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

ISO/IEC JTC1/SC29/WG11
MPEG2018/M42309
April 2018, San Diego, USA

Source	**Poznań University of Technology**
Status	**Input**
Title	**Multiview compression of lightfield images**
Author	**Krzysztof Wegner, Olgierd Stankiewicz, Tomasz Grajek**

Abstract

In this document we present results of lightfield image compression with the use 3D-HEVC. Grids of 27×27 and 37×37 views are compressed.

1. Introduction

One of the goals of the MPEG-I exploration is to study the limits of the usage of 3D-HEVC in the context of lightfield compression.

3D-HEVC is an extension of MPEG-H part 2 (HEVC) standard, designed for efficient compression of a small number (three, four) of linearly arranged views. It uses interview prediction mechanism in order to exploit interview correlation for the sake of compression. Efficient interview motion prediction with advanced residual prediction allows for almost 40% bitrate reduction in comparison to simulcast (independent) transmission of three-four views.

In our previous document [m41805] we have proposed some interview prediction structures for efficient lightfield image compression. In that previous document we have studied grids of 3×3, 5×5 and 7×7 views. In this work we have considered much broader scenarios, e.g. compression of grid of 27×27 views (729 views in total) and 37×37 views (1369 views in total), which pushes the 3D-HEVC to its limits.

2. Lightfield image acquisition and formats

There are two commonly used ways of capturing lightfield images: multi-camera array and camera equipped with microlenses (Fig 1).

Fig. 1. Lightfield images acquisition systems.

Each of those ways leads to a different representation of the acquired lightfield (Fig. 2). The multi-camera approach leads to view-based representation, where each view represents entire scene as seen from a different viewpoint. The usage of camera with microlenses leads to completely different lightfield representation – so called microimages (angular) representation, in which each microimage represents only a fraction (part) of the observed scene.

Importantly, each of these representations can be converted into the other one.

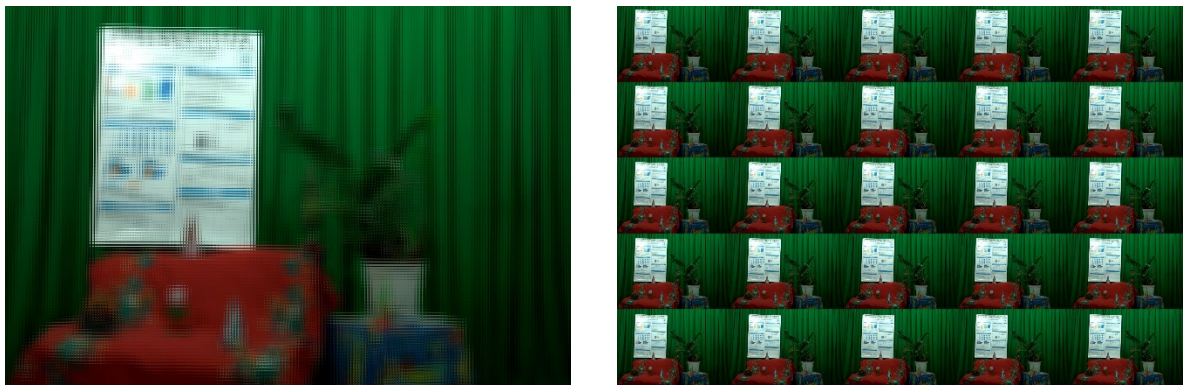

Fig. 2. Lightfield images in two commonly used representations:
microimages (left) multiview (right).

3. 3D-HEVC compression technique

3D-HEVC is designed to cope with multiview images and thus in order to use it for lightfield compression, we have used lightfield images in multiview representation format.
For the experiments we have used reference implementation of 3D-HEVC developed by JCT-3V (Joint Collaborative Team on 3D Video Coding Extension Development). The version of 3D-HEVC test model HTM, that has been used in the experiments, can be accessed at

https://hevc.hhi.fraunhofer.de/svn/svn_3DVCSoftware/tags/HTM-13.0

Limitations of the current version of the 3D-HEVC are well known [w17133] and can be summarized as follows:

- Current specification of 3D-HEVC limitation number of simultaneously compressed views to only 64.
- Number of Picture Parameter Sets that can be included in a single bitstream is limited to only 16.

In order to allow some degree of freedom in configuring 3D-HEVC we have removed some limitations in design. The following source code modification have been made to 3D-HEVC software to allow lightfield images compression in the scenario considered in this document.

```
In TLibCommon/TypeDef.h, line 57,
#define HEVC_EXT 2 // 3D-HEVC mode // define 2 for 3D-HEVC mode.

In TLibDecoder/TDecCAVLC.cpp,
line 235: // assert (uiCode <= 15); // comment it out
line 771: // assert (uiCode <= 15); // comment it out

In TLibCommon/TComSlice.h,
line 3039: // assert (psId < m_maxId); // comment it out
```

We have used 3D-HEVC in its most efficient mode HEVC_EXT 2, where sub coding unit level tools are enabled.

4. Other lightfield compression technology

For comparison purposes, we have also used one of the state-of-the-art compression techniques - HEVC with a special tools for compression of lightfield images in microimages representation, the so-called locally linear embedding-based prediction [LLEBP3D].

5. Used lightfield dateset

In order to study/compare 3D-HEVC with the other method on a commonly used LF acquired by cameras with microlenses. We have taken two lightfield images: “DemiCut” and “PlaneAndToy and converted it into a multiview representation (Fig. 3). This way we have obtained 37×37 view for DemiCut and 27×27 views for Plane And Toy lightfield images.

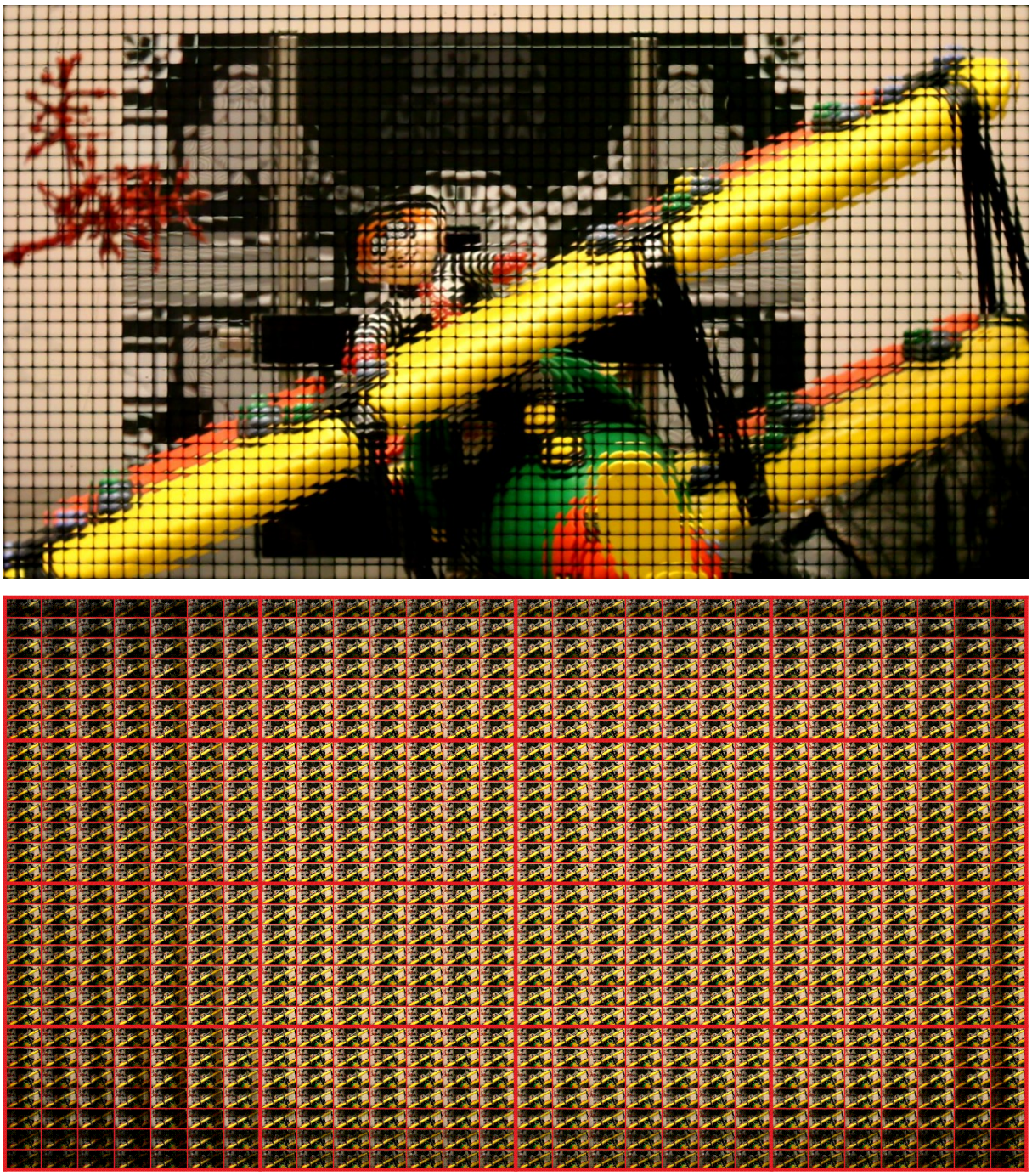

Fig. 3. Conversion of the PlaneAndToy microlens image into a 27x27 views lightfield.

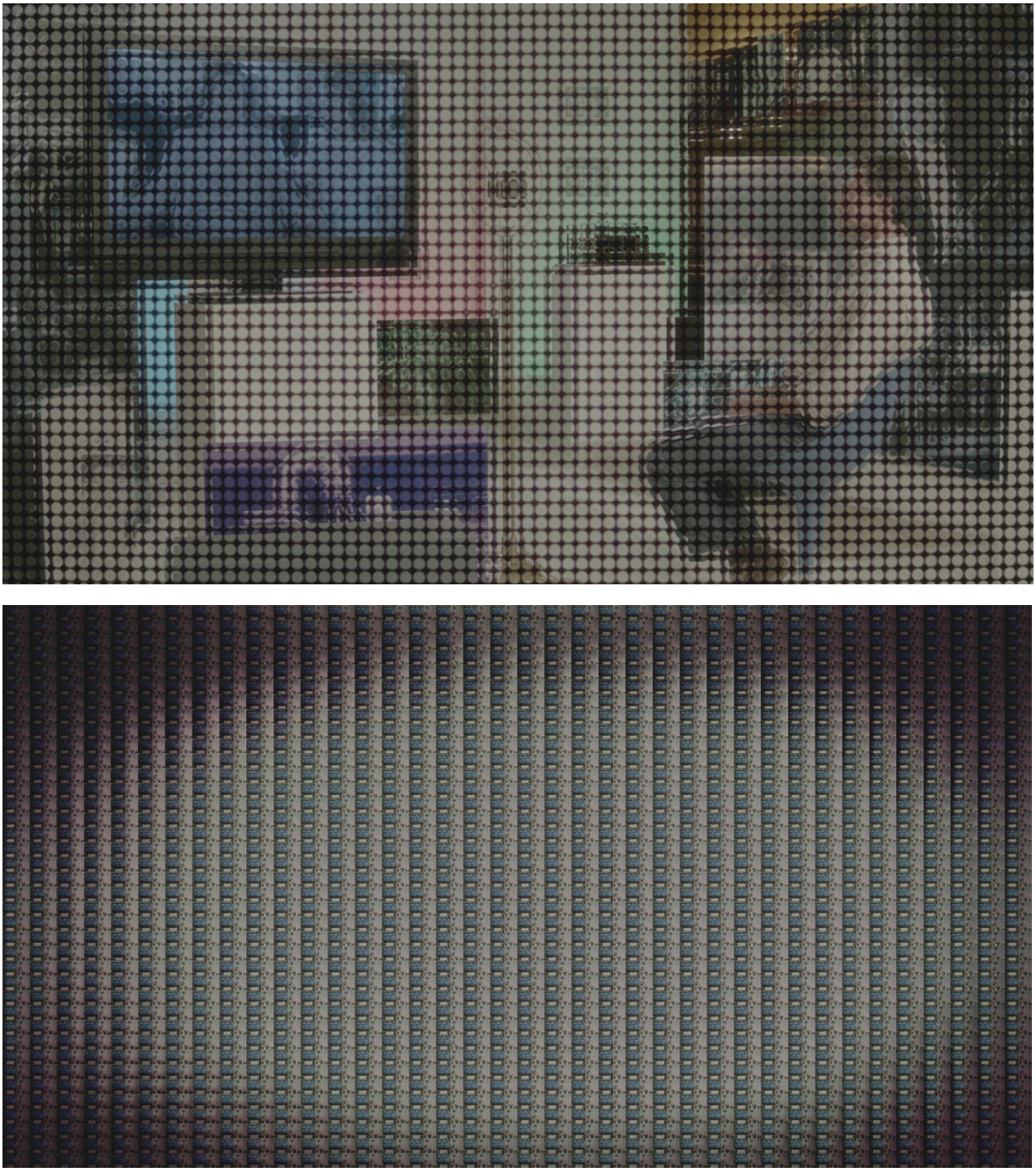

Fig. 4. Conversion of the DemiCut microlens image into a 37x37 views lightfield.

6. 3D-HEVC interview prediction structure.

In the study we have used three 3D-HEVC configurations:

- Simulcast scenario. Each view was encoded independently from the others with HEVC codec.
- Interview prediction from center view only. Center view from the whole grid (27×27 or 37×37) is being encoded independently form others. All surrounding views use prediction from center view only. Fig. 5 presents subset (9x9 view) of a prediction structure applied in this case. Numbers in rectangles represent a layer number – i.e. order in which views are being compressed. All views are encoded with appropriately configured 3D-HEVC.

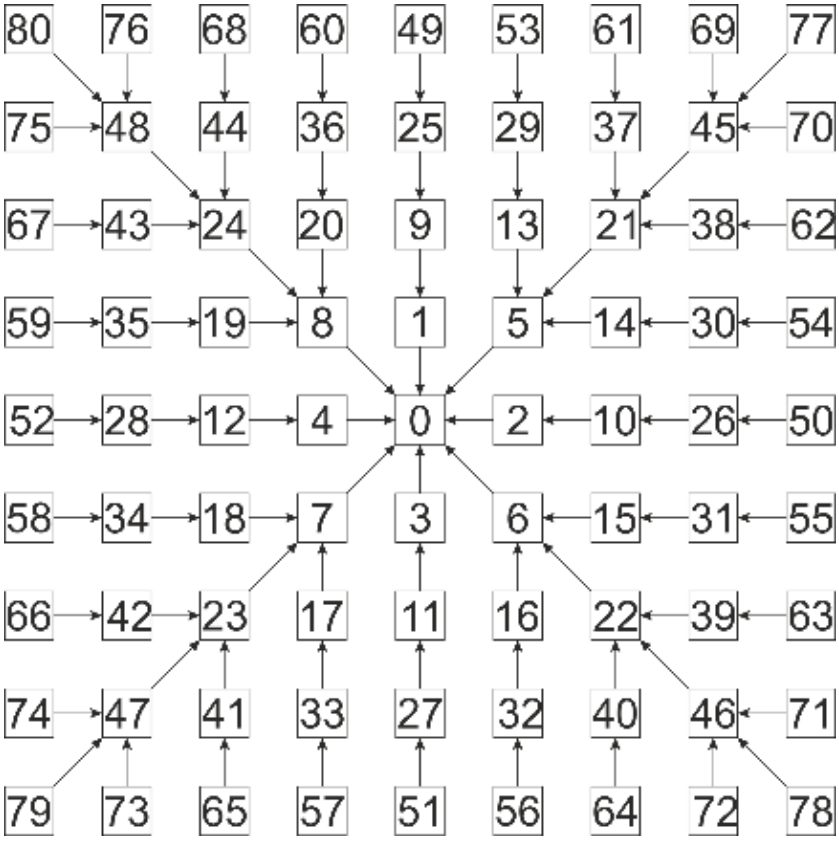

Fig. 5. Prediction structure for encoding 9x9 grid of views.

- Interview prediction along grid. Center view from the grid is being encoded independently form others. Views above, below, on the left and on the right of the center view are being predicted based on the center view, while corner views can use neighboring views but not center view. Figure 6 presents subset (9x9 view) of a prediction structure applied in this case.

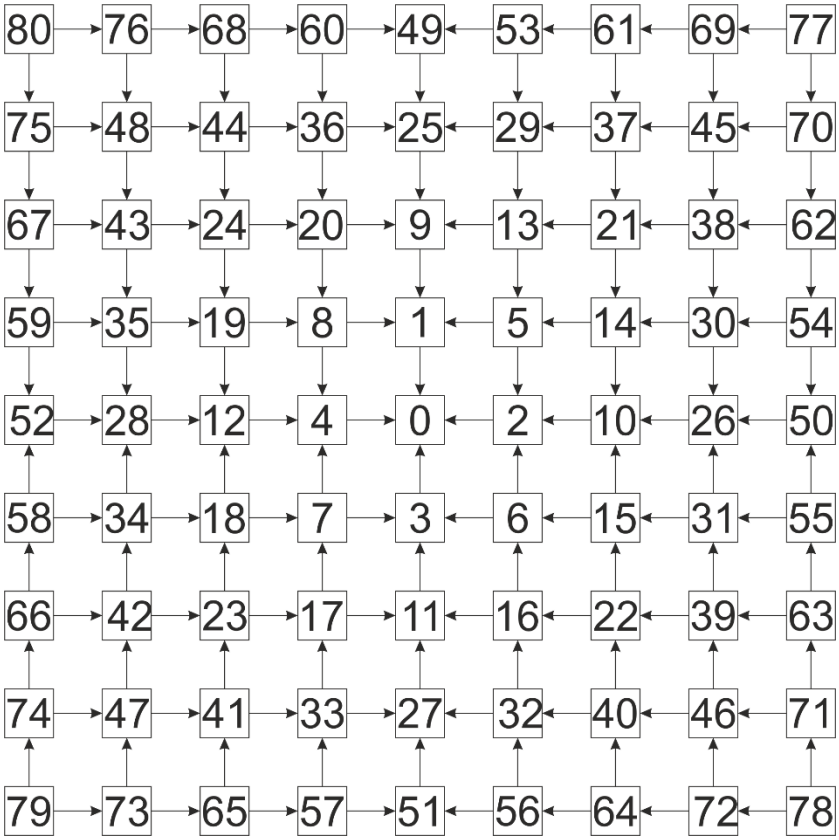

Figure 6. Example of interview prediction structure along the grid for grid of 9x9 views.

7. Results

All of the selected images have been encoded with the following codec:

- 3D-HEVC A – 3D-HEVC with “from center direction” interview prediction structure
- 3D-HEVC B – 3D-HEVC with “along the grid” interview prediction structure
- Simulcast – HEVC-based separate compression of all the images.

Table 1. Comparison of BD-Rate metric for 27x27 PlayandToys lightfield image. 3D-HEVC A - “from center direction” prediction structure, 3D-HEVC B - “along the grid” interview prediction structure, Simulcast - HEVC-based separate compression of all the images.

	3D-HEVC A	3D-HEVC B	Simulcast	3D-HEVC A without VPS	3D-HEVC B without VPS
3D-HEVC A	0.00	0.12	226.65	1487.63	1510.68
3D-HEVC B	-0.12	0.00	226.30	1486.66	1509.67
Simulcast	-69.39	-69.35	0.00	319.33	326.08
3D-HEVC A without VPS	-93.70	-93.70	-76.15	0.00	1.39
3D-HEVC B without VPS	-93.79	-93.79	-76.53	-1.37	0.00

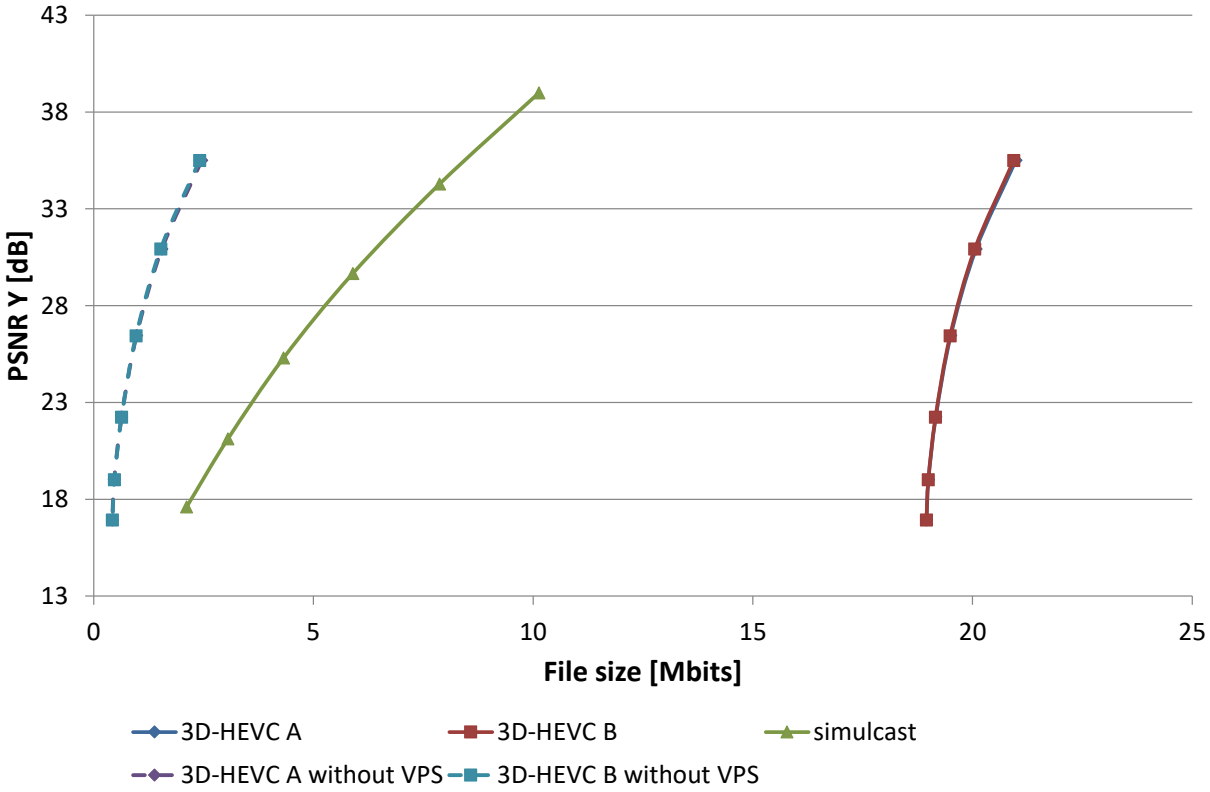

Fig. 7. Results for Play and Toy sequence.

Table 2. Comparison of BD-Rate metric for 37x37 DemiCut lightfield image. 3D-HEVC A - "from center direction" prediction structure, 3D-HEVC B - "along the grid" interview prediction structure, Simulcast - HEVC-based separate compression of all the images.

	3D-HEVC A	3D-HEVC B	Simulcast	3D-HEVC A without VPS	3D-HEVC B without VPS
3D-HEVC A	0.00	0.01	1319.14	5745.87	5765.79
3D-HEVC B	-0.01	0.00	1320.30	5753.35	5768.01
Simulcast	-92.95	-92.96	0.00	295.27	296.39
3D-HEVC A without VPS	-98.29	-98.29	-74.70	0.00	0.20
3D-HEVC B without VPS	-98.30	-98.30	-74.77	-0.20	0.00

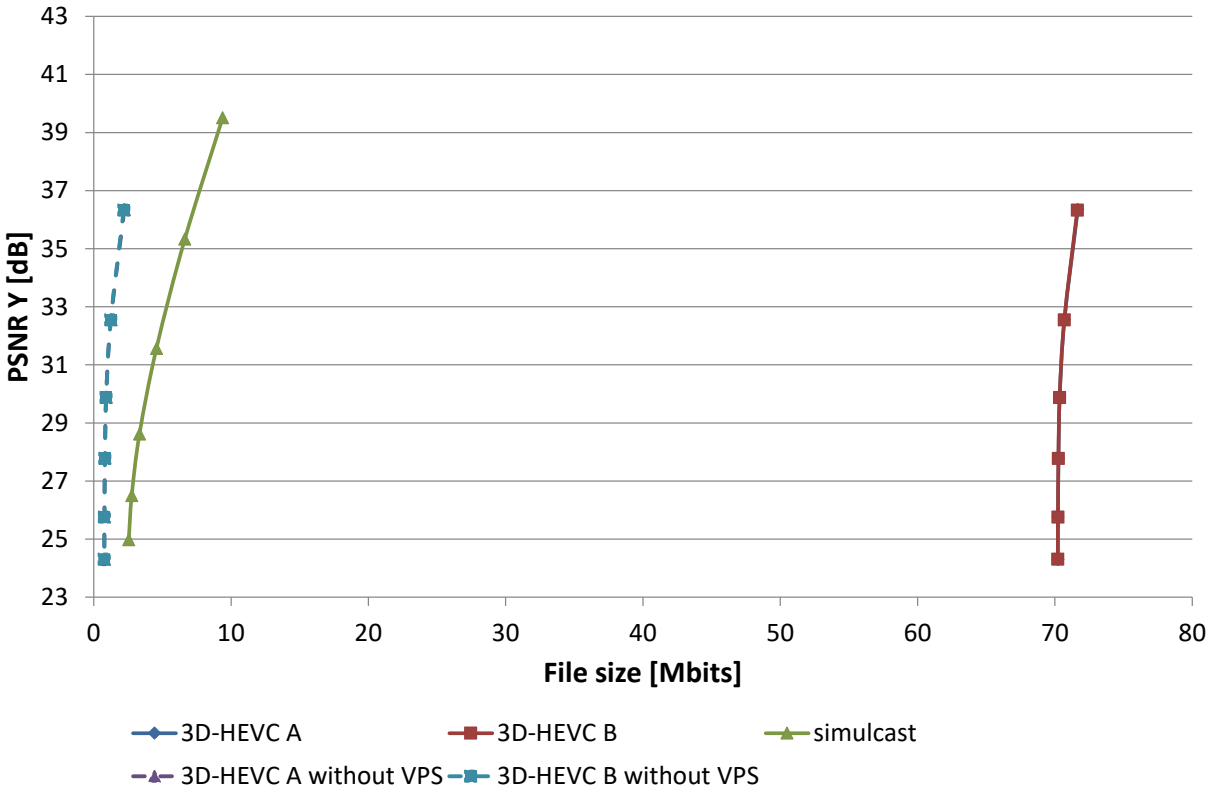

Fig. 8. Results for DemiCut sequence.

8. Conclusion

We have shown that with small changes to the software it is possible to encode moderate sizes of multiview lightfield images (grid of 27×27 or 37×37 views).

Definition of the prediction structure needed to be transmitted in Video Parameter Set (VPS). Unfortunately, VPS for even such moderate size of multiview lightfield image is so enormous big that it consumes all possible coding gains from interview prediction.

It is important for the future MPEG-I compression standard to develop more efficient way to transmit interview prediction structure to the decoder.

Acknowledgement

This work was supported by The National Centre for Research and Development in the LIDER Programme (LIDER/34/0177/L-8/16/NCBR/2017).

References

[1] [m41805] Krzysztof Wegner, Luis F. R. Lucas, Sergio M. M. Faria, Tomasz Grajek "Results of exploration experiment on dense Light Fields compression" ISO/IEC JTC1/SC29/WG11, MPEG2017/M41805, Macao, China, October 2017.

[2] [w17133] "Limitations of multi-view extensions of HEVC and fixes for MPEG-I Phase 2" ISO/IEC JTC1/SC29/WG11 MPEG2017/N17133, Macau, China, October 2017.

[3] [LLEBP3D] LFR Lucas, C Conti, P Nunes, L Ducla Soares, NMM Rodrigues, C.L. Pagliari, E.A.B. da Silva, S.M.M. de Faria “Locally linear embedding-based prediction for 3D holoscopic image coding using HEVC”,, 22nd European Signal Processing Conference (EUSIPCO), pp. 11-15, 2014.