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

ISO/IEC JTC1/SC29/WG11
MPEG2017/M41699
October 2017, Macau, China

Source	**Poznań University of Technology**
Status	**Input**
Title	**Do we need multiview profiles for future video coding generations ?**
Author	**Jarosław Samelak, Olgierd Stankiewicz, Marek Domański**

Abstract

The existing video coding standards, like MPEG-2, AVC, HEVC, provide multiview profiles. It was demonstrated that application of the multiview coding technology provides some gains over simulcast video coding of multiple views. Neverthelss, the multiview video coding technology was adopted by industry in the limited number of applications only. On the other hand, the frame compatible approach to compress stereoscopic video was quite common recently. Moreover, a new technology of Screen Content Coding has been adopted recently, and this technology seems to be successful in real-word applications. Screen Content Coding provides a tool of Intra Block Copy. In this paper, we show that this tool may be efficiently used for multiview video coding, when the quarter-pel accuracy of vectors is applied. We demonstrate that intraframe coding efficiency of a standard intraframe HEVC codec augmented with the Intra Block Copy tool is similar to that of Multiview HEVC under the assumption that the accuracy of vectors is the same for both codecs. The HEVC codec augmented with the Intra Block Copy tool has the simple single-loop architecture that is compliant with that of the widely used single-view codecs. Therefore, we ask the question, if we need multiview profiles in the future video codec generation.

1 Introduction

The existing video coding standards, like MPEG-2, AVC [4], HEVC [1], provide multiview profiles that standardize compression technologies for multiview video captured using many cameras located on a line and having their optical axes co-planar and parallel.

The state-of-the-art multiview video coding technology is Multiview HEVC (MHEVC) [1,2,3]. Such codecs, similarly as the AVC-based multiview codecs [4], relay on multi-loop structure, and they produce multi-layer bitsreams. Therefore, the codec architecture is more complex than that for classic single-view video codecs. For the multiview video codecs, their compression performance outperforms the performance of 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 if the real camera locations are relatively close to each other and the optical axes of the real cameras are parallel.

Unfortunately, the proliferation 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.

Moreover, a new technology of Screen Content Coding [1,5] 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 [1,5,6] 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 [7,8]. 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 [7,8].

In this paper, we show that the Intra Block Copy tool may be easily adopted to efficient coding of multiview video. The only significant modification is the increase of the translation vector accuracy, from full-pel to quarter-pel. We demonstrate that intraframe coding efficiency of a standard intraframe HEVC codec augmented with the Intra Block Copy tool is similar to that of Multiview HEVC under the assumption that the accuracy of vectors is the same for both codecs. The HEVC codec augmented with the Intra Block Copy tool has the simple single-loop architecture that is compliant with that of the widely used single-view codecs. Therefore, we ask the question, if we need multiview profiles in the future video codec generation.

2 Multiview coding using Intra Block Copy tool

Our idea of using the Intra Block Copy tool is the following. At all time instants, we merge all views into one frame, i.e. 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.

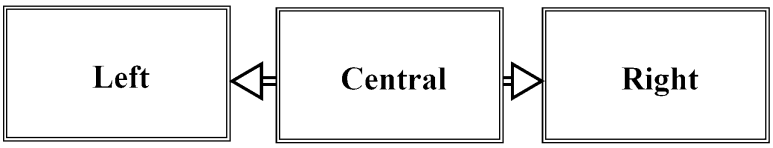

Figure 1: Multiview video coding: Typical directions of the inter-view prediction for 3-view video.

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).

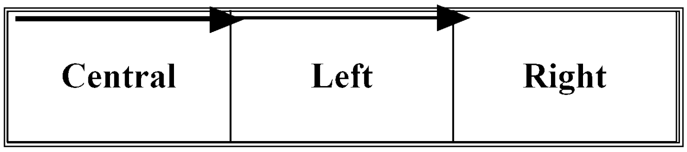

Figure 2: Intra Block Copy: Merging 3 frames into one compound frame with 3 tiles. The arrows denote the inter-tile prediction and the respective translation vectors.

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 Experimental results for intra-frame 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 [9, 10] and HTM 16.2 [11], 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 experiment is aimed at natural multiview video content, i.e. not computer-generated. Therefore the standard multiview test video sequences have been used as described in the Common Test Conditions for multiview video coding [14]:

- Poznan Hall2 [17],

- Poznan Street [17],
- Kendo [15],
- Baloons [15],
- Newspaper [16].

The results are included in Table 1. The respective BD-rates [18] are given in Table 2.

Table 1. Coding performance comparison: Multiview HEVC versus HEVC-SCC with quarter-pel translation vector accuracy.

		MV-HEVC (HTM-16.2)				Improved HEVC-SCC (HM-16.9)			
Seq	QP	Bitrate [kbps]	PSNR Y	PSNR Cr	PSNR Cb	Bitrate [kbps]	PSNR Y	PSNR Cr	PSNR Cb
Poznan_Hall2	25	15435.71	42.88	48.29	47.87	13799.27	42.74	48.16	47.74
	30	6480.49	41.63	47.18	46.81	5930.52	41.51	47.05	46.67
	35	3206.01	40.17	45.57	45.27	2958.95	40.03	45.44	45.11
	40	1734.47	38.34	44.33	44.07	1600.99	38.19	44.21	43.91
Poznan_Street	25	55164.38	41.41	46.92	46.14	51197.80	41.12	46.77	45.97
	30	22952.34	38.52	45.25	44.42	20542.02	38.24	45.10	44.24
	35	10117.16	36.05	43.32	42.50	9025.41	35.81	43.18	42.32
	40	5007.75	33.67	42.03	41.25	4446.89	33.45	41.88	41.07
Kendo	25	8381.30	44.93	45.17	45.28	7844.41	44.76	45.06	45.11
	30	4632.23	42.92	44.21	43.64	4341.71	42.71	44.08	43.45
	35	2610.01	40.48	42.95	41.67	2448.91	40.26	42.83	41.51
	40	1544.45	37.73	41.98	40.18	1448.44	37.51	41.86	40.03
Balloons	25	12482.84	44.15	43.69	44.09	11727.96	43.96	43.54	43.90
	30	7062.19	41.92	42.21	42.22	6654.86	41.70	42.05	42.03
	35	4050.80	39.22	40.54	40.21	3818.11	38.99	40.39	40.03
	40	2395.88	36.16	39.34	38.77	2244.07	35.91	39.19	38.62
Newspaper_CC	25	21258.71	42.01	44.01	43.98	19801.41	41.74	43.81	43.78
	30	10802.27	39.20	42.17	42.10	9956.00	38.93	41.98	41.90
	35	5605.80	36.45	40.28	40.13	5166.59	36.22	40.09	39.92
	40	3069.35	33.73	38.95	38.72	2832.75	33.51	38.78	38.54

Table 2. BD-rates for the comparison between HEVC-SCC with quarter-pel translation vector accuracy and Multiview HEVC. A negative number means that, for the same value of luma PSNR, the modified HEVC-SCC provides lower bitrate.

Sequence	BD-rate
Poznan_Hall2	-2.32%
Poznan_Street	-2.68%
Kendo	-1.56%
Balloons	-1.42%
Newspaper_CC	-2.15%

The experiment demonstrated that the HEVC-SCC with quarter-pel translation vector accuracy provides slightly lower bitrates than 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.

4 Conclusions regarding future generations of video codecs

The abovementioned experiment was performed for intraframe coding that mostly benefits from inter-view predicition. The results demonstrate that the efficiency of the inter-view prediction is virtually the same for Multiview HEVC and for HEVC augmented by Intra Block Copy tool using the same resolution of translation/displacement vectors. It is worth to add that the latter codec has simpler single-loop structure and is nearly compliant with standard HEVC Screen Content Codec. Some other similar results are included in another paper of the authors (to be published).

The result was obtained for rectified multiview video clips acquired using cameras with parallel optical axes, i.e. for the application scenario, for which Multiview HEVC was designed. Therefore, a question arises: Do we need to develop multiview video codecs for future generations of video compression techniques? The results of the abovementioned experimental suggest the negative answer. Instead of that it is worth to have the Intra Block Copy tool with the vectors having the same resolution as for motion compensated prediction.

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] "Coding of audio-visual objects, Part 10: Advanced Video Coding", ISO/IEC Int. Standard 14496-10: 2014 and "Advanced video coding for generic audiovisual services", ITU-T Rec. H.264 (V9) (2014).

[5] 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.

[6] 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.

[7] 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.

[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.