

Multi-view Video Coding Based on High Efficiency Video Coding

Kwan-Jung Oh, Jaejoon Lee, and Du-Sik Park

Advanced Media Lab, Samsung Electronics Co., Ltd
San 14, Nongseo-dong, Giheung-gu, Yongin-si, Gyeonggi-do, 446-712 Korea
{kwanjung.oh, jaejoon1.lee, dusikpark}@samsung.com

Abstract. Multiview video coding is one of the key techniques to realize the 3D video system. MPEG started a standardization activity on 3DVC (3D video coding) in 2007. 3DVC is based on multiview video coding. MPEG finalized the standard for multiview video coding (MVC) based on H.264/AVC in 2008. However, High Efficiency Video Coding (HEVC) which is a 2D video coding standard under developing outperforms the MVC although it does not employ interview prediction. Thus, we designed a new multiview video coding method based on HEVC. Interview prediction was added into HEVC and some coding tools were refined to be proper to MVC. The encoded multiple bitstreams are assembled into one bitstream and it is decoded into multiview video at decoder. From experimental results, we confirmed that the proposed MVC based on HEVC is much better than H.264/AVC, MVC, and HEVC. It achieves about 59.95% bit saving compared to JMVC simulcast at the same quality.

Keywords: Multiview video coding, High efficiency video coding, HEVC, 3D video coding, 3DVC.

1 Introduction

The successive development of multimedia systems and network has been contributed to the standardization of video compression. In the last three decades, several international video coding standards have been established, for example, H.261/H.263 for video telephony and MPEG-1 and MPEG-2 for video CD and digital TV [1], [2]. After that, the MPEG-4 part 2, object oriented video coding standard, and the most popular video coding standard, H.264/AVC (Advanced Video Coding), were produced [3]. Although several further extensions such as FReXt (Fidelity Range Extension) [4] and SVC (Scalable Video Coding) [5] were added to the H.264/AVC, many researchers have desired to develop a new video coding standard more efficient than H.264/AVC. A formal joint Call for Proposals (CfP) [6] on a next-generation video coding technology was issued in January 2010 by ITU-T VCEG (video coding experts group) and ISO/IEC MPEG (moving picture experts group), and proposals were evaluated at the first meeting of the MPEG & VCEG, JCV-VC (Joint Collaborative Team on Video Coding) [7], which took place in April 2010.

With the success of three-dimensional (3D) movies and various 3D display devices, several international 3D video coding scenarios also have been standardized with the 2D video coding standards. MPEG-2 MVP (multiview profile) [8], MPEG-4 MAC (multiple auxiliary component) [9], and H.264/AVC MVC (multiview video coding) was standardized for multiview video coding. MVC was standardized in July 2008 [10] and currently MVC suitability for interlaced multiview video is investigated. MPEG started a standardization activity on 3DVC (3D video coding) in 2007. The MPEG 3DVC announced a version on 3D video which supports various 3D displays and its new coding standard will be developed in two years [11]. Related works on 3D video coding include depth coding, virtual view synthesis (rendering), depth estimation as well as multiview video coding. In April 2011, a CfP on 3DVC was issued and various 3D video coding technologies have been studied.

In this paper, we propose a multiview video coding based on high efficiency video coding. We believe that the proposed approach will be the basis of the 3D video coding standard since MVC and HEVC are the most powerful coding standards for multiview video coding and 2D video coding. The rest of this paper is organized as follows. In Section 2, HEVC is introduced. Implementation of the proposed method is explained in Section 3. In Section 4, experimental results are given. Section 5 presents conclusions.

2 High Efficiency Video Coding

HEVC [12] is a new video coding standard and this is currently under developing in JCT-VC established by ISO/IEC MPEG and ITU-T VCEG. It aims to substantially improve coding efficiency compared to H.264/AVC and is targeted at next-generation HDTV (Ultra HDTV, 7680x4320) displays as well as improved visual quality in terms of noise, color gamut, and high dynamic range [13]. The first formal HEVC test model, HM, was established in the third JCT-VC meeting held in Guangzhou, China and it provides more flexibility than H.264/AVC. The basic coding unit (CU) has a similar role to the macroblock in H.264/AVC. CU can be further split into prediction unit (PU). Transform unit (TU) is defined for transform and quantization. The overall coding structure is characterized by various sizes of CU, PU, and TU in a recursive manner, once the size of the largest coding unit (LCU) and the hierarchical depth of CU are defined [13].

The CU allows content-adaptive recursive splitting into four equally sized blocks, starting from 64x64 to 8x8 for luma samples. Both skipped CU and non-skipped CU types are allowed. The skipped CU is considered to be an inter prediction mode without coding of motion vector differences and residual data. The non-skipped CU is assigned to one of two prediction modes, intra and inter. Fig. 1 shows the example of CU structure.

The PU is the elementary unit used for carrying the information related to the prediction. In general, it is not restricted to being square in shape, in order to facilitate partitioning which matches the boundaries of real objects in the picture. Each CU may contain one or more PUs. PU types can be skip, intra, and inter. Fig. 2 shows the four types of PU structure.

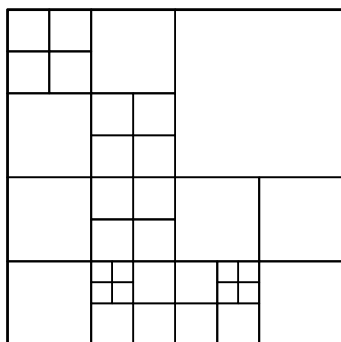


Fig. 1. Example of Coding Unit Structure

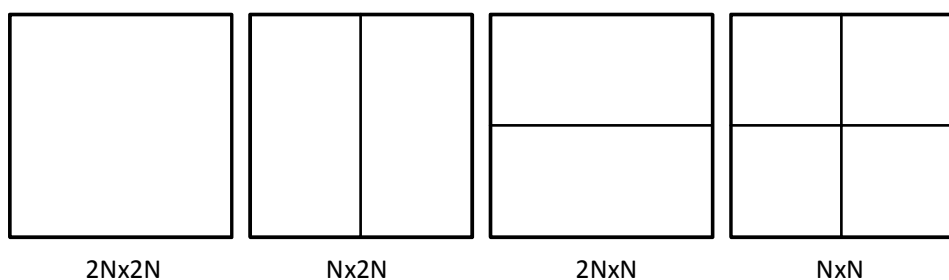


Fig. 2. Four Types of Prediction Unit Structure

The third type of unit is the TU, which is the unit for transform and quantization. It should be noted that the size of TU may be larger than that of the PU but not exceed that of the CU. It is always square and it may take a size from 4×4 up to 32×32 luma samples. Each CU may contain one or more TUs, where multiple TUs may be arranged in a quadtree structure, as illustrated in Fig. 3. Fig. 4 shows the relationship between CU, PU, and TU.

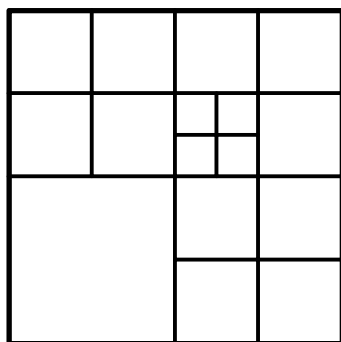


Fig. 3. Example of Transform Unit Structure

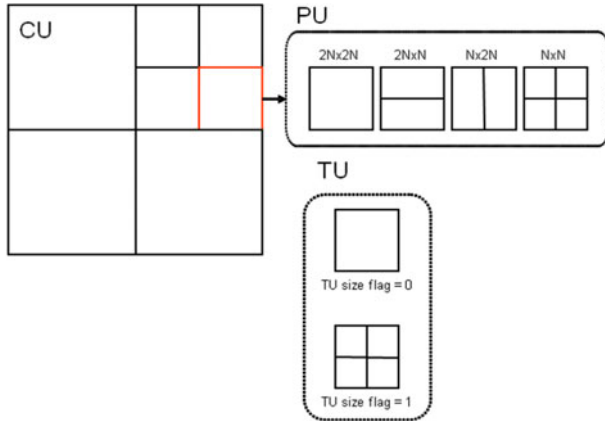


Fig. 4. Relationship between CU, PU, and TU

In addition to new coding structure, HEVC supports various new techniques such as, angular intra prediction, AMVP (advanced motion vector prediction), IBDI (internal bit depth increase), variable size transform/quantization, ALF (adaptive loop filter), and etc [14]. Currently, two configurations are suggested for typical coding tools. They are for high efficiency and low complexity, respectively as in Table. 1.

Table 1. Structure of tools in HM configurations

High Efficiency Configuration	Low Complexity Configuration
Coding Unit tree structure (8x8 up to 64x64 luma samples)	
Prediction Units	
Transform unit tree structure (3 level max.)	Transform unit tree structure (2 level max.)
Transform block size of 4x4 to 32x32 samples (always square)	
Angular Intra Prediction (34 directions max.)	
DCT-based interpolation filter for luma samples (1/4-sample, 8-tap)	
DCT-based interpolation filter for luma samples (1/8-sample, 4-tap)	
Coding Unit based Skip & Prediction Unit based merging	
Advanced motion vector prediction	
CABAC	Low complexity entropy coding phase 2
Internal bit-depth increase (2 bits)	X
X	Transform precision extension (2 bits)
Deblocking filter	
Adaptive loop filter	X

3 Implementation Multiview Video Coding Using HEVC

H.264/AVC MVC manages two reference software JMVM (joint multiview video model) [15] and JMVC (joint multiview video coding) [16] where several tools in JMVM are removed, such as illumination compensation, motion skip mode, and some tools related to SVC. However, the basic coding architectures of JMVM and JMVC are the same. We implement the interview prediction in JMVC to HEVC. For that, MVC related syntaxes, sequence parameter set MVC extension syntax and NAL unit header MVC extension syntax, are added to HEVC. Above two syntaxes are defined in Table 2 and Table 3. The proposed MVC based on HEVC supports hierarchical-B coding structure in temporal direction and I-B-P suture for view direction. The basic view-temporal coding structure is same with that of the JMVC as shown in Fig. 5.

Table 2. Sequence parameter set MVC extension syntax

Syntax	C	Descriptor
num_views_minus1	0	ue(v)
for(i = 0; i <= num_views_minus1; i++)		
view_id[i]	0	ue(v)
for(i = 1; i <= num_views_minus1; i++) {		
num_anchor_refs_10[i]	0	ue(v)
for(j = 0; j < num_anchor_refs_10[i]; j++)		
anchor_ref_10[i][j]	0	ue(v)
num_anchor_refs_11[i]	0	ue(v)
for(j = 0; j < num_anchor_refs_11[i]; j++)		
anchor_ref_11[i][j]	0	ue(v)
}		
for(i = 1; i <= num_views_minus1; i++) {		
num_non_anchor_refs_10[i]	0	ue(v)
for(j = 0; j < num_non_anchor_refs_10[i]; j++)		
non_anchor_ref_10[i][j]	0	ue(v)
num_non_anchor_refs_11[i]	0	ue(v)
for(j = 0; j < num_non_anchor_refs_11[i]; j++)		
non_anchor_ref_11[i][j]	0	ue(v)
}		

Table 3. NAL unit header MVC extension syntax

Syntax	C	Descriptor
nal_unit_header_mvc_extension() {		
non_idr_flag	All	u(1)
priority_id	All	u(6)
view_id	All	u(10)
temporal_id	All	u(3)
anchor_pic_flag	All	u(1)
inter_view_flag	All	u(1)
reserved_one_bit	All	u(1)
}		

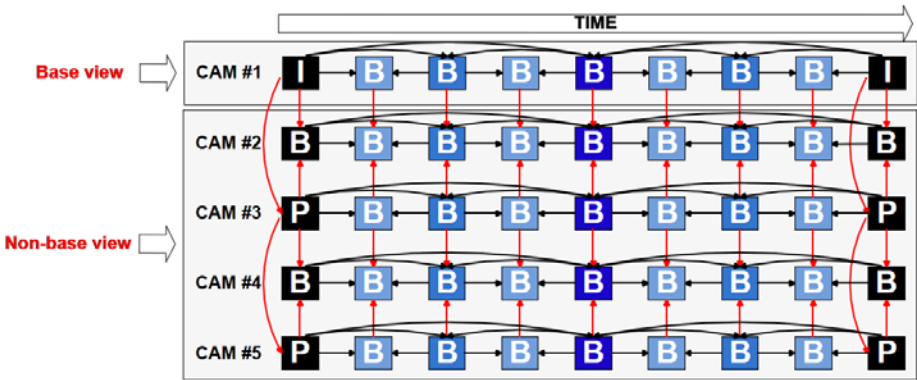


Fig. 5. View-temporal Coding Structure for MVC

For the base view in Fig. 5, we encode the sequence parameter set MVC extension syntax as prefix for slice header as show in Fig. 6. The base view is compatible with HEVC by skipping white colored NAL units such as Subset SPS, additional PPS, and prefixes.

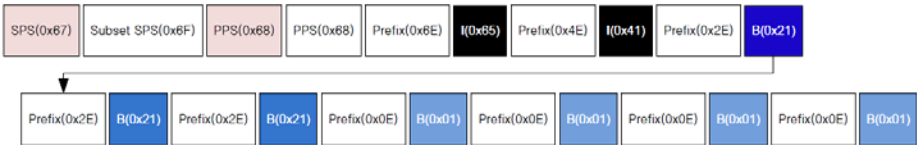


Fig. 6. Arrangement of NAL Units for Base View

For the non-base views, a new slice header (0x14) containing a prefix (0x2E) information and slice header information (0x01) is defined. Thus, it does not need to send a prefix defined in base view. Fig. 7 shows an arrangement of NAL units for P view and B view.

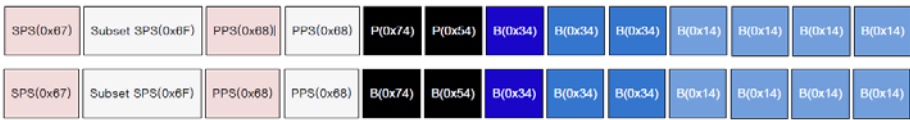


Fig. 7. Arrangement of NAL Units for Non-base View

Multiple encoded bitstreams are unified one bitstream by assembler. Fig. 8 shows the NAL units arrangement of assembled bitstream in case of five views. Fig. 9, Fig. 10, and Fig. 11 show how encoder, assembler, and decoder are working.

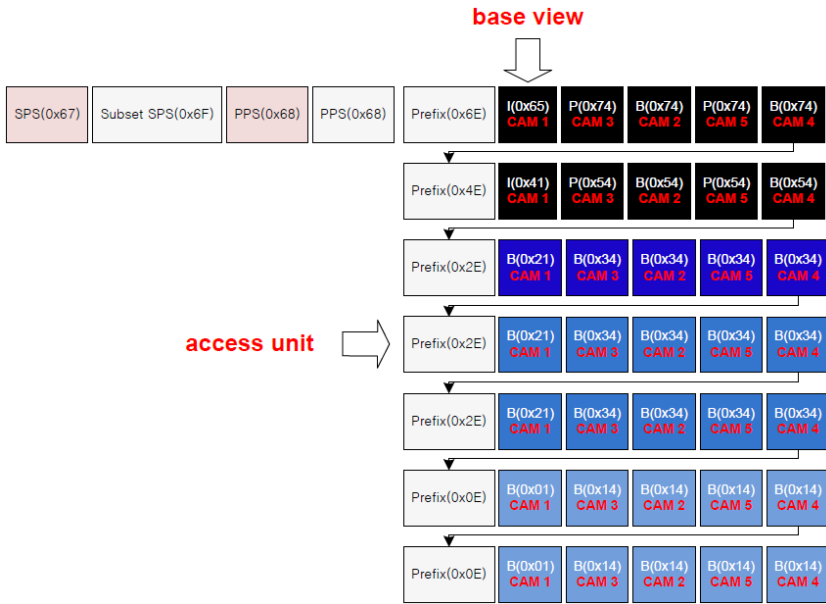


Fig. 8. NAL Unit Arrangement in Assembled Bitstream

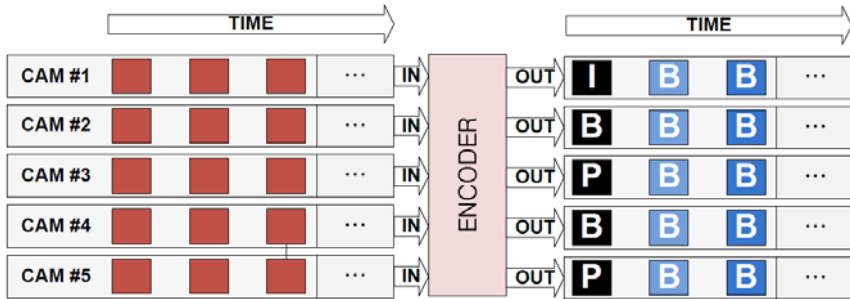


Fig. 9. Encoder

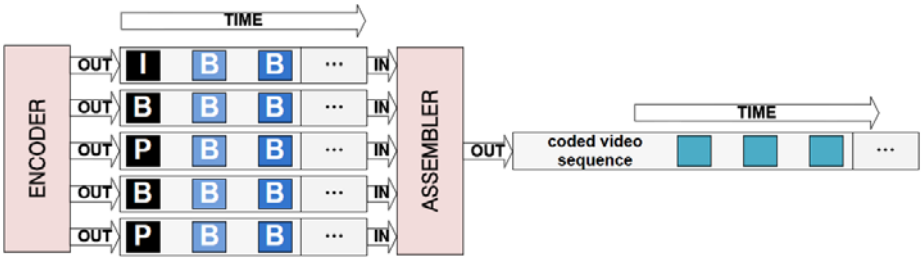


Fig. 10. Bitstream Assembler

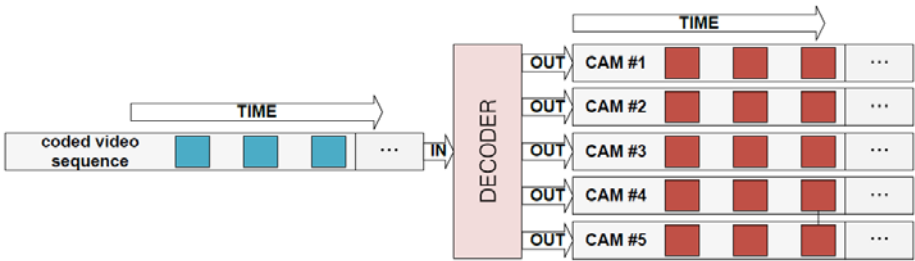


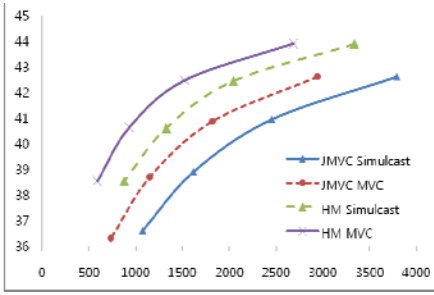
Fig. 11. Decoder

4 Experimental Results

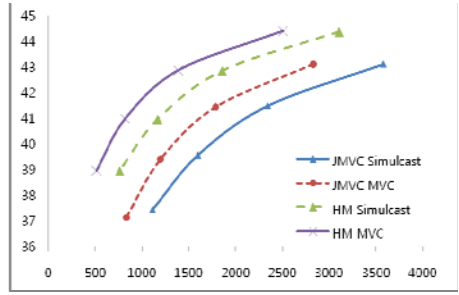
The proposed multiview video coding using high efficiency video coding was implemented based on HM 2.0. To verify the efficiency of the proposed method, we tested JMVC, HM simulcast, and proposed method compared to JMVC simulcast. Encoding of 61 frames for each view with QP 24, 28, 32, 36 was performed for 3 views case with the same test conditions as in MPEG 3DVC CfP [17]. The efficiency of the proposed algorithm was evaluated with BDBR (Bjontegaard delta bit rate) metric [18]. Table 4 shows the experimental results and Fig. 12 shows RD curves.

Table 4. Experimental results in terms of BDBR

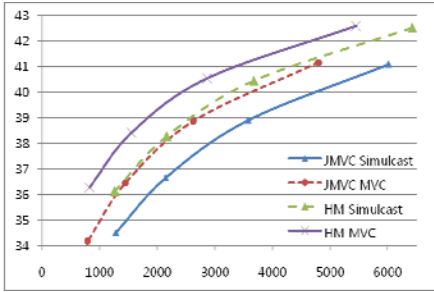
Sequence	JMVC	HM Simulcast	Proposed (HM MVC)
Balloons	-25.38 %	-41.99 %	-59.37 %
Kendo	-22.43 %	-44.47 %	-60.92 %
Lovebird1	-27.40 %	-30.33 %	-50.96 %
Newspaper	-18.97 %	-36.27 %	-52.38 %
GT_Fly	-39.03 %	-30.02 %	-68.57 %
Poznanhall2	-12.89 %	-54.96 %	-66.54 %
Poznanstreet	-35.56 %	-34.62 %	-58.34 %
Undo_Dancer	-40.12 %	-30.02 %	-61.95 %
Average	-28.06 %	-37.24 %	-59.95 %



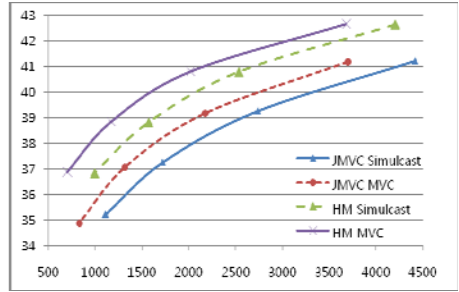
(a) Balloons



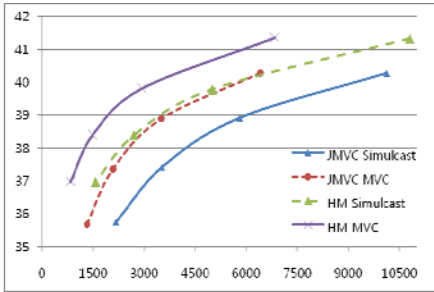
(b) Kendo



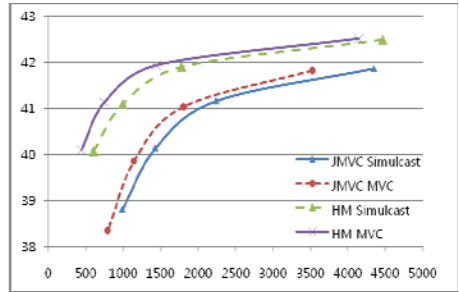
(c) Lovebird1



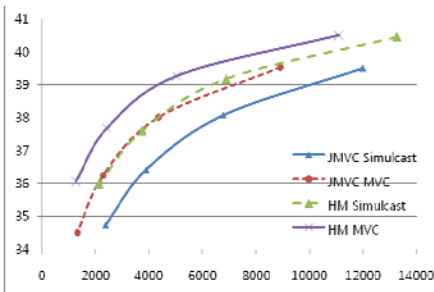
(d) Newspaper



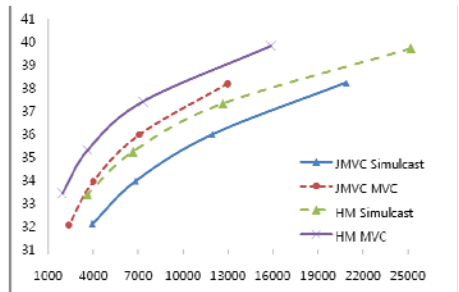
(e) GT_Fly



(f) Poznanhall2



(g) Poznanstreet



(h) Undo_Dancer

Fig. 12. RD Curves

5 Conclusions

In this paper, multiview video coding using high efficiency video coding is proposed. Interview prediction based on JMVC was added to HM. In addition, an assembler to combine multiview bitstreams into one bitstream was implemented. To verify the efficiency of the proposed method, four coding approaches were tested; JMVC simulcast, JMVC, HM simulcast, and the proposed method. From the experimental results, we confirmed that the proposed HM MVC outperforms the other approaches. It achieves about 59.95% bit saving compared to JMVC simulcast at the same quality. We expect that the proposed method will be a basement of MPEG 3DVC standard.

References

1. Hang, H.M., Woods, J.W.: Handbook for Visual Communications. Academic Press (1995)
2. Rao, K.R., Hwang, J.J.: Techniques and for Image, Video, and Audio Coding. Prentice-Hall (1996)
3. Richardson, I.E.: H.264 and MPEG-4 Video Compression: Video Coding for Next Generation Multimedia. Wiley (2003)
4. Sullivan, G.J., Topiwala, P., Luthra, A.: The H.264/AVC Advanced Video Coding Standard: Overview and Introduction to the Fidelity Range Extensions. In: SPIE Conference on Applications of Digital Image Processing, vol. 5558(1), pp. 454–474 (2004)
5. Schwarz, H., Marpe, D., Wiegand, T.: Overview of the Scalable Video Coding Extension of the H.264/AVC Standard. IEEE Trans. CSVT 17(9), 110–120 (2007)
6. ITU-T VCEG and ISO/IEC MPEG VCEG-AM91 & N11113: Joint Call for Proposals on Video Compression Technology (2010)
7. ITU-T VCEG and ISO/IEC MPEG VCEG-AM90 & N11112: Terms of Reference of the Joint Collaborative Team on Video Coding Standard Development (2010)
8. Chen, X., Luthra, A.: MPEG-2 multiview profile and its application in 3D TV. In: SPIE-Multimedia Hardware Architectures, vol. 3021, pp. 212–223 (1997)
9. Karim, H.A., Worrall, S., Sadka, A.H., Kondoz, A.M.: 3-D video compression using MPEG-4-multiple auxiliary component (MPEG4- MAC). In: IEE 2nd International Conference on Visual Information Engineering (2005)
10. ISO/IEC JTC1/SC29/WG11 w9978: Text of ISO/IEC 14496-10:2008/FDAM 1 Multiview Video Coding (2008)
11. ISO/IEC JTC1/SC29/WG11 N10357: Vision on 3D Video (2009)
12. Joint Collaborative Team on Video Coding (JCT-VC) of ITU-T SG16 WP3 and ISO/IEC JTC1/SC29/WG11 JCTVC-A124: Video coding technology proposal by Samsung (and BBC) (2010)
13. High efficiency video coding, <http://www.h265.net>
14. ITU TSB: Joint Collaborative Team on Video Coding. ITU-T Retrieved 21 (2010)
15. ISO/IEC JTC1/SC29/WG11 JVT-AA207: Joint multi-view video model (JMVM) 8.0 (2008)
16. ISO/IEC JTC1/SC29/WG11 JVT-AD207: WD 4 reference software for MVC (2009)
17. ISO/IEC JTC1/SC29/WG11 N12036: Call for Proposals on 3D Video Coding Technology (2011)
18. ITU-T SG16 Q.6 VCEG-AE07: An Excel Add-in for Computing Bjøntegaard Metric and its Evolution (2007)