Joint Collaborative Team on Video Coding (JCT-VC)
of ITU-T SG16 WP3 and ISO/IEC JTC1/SC29/WG11
7th Meeting: Geneva, CH, 21-30 November, 2011

Document: JCTVC-G582

Title:	**Multiview HEVC – experimental results**		
Status:	Input Document to JCT-VC		
Purpose:	Information		
Author(s) or Contact(s):	Marek Domański Tomasz Grajek Damian Karwowski Krzysztof Klimaszewski Jacek Konieczny Maciej Kurc Adam Łuczak Robert Ratajczak Jakub Siast Olgierd Stankiewicz Jakub Stankowski Krzysztof Wegner	Tel: Email:	+48 6653900 dkarwowski@multimedia.edu.pl jstankowski@multimedia.edu.pl
Source:	Poznan University of Technology		

Abstract

This documents presents an approach of providing multiview compression capability in HEVC in similar way to MPEG-4 AVC/H.264 Annex H. The results for experimental implementation of HEVC-based multiview codec and prospective performance of multiview prediction in HEVC-based multiview codecs are described. Codec has been implemented using HEVC reference software (HM 3.0), by application the compression scheme similar to Multiview Video Coding technology (MVC). Coding efficiency of HEVC-based multiview coder was evaluated and compared to efficiency of simulcast HEVC. Performance of the proposed encoder was tested with the sequences provided in Call for Proposals (CfP) on 3D Video Coding (3DVC). The average compression gain of using multiview prediction in video encoder is 22.7% in 2-view case and 30.5% in 3-view case, relative to simulcast scenario.

1 Introduction

In order to represent multiview video signal in an efficient way, Multiview Video Coding standard (MVC) has been developed as the multiview extension (annex H) to MPEG-4 AVC/H.264 video coding technology [1]. In MVC, motion compensated prediction has been adopted to perform both temporal (inter-frame) and inter-view prediction. Inter-view prediction allows exploiting the inter-view correlation that exists in multiview video signal and reducing the bitstream representing the side views.

In the HEVC the motion compensated prediction mechanism has been considerably developed. Significant number of new coding tools has been also introduced when compared to MPEG-4 AVC/H.264. Due to substantial differences in motion compensated prediction between MPEG-4 AVC/H.264 and HEVC, the potential performance of multiview prediction in HEVC-based coder has been evaluated.

2 Implementation overview

The multiview HEVC coder has been implemented based on HEVC reference software. Implementation was done based on HM 3.0 software. Relative to the original HEVC coder, mechanisms for inter-view prediction were embedded. This way, the modified codec works similar to MVC. Inter-view prediction was implemented by modifying the scheme of reference lists construction. For dependent views, the inter-view reference frames are inserted to the reference lists and precede the temporal reference frames. (Figure 1.)

Additional features of described implementation are as follows:

- the base view bitstream is compliant with the one produced by original HM software and can be decoded by the reference decoder;
- implementation of multiview HEVC causes minimal interference in HM software structure and can be easily integrated into any higher version of HEVC reference software;
- no additional complexity was introduced (merely additional reference frames are added to reference lists).

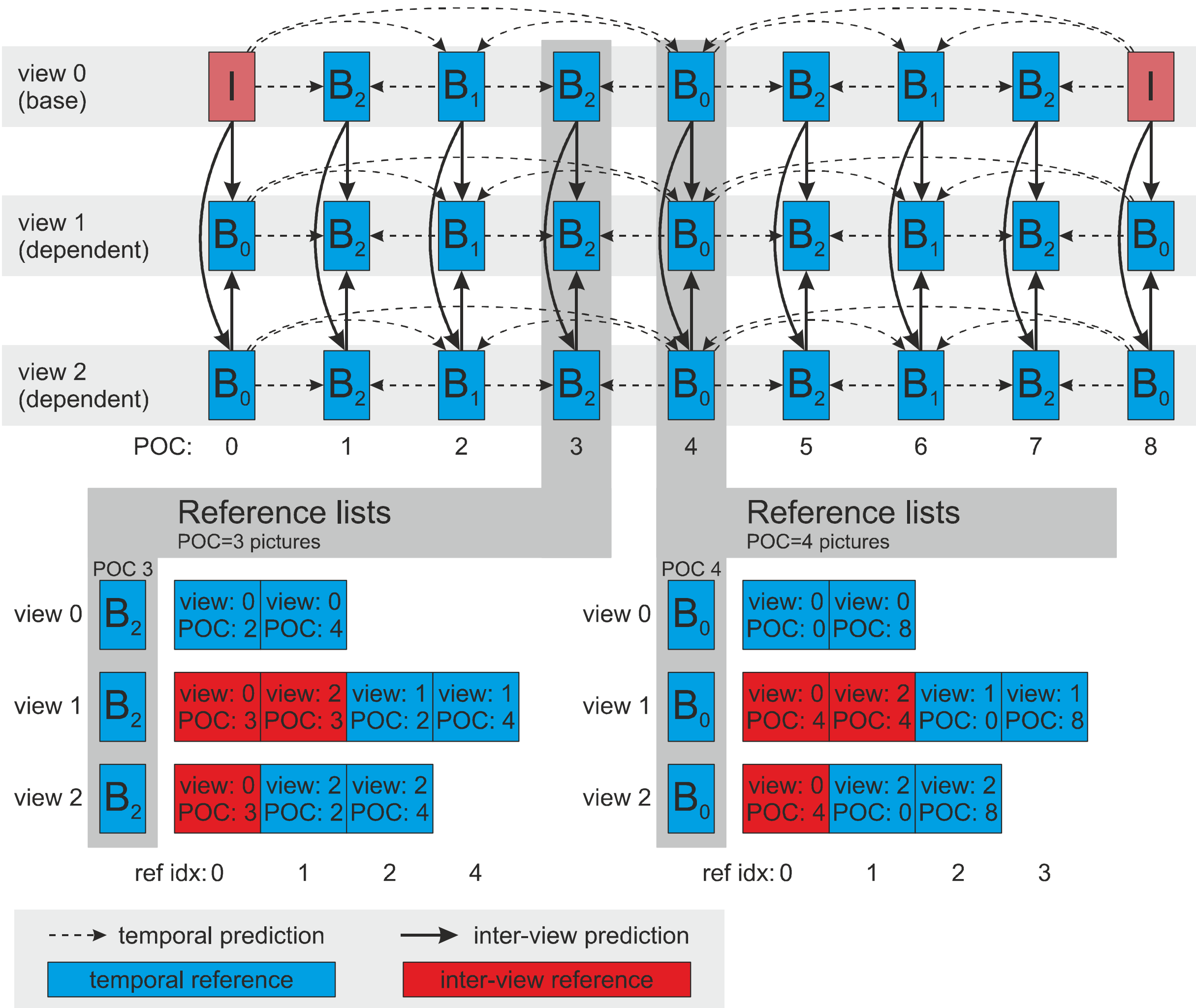

Figure 1.: The reference lists construction (example for the case of one base view and two dependent views). Temporal prediction has been simplified.

3 Coding scenarios and test sequences

The performance evaluation was performed under the conditions described in the Call for Proposals (CfP) on 3D Video Coding (3DVC) document [2] and in Common Test Conditions and software reference configurations document [3]. The CfP defines two coding scenarios: 2-view and 3-view case.

The 2-view case corresponds to coding a stereoscopic pair of views, intended to be displayed on a stereoscopic monitor. In this scenario, there is one base view and one dependent view. The dependent view is encoded using the base view as the reference and the base view is referenced for both anchor and non-anchor frames. Figure 2. illustrates the 2-view case.

The 3-view case has been intended to use in auto-stereoscopic displays. In this scenario, there are base view and two dependent views. The first dependent view is encoded using the base view as a reference and the base view is referenced for both anchor and non-anchor frames. The second dependent view is encoded using the base view and first dependent view as a reference, and those views are referenced for both anchor and non-anchor frames. Figure 3. illustrates the 3-view case.

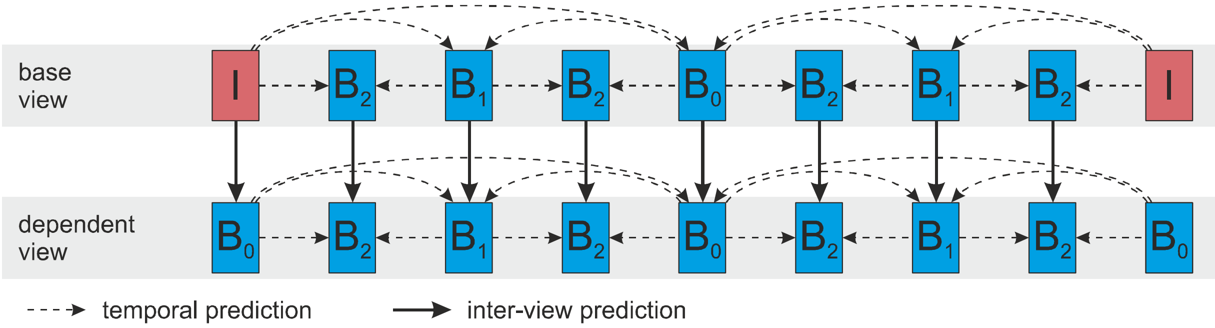

Figure 2.: The 2-view case prediction scheme. Temporal prediction has been simplified.

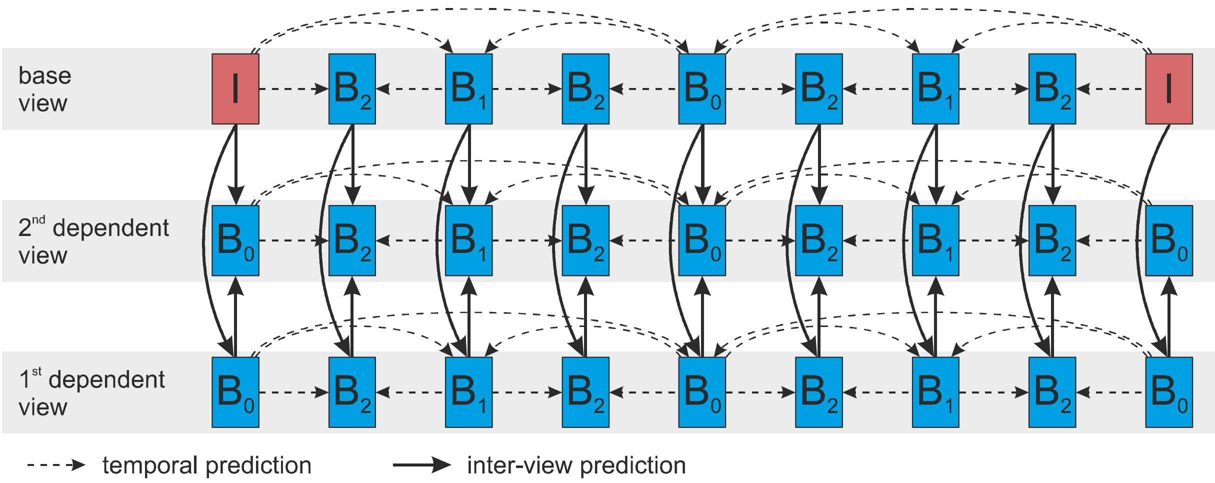

Figure 3.: The 3-view case prediction scheme. Temporal prediction has been simplified.

Coding performance was evaluated using a set of multiview sequences described in Call for Proposals (CFP) on 3D Video Coding [3]. The same set of sequences was used to evaluate the performance of 3D Video Coding propositions.

Test sequences are divided into two classes:

- Class A, containing two natural sequences (Poznan_Hall2, Poznan_Street) and two computer rendered sequences (Undo_Dancer, GT_Fly) in Full HD resolution.
- Class C, containing four natural sequences in XGA resolution.

Detailed information about test sequences is in Table 1.

Table 1. Test sequences

Class A Test Sequence	**Resolution**	**Framerate** [FPS]	**Length** [frames]	**2-view case** [views]	**3-view case** [views]	**Provider**
Poznan_Hall2	1920x1088	25	200	7-6	7-6-5	PUT
Poznan_Street	1920x1088	25	250	4-3	5-4-3	
Undo_Dancer	1920x1088	25	250	2-5	1-5-9	Nokia
GT_Fly	1920x1088	25	250	5-2	9-5-1	

Class C Test Sequence	**Resolution**	**Framerate** [FPS]	**Length** [frames]	**2-view case** [views]	**3-view case** [views]	**Provider**
Kendo	1024x768	30	300	3-5	1-3-5	Nagoya
Balloons	1024x768	30	200	3-5	1-3-5	
Lovebird1	1024x768	30	240	6-8	4-6-8	ETRI / MPEG Korea Forum
Newspaper	1024x768	30	300	4-6	2-4-6	GIST

4 Results

Coding results are shown as a comparison between independent coding of every view using HEVC (simulcast scenario) and join coding of views using HEVC-based multiview encoder that exploits inter-view prediction. All the results are gathered for coding conditions corresponding to random access high efficiency setup (RA-HE). In summary, the average BD-rate gain[4] of 22.7% and 30.5% were achieved for 2-view and 3-view case respectively.

Table 2. Performance of multiview HEVC compared to simulcast HEVC (HM 3.0)

	RA-HE 2-view case			RA-HE 3-view case		
	Y BD-rate	U BD-rate	V BD-rate	Y BD-rate	U BD-rate	V BD-rate
Class A	-26,8%	-27,0%	-26,9%	-35.5%	-35.4%	-35.1%
Class C	-18.6%	-15.1%	-16.4%	-25.5%	-22.1%	-23.5%
Overall	-22.7%	-21.1%	-21.7%	-30.5%	-28.8%	-29.3%
	-22.6%	-21.1%	-21.7%	-30.4%	-28.8%	-29.3%
Enc Time	113%			126%		
Dec Time	96%			97%		

Table 3. Performance of multiview HEVC compared to simulcast HEVC (HM 3.0) (results for sequences)

	RA-HE 2-view case			RA-HE 3-view case		
	Y BD-rate	U BD-rate	V BD-rate	Y BD-rate	U BD-rate	V BD-rate
Poznan_Hall2	-19.0%	-16.7%	-16.8%	-23.4%	-19.1%	-21.0%
Poznan_Street	-24.2%	-23.7%	-23.5%	-33.2%	-33.6%	-31.3%
Undo_Dancer	-30.5%	-33.3%	-32.8%	-41.3%	-44.0%	-43.0%
GT_Fly	-33.5%	-34.5%	-34.5%	-44.0%	-44.9%	-45.0%
Kendo	-15.9%	-9.1%	-12.0%	-24.4%	-19.4%	-21.9%
Balloons	-19.2%	-15.3%	-17.7%	-27.1%	-24.0%	-25.9%
Lovebird1	-18.6%	-17.6%	-17.7%	-28.2%	-25.6%	-26.1%
Newspaper	-20.4%	-18.5%	-18.3%	-22.1%	-19.4%	-20.3%
Overall	-22.7%	-21.1%	-21.7%	-30.5%	-28.8%	-29.3%

Table 3. Performance of multiview HEVC compared to simulcast HEVC (HM 3.0) (3-view case, results for sequences and views)

	RA-HE base view			RA-HE 1st dependent view			RA-HE 2nd dependent view		
	Y BD-rate	U BD-rate	V BD-rate	Y BD-rate	U BD-rate	V BD-rate	Y BD-rate	U BD-rate	V BD-rate
Poznan_Hall2	0.0%	0.0%	0.0%	-22.7%	-16.6%	-21.8%	-46.6%	-42.0%	-41.8%
Poznan_Street	0.0%	0.0%	0.0%	-36.6%	-37.3%	-31.9%	-64.2%	-64.7%	-63.5%
Undo_Dancer	0.0%	0.0%	0.0%	-50.2%	-52.9%	-51.9%	-73.2%	-75.6%	-74.6%
GT_Fly	0.0%	0.0%	0.0%	-52.9%	-54.1%	-54.3%	-78.3%	-79.1%	-79.2%
Kendo	0.0%	0.0%	0.0%	-21.8%	-14.3%	-19.3%	-51.8%	-46.1%	-48.0%
Balloons	0.0%	0.0%	0.0%	-30.4%	-25.6%	-27.3%	-51.5%	-47.8%	-50.7%
Lovebird1	0.0%	0.0%	0.0%	-36.3%	-32.2%	-33.0%	-52.1%	-48.7%	-49.5%
Newspaper	0.0%	0.0%	0.0%	-16.0%	-13.8%	-15.3%	-50.3%	-45.7%	-46.9%
Overall	0.0%	0.0%	0.0%	-33.4%	-30.9%	-31.9%	-58.5%	-56.2%	-56.8%

5 Conclusions

In this contribution, results for experimental implementation of HEVC-based multiview coder were presented. The inter-view prediction gain is 22.7% and 30.5% for 2-view and 3-view case respectively.

6 References

[1] International Standard ISO/IEC 14496-10:2009, Information technology – Coding of Audio-Visual Objects, Part 10, Advanced Video Coding, Annex H - Multiview video coding.

[2] Call for Proposals on 3D Video Coding Technology, ISO/IEC JTC1/SC29/WG11, MPEG2011/N12036, Geneva, Switzerland, March 2011.

[3] F. Bossen, “Common test conditions and software reference configurations”, JCTVC-F900, Torino, July 2011

[4] “Improvements of the BD-PSNR model,” ITU-T SG16 Q.6 Document, VCEG-AI11, Berlin, July 2008.

7 Patent rights declaration(s)

Poznan University of Technology may have current or pending patent rights relating to the technology described in this contribution and, conditioned on reciprocity, is prepared to grant licenses under reasonable and non-discriminatory terms as necessary for implementation of the resulting ITU-T Recommendation | ISO/IEC International Standard (per box 2 of the ITU-T/ITU-R/ISO/IEC patent statement and licensing declaration form).