
Citation:

Malik, B and Akbari, S and Kor, A (2017) "HEVC based Mixed-resolution Stereo Video Coding for Low Bitrate Transmission." In: 2017 IEEE International Conference on Imaging Systems and Techniques (IST). IEEE, Beijing/china, pp. 289-293. ISBN 978-1-5386-1620-8

Link to Leeds Beckett Repository record:

<https://eprints.leedsbeckett.ac.uk/id/eprint/4260/>

Document Version:

Book Section (Accepted Version)

Creative Commons: Attribution 4.0

Date of Conference: 18-20 Oct. 2017

The aim of the Leeds Beckett Repository is to provide open access to our research, as required by funder policies and permitted by publishers and copyright law.

The Leeds Beckett repository holds a wide range of publications, each of which has been checked for copyright and the relevant embargo period has been applied by the Research Services team.

We operate on a standard take-down policy. If you are the author or publisher of an output and you would like it removed from the repository, please [contact us](#) and we will investigate on a case-by-case basis.

Each thesis in the repository has been cleared where necessary by the author for third party copyright. If you would like a thesis to be removed from the repository or believe there is an issue with copyright, please contact us on openaccess@leedsbeckett.ac.uk and we will investigate on a case-by-case basis.

HEVC based Mixed-resolution Stereo Video Coding for Low Bitrate Transmission

Bruhanth Mallik, Akbar Sheikh Akbari, Ah-Lian Kor
School of Computing, Creative Technology & Engineering,
Faculty of Arts, Environment & Technology, Leeds Beckett University, U.K.
b.mallik6347@student.leedsbeckett.ac.uk

Abstract—This paper presents a mixed resolution stereo video coding model for High Efficiency Video Codec (HEVC). The challenging aspects of mixed resolution video coding are enabling the codec to encode frames with different frame resolution/size and using decoded pictures having different frame resolution/size for referencing. These challenges are further enlarged when implemented using HEVC, since the incoming video frames are subdivided into coding tree units. The ingenuity of the proposed codec's design, is that the information in intermediate frames are down-sampled and yet the frames can retain the original resolution. To enable random access to full resolution decoded frame in the decoded picture buffer as reference frame a down-sampled version of the decoded full resolution frame is used. The test video sequences were coded using the proposed codec and standard MV-HEVC. Results show that the proposed codec gives a significantly higher coding performance over the MV- HEVC codec.

Keywords— *HEVC; stereo; mixed-resolution; low bitrate; video compression*

I. INTRODUCTION

Stereo video production is one of the simplest and cost-effective ways of making 3D videos, using video pairs acquired simultaneously by two identical cameras, which are geometrically aligned with parallel axes. The elementary function of a stereo video codec is to efficiently compress stereoscopic videos for prompt transmission and compact storage [1]. During the last decade, multiview video coding (MVC) has been investigated by many researchers and several standard multiview video codecs such as H.264/MVC and MV-HEVC, have been developed [2], [4]. The limitations of these techniques are the inherent computational complexity and high bandwidth requirement for multiple views. To efficiently code stereo and multiview videos, various coding algorithms have been proposed based on application of flexible group of pictures, motion homogeneity estimation and adaptive search window sizes [5]. Encoding stereo videos for low bitrate transmission requires use of greater quantization values, which leads to distortions and blocking artefacts, due to loss of high frequency information within the videos. Examining low bitrate HEVC bitstreams have revealed that the slice data overhead, due to large frame sizes, uses most of the bits rather than the actual visual information contained in the frame [3], [4].

Mixed-resolution video coding is an emerging technique for coding video sequences at high compression ratios. The success of the mixed-resolution video coding, also known as

asymmetric spatial inter/intra view frames resolution video codecs, is due to human visual system's tolerance to suppressed high frequency components of the videos [1], [6], [7]. Subjective quality assessments have confirmed that asymmetric frame resolution and quality stereoscopic image/video compression do not show any statistically significant change with regard to eye dominance of the subject [8], [9]. A subjective study on the quality of spatially and temporally down-sampled stereo video frames, showed that spatially down-sampled videos are preferred and potentially are more suitable for stereoscopic video coding at low bitrate [10] [11]. Various asymmetric mixed resolution stereoscopic video coding algorithms have been reported in the literature [12]-[17]. Combination of down sampling and inter-view prediction in asymmetric stereo video coding has proven to outperform symmetric codecs at low bitrate [12], [13]. Motion/disparity estimation/compensation of a down-sampled frame from a full resolution frame was investigated in [14], [15]. The findings show their higher coding performance than that of the anchor codec. The asymmetric stereo video codec proposed in [16], encodes right view frames by discarding the chrominance channel and controlling the luminance quality by applying thresholds at just above the noticeable distortion level. However, subjective evaluation of its video frames indicated that reducing the quality in terms of PSNR for one of the views to just above noticeable distortion threshold, does not achieve the desirable video quality at low bitrates. An adaptive spatial mixed resolution technique using MV-HEVC was proposed in [17]. It uses the correlation between frequency power spectrums of the adjacent views and quality metrics of the encoded down sampled video frames, to select the appropriate down sampling factor. In [18], application of the HEVC for mixed resolution stereoscopic video coding, with the aim of reducing quality-imbalance problem of the decoded video sequences to less than 2 dBs has been reported. The prediction structure of this codec shows that the left view video frames were coded without reference to the right view frames, whereas for the right view video frames, it could use the adjacent frame of the neighboring view in addition to the intra-view predication similar to left view prediction frames. In addition, the first frame of the right view in each group of pictures, has been coded in full resolution. This codec reduced the quality-imbalance to less than 2 db. However, due to not fully exploiting inter-view frame correlation and coding the first frame of the right view in each GOP with full resolution, the full capability of the mixed-resolution stereo video coding using HEVC has not been efficiently utilized.

This paper presents a HEVC based mixed resolution coding architecture, using spatial resolution scaling, for stereoscopic videos. The proposed codec applies an interleaving algorithm to stereo video frames for generating a single stream mixed spatial resolution video sequence. Experimental results using four standard video sequences show that the proposed video codec generates significantly higher coding performance than the anchor MV-HEVC codec, the state of art mixed resolution stereo video codec, and it fulfills the 2db quality-imbalance criteria. The remainder of this paper is organized as follows: Section 2 presents the framework of the proposed technique by introducing the structure of the mixed resolution stereo videos, the intermediate frames down sampling method, frame interleaving algorithm and the codec's design to encode mixed resolution stereo videos. Section 3 presents the experimental results and finally, the paper is concluded in Section 4.

II. PROPOSED HEVC BASED MIXED-RESOLUTION STEREO VIDEO CODEC

The primary objective of the stereo video codec proposed in this paper is to achieve high compression ratio without compromising the quality of the stereo videos for low bitrate transmission. In the proposed stereo video coding scheme, the HEVC standard video codec has been modified and configured to encode frames having different resolutions in the interleaved mixed resolution stereo videos. The frames of each stereo Group of Pictures (GoP) are reordered into a single video stream by an interleaving algorithm. While the first frame of the left view retains its original resolution, the remaining frames are filtered and down-sampled by a factor of two across both axes. The HEVC based mixed resolution stereo video coding (HEVC-MRSVC) scheme encodes the full resolution key frames from stereo video sequence as I-frames, whereas the low-resolution intermediate (non-key) frames are encoded as P- or B-frames. The structure of the mixed resolution stereo video sequence is shown in Fig. 1. The intermediate full resolution P- and B-frames are first filtered using a Blackman 2D FIR low-pass filter to mitigate aliasing artefacts due to down sampling [19]. The Blackman 2D FIR filter's coefficients are given in Table 1. The filtered intermediate frames are then both horizontally and vertically down-sampled by a factor of two to maintain the same aspect ratio as the full resolution key-frames (I-frames). The filtered and down-sampled low resolution intermediate frames are superimposed onto the top left corner of a full resolution sized frame that has every pixel set to value zero, as illustrated in Fig 2.a. As a result, the intermediate frames and the key frames of the mixed-resolution video sequence have the same frame size. The pre-processed intermediate frames have scene information in the form of pixel values, of lower resolution compared to the key frames and these pixels are placed in the top-left quadrant of the frame. Since video codecs exploit the correlation within a frame, the regions with zero pixel values will be encoded with minimum signaling and bitstream. The intermediate frames of the encoded and reconstructed mixed-resolution video are up-sampled by calculating the missing samples through Bi-cubic interpolation and filtered to retrieve the original resolution, as in Fig 2.b. Thus, in the proposed mixed resolution stereo video coding scheme, the key issue of coding mixed resolution videos has been addressed by the

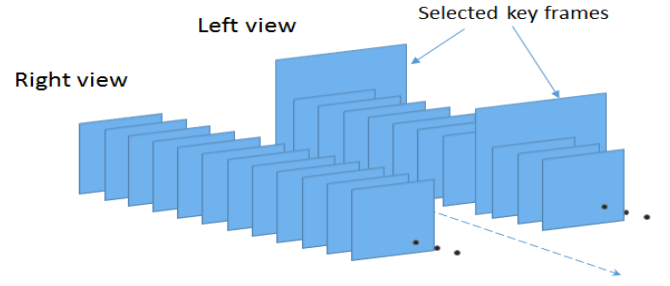


Fig. 1. Structure of the mixed resolution stereo video sequence

TABLE I. BLACKMAN 2D FIR FILTER COEFFICIENTS

0.0381	0.1051	0.0381
0.1051	0.4273	0.1051
0.0381	0.1051	0.0381

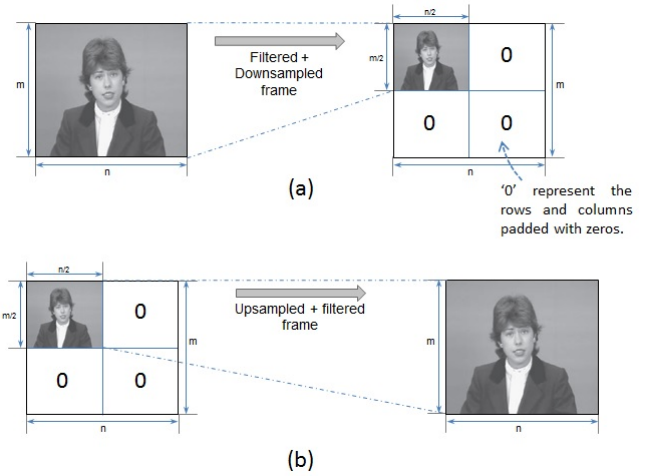


Fig. 2. Intermediate frames, a) input B- or P-frame and its filtered, down-sampled and superimposed frame, b) decoded superimposed frame and its up-sampled frames.

intermediate frames resolution down-sampling and re-mapping procedure.

Encoding of a frame interleaved monoscopic stereo video has a lower decoding complexity compared to encoding stereo views with multi-layer approach codec (for example MV-HEVC) [20]. Further, frame interleaving enables monoscopic HEVC to encode temporally reordered stereo video frames as a single video. The frame interleaving algorithm in the proposed stereo video coding scheme, rearranges the mixed-resolution stereo video frames, such that two consecutive frames of each view are always adjacent to each other, forming a single stream video sequence, as represented by the dashed arrows in Fig 3. In HEVC-MRSVC scheme, the HEVC standard is used to encode and decode the frame interleaved mixed resolution stereo video frames. The HEVC standard uses a hybrid coding structure of prediction and transformation similar to its predecessors. In order to improve coding efficiency and support a wide range of video application formats (including Ultra-HD and 4k resolution videos), HEVC has been designed

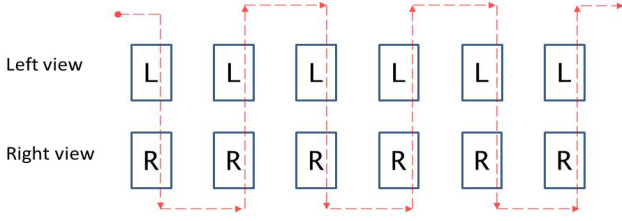


Fig. 3. Frame interleaving algorithm's contour to reorder stereo frames.

with improved advanced coding tools. The codec's high-level syntax has been improved, to be more flexible in order to cater for diverse video applications and enhance parallel processing capabilities. Unlike its predecessors, HEVC's video coding layer does not use macroblocks, instead, the video frames are partitioned into coding tree units (CTUs). Each CTU contains a luma coding tree block, the corresponding chroma coding tree block and the syntax elements. When implemented using HEVC, these challenges with mixed resolution video coding will be even more complex, since the incoming video frames are subdivided into coding tree units by the codec's novel quad-tree based block partitioning concept. In addition, the HEVC standard has not been designed to code varying frame resolution within the video sequence [3], [4]. To address these limitations, the standard HEVC's algorithm is modified to efficiently encode frame interleaved mixed-resolution stereo videos. The low resolution intermediate frames used in the proposed HEVC-MRSVC scheme have been set to possess the same aspect ratio and frame size as their full resolution counterpart, with information saved at a lower resolution. The reference frame structure of the proposed HEVC-MRSVC scheme is shown in Fig 4. The proposed HEVC-MRSVC's reference frame structure allows low-resolution intermediate frames to use other low resolution intermediate frames and down-sampled version of the copy of I-frames. Further, the proposed codec's reference frame structure is configured to avoid using low resolution intermediate frames as reference to full resolution I-frames. The copy of the reconstructed full resolution I-frames saved in decoded picture buffer (DPB) goes through the same down sampling process as shown in Fig 2.a. Thus, by making these design modification to the standard HEVC, the codec has been equipped to encode mixed-resolution videos.

To implement the proposed HEVC-MRSVC video codec, JCT-VC HEVC software version HM16.12 was used. In addition to the modifications to the standard HEVC, the configuration parameters were changed to encode frame interleaved mixed-resolution stereo videos with the reference frame structure as shown in Fig 4. Since the stereo video frames are temporally reordered by the proposed frame interleaving algorithm, the group of pictures (GoP) size is set at 16 frames and the intra frame period at 48. The goal of motion/disparity estimation/compensation is to reduce the energy of the difference block. This is achieved by finding the same scene in either the neighboring view or previous frame. In case of neighboring views, the scene location is a function of distance of the camera from the scene and inter-camera angle

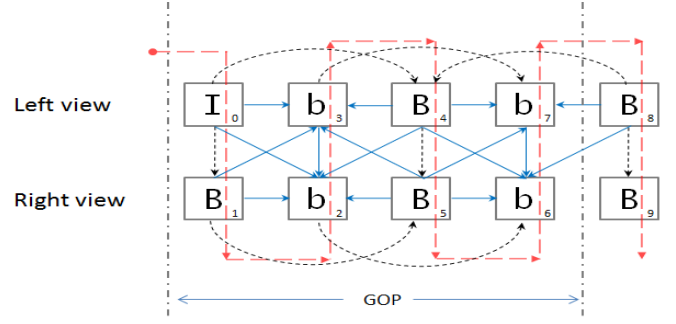


Fig. 4. Reference frame structure of the proposed mixed-resolution stereo video codec.

[21]. For the standard test videos, used in this study, motion vector search range was set at 96 to mitigate the effect of the inter-camera angle and camera distance from the scene for disparity estimation/compensation.

III. EXPERIMENTAL RESULTS

The coding performance of the proposed HEVC based mixed-resolution stereo video codec (HEVC-MRSVC) is compared with the standard MV-HEVC codec using views 4-5, 1-5, 1-3, and 1-3, of the standard multiview test video sequences "Poznan_Street", "Undo_Dancer", "Balloons" and "Kendo", respectively. Experimental results for the anchor MV-HEVC codec are available in JCT3V-G1100 common test condition documentation [22] which are used for the purpose of evaluation in this paper. Figs. 5–8 show the resulting average PSNR for the Y-frames at Quantization Parameters (QP) of 20, 25, 30, 35 and 40. Figs. 5-6 show the resulting average Y-PSNR for "Balloons" and "Kendo" test videos, respectively. The "Balloons" and "Kendo" data sets have been recorded under artificial lighting conditions. These videos have progressively changing backgrounds and a number of fast moving objects in the foreground. From Fig. 5, it can be seen that the proposed codec, on an average, has higher luminance quality metric, i.e. Y-PSNR of up to 1.8dB than the anchor MV-HEVC codec. The proposed HEVC-MRSVC codec gives 1dB higher PSNR than the anchor codec at 450 kbps and 2.33dB higher PSNR at 214 kbps. From Fig. 6, it can be noted that the proposed codec achieves a coding gain of 0.8 dB compared to the anchor codec at 400 kbps. The coding performance of the proposed codec increases as the QP increases and it reaches 1.8dBs at 250 kbps. By looking at the videos and results, it can be inferred that the proposed codec provides superior coding performance compared to the anchor codec when videos contain large number of moving objects in both the foreground and background. The resulting average PSNR for coding "Poznan_Street" video sequence using the proposed HEVC-MRSVC and the anchor codecs are presented in Fig. 7. The "Poznan_Street" data set is an outdoor-recorded video sequence captured under natural lighting. It contains multiple moving objects with a stationary background and fixed camera position. In the figure, it can be seen that the proposed codec's Y-frames attain 0.8 (at 1300kbps) to 2.0 (at 60 kbps) dBs higher PSNR than the anchor codec's frames. By looking at this video, it can be seen that the video has huge

still background with small moving areas. Since the still areas are coded by larger CTUs, which require small number of bits to code, the coding performance of the proposed codec is limited by the small moving areas and the results do not exhibit significant gain. Fig. 8 shows the resulting coding performance for “Undo_Dancer” test video. “Undo_dancer” is a computer graphic animation with scene changes representing camera motion, object motion, with both highlighted background and foreground details. As it can be seen in Fig. 8, the proposed codec gives higher coding performance of almost 3 dB, than the anchor codec. This can be explained by the nature of the video, which is a computer graphic with less light noise and depth details.

The state of art symmetric mixed resolution stereoscopic video coding (SMRSVC) model proposed in [18] mainly focuses on mitigating the quality imbalance between the stereo views, and its overall coding performance is very close to that of the anchor MV-HEVC codec. The performance of the proposed HEVC-MRSVC codec, when coding “Poznan_Street” test videos, depicted in Fig. 7, shows that our proposed codec significantly outperforms the anchor codec, hence, the state of the art symmetric mixed resolution stereoscopic video codec (SMRSVC). In addition, our proposed HEVC-MRSVC codec demonstrates an average PSNR difference of 1.3 dBs, which is well within the 2dB criteria required for viewers to be unable to differentiate the quality differences between the two views [16].

To invoke a sense of the achieved visual quality and enable the reader to compare the resultant videos with the anchor codec’s video, intermediate left view frame 72 in “Poznan_Street” video sequence of the proposed and anchor codecs are shown in Fig. 9. It can be seen that the proposed codec’s image, shown in Fig. 9.a, exhibits higher visual quality with less blocking artefacts than that of anchor codec’s image, shown in Fig. 9.b. The anchor MV-HEVC’s image suffers from noticeable blocking artefacts with blurry edges on moving objects, e.g. edges of the car on the right side of the image and head of the person on the left side of the image.

IV. CONCLUSIONS

In this paper, a HEVC based mixed-resolution stereo video codec (HEVC-MRSVC) was presented. The proposed codec converts the stereo video frames to a single stream mixed spatial resolution video sequence by applying an interleaving algorithm. The HEVC standard video codec was modified and configured to encode frames having different resolutions in the interleaved mixed resolution stereo videos. The proposed HEVC-MRSVC encodes the I-frames in full resolution, and the intermediate frames in low resolution. The low-resolution intermediate frames are superimposed on full resolution sized frames to facilitate easy implementation. Experimental tests on standard video sequences show that the proposed video codec significantly outperforms the anchor codec at low bitrate transmission both objectively and subjectively, while containing the quality-imbalance between the decoded video frames of the stereo pair below the required 2 dB criteria.

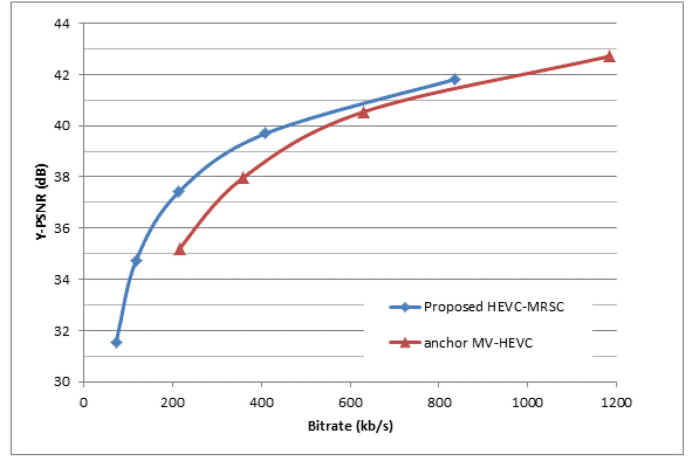


Fig. 5. PSNR vs bitrate for MV-HEVC anchor codec and the proposed HEVC-MRSVC codec for coding “Balloons” test video

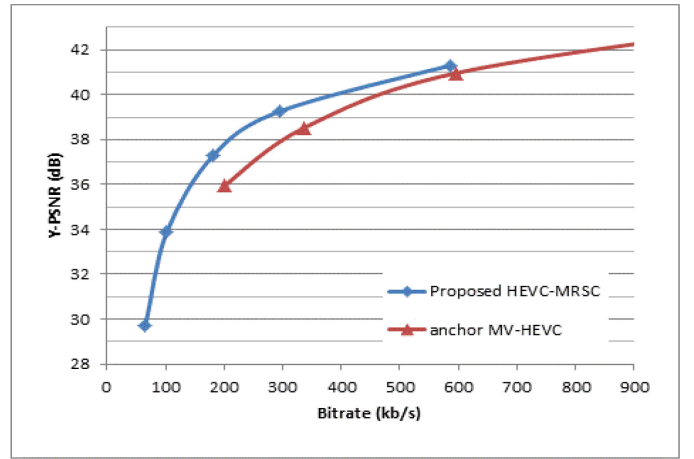


Fig. 6. PSNR vs bitrate for MV-HEVC anchor codec and the proposed HEVC-MRSVC codec for coding “Kendo” test video

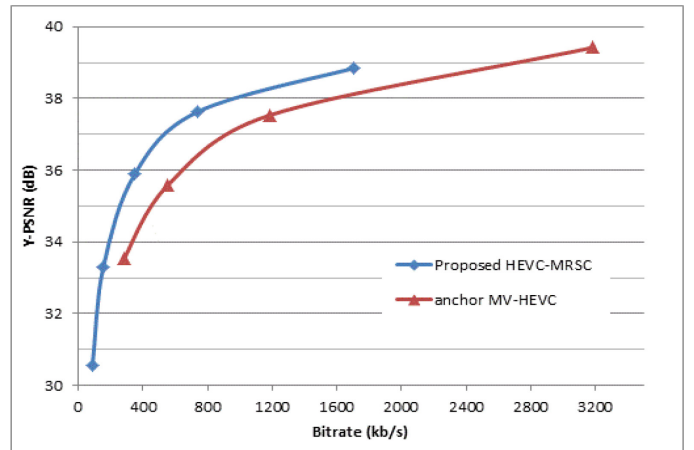


Fig. 7. PSNR vs bitrate for MV-HEVC anchor codec and the proposed HEVC-MRSVC codec for coding “Poznan_Street” test video

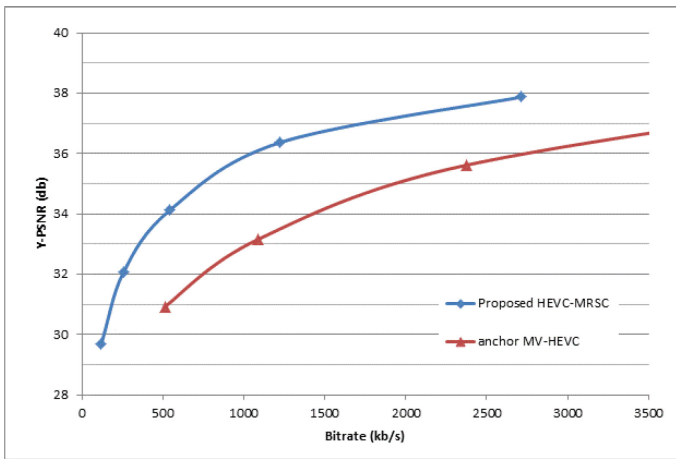


Fig. 8. PSNR vs bitrate for MV-HEVC anchor codec and the proposed HEVC-MRSVC codec for coding "Undo_Dancer" test video



(a) Proposed codec



(b) MV-HEVC

Fig. 9. Decoded frame number 72 from "Poznan_Street" videos at 545kbps bitrate of a) the proposed codec and b) the MV-HEVC standard codec.

REFERENCES

- [1] M. G. Perkins, "Data compression of stereopairs," in *IEEE Trans. on Communications*, vol. 40, no. 4, pp. 684-696, 1992.
- [2] A. Vetro, T. Wiegand and G.J. Sullivan, "Overview of the stereo and multiview video coding extensions of the H. 264/MPEG-4 AVC standard," in *Proc. of the IEEE*, vol. 99, no. 4, pp. 626-642, April 2011.
- [3] "High efficiency video coding," ITU-T Recommendation H.265 and ISO/IEC 23008-2, April 2013 (and subsequent editions).
- [4] G.J. Sullivan, J. Ohm, T.K. Tan and T. Wiegand, "Overview of the High Efficiency Video Coding (HEVC) Standard," in *IEEE Trans. on Circuits and Systems for Video Technology (TCSVT)*, vol. 22, no. 12, pp. 1649-1668, December 2012.
- [5] P. Merkle, H. Brust, K. Dix, A. Smolic and T. Wiegand. "Stereo video compression for mobile 3D services," in *3DTV Conf. The True Vision-Capture, Transmission and Display of 3D Video*, pp.1-4, 2009.
- [6] I. Dinstein, G. Guy, J. Rabany, J. Tzelgov and A. Henik "On stereo image coding," in *IEEE 9th Int. Conf. Pattern Recognition*, pp. 357-359, 1988.
- [7] P. D. Gunatilake, M. W. Siegel and A. G. Jordan. "Compression of stereo video streams," in *Signal Processing of HDTV International Workshop on HDTV'93*, pp. 173-185, 1994.
- [8] L. B. Stelmach and W. J. Tam, "Stereoscopic image coding: Effect of disparate image-quality in left- and right-eye views," in *Signal Process. Image Commun.*, vol. 14, pp. 111-117, 1998.
- [9] P. Seuntjens, L. Meesters, W. Ijsselstein, "Perceived quality of compressed stereoscopic images: Effects of symmetric and asymmetric JPEG coding and camera separation," in *ACM Transactions on Applied Perception (TAP)*, vol. 3, no. 2, 95-109, 2006.
- [10] G. Saygili and C. G. Gürler. "Evaluation of asymmetric stereo video coding and rate scaling for adaptive 3D video streaming," in *IEEE Trans. on Broadcasting*, vol. 57, no. 2, pp. 593-601, 2011.
- [11] A. Aksay, C. Bilen, E. Kurutepe, T. Ozelebi, G.B. Akar, M. R. Civanlar and A. M. Tekalp, "Temporal and spatial scaling for stereoscopic video compression" in *14th European Signal Processing Conf., Proc. EUSIPCO*, Florence, Italy, pp. 1-5, September 2006.
- [12] H. Brust, A. Smolic, K. Mueller, G. Tech and T. Wiegand, "Mixed resolution coding of stereoscopic video for mobile devices" in *3DTV Conference: The True Vision-Capture, Transmission and Display of 3D Video*, May 2009
- [13] C. G. Gürler and K. T. Bağcı. "Adaptive stereoscopic 3D video streaming", in *17th IEEE Int. Conf. Image Processing (ICIP)*, pp. 2409-2412, 2010.
- [14] C. Fehn, P. Kauff, S. Cho, H. Kwon, N. Hur and J. Kim "Asymmetric coding of stereoscopic video for transmission over T-DMB," in: *Proc. 3DTV Conf.*, Kos Island, Greece, pp. 1-4, May 2007
- [15] Y. Chen, Y.K. Wang, M.M. Hannuksela and M. Gabbouj "Picture-level adaptive filters for asymmetric stereoscopic video," in *Proc. IEEE Int. Conf. Image Processing*, San Diego, California, USA, pp. 1944-1947, October 2008.
- [16] F. Shao, G. Jiang, M. Yu, K. Chen and Y.S. Ho, "JND-based asymmetric coding of stereoscopic video for mobile 3DTV applications," in *Proc. Int. Conf. Image and Signal Processing*, Shanghai, China, pp. 186-189, October 2011.
- [17] P. Aflaki, M. M. Hannuksela and M. Gabbouj, "Adaptive Spatial Resolution Selection for Stereoscopic Video Compression with MV-HEVC: A Frequency Based Approach," in *IEEE Int. Symp. Multimedia (ISM)*, pp. 267-270, 2014.
- [18] K. L. Chung and Y. H. Huang, "Efficient multiple-example based super-resolution for symmetric mixed resolution stereoscopic video coding" *Journal of Visual Communication and Image Representation*, vol. 39, 65-81, 2016.
- [19] A.S. Akbari and P.B. Zadeh, "Wavelet Based Image Enlargement Technique," in *Int. Conf. Global Security, Safety, and Sustainability. Springer International Publishing*, pp. 182-188, 2015.
- [20] B. Mallik, A. Sheikh Akbari and P.B. Zadeh, "HEVC based Stereo Video codec," in *2nd IET International Conference on Intelligent Signal Processing*, London, UK, 2015.
- [21] S. Bouyagoub, A. Sheikh Akbari, D. Bull and N. Canagarajah, "Impact of camera separation on performance of H. 264/AVC-based stereoscopic video codec" *IET Electronics letters*, vol. 46, no. 5, pp. 345-346, 2010.
- [22] K. Muller and A. Vetro, "Common Test Conditions of 3DV Core Experiments", ITU-T SG 16 WP 3 and ISO/IEC JTC 1/SC 29/WG 11, JCT3V G1100, pp. 1-7, 2014.