNR related to the original uncompressed views in case of HEVC-simulcast, MV-HEVC and 3D-HEVC coding with various constant QP values.

5 Exploration Experiment 3 A2: Evaluation of quality of virtual views synthesized between the coded views

The goal of EE3 A2 is to evaluate necessary subset of views that is needed to be transmitted in order to reproduce super-multiview sequence with high quality.

Four subsets of views (selected from all views) are selected, according to EE3 description [2]. Those are presented in Table 1. Those views were fed into a 3D-HEVC codec. The decoded reconstructed views and associated depth maps are used to synthesize the remaining (not transmitted) views. Such synthesized views, together with the coded views, are used (Fig. 3) for evaluation of the quality.

Table 1. Specification of reference view position and virtual view position.

Sequence name	Subset	Positions of reference view to be transmitted and that are used to synthesize from	Positions of virtual views to be generated
Bee	3-views	Every 32nd view from 94 to 158, 3 views in total 58,90,122	Every view from: 59-89 and from 91-121
	5-views	Every 16th view from 94 to 158, 5 views in total 58,74,90,106,122	Every view from: 59-73, 75-89,91-105 and 107-121
	9-views	Every 8th view from 94 to 158, 9 views in total 58,66,74,82,90,98,106,114,122	Every view from: 59-65, 67-73, 75-81, 83-89, 91-97, 99-105, 107-113 and 115-121
	17-views	Every 4th view from 94 to 158, 17 views in total 58,62,66,70,74,78,82,86,90, 94,98,102,106,110,114,118,122	Every view from: 59-61, 63-65, 67-69, 71-73, 75-77, 79-81, 83-85, 87-89, 91-93, 95-97, 99-101, 103-105, 107-109, 111-113, 115-117, 119-121

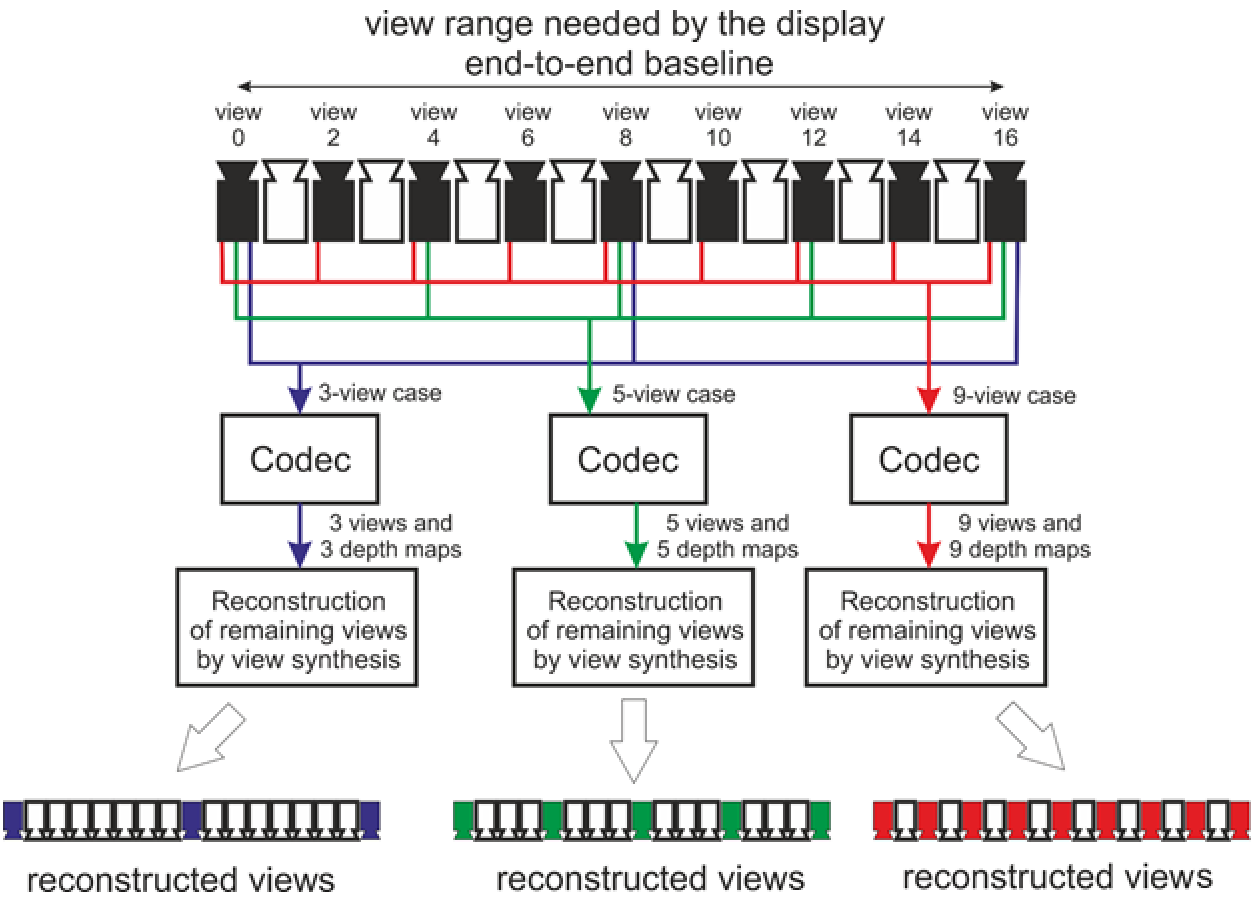

Figure 3. Overview of EE3 A2 experiment for 3-view, 5-view and 9-view cases.

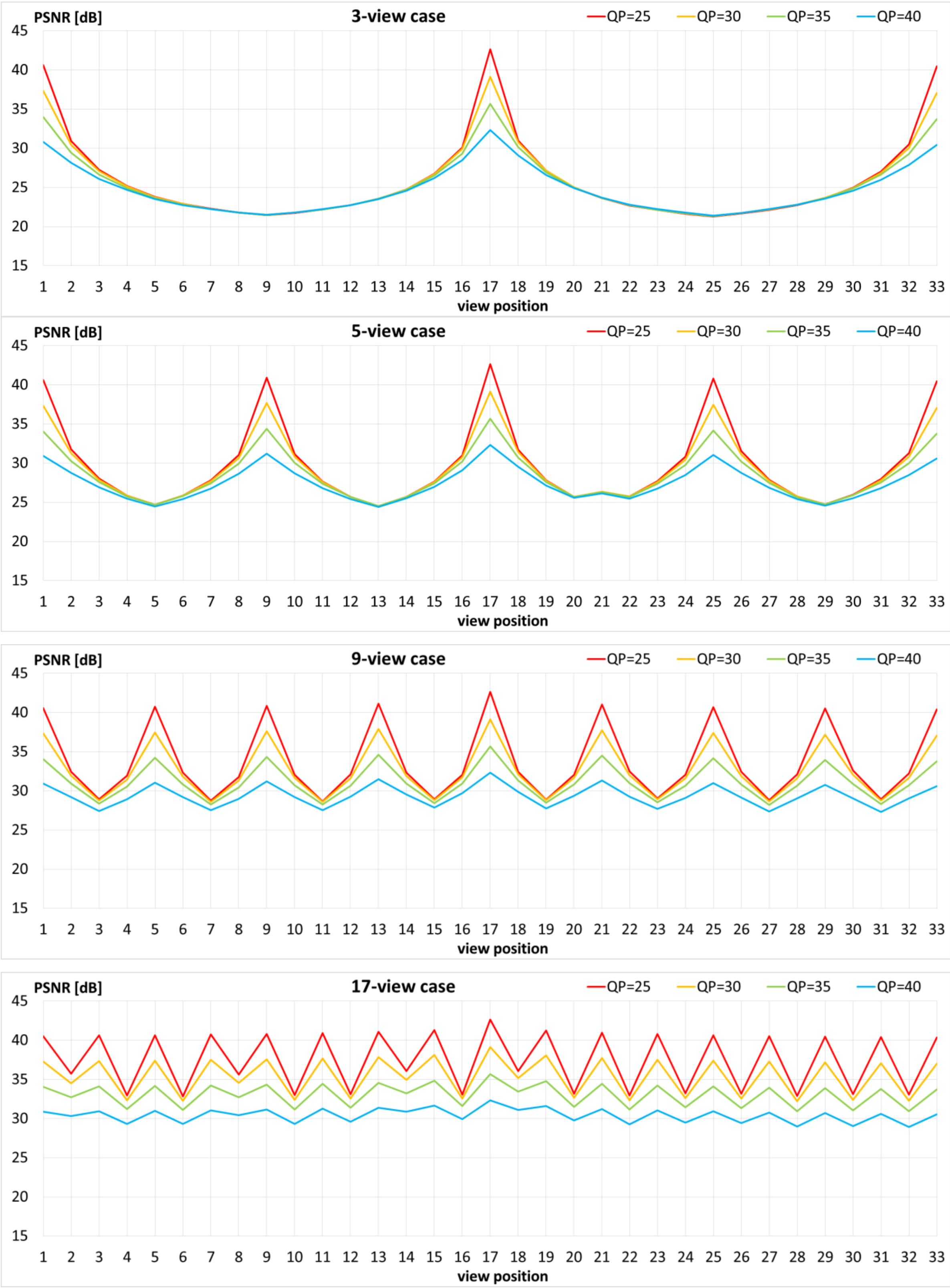

Figure 4. Quality of views coded and synthesized in EE3 A2, measured as luminance PSNR related to the original uncompressed views, in case of 3D-HEVC coding with various constant QP values and in various view subsetting scenarios.

6 Conclusions

In this documents, the results for Bee super-multiview test sequence are presented. Two experiments have been performed.

In the first experiment (EE3 A1) it is shown that MV-HEVC and 3D-HEVC provide substantial gains (about 77% in BD-rate) versus HEVC-simulcast in a case when all views are coded. Those gains are practically the same for MV-HEVC and for 3D-HEVC. It can be seen that usage of coding techniques that employ inter-view prediction leads to degradation of quality in the side views, which is about 1÷2 dB.

In the second experiment (EE3 A2) with view subsetting, four scenarios were used, in which, respectively, 3-views, 5-views, 9-views and 17-views are used for coding, whereas remaining views are synthesized at the decoder with use of transmitted depth maps corresponding to the transmitted video. The results show that the peak quality is always attained in coded views. As for the synthesized views, the quality deteriorates with the increase of the distance to the nearest coded views (the left nearest and the right nearest). The quality drop is up to about 10 dB. Also in this experiment, quality of the side views is lower than quality of the base view by about 1÷2 dB.

Acknowledgement

This work was supported by the research project financed by National Science Centre, Poland, according to the decision 2012/05/B/ST7/01279.

References

[1] “Purpose for FTV Exploration” ISO/IEC JTC1/SC29/WG11 Doc. MPEG2014/N14546, Sapporo, Japan, July 2014

[2] “Description of Exploration Experiments on Free-viewpoint Television (FTV)” ISO/IEC JTC1/SC29/WG11 Doc. MPEG2014/N14551, Sapporo, Japan, July 2014.

[3] G. Tech, K. Wegner, Y. Chen, M. Hannuksela, J. Boyce, „MV-HEVC Draft Text 3”, JCT3V Document, JCT3V-C1004, Geneva, Switzerland, Jan. 2013.

[4] G. Tech, K. Wegner, Y. Chen, S. Yea, "3D-HEVC Draft Text 5", JCT3V Document, JCT3V-I1001, Sapporo, Japan, 2014.

[5] T. Senoh, A. Ishikawa, M. Okui, K. Yamamoto, N. Inoue, "FTV AHG: EE1 and EE2 Results with Bee by NICT", IEC JTC1/SC29/WG11 Doc. MPEG2013/ M32995, Valencia, Spain, April 2014.