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/ m37893
February 2016, San Diego, US

Source Poznan University of Technology
Status Contribution to the 114th MPEG Meeting
Title [FTV AHG] Technical Description of Poznan University of Technology proposal for Call for Evidence on Free-Viewpoint Television

Author Marek Domański (domanski@et.put.poznan.pl),
Adrian Dziembowski (adziembowski@multimedia.edu.pl)
Adam Grzelka (agrzelka@multimedia.edu.pl),
Łukasz Kowalski (lkowalski@multimedia.edu.pl),
Dawid Mieloch(dmieloch@multimedia.edu.pl),
Jarosław Samelak (jsamelak@multimedia.edu.pl),
Olgierd Stankiewicz (ostank@multimedia.edu.pl),
Jakub Stankowski (jstankowski@multimedia.edu.pl),
Krzysztof Wegner (kwegner@multimedia.edu.pl).

1. Introduction

This documents presents a technical description of compression technology prepared at Poznan University of Technology in response to “Call for Evidence on Free-Viewpoint Television: Super-Multiview and Free Navigation” [1].

The proposed technology is based on 3D-HEVC and is designed for coding of Free Navigation (FN) materials. The submitted results are for all four FN sequences.
For the view synthesis, a proprietary software was used instead of VSRS [2].

2. Coding technology

The proposed coding technology is based on 3D-HEVC technology. However, it is not compatible with 3D-HEVC standard, e.g. proposed bitstreams cannot be decoded with the use of 3D-HEVC decoder. The syntax changes are mostly at high level, but also there are some minor adjustments at low-level.

The most important improvement is related to supported arrangement of the cameras. In 3D-HEVC standard, a number of specific tools was introduced and others were simplified with explicit 1D parallel view arrangement assumption. Such simplification states that the views are linearly aligned. Therefore, disparity vectors are restricted only to horizontal direction, and they can be derived from depth data through simple linear equation. This simplifies various processes in both encoder and the decoder, e.g. view synthesis/depth-based prediction.

In our proposal, we allow coding of materials acquired with the use of non-linear camera arrangements (e.g. arc). For that, we have extended the coded for a generic case, thus removing restriction only for horizontal direction, i.e. applied general derivation of disparity vector based on depth data. In order to achieve that, we use the approach from [3] and we have modified several syntax elements and tools:

- Transmission of camera parameters in VPS. In our proposal full camera parameters are transmitted, 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. Because this is a preprocessing step, for our proposal we provide PSNR values which are measured against the color-corrected views.

For the implementation and experiments, we use HTM 13.0 software [5] for 3D-HEVC and MV-HEVC. 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 but View Synthesis Optimization (VSO) for Depth Coding switched off. Such fact degrades the performance of our proposal related to the anchors, where, in some cases (e.g. Soccer Linear) VSO was switched on.

3. View synthesis technology

For generation of virtual views we have used a proprietary View Synthesis software which in general is similar to VSRS [2]. The main differences are as follows:

- Multiple-view-based synthesis. In VSVS only 2 neighboring views are used to synthesize a given view. In our software, information from all available views is merged together i.e. to fill the disocclusions.
- Modified hole-filling scheme which fixes some issued related to “non-linear” occlusions (in arc camera arrangement occlusions are more complex than in 2D linear camera arrangement).
- Higher synthesis precision. In VSRS sub-pixel precision is very minimalistic and simple. In the proposal a more complex scheme is employed.
- Color correction. Mismatch of color of given object/pixel in distinct views is corrected. Because this is performed pixel-wise, it not only performs classical color correction, results from acquisition of the scene by uncalibrated cameras, but also reduces flickering resulting from non-lambertian reflections of light over the objects (e.g. specular reflections).

4. Submitted results

We submit full set of results for Free Navigation (FN) scenario for all four sequences:

- BigBuckBunny
- PoznanBlocks
- SoccerArc
- SoccerLinear

The sequences has been encoded at four rate points RP1 to RP4. Basing no reconstructions, and also the original sequences (RP0), virtual views has been generated in between the original views, according to the CfE.

The virtual views are available for making sweeps and viewing during the 114th MPEG meeting.

We also attach PSNR values of the submitted results in form of XLS file. The averaged results, presented as Bjøntegaard [6] deltas are presented in Table 1.

Table 1. Bjøntegaard deltas of the proposed bitstream versus the anchors, calculated basing on the reconstructed videos. Negative values depicts bitrate reductions for the same PSNR quality.

Sequence	Bjontegaard delta, versus anchors
BigBuckBunny	-5.21%
PoznanBlocks	-16.25%
SoccerArc	0.78%
SoccerLinear	24.02%
Average (w/o Soccer Linear)	**-6.89%**
Average (all)	**0.84%**

For interpretation of the results, few things has to be noticed. During coding of the sequences with the proposed coded, View Synthesis Optimization (VSO) for Depth Coding was switched off. Such fact degrades the performance of our proposal related to the anchor, where, in some cases (e.g. Soccer Linear) VSO was switched on. Therefore, we provide average Bjøntegaard rates with and without Soccer Linear sequence. As it can be seen in Table 1, if we compare sequences where VSO settings were equal (and thus the comparison is fair), our proposal allows for about 5.89% of bitrate reduction. If we add SoccerLinear sequence, which itself introduces loss of about 24%, then the average Bjøntegaard delta is 0.84%.

Moreover it can noticed that our proposal works the best for sequences with arc camera arrangement, like BigBuckBunny or PoznanBlocks, and it doesn't influence coding of sequences with near linear camera arrangement, like Soccer Arc.

5. Acknowledgement

This research project was supported by public funds in the course of DS "Działalność Statutowa" project.

6. 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] O. Stankiewicz, K. Wegner, M. Tanimoto, M. Domański, "Enhanced Depth Estimation Reference Software (VSRS) for Free-viewpoint Television," ISO/IEC JTC1/SC29/WG11, MPEG2013/M31520, Geneva, CH, Oct. 2013.

[3] J. Stankowski, Ł. Kowalski, J. Samelak, M. Domański, T. Grajek, K. Wegner „3D-HEVC extension for circular camera arrangements," 3DTV-CON, Lisbon 2015.

[4] M. Domanski, M. Bartkowiak, A. Dziembowski, T. Grajek, A. Grzelka, A. Luczak, D. Mieloch, Jaroslaw Samelak, Olgierd Stankiewicz, Jakub Stankowski Krzysztof Wegner, "New Results in Free-viewpoint Television Systems for Horizontal Virtual Navigation", submitted to International Conference on Multimedia and Expo 2016 (ICME 2016), 2016.

[5] 3D HEVC reference codec online
https://hevc.hhi.fraunhofer.de/svn/svn_3DVCSoftware/tags/HTM-13.0

[6] G. Bjøntegaard, "Calculation of average PSNR differences between RD-curves," VCEG M33, Austin, Apr. 2001.