Efficient frame-compatible stereoscopic video coding using HEVC Screen Content Coding

Jarosław Samelak, Jakub Stankowski, Marek Domański
Poznań University of Technology, Chair of Multimedia Telecommunications and Microelectronics, Poland
{jsamelak, jstankowski}@multimedia.edu.pl, marek.domanski@ put.poznan.pl

Abstract - **The paper presents application of the emerging HEVC Screen Content Coding for frame-compatible compression of stereoscopic video. Such a solution may be an alternative to the Multiview HEVC, which is the state-of-the-art dedicated technique for multiview video compression. The paper provides an extensive description of main differences between both compression techniques. Authors also present adaptation of the Screen Content Coding to compress stereoscopic video as fast and efficiently as possible. The paper reports experimental results of the comparison between HEVC Screen Content Coding and Main profiles for frame-compatible compression of stereoscopic video. The advantages and disadvantages of the proposed technique are enumerated in the conclusions.**

Keywords - HEVC, Screen Content Coding, Multiview, stereovision, video compression

I. INTRODUCTION

Screen Content Coding (SCC) [1] is a new extension of High Efficiency Video Coding (HEVC) [2,7] that has been developed for efficient compression of video containing mostly computer-generated content, as for many applications, such as screen sharing, wireless and remote displays, remote control or cloud gaming. Efficient compression of such content is achieved using new coding tools, such as:

- Adaptive Motion Vector Resolution [1] – allows to control the resolution of motion vectors.
- Adaptive Color Transform [1] – allows to encode an image block in the RGB color space directly or with conversion to the YCbCr color space.
- Intra Block Copy (IBC) [3] – searches for similar block of points within a large part of the same picture.
- Palette Mode [4] – if the content is composed of limited number of colors, Palette Mode enumerates them and represents as indexes.

The abovementioned tools provide efficient compression of noise-free rendered content with repeated patterns, large single-colored areas and unnatural motion.

II. APPLICATION OF SCREEN CONTENT CODING TO STEREOSCOPIC VIDEO

Apart from Screen Content Coding, HEVC contains other extensions designed for different purposes [5]. One of them is Multiview extension [6], dedicated for compression of video acquired with a set of cameras. The main coding tool included in Multiview extension is the Disparity-Compensated Prediction (DCP), which exploits the similarities between encoded views to improve the overall compression capability. The experiments demonstrate that the inter-view DCP is mostly chosen as the optimum prediction mode for the video portions where intra-frame coding would be used otherwise.

Therefore, the authors of the paper suggest that the SCC Intra Block Copy tool can substitute the Disparity-Compensated Prediction if all the views from multiview video compose a single image, as presented in Fig. 1. In such a case, Screen Content Coding encoder would match the corresponding blocks of points in different views, as it is done by the DCP.

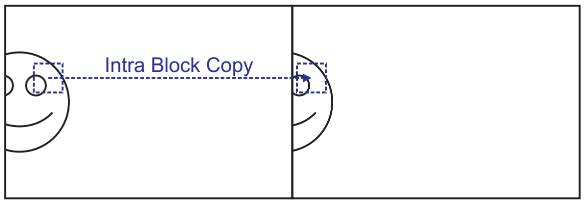

Figure 1. Simulating Disparity-Compensated Prediction using Intra Block Copy.

Unfortunately, the Intra Block Copy is significantly more constrained as compared to DCP. First of all, the block matching precision is limited to full-pel, in contrary to the quarter-pel precision in Disparity-Compensated Prediction. Moreover, the block sizes in the case of Intra Block Copy are restricted to intra prediction CU sizes, while in DCP the block matching can be performed for the size of any available Prediction Unit. The aforementioned differences result in Intra Block Copy being less efficient but faster than the Disparity-Compensated Prediction. The results of comparison between Multiview HEVC and Screen Content Coding were reported in [7]. This paper focuses on special type of multiview content – stereoscopic video.

A stereoscopic video is composed of two views, each of which is destined for one of spectator's eyes in order to create impression of depth. The stereoscopic video would be the most efficiently compressed with a dedicated multiview extension. In 3D digital television, stereoscopic video is usually transmitted in a frame-compatible format with both views decimated vertically or horizontally, as presented in Fig. 2. Then, such a video is compressed with main profile of the encoder with additional signalization in SEI (Supplemental Enhancement Information). Obviously, such a solution is not the most efficient one, but it allows to use standard infrastructure, and to avoid implementation of multiview codec on the device.

Figure 2. An exemplary stereoscopic image with views decimated horizontally.

Screen Content Coding has a chance to be much more popular than the Multiview extension, because it would be beneficial in many important technologies, such as cloud gaming, screen casting, desktop sharing or remote control. Therefore, with little effort, Screen Content Coding could be applied also for stereoscopic video compression, as presented in Fig. 3.

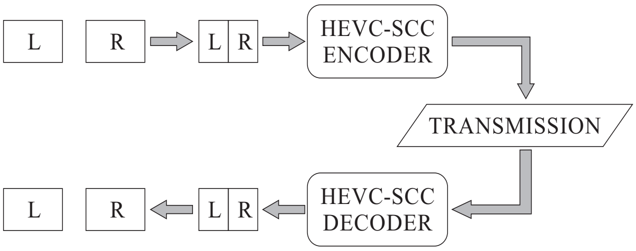

Figure 3. Proposal solution for an SCC-based stereoscopic video codec. L – left video, R – right video

In the first step, both views are decimated and merged into one frame-compatible sequence. Next, the prepared input sequence is compressed with Screen Content Coding. At the decoder side, the reconstructed video has to be split back into separate views and interpolated to achieve the original resolution of the video.

III. Experiment Conditions

Sections IV and V present a number of experiments conducted to verify the efficiency of the proposed technique. Two HEVC-based codecs were used in the experiments: HEVC Main and HEVC Screen Content Coding. Each codec was compiled from the appropriate reference software, as shown in Table I. The Screen Content Codec is implemented on the top of the same version of HEVC software (HM-16.9) as used for HEVC Main, therefore the results are not influenced by any differences other than tools included in Screen Content Coding.

TABLE I. Used encoders and corresponding software.

Encoder	Software
HEVC Main	HM-16.9 [8]
HEVC Screen Content Coding	HM-16.9+SCM-8.0 [9]

The tests were performed on 100 frames of 2 views obtained from 6 commonly used multiview sequences. Table II contains information about selected views and resolutions of each sequence, while Fig. 4 presents exemplary frames. The sequences were acquired with linear camera setup and rectified. Most of them (except for Fig. 4b) contain camera-captured content.

TABLE II. Numbers of used views and their resolutions

Sequence	Resolution	Views (left, right)
Balloons [12]	1024 × 768	3, 4
Big Buck Bunny Butterfly [13]	1280 × 768	45, 50
Kendo [12]	1024 × 768	3, 4
Newspaper [14]	1024 × 768	4, 6
Poznan_Hall2 [15]	1920 × 1088	6, 5
Poznan_Street [15]	1920 × 1088	4, 3

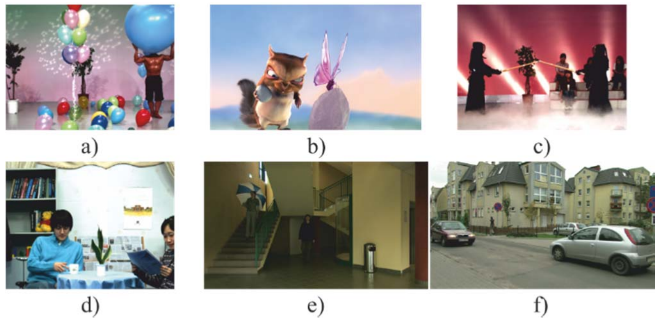

Figure 4. Test sequences: a) Balloons [12], b) Big Buck Bunny Butterfly [13], c) Kendo [12], d) Newspaper [14], e) Poznan_Hall2 [15], f) Poznan_Street [15].

The experiment in Section IV was conducted in the *Random Access* configuration with intra period equal to 24, while the experiment in Section V was additionally performed in the *All Intra* configuration. The encoders were set up with respect to Common Test Conditions [10][11], using the appropriate configuration files provided within the reference software.

The goal of each experiment was to measure the compression efficiency and the encoding time for the two variants of HEVC encoding. The efficiency was calculated using Bjøntegaard metric for luma PSNR [16]. For the encoding time, an average per cent time saving was calculated, related to the chosen reference encoder. To achieve a high degree of confidence for the time results, the experiments were repeated 10 times and the results were averaged.

All experiments were performed on a PC with Intel Xeon 3GHz CPU.

IV. Configuring Screen Content Coding

The Screen Content Coding profile was designed exclusively for video that consists mostly of the computer-generated content. In the proposed solution, authors apply it to natural video, captured by a pair of cameras. Therefore, some of the tools contained in the Screen Content Coding may be useless. This Section compares performance of these tools in terms of speed and efficiency. The goal of this experiment is to find the optimal configuration of the Screen Content Coding for compression of camera-captured stereoscopic video.

The authors have selected three Screen Content Coding tools that should be configured differently than in the Common Test Conditions [11]: Intra Boundary Filter (disabled by default), Hash-Based Motion Estimation [1], Palette Mode (enabled by default). Encoder benefits from using these tools only if the content is mostly rendered, therefore it may be profitable to toggle their usage in the case of natural video. In order to verify this statement, an experiment that compares the performance of Screen Content Coding with and without the proposed modifications was performed. The default configuration, proposed in Common Test Conditions, was used as a reference. Figs. 5 and 6 present the comparison in the bitrate and the encoding time.

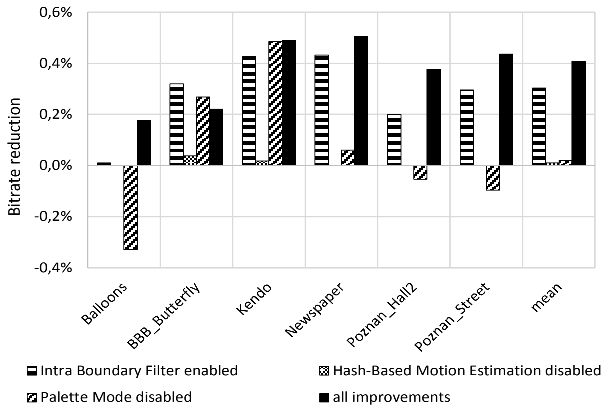

Figure 5. Bitrate reduction against default Screen Content Coding configuration.

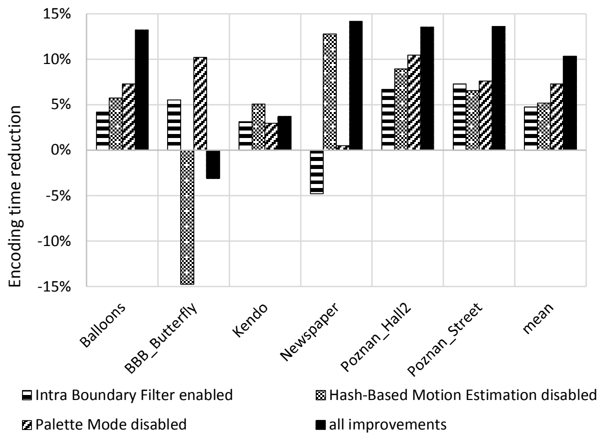

Figure 6. Encoding time reduction against default Screen Content Coding configuration.

As expected, disabling Hash-Based Motion Estimation and Palette Mode did not decrease the efficiency, which means these tools are not efficient for natural content. Enabling the Intra Boundary Filter slightly reduced the bitrate. Adopting SCC configuration to the camera-captured content allows to reduce the encoding time by about 10% (the average estimated using 6 test video sequences).

The proposed modifications of the Screen Content Coding turned out to be beneficial both for the encoding time and the compression efficiency. Therefore, for the remaining experiment, Intra Boundary Filter was enabled, while Hash-Based Motion Estimation and Palette Mode were disabled. The changes were made only in the configuration – the reference Screen Content Coding software remained unmodified.

V. Evaluation of the Proposed Solution

This Section provides experimental results of comparison between frame-compatible Screen Content Coding and HEVC Main profile for compression of stereoscopic video. The goal of the following tests is to determine whether the proposed solution provides better compression efficiency than the most common approach to stereoscopic video coding.

The tests were conducted in the *All Intra* and *Random Access* configurations, following test conditions described in Section III. For both configurations, following coding scenarios were performed:

- HEVC Main simulcast – both views encoded independently using HEVC Main profile. This scenario was used as a reference, because it is the simplest method of encoding stereoscopic video.
- HEVC Main Side-by-Side – both views joined together horizontally and encoded as a frame-compatible sequence using HEVC Main profile. This scenario was performed to verify if joining the views itself improves the compression efficiency.
- HEVC Screen Content Coding simulcast – each view encoded independently using Screen Content Coding. This scenario was performed to verify if using Screen Content Coding itself improves the compression efficiency.
- HEVC Screen Content Coding Side-by-Side – both views joined together horizontally and encoded as a frame-compatible sequence using Screen Content Coding. It is the proposed solution.

For the simulcast encoders, the input views were decimated horizontally, as it is done while creating frame-compatible stereoscopic video. Therefore, the amount of data to compress is the same for all encoders and the decimation does not influence the results.

Figs. 7-8 and Figs. 9-10 present experimental results for *All Intra* and *Random Access* configurations, respectively.

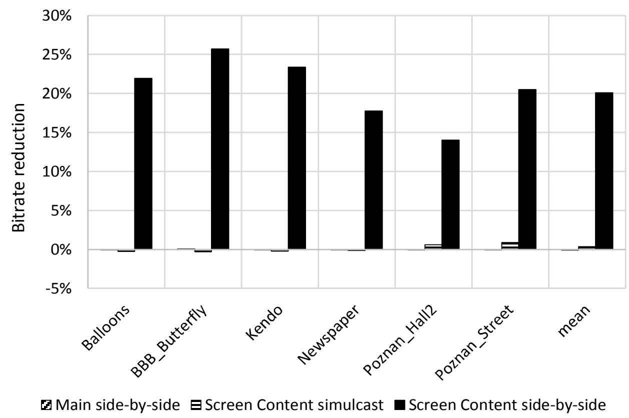

Figure 7. Bitrate reduction against Main simulcast – All Intra coding scenario.

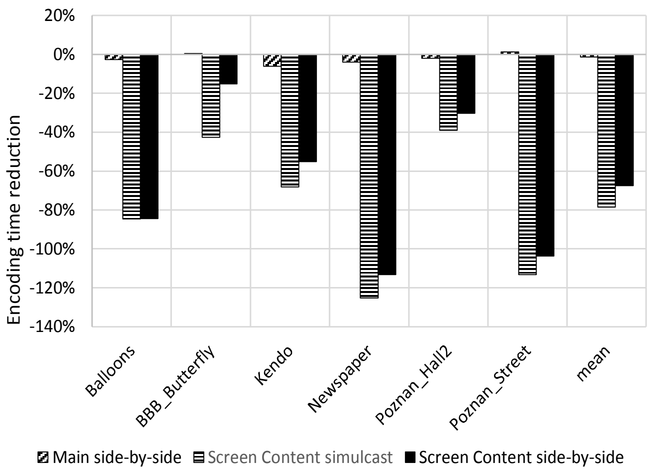

Figure 8. Encoding time reduction against Main simulcast – All Intra coding scenario.

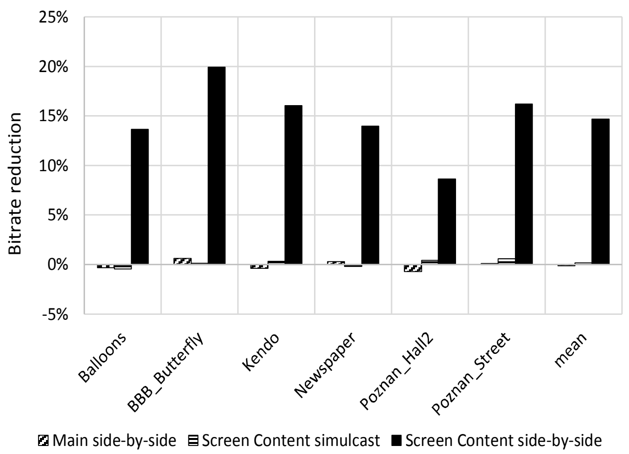

Figure 9. Bitrate reduction against Main simulcast – Random Access coding scenario.

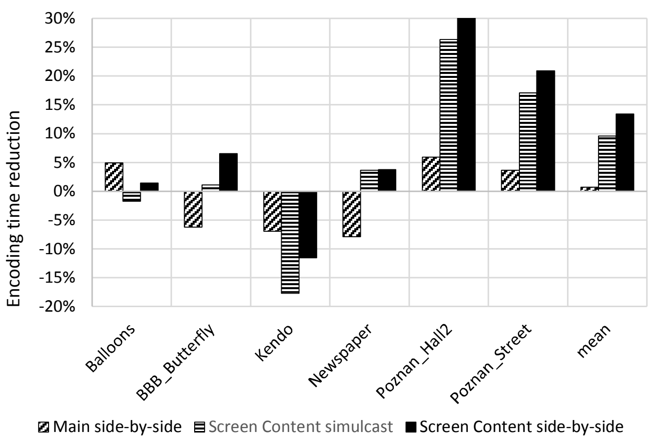

Figure 10. Encoding time reduction against Main simulcast – Random Access coding scenario.

Neither Main Side-by-Side nor Screen Content Coding simulcast improve the compression efficiency compared to Main simulcast. On the other hand the proposed solution, which is a combination of frame-compatibility and Screen Content Coding, significantly reduces the bitrate for both configurations. This proves that the Intra Block Copy tool utilizes the similarities between views that compose a single frame.

For *All Intra* configuration, Screen Content Coding reduces the bitrate on average by 20% at the cost of nearly 65% slower compression. In case of *Random Access*, both bitrate and time are reduced by roughly 15%. Thus, Screen Content Coding can be efficiently applied to compression of the frame-compatible stereoscopic video in Side-by-Side format.

VI. CONCLUSIONS

In the paper, the authors presented an idea of using Screen Content Coding for compression of frame-compatible stereoscopic video in Side-by-Side. This format was chosen as it is mostly used for stereoscopic television nowadays.

In the preliminary experiment, the configuration of the Screen Content Codec was optimized for compression of natural, camera-captured content. The proposed modifications allowed to reduce the encoding time by ~10%, as compared to the standard configuration recommended for computer-generated content.

Then, Screen Content Coding was compared to HEVC Main profile, which is the most popular solution for compression of the frame-compatible stereoscopic video. Obviously, the experiment demonstrated that the application of Main profile for the frame-compatible stereoscopic video provides no improvement over simulcast. On the other hand, Screen Content Coding with Intra Block Copy tool exploits the inter-view similarities, thus resulting in about 15% bitrate and encoding time reductions versus the abovementioned commonly used solution. Therefore, the devices supporting Screen Content Coding can be easily configured for compression of the natural, camera-captured stereoscopic video. Such functionality may be beneficial in many applications where stereoscopic-video transmission will be added to the HEVC-based television or the video-over-the-top systems.

Obviously, the application of Multiview HEVC provides even more efficient compression of stereoscopic video than the approaches studied in this paper. Nevertheless this state-of-the-art approach needs two-layer video transmission that is different from the single-layer video transmission currently used everywhere, and needs modified infrastructure. For the sake of brevity, the more in-depth comparison of applications Screen Content Coding and Multiview Video Coding to stereoscopic video coding must be left beyond the scope of the paper. These issues will be included into another work of the authors that will be published soon.

ACKNOWLEDGEMENT

Research project was supported by public funds under Project 08/84/DSPB/0190 Poznań University of Technology, Chair of Multimedia Telecommunications and Microelectronics.

REFERENCES

[1] J. 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.

[2] G. J. Sullivan, J. Ohm, W. J. Han, and T. Wiegand, "Overview of the High Efficiency Video Coding (HEVC) Standard," in *IEEE Transactions on Circuits Systems for Video Technology*, vol. 22, no. 12, pp. 1649-1668, Dec. 2012.

[3] 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.

[4] W. Pu *et al.*, "Palette Mode Coding in HEVC Screen Content Coding Extension," in *IEEE Journal on Emerging and Selected Topics in Circuits and Systems*, vol. 6, no. 4, pp. 420-432, Dec. 2016.

[5] 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.

[6] 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.

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

[9] 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.

[10] 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.

[11] 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.

[12] 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.

[13] P. T. Kovacs, "[FTV AHG] Big Buck Bunny light-field test sequences," *MPEG M35721*, Geneva, Feb. 2015.

[14] 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.

[15] 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.

[16] G. Bjøntegaard, "Calculation of Average PSNR Differences between RD-curves", *ITU-T SG16, Doc. VCEG-M33*, Austin, USA, Apr. 2001.