**INTERNATIONAL ORGANISATION FOR STANDARDISATION**

**ORGANISATION INTERNATIONALE DE NORMALISATION**

**ISO/IEC JTC1/SC29/WG11**

**CODING OF MOVING PICTURES AND AUDIO**

**ISO/IEC JTC1/SC29/WG11 MPEG2016/N16318**

**June 2016, Geneve, CH**

|  |  |
| --- | --- |
| Status | Approved |
| Title: | Summary of the results of the Call for Evidence on Free-Viewpoint Television: Super-Multiview and Free Navigation |
| Source: | Test, Requirements |
| Authors | Vittorio Baroncini, Masayuki Tanimoto, Olgierd Stankiewicz |

# Introduction

MPEG has issued a Call for Evidence (CfE) on Free-Viewpoint Television: Super-Multiview and Free Navigation [1] in order to evaluate the merits of proposed 3D data formats and associated compression technology. This document provides a summary of the evaluation of the responses to the call received in the context of two investigated target application scenarios: Super-Multiview (SMV) displays and Free Navigation (FN), both described in the CfE [1].

Super-Multiview (SMV) displays render hundreds of linearly or angularly arranged, horizontal parallax ultra-dense views. Therefore, vast number of input views need to be transmitted. As such, it puts very demanding requirements to the codec for both throughput and compression efficiency, even considering the fact that the views are highly correlated, which was observed in previously during the experiments.

Free Navigation (FN) on the other hand enables users to view a scene by freely changing the viewpoints as we do naturally in the real world. FTV is the ultimate 3D-TV experience with infinite number of selectable views. Infinite number of views cannot be acquired and transmitted and thus only a limited set of input views is acquired and coded, while the content presented to the viewer is synthesized.

Although there exist commonalities between SMV and FN, these two categories have been evaluated in a different way: SMV aims at high compression exploiting the essential information embedded in all camera views, while improved view synthesis is an additional cornerstone for FN in large baseline arbitrary camera arrangements.

# Application Scenario #1: Super-Multiview Display (SMV)

## Methodology

### Test Material

The data set specified in the respective sections of the Call for Evidence (CfE) [1] was used for evaluation. The source is a large number of views as required by the display. The number of views is 80 views and they are arranged in a dense 1D array (linear or arc) (Fig. 1).

Sequences that are shorter than 10 seconds have been extended twice in a ping-pong loop to last exactly 10 seconds. This has been done immediately after decoding (e.g. before creation of the sweeps).

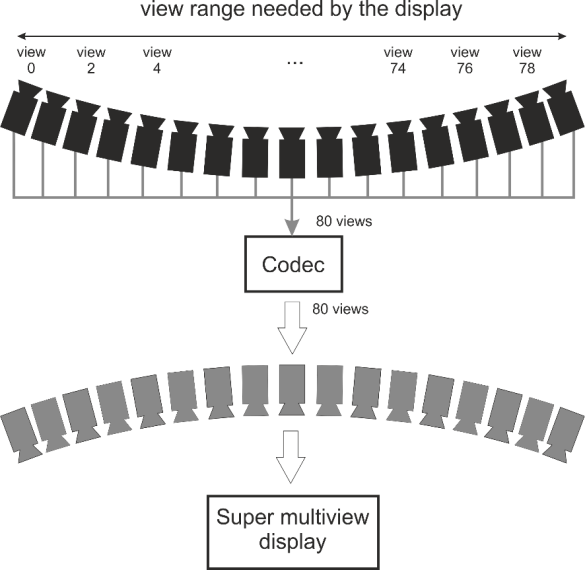


Fig. 1. Super-Multiview (SMV) display application scenario.

Only one response to the CfE in SMV category has been received [6][7]. This response was partial and included only two synthetic sequences – Big Buck Bunny flowers and Big Buck Bunny butterfly. Therefore, from in total four sequences mentioned in the CfE, only these sequence have been used. Their description, along with the exact positions of the views (view numbers), considered as input and output of the codec, is shown in Table 1.

Table 1. Summary of the used sequence.

|  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- |
| Source | Seq. Name | Number of Views | Resolution (pel) | Frame rate (fps) | Length | Cam Arrangement (1D parall, 1D arc, 2D parall, 2D arc, Sphere, Arbitrary) | Transmitted views positions | Transmitted frame range |
| Holografika | Big Buck Bunny flowers | 91 | 1280x768 | 24 | 5 sec  120 frames | 45 degree arc convergent | 5-84 | 0-120 |
| Holografika | Big Buck Bunny butterfly | 91 | 1280x768 | 24 | 5 sec  120 frames | 45 degree arc convergent | 5-84 | 0-119 |

### Anchor and submission configuration

In the anchor configuration all of the input views have been directly compressed with the use of modified 3D‑HEVC encoder version 13.0 [1], and then transmitted and decoded for the display. The details about the anchor configuration can be found in the CfE.

Four rate-points R1 to R4 (Table 2) were specified both for anchors and the proponents.

Table 2. Specification of rate-points for SMV case.

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| Seq. Name | Maximum bitrate for all data for recreation of required  number of output views [kbps]  (QP values used for anchor bitstreams) | | | |
| Rate-point 1 | Rate-point 2 | Rate-point 3 | Rate-point 4 |
| Big Buck Bunny flowers | 5513.0 (35) | 3298.5 (40) | 1905.0 (45) | 1156.3 (50) |
| Big Buck Bunny butterfly | 1665.9 (37) | 1245.3 (40) | 862.9 (44) | 563.9 (50) |

### Preparation of the material for subjective evaluation

The materials to be evaluated were prepared according to the methodology described is the CfE [1] as follows. Video clips of the decoded views were combined to create sweeps through all of the transmitted views. The starting position of the sweeps were selected randomly by the test chair as specified in Table 3.

Table 3. Starting positions of the sweeps, selected randomly by the test chair.

|  |  |  |
| --- | --- | --- |
| No. | Seq. Name | Starting position of the sweeps |
| 1 | Champagne Tower | 75 |
| 2 | Pantomime | 15 |
| 3 | Big Buck Bunny flowers | 55 |
| 4 | Big Buck Bunny butterfly | 45 |

Sweeps were constructed at a speed of one frame per view (Fig. 2). This has been performed both for the anchor and the (only) contribution.

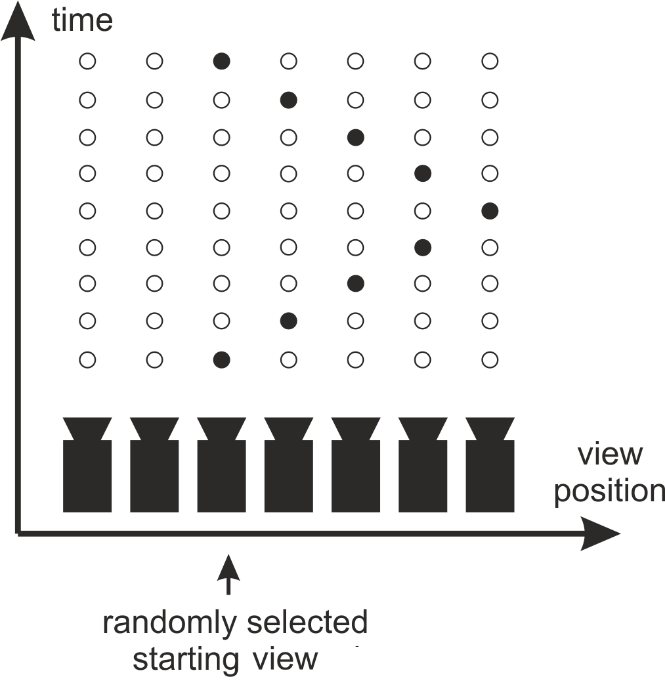


Fig. 2. Super-Multiview (SMV) display application scenario.

### Subjective evaluation procedure

Both the response and the anchor for Super-Multiview scenario were evaluated through formal subjective testing. The viewing was performed on June 1st 2016, starting from 8:30am.

The created sweeps were evaluated subjectively with ACR-HR method (Absolute Category Rating with Hidden References) [10] with the use of 11-point MOS (Mean Opinion Score) scale from 0 (bad) to 10 (excellent).

The viewers, 12 in total, were asked to evaluate the quality of the shown video, related to their own developed absolute quality scale. Each session was preceded with a training session in order stabilize and calibrate the scores given by the viewers.

The presented data consisted in:

* the references (uncompressed original input data- anchor R0),
* the anchors (input data compressed with HTM13) at four rate-points (anchor R1-R4),
* the response, each at four rate-points (response R1-R4).

Additionally, randomly selected test points have been repeated in order to allow consistency check. The presentation order was randomized.

The attained vote scores were analyzed statistically which yielded the mean results and 95% confidence intervals.

## Received responses

There was only one response received [6][7]. It conforms all of the requirements of the CfE, but includes only results for two sequences: Big Buck Bunny flowers and Big Buck Bunny butterfly.

The idea of the response consists in coding of a limited set of views (e.g. 28 out of 80) with corresponding depth maps estimated at the encoder (from the full set of views, e.g.80). The estimated depths are transmitted along with the views. At the decoder, the remaining (not coded) views are being synthesized with the use of depth-image based rendering (DIBR). For that a new algorithm called DEVS was used. For the compression of selected views with depth maps unmodified HTM version 13 were used.

## Results

### Objective results

Average PSNR of all coded views have been measured. Total bitrate of all data, including ythe depth maps and supplementary information, is used.





Fig. 3. Results of subjective evaluation of Super-Multiview (SMV) application scenario.

Table 4. Summary of the objective results for Super-Multiview (SMV) application scenario. Bjøntegaard deltas (BD-RATE) of the response versus anchors (Decoded view PSNR / Decoded video bitrate). Negative values depict gains of the response.

|  |  |
| --- | --- |
| **Sequence** | **NICT** |
| Big Buck Bunny flowers | -3,2% |
| Big Buck Bunny butterfly | -17.5% |
| **Average (all)** | **-10.4%** |

### Subjective results





Fig. 4. Results of subjective evaluation of Super-Multiview (SMV) application scenario.

For some of the rate-point is was observed statistically significant differences from the anchor e.g. R1 and R2 of BBB butterfly. For others, although, there is a tendency that no statistically significant conclusion can be made.

Is order to attain summary of the compression gains that can be attained with the use of the proposed technology Bjøntegaard deltas over the Mean Opinion Scores (MOS) have been calculated (Table 4).





Fig. 5. Results of subjective evaluation of Super-Multiview (SMV) application scenario.   
On horizontal axis there is actual bitrate of the submitted bitstreams. Confidence intervals of given points have been marked as vertical bars. Also, the attained Bjøntegaard deltas have been visualized as horizontal arrows.

Table 4. Summary of the subjective results for Super-Multiview (SMV) application scenario. Bjøntegaard deltas (BD-RATE) of the response versus anchors (11-point MOS / Decoded video bitrate). Negative values depict gains of the response.

|  |  |
| --- | --- |
| **Sequence** | **NICT** |
| Big Buck Bunny flowers | -20.5% |
| Big Buck Bunny butterfly | -12.0% |
| **Average (all)** | **-16.3%** |

# Application Scenario #2: Free Navigation (FN)

## Methodology

### Test Material

The data sets specified in the respective sections of the Call for Evidence (CfE) [1] were used for the evaluation. The source is a sparse number of 7 views with arbitrary positioning and wide baseline distance between each view. The input views, along with all supplemental data, such as depth, are transmitted. At the output of considered system arbitrary view positions in 3D space can be rendered (Fig. 6).

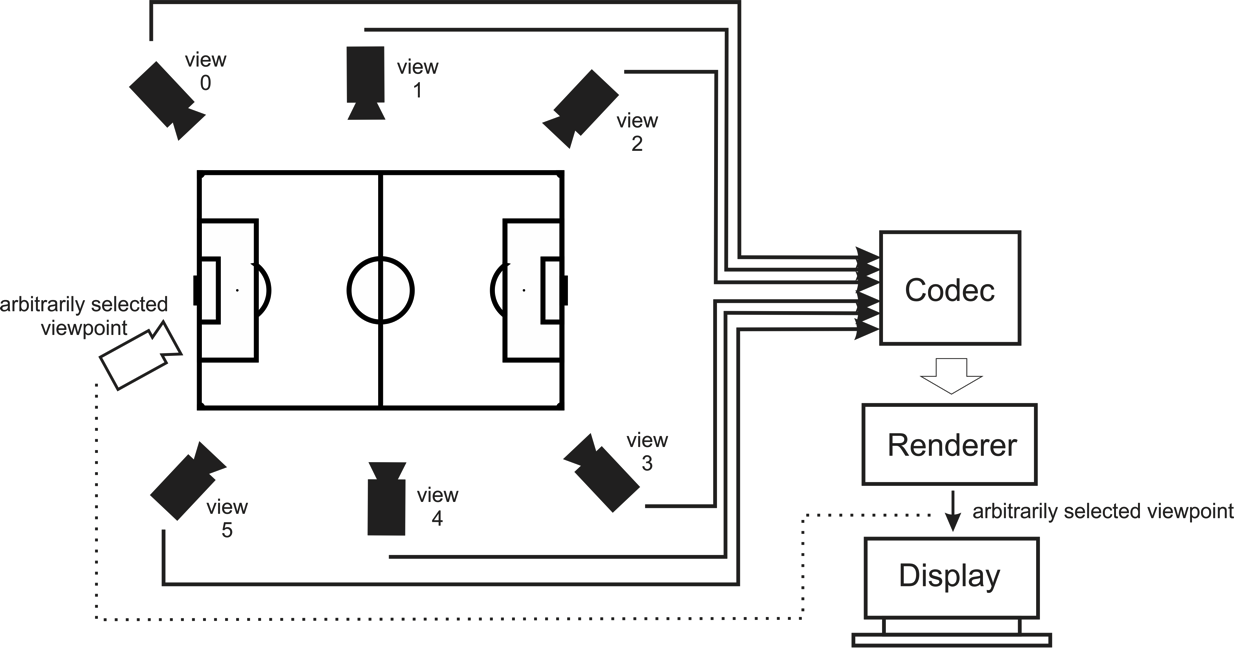


Fig. 6. Free Navigation (FN) application scenario.

Sequences that has been used for evaluation are summarized in Table 5. Seven views (as specified in CfE) of each sequence were coded and used for evaluation. Their positions are also shown in Table 5.

Sequences that are shorter than 10 seconds have been extended twice in a ping-pong loop to last exactly 10 seconds. This has been done immediately after decoding (e.g. before creation of the sweeps).

Table 5. Summary of the sequences used.

|  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| No. | Source | Seq. Name | Number of Views | Resolution (pel) | Frame rate (fps) | Length | Camera Arrangement | Depth Data Available | Views positions to be transmitted | Frame range to be transmitted |
| 1 | UHasselt | Soccer-Linear 2 | 8 | 1392x1136 | 60 | 10 sec 600 frames | 1D parallel | Yes | 1-7 | 0-599 |
| 2 | UHasselt | Soccer-Arc 1 | 7 | 1920x1080 | 25 | 22 sec 550 frames | 120 deg. Corner, arc | Yes | 1-7 | 0-249 |
| 3 | Poznan University of Technology | Poznan Blocks | 10 | 1920x1080 | 25 | 40 sec 1000 frames | 100 deg. arc around the scene | Yes | 2-8 | 0-249 |
| 4 | Holografika | Big Buck Bunny  Flowers  noBlur | 91 | 1920x1080 | 24 | 5 sec  121 frames | 45 deg, arc | Yes, ground truth depth | 6,19,32,45,58,71,84 | 0-120 |

### Output virtual views and view synthesis

The views, described in Table 6, coded along with the depth data, were reconstructed and used for view synthesis of intermediate virtual views, placed in–between of the input views.

The numbers of required virtual views rendered between each pair of transmitted views are provided in Table 6. Exact camera parameters for each individual virtual viewpoint were taken from the attachment of the CfE.

Table 6. Number of virtual views between each pair of cameras.

|  |  |
| --- | --- |
| Sequence | Number of required virtual views between each pair of cameras |
| Soccer Linear 2 | 14 |
| Soccer Arc 1 | 22 |
| Poznan Blocks | 12 |
| Big Buck Bunny Flowers | 12 |

### Anchor and submission configuration

In the anchor configuration all of the input views along with the corresponding depth data have been compressed using modified 3D-HEVC encoder (based on HTM version 13.0). After decoding, the requested view positions were synthesized using VSRS version 4.1 with enabled depth-based depth blending technique. The details about the anchor configuration are in the CfE.

In anchor configuration seven input views together with their corresponding depth data were encoded at 4 rate-points, as specified in Table 7.

Table 7. Specification of rate-point for Free Navigation (FN) scenario.

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| No. | Seq. Name | Maximum bitrate for all data required for recreation of required number of output views [kbps] (QP texture/ QP depth) | | | |
| Rate-point 1 | Rate-point 2 | Rate-point 3 | Rate-point 4 |
| 1 | Soccer Linear 2 | 3952.8 (30/39) | 1287.9 (37/43) | 506.0 (44/47) | 362.8 (47/50) |
| 2 | Soccer Arc | 5462.9 (30/39) | 2284.2 (37/43) | 899.1 (44/47) | 592.0 (47/50) |
| 3 | Poznan Blocks | 5927.8 (30/39) | 3187.95 (35/42) | 1559.6 (40/45) | 823.4 (45/48) |
| 4 | Big Buck Bunny Flowers | 1769.7 (37/43) | 1267.8 (40/45) | 816.9 (44/47) | 561.2 (47/50) |

### Preparation of the material for subjective evaluation

The materials for the subjective evaluation were prepared according to the methodology described is the CfE [1] as follows. Video clips of the rendered views were combined to create sweeps through all of the rendered and reconstructed views. The starting positions of the sweeps were selected randomly by the test chair (Table 8).

Table 8. Starting positions of the sweeps selected randomly by the test chair.

|  |  |  |
| --- | --- | --- |
| No. | Seq. Name | Starting position of the sweeps |
| 1 | Soccer Linear 2 | 60 |
| 2 | Soccer Arc | 110 |
| 3 | Poznan Blocks | 20 |
| 4 | Big Buck Bunny Flowers | 35 |

The sweeps (Fig. 7) were constructed at a speed of one frame per view. This has been performed for the anchor and the responses.

Virtual views

virtual views

time

coded  
 views

Virtual path (sweep) with randomly selected starting position

Fig. 7. Free Navigation (FN) evaluation procedure.

### Subjective evaluation procedure

The submissions and the anchor for Free Navigation scenario were evaluated through formal subjective testing. The viewing was performed on May 31st 2016, starting from 2pm.

The created sweeps were evaluated subjectively with ACR-HR method (Absolute Category Rating with Hidden Reference) [10] with the use of 11-point MOS (Mean Opinion Score) scale from 0 (bad) to 10 (excellent).

The viewers, 30 in total, were asked to evaluate the quality of the shown video, related to their own developed absolute quality scale. Each session was preceded with a training session in order stabilize and calibrate the scores given by the viewers.

The presented data consisted of sweeps prepared from the following data:

1. Anchors

* decoded data at four rate-points (anchor R1-R4),
* uncompressed original input data (anchor R0).

1. Responses

* decoded data at four rate-points (response R1-R4),
* uncompressed original input data, synthesized with the submitted view synthesis technology (response R0 if available, anchor R0 otherwise).

In order not to exceed the human focus time, the whole test was divided into two parts. Each part were repeated 3 times resulting in 6 session it total. Additionally, randomly selected test points have been repeated in order to allow consistency check. The presentation order was randomized per session in order to prevent any contextual affect appear.

The attained scores were analyzed statistically which yielded the mean results and 95% confidence intervals.

## Received responses

There were three responses received which were evaluated in the context of the CfE:

* Poznan University of Technology [2][3].
* Zhejiang University [4][5].
* Hasselt University [8].

### Poznan University of Technology

Response from Poznan University of Technology is based on 3D-HEVC technology (modified HTM version 13.0). However, it is not compatible with 3D-HEVC standard, e.g. proposed bitstreams cannot be decoded with the use of 3D-HEVC decoder. Most of the high level syntax does not changes, but also there are some adjustments at low-level. Several syntax elements and tools have been modified:

* Transmission of camera parameters in VPS, including extrinsic and intrinsic parameters, like rotation matrices, translation and distortion parameters.
* Modification of Disparity Compensated Prediction (DCP). In our codec instead of disparity (along restricted horizontal direction) we use depth-based compensation.
* Modification of Neighboring Block Disparity Vector (NBDV). Instead of disparity restricted to horizontal direction, we use a vector.
* Modification of Depth-oriented NBDV (DoNBDV). In 3D-HEVC the disparity for a given block is set to the value that corresponds with the maximum value of four corner depth samples value of virtual depth map block. In the proposed method, the disparity is calculated based on half of the maximum depth sample value and the position of selected corner of the block.
* Modification of View Synthesis Prediction (VSP). In 3D-HEVC view synthesis in prediction is restricted to horizontal translation only. In our codec, full DIBR scheme is performed.
* Modification of Inter-view Motion Prediction (IvMP). In 3D-HEVC motion vectors are purely 2D. Because other views lay on the same plane, motion vectors after projection to other view remain the same. In the proposed extension, during the prediction, we accordingly rotate motion vector in 3D space.
* Color correction. As an initial preprocessing step the input views are color corrected.

The configuration is resembling those of anchors in the CfE [1], i.e. Main Profile, GOP size = 8, intra period = 24, hierarchical GOPs on, 4 reference frames, Neighboring Block Disparity Vector on, Depth oriented NBDV on, View Synthesis Prediction on, Inter-view Motion Prediction on, Illumination Compensation on.

One difference worth noticing is that View Synthesis Optimization (VSO) has been switched off.

### Zhejiang University.

Response from Zhejiang University is based on 3D-HEVC technology (modified HTM version 13.0). However, it is not compatible with 3D-HEVC standard, e.g. proposed bitstreams cannot be decoded with the use of 3D-HEVC decoder. Several coding tools have been modified and improved:

* Disparity Compensated Prediction: depth-based compensation, 2D disparity.
* Neighboring Block Disparity 2D-Vector.
* 2D disparity based View Synthesis Prediction (VSP).
* Adjust DV borrowed from neighboring block according to position of blocks.
* Adjust Depth borrowed from neighboring block according to position of blocks.
* Projection of motion vector as Inter-view MV predictor.
* Depth RDO for non-linear Camera set.

View Synthesis Optimization (VSO) have been adjusted to non-linear camera arrangement.

### Hasselt University.

Response from Hasselt University includes results only for Soccer Linear 2 sequence.

In the proposed approach depth data is not necessary to be transmitted and it is estimated at the decoder. Efficient plane sweeping depth estimation is proposed to estimate depth from the decoded video. Because the depth is not transmitted the bitrates of submitted bitstreams are adequately smaller than those of the anchors and thus smaller than the required rate-points R1-R4.

For the view synthesis, the response uses proprietary view synthesis software.

### Comparison of the submitted technologies

Two of the responses submitted (3.2.1 and 3.2.2) are the extensions of the 3D-HEVC technology for non-linear camera arrangement. Some of the tools are common in the proposals. There are also some tools that are disjoint and are implemented exclusively in only one of the responses. Merging of the tools from both responses can possibly result in higher compression performance than any of them independently.

The last response (3.2.3) is not proposing any new coding technology but rather focuses on decoder side depth estimation and rendering.

Summary of proposed tools and technologies is made in Table 9.

Table 9. Comparison of the tools submitted in the responses to the CfE in FN category.

|  |  |  |  |
| --- | --- | --- | --- |
| **Tools** | **Poznan University** | **Zhejiang University** | **Hasselt University** |
| Color correction: preprocessing. | Y | - | - |
| Camera parameter transmission extension:  extrinsic and intrinsic parameters, like rotation matrices, translation and distortion parameters. | Y | - | - |
| Disparity Compensated Prediction:  depth-based compensation，2D disparity. | Y | Y | - |
| Neighboring Block Disparity 2D-Vector | Y | Y | - |
| Modification of Depth-oriented NBDV (DoNBDV):  disparity is half of the maximum depth sample value and the position of selected corner of the block. | Y | - | - |
| 2D disparity based View Synthesis Prediction (VSP). | Y | Y | - |
| Adjustment of DV borrowed from neighboring block:  according to position of blocks. | - | Y | - |
| Adjustment of Depth from neighboring block:  according to position of blocks. | - | Y | - |
| Projection of motion vector as Inter-view MV predictor | Y | Y | - |
| Depth RDO for non-linear camera arrangement  View Synthesis Optimization for non-linear camera arrangement | - | Y | - |
| Proprietary view synthesis software:  e.g. based on multiple views, higher precision, hole filling, color correction | Y | - | Y |
| Depth estimation at the decoder side | - | - | Y |

## Results

### Objective results

Average PSNR values of all coded have been measured and used to calculate Bjøntegaard bitrate deltas. Total bitrate of all data including a depth maps and supplementary information was used.

Table 10. Summary of the objective results for Free Navigation (FN) application scenario. Bjøntegaard deltas (BD-RATE) of the proposal versus anchors (Decoded view PSNR / Decoded video bitrate). Negative values depict gains of the proposal.

|  |  |  |  |
| --- | --- | --- | --- |
|  | **Bjøntegaard** **delta, versus anchors (Decoded view PSNR / Decoded video bitrate)** | | |
| **Sequence** | **UHasselt\*** | **Poznan** | **Zheijang** |
| Big Buck Bunny flowers | - | -5.21% | -4.62% |
| PoznanBlocks | - | -16.25% | -9.77% |
| SoccerArc | - | -1.14% | -11.93% |
| SoccerLinear | -31,6% | +2.20% | -0.18% |
| **Average (nonlinear)** | - | **-7.53%** | **-8.77%** |
| **Average (all)** | - | **-5.10%** | **-6.62%** |

\*- this proposal consists in not sending the depth data and thus the bitrates of submitted bitstreams are lower, which corresponds to the bitrate reduction.

Table 11. Objective results for Free Navigation (FN) application scenario of Zhejiang University response. Bjøntegaard deltas (BD-RATE) of the proposal versus anchors (decoded and synthesized views PSNR / Total bitrate).

|  |  |
| --- | --- |
|  | **Bjøntegaard** **delta, versus anchors (Decoded and synthesized view PSNR / total bitrate)** |
| **Sequence** | **Zheijang** |
| Big Buck Bunny flowers | -2.08% |
| PoznanBlocks | -11.84% |
| SoccerArc | -17.60% |
| SoccerLinear | -0.04% |
| **Average (nonlinear)** | **-10.51%** |
| **Average (all)** | **-7.89%** |

We consider Soccer linear 2 independently as it is a linearly arranged sequence, and is hence outside the scope of the current FTV activity, which targets multi-camera setups beyond linear arrangements.

### Subjective results

For some responses for some of the sequences there were observed statistically significant differences from the anchor e.g. BBB Flowers. For others although there is a strong tendency that no statistically significant conclusions can be made.

In order to summarize the compression gains that we reached by the proposed technology, Bjøntegaard deltas over the mean opinion scores have been calculated. (Table 11).





Fig. 8a. Results of subjective evaluation of Free Navigation (FN) application scenario. The bars related to four coded rate-point R1-R4, as well as synthesis from the uncompressed original (R0) are marked.





Fig. 8b. Results of subjective evaluation of Free Navigation (FN) application scenario. The bars related to four coded rate-point R1-R4, as well as synthesis from the uncompressed original (R0) are marked.





Fig. 9a. Results of subjective evaluation of Free Navigation (FN) application scenario.   
On horizontal axis there is actual bitrate of the submitted bitstreams. Confidence intervals of given points have been marked as vertical bars. Also, the exemplary attained Bjøntegaard deltas have been visualized as horizontal arrows.





Fig. 9b. Results of subjective evaluation of Free Navigation (FN) application scenario.   
On horizontal axis there is actual bitrate of the submitted bitstreams. Confidence intervals of given points have been marked as vertical bars. Also, the exemplary attained Bjøntegaard deltas have been visualized as horizontal arrows.

Table 12. Summary of the subjective results for Free Navigation (FN) application scenario. Bjøntegaard deltas (BD-RATE) of the responses versus anchors (11-point MOS / Decoded video bitrate). Negative values depict gains of the responses.

|  |  |  |  |
| --- | --- | --- | --- |
|  | **Bjøntegaard** **delta, versus anchors (MOS / Decoded video bitrate)** | | |
| **Sequence** | **UHasselt** | **Poznan** | **Zheijang** |
| Big Buck Bunny flowers | - | -43.3% | -19.2% |
| Poznan Blocks | - | -36.9% | -4.9% |
| Soccer Arc | - | -70.1% | -69.6% |
| Soccer Linear 2 | -46.51% | -28.5% | 27.0% |
| **Average (nonlinear)** | - | **-50.1%** | **-31.3%** |
| **Average (all)** | - | **-44.7%** | **-16.7%** |

The differences in the achieved results in subjective and objective evaluation come from the fact that in objective evaluation only the decoded views are considered, while in the subjective evaluation also the rendering capability was assessed.

We consider Soccer linear 2 independently as it is a linearly arranged sequence, and is hence outside the scope of the current FTV activity, which targets multi-camera setups beyond linear arrangements.

# Concluding remarks

The group has evaluated the responses to the CfE. The obtained results show significant improvement and strongly suggest that there is technology better adapted to the considered application scenarios of SMV and FN than the currently standardized solutions.

In Super-Multiview (SMV) application scenario the observed overall Bjøntegaard bitrate reduction is about 10% objectively. Such results have to be although considered with the fact in mind, that the only response in this category contained only 2 sequences.

In Free Navigation (FN) application scenario the observed overall Bjøntegaard bitrate reduction is on average from about 7% to 10% depending on the proposal, for the set of sequences with nonlinearly arranged cameras, which constitute the main scope of the FTV group. Also subjective evaluation was performed for the responses and in some of the cases statistically significant improvements were observed.

For the sequence with linearly arranged cameras, one response was received which allowed for even higher bitrate reductions thanks to the decoder-side depth estimation. The achieved results are 32% of bitrate reduction with respect to PSNRs.

As a side product of the evaluation of the responses it was observed that there is a room for improvement of the quality of the views rendered from the uncompressed data.

# References

[1] “Call for Evidence on Free-Viewpoint Television: Super-Multiview and Free Navigation”, ISO/IEC JTC1/SC29/WG11 MPEG2015/N15348, June 2015, Poland, Warsaw.

[2] Marek Domański, Adrian Dziembowski, Adam Grzelka, Łukasz Kowalski, Dawid Mieloch, Jarosław Samelak, Olgierd Stankiewicz , Jakub Stankowski , Krzysztof Wegner, „[FTV AHG] Technical Description of Poznan University of Technology proposal for Call for Evidence on Free-Viewpoint Television”, ISO/IEC JTC1/SC29/WG11, MPEG2016/m37893, San Diego, US, February 2016.

[3] Marek Domański, Adrian Dziembowski, Adam Grzelka, Łukasz Kowalski, Dawid Mieloch, Jarosław Samelak, Olgierd Stankiewicz, Jakub Stankowski, Krzysztof 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, Geneve, CH, May 2016.

[4] Lu Yu, Qing Wang, Ang Lu, Yule Sun, “Response to Call for Evidence on Free-Viewpoint Television: Zhejiang University”, ISO/IEC JTC1/SC29/WG11, MPEG2016/m37608, San Diego, US, February 2016.

[5] Ang Lu, Yule Sun, Lu Yu, "Extended results of ZJU’s CfE on Free-Viewpoint Television" ISO/IEC JTC1/SC29/WG11 MPEG2016/M38549, Geneva, Switzerland, May 2016.

[6] Takanori Senoh, Akio Ishikawa, Makoto Okui, Kenji Yamamoto, Naomi Inoue, “FTV AHG: CfE Response fDEVS” , ISO/IEC JTC1/SC29/WG11, MPEG2016/ m37953, San Diego, US, February 2016.

[7] Takanori Senoh, Kenji Yamamoto, Nobuji Tetsutani, Hiroshi Yasuda, "FTV AHG: Flowers results with CfE Response fDEVS".

[8] Lode Jorissen, Patrik Goorts, Gauthier Lafruit, “[FTV AhG] Soccer Light Field Interpolation Applied on Compressed Data”, ISO/IEC JTC1/SC29/WG11, MPEG2016/m37674, San Diego, US, February 2016.

[9] G. Bjøntegaard, “Calculation of Average PSNR Differences between RD-curves”, ITU-T SG16, Doc. VCEG-M33, April 2001.

[10] [ITU-T Rec. P.910 : Subjective video quality assessment methods for multimedia applications](https://www.itu.int/rec/T-REC-P.910/en), 2008.

[11] Vittorio Baroncini, Masayuki Tanimoto, Olgierd Stankiewicz, "Results of the Call for Evidence on Free-Viewpoint Television: Super-Multiview and Free Navigation" ISO/IEC JTC1/SC29/WG11 MPEG2016/N16128, San Diego, US, February 2016.