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

ISO/IEC JTC1/SC29/WG11
MPEG2019/m46343
January 2019, Marrakech, Morocco

Source **Poznań University of Technology, Poznań, Poland**

Status **Information**

Title **Unified Screen Content and Multiview Video Coding - Experimental results**

Author **Marek Domański,**

Contact marek.domanski@put.poznan.pl,

Abstract

The HEVC standard provides separate screen-content and multiview profiles. Application of the multiview coding technology provides some gains over simulcast video coding of multiple views. Unfortunately, the multiview video coding technology was adopted in the limited number of applications only. On the other hand, the frame compatible approach to compress stereoscopic video was quite common recently. Moreover, the technology of Screen Content Coding seems to be successful in real-word applications. In this paper, we provide a modification HEVC Screen Content Coding that exhibits roughly the same coding efficiency as HEVC SCC and MV-HEVC for graphics and frame-compatible multiview content, respectively.

1 Introduction

The state-of-the-art multiview video coding technology is Multiview HEVC (MHEVC) [1,2,3]. Such codecs have multi-loop structure, thus producing multi-layer bitstreams. Their architecture is more complex than that for single-view HEVC codecs. The multiview video MV-HEVC codecs outperform simulcast coding, especially for rectified multiview video. In extreme cases, the bitrate reduction may reach 50% with respect to the simulcast, but usually it is about 15-30%, even for parallel optical axes and the dense camera locations.

Unfortunately, practical adoption of the multiview video codecs is still quite limited. One of the reasons may be their specific, more complex architecture. On the other hand, the frame compatible approach to compress stereoscopic video was quite common recently. For the frame-compatible stereoscopic video coding, in this approach, the left and right views are merged into a single frame, usually after decimation. Then the new video is compressed using standard single-view encoders. Such an approach is dominant for stereoscopic television transmission.

The technology of Screen Content Coding [1,4] has been adopted and standardized recently, and this technology is considered to have significant application potential in the near future. Screen Content Coding is a set of additional coding tools that extend the HEVC while its basic single-loop encoder/decoder architecture remain unchanged. Application of these tools does not increase the number of layers in the encoded bitstreams. Therefore, the implementations of Screen Content Coding are relatively simple. Among the others, Screen Content Coding provides a tool of Intra Block Copy [1,4,5] that allows efficient intra-frame prediction when repetitive patterns exist in a single frame.

As already mentioned, merging of views into one frame is a well-known approach to compression of multiview video that is widely used especially for stereoscopic video. Such video may be compressed using standard single-view techniques. The video frames that comprise several views, contain also a repetitive pattern that may be efficiently predicted using the Intra Block Copy tool. That approach was already successfully used to compress multiview video [6,8]. Nevertheless, in the existing standard, the translation vectors have their values limited to integer numbers. Such limitation reduces the efficiency of the intraframe prediction, therefore the direct application of Screen Content Coding extension of HEVC, although quite efficient, is not as efficient as Multiview HEVC [6,8].

In this paper, we show that the Screen Content Coding technology may be easily adopted to efficient coding of frame-compatible multiview video. The major modification is the increase of the translation vector accuracy, from full-pel to quarter-pel, accompanied with some minor modifications mentioned further in the paper. We demonstrate that both intra- and interframe coding efficiency of this modified HEVC Screen Content codec is very similar to that of Multiview HEVC. In that way we propose Unified Screen Content and Multiview HEVC.

The general idea was already provided in our former contributions [6,7,8] but in this contribution the Screen Content Codec with quarter-pel vector accuracy is augmented the Unified Screen Content and Multiview Codec by some minor improvements that are described below.

2 Multiview coding using Intra Block Copy tool

Our idea of using the Intra Block Copy tool is very straightforward [7]. This idea was already presented but with experimental results for the intra mode only [At all time instants, we concatenate the views into one horizontal vector of views that is packed into one frame of a larger format. In that way we transform a mutiview video into a single-view video that can be compressed using standard HEVC encoders.

A good practice of multiview video coding is to start inter-view prediction from a central view (cf. Fig.1). Such approach usually provides better prediction than that starting from the leftmost or rightmost view.

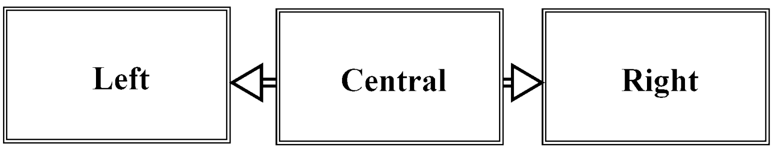

Figure 1: Multiview video coding: Typical directions of the inter-view prediction for 3-view video.

Application of Intra Block Copy is possible after merging all views into one compound frame. This frame may be a single tile, but each view may be assigned as a tile. In order to provide the same prediction scheme as for multiview coding, the central view should be the leftmost part of the compound frame (cf. Fig. 2).

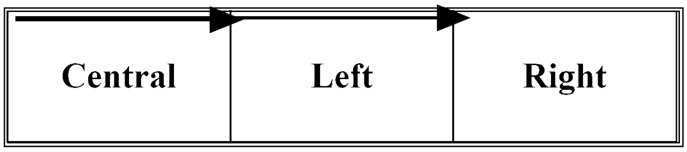

Figure 2: Intra Block Copy: Merging 3 frames into one compound frame with 3 tiles. The arrows denote the inter-tile prediction and the respective translation vectors.

The respective translation vector search starts always with the translation vector corresponding to the inter-view disparity vector for co-located blocks in the neighboring view. The translation vector estimation may be implemented in a way similar to motion vector search, or to disparity vector search. Also, we assume that the translation vectors may exhibit values being multiples of ¼, i.e. they exhibit the quarter-pel accuracy. We change the single-layer bitstream syntax accordingly. This change is not critical as the proposed translation vector format is the same as that for motion vectors used in the temporal interframe prediction.

3. Unified Screen Content and Multiview Codec (USCMC)

3.1. Tile encoding

In case of multiview frame-compatible video, the encoder starts with compressing first rows of each view, then the second rows etc. When applying Intra Block Copy, the area that can be used for matching the most similar blocks of points is limited to the previously compressed part of the slice. This means that the IBC cannot use as a reference blocks of points from different view, located lower than the currently analysed block.

In USCMC, the compression is configured in tiles [1]. The slice is divided into three tiles, each of which contains the content from a single camera (one view). Now, the leftmost tile is entirely compressed before the encoder starts to analyse the remaining tiles. This way, whole leftmost tile is available as a reference for Intra Block Copy applied to the remaining tiles. Such an approach reflects the coding order in Multiview HEVC, in which the middle view is compressed as first and then it becomes a reference for compression of the remaining views.

3.2 Intra Block Copy vectors precision

The output of the tool is a vector that points from the currently analysed block of points to the most similar block within the same picture. In USCMC, Intra Block Copy vectors have full-pel precision. The authors introduced to IBC a quarter-pel precision to follow the precision of disparity vectors in Multiview HEVC.

3.3 Starting point for block matching in Intra Block Copy

The goal of using Intra Block Copy, frame-compatibility and tile encoding (described above) is to match blocks of points from one view with corresponding blocks of points from another view. The resulting vector is expected to be very long, since it will point to a different tile. Obviously, long vectors are less efficient to compress than short ones. Additionally, it will take a lot of time for the Intra Block Copy to find the optimal match because it starts the search from the area nearby the analysed block [5].

In USCMC, the starting point for the IBC is the position of the collocated block in the leftmost tile, which reduces the length of the resulting vector, as well as the time to find it. For camera-captured content, IBC search algorithm is replaced by the technique used in Motion Compensated Prediction.

3.4 SAO and deblocking filtering per tile

In authors' proposal, the deblocking and SAO filters would be applied after compression of each three views, because they compose a single slice. Therefore, the leftmost tile would suffer from the encoding artefacts at the time it was used as a reference for compression of the remaining tiles, resulting in lower compression efficiency. Because of that, the proposed encoder applies SAO and deblocking filter after compression of each tile. Additionally, filtering across the tile boundaries was disabled, because the sharp edge between views is intended and should be preserved.

3.5 Different Quantization Parameter for side views

The USCMC codec was equipped with the possibility to define different Quantization Parameter for each tile. The information about applying different QP to the side tiles was included in the bitstream within the VPS extension. It is encoded as difference between original QP (applied to the leftmost tile), and the desired QP for the remaining tiles.

3.6 Reference tile border extension

In the proposed solution, prediction of blocks close to the right border of a tile may be inefficient because of the border between the reference (leftmost) tile and its neighbour. In order to avoid this issue, the reconstruction of the reference tile is separated from the whole image and its borders are extended.

3.7 Screen Content Coding tools configuration

The evaluation of the SCC tools in camera-captured video compression was already presented in [7,8]. As a result of this work, the authors made following changes in the Screen Content Coding configuration valid for Multiview content :

- enabled Intra Boundary Filter,
- disabled Hash-Based Motion Estimation,
- disabled Palette Mode,
- disabled Colour Transform.

4 Conditions of the experiments

All codecs are based on the same version of HEVC (HM-16.9 [9]), therefore the results are not influenced by any differences other than implemented improvements, Screen Content Coding or Multiview extension. The encoders were configured with respect to the appropriate Common Test Conditions [12, 13, 14], with some changes in the MV-HEVC configuration, in which the vertical disparity search range was set to 64 and Early Skip Detection was turned on to make it consistent with HEVC-SCC configuration. Obviously, the configuration of improved HEVC-SCC was also modified according to the improvements *A-C*, proposed in Section III.

Table 1. The encoders and corresponding software.

Encoder	Software
MV-HEVC	HTM-16.2 [HTM]
HEVC-SCC (frame-compatible)	HM-16.9+SCM-8.0 [HM-SCC]
USCMC - improved HEVC-SCC (frame-compatible)	HM-16.9+SCM-8.0 + authors' improvements

The experiments were conducted in two different coding scenarios: All Intra (only intra-frame and inter-view prediction allowed) and Random Access (inter-frame prediction allowed as well, intra period equal to 24). Each time, both compression efficiency and encoding time were measured. All experiments were performed on a PC with Intel Xeon 3GHz CPU.

5 Experimental results for multiview video coding

The goal of the experiment is to compare the coding efficiency of:

- single-view HEVC augmented by Intra Block Copy with quarter-pel accuracy of the translation vectors,
- multiview HEVC.

For the sake of simplicity, the experiment is limited to intraframe coding only. This is because, in muliview video coding, the inter-frame prediction is beneficial at most for the intraframe coding

In order to provide fair comparisons, the corresponding versions of HEVC with Screen Content Coding (SCC) and MHEVC software have been used, i.e. HM-16.9 + SCM 8.0 and HTM 16.2, respectively. Please note that HTM 16.2 software is developed on top of HM 16.9 software.

The experimental conditions and codec configurations were set according to the respective Common Test Conditions documents [12-14]. The basic difference was that the QP parameters were set equal for all views in both codecs.

The software of HM-16.9 + SCM 8.0 was augmented by the quarter-pel accuracy of the translation vectors. For experiments, Intra Boundary Filter was enabled, while Hash-Based IBC Search and Palette Mode were disabled.

The results are included in Tables 2 and 3. The respective BD-rates [18] are given in these tables.

Table 2. Experimental results against MV-HEVC – All Intra

Sequence	Bitrate increase [%]		Encoding time increase [%]	
	HEVC-SCC	USCMC	HEVC-SCC	USCMC
Poznan Hall 2	27.36	-1.67	+32	-5
Poznan Street	29.05	-0.55	+67	-12
Kendo	22.34	0.06	+11	-4
Balloons	18.27	0.00	+24	-1
Newspaper	18.56	-0.35	+37	-3
Average	23.12	-0.50	+34	-5

Table 3. Experimental results against MV-HEVC – Random Access

Sequence	Bitrate increase [%]		Encoding time increase [%]	
	HEVC-SCC	USCMC	HEVC-SCC	USCMC
Poznan Hall 2	25.43	-0.63	+80	+0
Poznan Street	30.82	0.51	+99	+2
Kendo	18.39	0.37	+56	+1
Balloons	16.19	0.33	+92	-1
Newspaper	17.74	0.46	+94	+0
Average	21.72	0.21	+84	+0

The experiment demonstrated that the USCMC provides virtually the same bitrates that Multiview HEVC for the same quality of compressed multiview video. This result has been obtained for QP-parameter values equal for all views. Adjusting of QP for individual views, as it was common in multiview coding experiments [14], may slightly modify this result.

6 Experimental results for screen content coding

The goal of the experiment is to prove that HEVC SCC and USCMC provide roughly the same coding efficiency for computer-generated graphical content, see Tables 4 and 5 - the results are calculated USCMC with the reference to HEVC SCC.

Table 4. All Intra (only intra-frame prediction) – results for screen content.

Sequence	Bitrate increase [%]	Encoding time increase [%]
ChinaSpeed	0.57	+13
SlideShow	0.67	+24

Table 4. Random Access (only intra-frame prediction) – results for screen content.

Sequence	Bitrate increase [%]	Encoding time increase [%]
ChinaSpeed	0.33	+12
SlideShow	0.43	+11

4 Conclusions regarding future generations of video codecs

For the forthcoming standardization of a new generation of video codecs, we propose to have one Unified Screen Content and Multiview Codec instead of independent Screen Content and Multiview profiles.

Acknowledgement

The work has been supported by the public funds under the DS project from Poznań University of Technology.

REFERENCES

[1] "High Efficiency Video Coding", ITU-T Rec. H.265, 4th Ed., December 2016 and ISO/IEC IS 23008-2, 3rd Ed., October 2017.

[2] G. J. Sullivan, J. M. Boyce, Y. Chen, J. R. Ohm, C. A. Segall and A. Vetro, "Standardized Extensions of High Efficiency Video Coding (HEVC)," in IEEE *Journal of Selected Topics in Signal Processing*, vol. 7, no. 6, pp. 1001-1016, Dec. 2013.

[3] G. Tech, Y. Chen, K. Müller, J. R. Ohm, A. Vetro and Y. K. Wang, "Overview of the Multiview and 3D Extensions of High Efficiency Video Coding," in *IEEE Transactions on Circuits and Systems for Video Technology*, vol. 26, no. 1, pp. 35-49, Jan. 2016.

[4] Xu, R. Joshi, and R. A. Cohen, "Overview of the Emerging HEVC Screen Content Coding Extension," in *IEEE Transactions on Circuits and Systems for Video Technology*, vol. 26, no. 1, pp. 50-62, Jan. 2016.

[5] X. Xu *et al.*, "Intra Block Copy in HEVC Screen Content Coding Extensions," in *IEEE Journal on Emerging and Selected Topics in Circuits and Systems*, vol. 6, no. 4, pp. 409-419, Dec. 2016.

[6] J. Samelak, J. Stankowski, M. Domański, "Experimental results for frame-compatible multiview video coding using HEVC SCC," *JCT-VC of ITU-T SG16 WP3 and ISO/IEC JTC1/SC29/WG11 26th Meeting: Doc. JCTVC-Z0041*, Geneva, CH, Jan. 2017.

[7] J. Samelak, O. Stankiewicz, M. Domański, *"Do we need multiview profiles for future video coding generations ?," JCT-VC of ITU-T SG16 WP3 and ISO/IEC JTC1/SC29/WG11 26th Meeting: Doc. JVET-H0044*, Macao, CN, Oct. 2017.

[8] J. Samelak, J. Stankowski, M. Domański, "Efficient frame-compatible stereoscopic video coding using HEVC Screen Content Coding", IEEE International Conference on Systems, Signals and Image Processing IWSSIP 2017, Poznań, Poland, 22-24 May 2017, IEEExplore.

[9] JCT-VC, HEVC reference software repository, *https://hevc.hhi.fraunhofer.de/svn/svn_HEVCSoftware/tags/ HM-16.9.* Web. 15 Dec. 2016.

[10] JCT-VC, HEVC Screen Content Coding reference software repository, *https://hevc.hhi.fraunhofer.de/svn/svn_HEVCSoftw are/tags/HM-16.9+SCM-8.0.* Web. 15 Dec. 2016.

[11] JCT-3V, Multiview HEVC reference software repository, *https://hevc.hhi.fraunhofer.de/svn/svn_3DVCSoftware/tags/ HTM-16.2.* Web. 15 Dec. 2016.

[12] F. Bossen, "Common Test Conditions and software reference configurations," *JCT-VC of ITU-T SG16 WP3 and ISO/IEC JTC1/SC29/WG11 12th Meeting: Doc: JCTVC-L1100*, Geneva, CH, Jan. 2013.

[13] H. Yu, R. Cohen, K. Rapaka, J. Xu, "Common Test Conditions for Screen Content Coding," *JCT-VC of ITU-T SG16 WP3 and ISO/IEC JTC1/SC29/WG11 21st Meeting: Doc. JCTVC-U1015r2*, Warsaw, PL, Jun. 2015.

[14] K. Müller, A. Vetro, "Common Test Conditions of 3DV Core Experiments," *JCT-3V of ITU-T SG16 WP3 and ISO/IEC JTC1/SC29/WG11 7th Meeting: Doc. JCT3V-G1100*, San José, US, Jan. 2014.
[15] M. Tanimoto, T. Fujii, N. Fukushima, "1D parallel test sequences for MPEG-FTV," *ISO/IEC JTC1/SC29/WG11, MPEG Doc. M15378*, Archamps, France, Apr. 2008.
[16] Y.S. Ho, E.K. Lee, C. Lee, "Multiview video test sequence and camera parameters," *ISO/IEC JTC1/SC29/ WG11 MPEG Doc. M15419*, Archamps, France, Apr. 2008.
[17] M. Domański, T. Grajek, K. Klimaszewski, M. Kurc, O. Stankiewicz, J. Stankowski, K. Wegner, "Poznań multiview video test sequences and camera parameters," *ISO/IEC JTC1/SC29/WG11 MPEG Doc. M17050*, Xian, China, Oct. 2009.
[18] G. Bjontegaard, "Calculation of Average PSNR Differences between RD-curves", *ITU-T SG16, Doc. VCEG-M33*, Austin, USA, Apr. 2001.