Coding of Multiple Video+Depth Using HEVC Technology and Reduced Representations of Side Views and Depth Maps

Marek Domański, Tomasz Grajek, Damian Karwowski, Jacek Konieczny, Maciej Kurc, Adam Łuczak, Robert Ratajczak, Jakub Siast, Olgierd Stankiewicz, Jakub Stankowski, Krzysztof Wegner
Poznań University of Technology, Chair of Multimedia Telecommunications and Microelectronics, Poznań, Poland

***Abstract*—During the last two decades, a new technology generation of video compression was introduced about each 9 years. Each new compression-technology generation provides halving of necessary bitrates as compared to the last previous generation. This increasing single-view compression performance is related to increasing compression performance of multiview video coding. For multiview video with associated depth maps, additional significant bitrate reduction may be achieved. The paper reports the original compression technology that was designed and developed at Poznań University of Technology in response to MPEG Call for Proposals on 3D Video Coding Technology. The main idea of this technique is to predict very efficiently the side views and the depth maps from the base view.**

Keywords- 3D video, compression, HEVC, AVC, H.264, MPEG, multiview video coding.

I. GENERATIONS OF VIDEO COMPRESSION TECHNOLOGY

During the last 20 years the video coding has evolved to the mature industrial technology with applications everywhere. In these two decades, three consecutive compression technology generations have been developed (Fig. 1). About 1994, the MPEG-2 video coding standard has been developed. This very successful standard is a representative of a whole generation of compression techniques that share a similar philosophy of algorithms and a similar level compression efficiency.

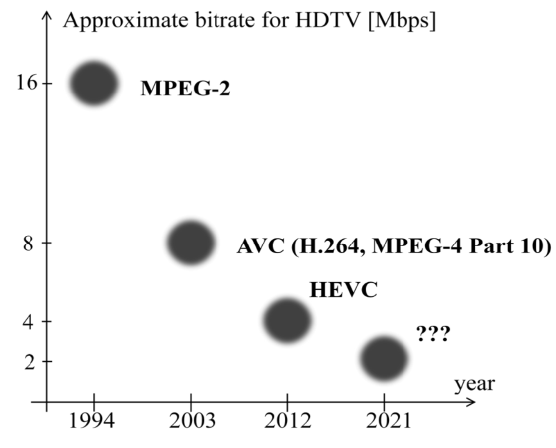


Fig.1. Generations of video compression technology.

Since then, two other technology generations of video coding have been developed. The Advanced Video Coding (AVC) [1] standard (ISO MPEG-4 Part 10 or ITU Rec. H.264) was developed about year 2003. It is the leading representative of the first of these two technology generations. For the latest generation, the representative is HEVC [2] (High Efficiency Video Coding) that is expected to became an official international standard soon.

When considering the three abovementioned representative video coding standards, some regularity is visible. For each next generation, for a given quality level, the bitrate is halved. The temporal intervals of about 9 years was between the consecutive technology generations of video coding. During such an interval the available computational power is increasing by a factor of about 20, according to the Moore law. After each such an interval, this increase may be consumed by the more sophisticated codecs of the next generation.

About each 9 years we have a new video compression generation that provides halved bitrates**.**

Of course, this “rule” was observed during two cycles only. It is probably too short time to establish a rule that may be used to forecast the future developments.

II. MULTIVIEW VIDEO CODING

For the multiview video coding, the respective techniques have been developed in all three abovementioned representative standards from the three compression generations. The multiview video coding techniques take advantage of inter-view predictions (Fig. 2) that is used in addition to the standard interframe prediction.

The best known technique is called just Multiview Video Coding (MVC), and is a part of the AVC (MPEG-4 AVC or H.264 Recommendation) standard. This technique provides bitrates reduced mostly by 15-25% as compared to independent compression of the views using AVC (called also as “simulcast AVC”).

On the other hand, the new generic video compression technology HEVC provides almost halved bitrates as compared to AVC. Therefore even independent compression of all views using HEVC (simulcast HEVC) appears to be significantly more efficient than the older but dedicated MVC technique [3].

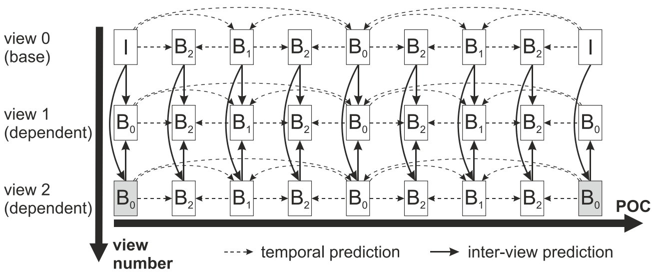


Fig.2. Interframe predictions by multiview video coding.

The new HEVC codec structure is quite similar to that of AVC. Therefore, the inter-view predictions can be easily adopted to HEVC [3-5]. We call the new codec Multiview-HEVC (MV-HEVC). For MV-HEVC, except of I-pictures in the base view, all other pictures may be B-pictures. In Fig. 2, the boxes marked in gray denote the B-pictures in MV-HEVC that would be P-pictures in MVC.

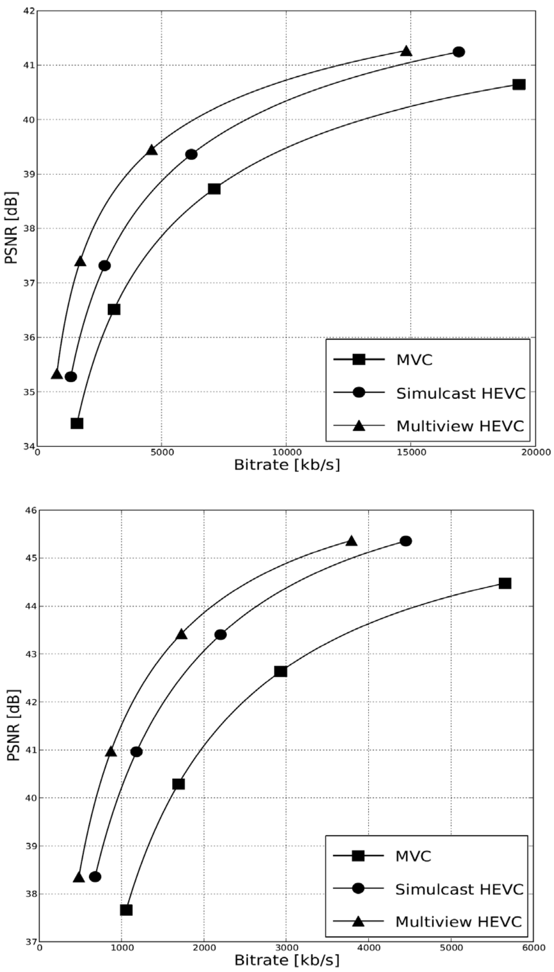


Fig. 3. Experimental R-D curves for the 3-view test sequences "Poznan_Street" (top) and "Kendo" (bottom).

For multiview video, the MV-HEVC codec provides the bitrates halved with respect to those in MVC (Fig. 3). This is roughly the same result as for standard HEVC and AVC used for single-view video.

III. 3D Video Coding Standardization

In 2011, in order to define the state-of-the-art 3D video coding technology, MPEG has announced Call for Proposals (CfP) on compression of the multiview video with associated depth maps (MVD). By the end of August 2011, in response to CfP, over 20 proposals have been submitted in the two categories: AVC- and HEVC-compatible. The proposals have been evaluated using subjective quality assessment of the decoded video clips. The proposal assessment was a huge 3D video subjective-quality evaluation experiment.

The results of subjective tests were disclosed during MPEG meeting in Geneva in the end of November 2011. In the HEVC-compatible class, the proposals from Fraunhofer Institute – H. Hertz Institute Berlin and from Poznań University of Technology were qualified as the best.

In this paper, we present the main idea of the original compression technology designed and developed at Poznań University of Technology in response to Call for Proposals on 3D Video Coding Technology [6].

IV. A new compression technique for 3D video

The new HEVC-compatible compression technology is intended to encode a limited number of video views together with the corresponding depth maps. For some applications, e.g. autostereoscopic displays, other views have to be synthesized (virtual views) in the receiver from the limited number of the decoded views and the corresponding depth maps (Fig. 4).

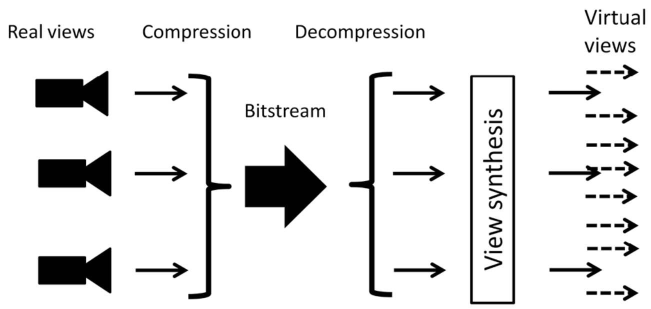


Fig. 4. 3D video compression scenario.

This technology is developed on the top of the MV-HEVC codec. In other words, the new 3D-video codec was built as an extension of the MV-HEVC codec. It comprises additional compression tools that exploit depth maps in order to efficiently encode the side views.

One of the views is coded in the HEVC syntax while in side views only the disoccluded regions are coded and transmitted, and the remaining parts are reconstructed at the receiver from the base view using view synthesis. The shapes of disoccluded regions are derived in the decoder from reconstructed depth maps. Therefore no side information on shape needs to be transmitted

In the proposal [6, 8] for 3D video compression technology, several additional compression tools have been used:

- Inter-view depth consistency refinement;
- Nonlinear depth representation;
- Depth-Based Motion Prediction (DBMP) –motion vectors are predicted from the central view by a 3D mapping [7].
- Depth-dependent adjustment of QP for texture ;
- Depth-Gradient-based Loopback Filter (DGLF) and Availability Deblocking Loopback Filter (ADLF) - additional in-loop filters that reduce the artifacts resulted from coding of disoccluded regions.

The base view is encoded using standard HEVC technology. The side views and their depth maps are efficiently predicted from the base view and its depth map. This prediction has been designed to be very efficient, therefore very small bitrates are allocated to the side views and their depth maps (Figs. 5 and 6).

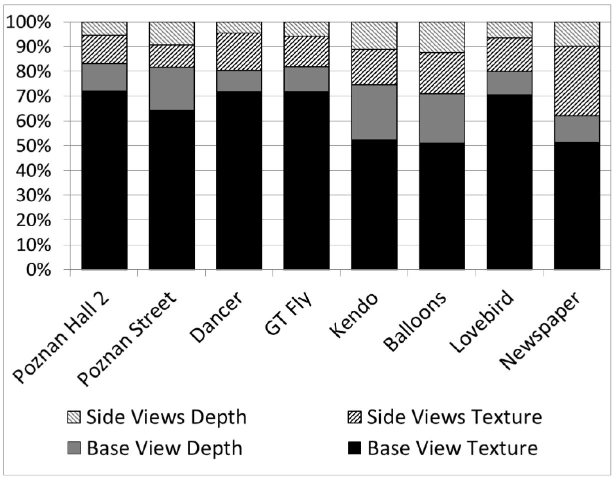


Fig. 5. Bitrate allocation for the 3-view test sequences at about 1.3 Mbps.

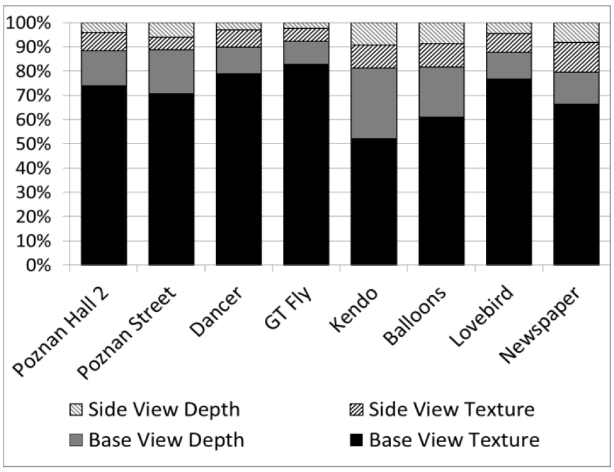


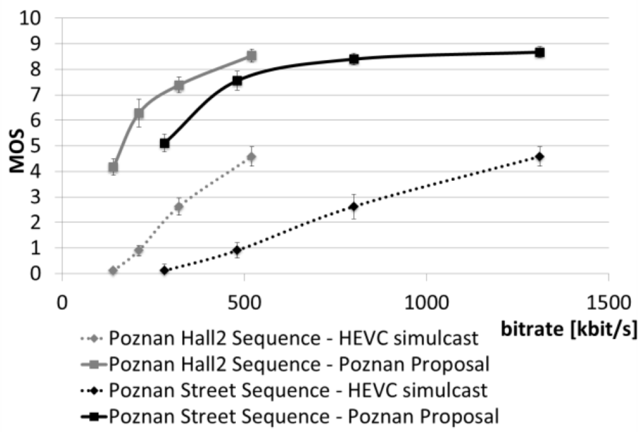
Fig. 6. Bitrate allocation for the 2-view test sequences at about 850 kbps.

The main idea of the proposed technology is to predict as much as possible from the base view. The inter-view prediction modes from MVC and MV-HEVC are used together view-synthesis based predictions that exploit 3D mapping for higher efficiency.

Most of the bitrate is spent on representation of the base view (50-80% of the total bitrate). The technology is capable for 3-view HD video compression at 6 Mbps.

V. COMPRESSION PERFORMANCE

During the proposal evaluation, the quality of the decoded video clips have been assessed in formal subjective testing by independent laboratories. For the selected bitrates, the quality of the decoded stereoscopic video was compared to the quality of the decoded video clips obtained from the simulcast HEVC codec. Subjective quality tests were performed using polarization stereoscopic displays as well as with the use of autostereoscopic displays. The Single Stimulus Impairment Scale (SSIS) test method was used with 11 quality levels. All tests were carried out with naive viewers. Mean Opinion Scores (MOS) and confidence intervals corresponding to 95% probability were computed. The results for some test 3D video clips are presented in Fig. 7.



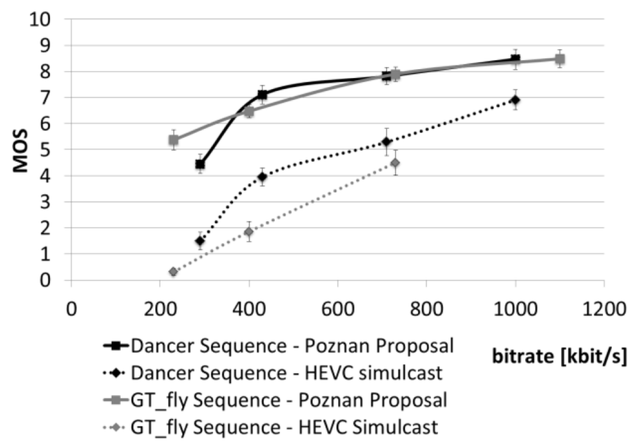


Fig. 7. The R-D curves for the Poznań 3D video codec obtained from the subjective quality tests.

VI. CONCLUSIONS

The original 3D video compression technology has been briefly described. The responses to MPEG Call for Proposals on 3D Video Coding Technology have proven, that this technology may be considered as the-state-of-the-art technology because all other submitted codecs proved either similar performance, or mostly worse performance. The proposed coding technology allows reduction of the bitrate of about 50-70% comparing to the HEVC simulcast. The proposed codec provides the bitrates that are about 20-50% lower than those obtainable by Multiview HEVC (MV-HEVC). Such substantial gain is achieved mostly by exploitation of the similarities between individual views and depth maps. The side views and the depth maps aree represented by very low number of bits leaving most of the bitrate to the base view that is encoded using standard HEVC technique.

ACKNOWLEDGMENT

This work has been supported by the public funds as a research project.

REFERENCES

[1] ISO/IEC 14496-10, International Standard, “Generic coding of audio-visual objects – Part 10: Advanced Video Coding,” 6nd Ed., 2010, [also:] ITU-T Rec. H.264, Edition 5.0 (version 11), 2010.

[2] Joint Collaborative Team on Video Coding (JCT-VC) of ITU-T SG16 WP3 and ISO/IEC JTC1/SC29/WG11, Doc. JCTVC-G1103, B. Bross, W.-J. Han, , J.-R. Ohm, G. J. Sullivan, T. Wiegand (Eds.), “High Efficiency Video Coding (HEVC) text specification Working Draft 5”, Geneva, November 2011.

[3] K. Wegner, O. Stankiewicz, K. Klimaszewski, M. Domański, ”Comparison of multiview compression performance using MPEG-4 MVC and prospective HEVC technology”, ISO/IEC JTC1/SC29/WG11 (MPEG), Doc. M17913, Geneva, Switzerland, July 2010.

[4] M. Domański, T. Grajek, D. Karwowski, K. Klimaszewski, J. Konieczny, M. Kurc, A. Łuczak, R. Ratajczak, J. Siast, O. Stankiewicz, J. Stankowski, K. Wegner, “Multiview HEVC – experimental results” JCT-VC (MPEG/VCEG) Doc. JCTVC-G582, Geneva, November 2011.

[5] J. Stankowski, M. Domański, O. Stankiewicz, J. Konieczny, J. Siast, K. Wegner, Extensions of the HEVC technology for efficient multiview video coding, IEEE Int. Conf. Image Processing, ICIP 2012, to be published.

[6] M. Domański, T. Grajek, D. Karwowski, K. Klimaszewski, J. Konieczny, M. Kurc, A. Łuczak, R. Ratajczak, J. Siast, O. Stankiewicz, J. Stankowski, K. Wegner, “Technical description of Poznań University of Technology proposal for Call on 3D Video Coding Technology”, ISO/IEC JTC1/SC29/WG11 (MPEG), Doc. M22697, Geneva, Switzerland, November 2011.

[7] J. Konieczny, M. Domański, “Extended inter-view direct mode for multiview video coding”, IEEE Int. Conf. Acoustics Speech Signal Proc., Prague, Czech Republic, May 2011.

[8] M. Domański, T. Grajek, D. Karwowski, K. Klimaszewski, J. Konieczny, M. Kurc, A. Łuczak, R. Ratajczak, J. Siast, O. Stankiewicz, J. Stankowski, K. Wegner, IEEE Int. Conf. Image Processing, ICIP 2012, to be published.