Joint Video Team (JVT) of ISO/IEC MPEG & ITU-T VCEG (ISO/IEC JTC1/SC29/WG11 and ITU-T SG16 Q.6)
24th Meeting: Geneva, CH, 29 June – 5 July, 2007

Document: JVT-X048
Filename: JVT-X048.doc

Title: **CE3: Depth Map Generation and Depth Map Coding for MVC**

Status: Input Document to JVT

Purpose: Proposal

Author(s) or Contact(s):

Yo-Sung Ho, Kwan-Jung Oh, Cheon Lee, Sang-Beom Lee, Sang-Tae Na
Gwangju Institute of Science and Technology (GIST)
1 Oryong-dong, Buk-gu, Gwangju, 500-712, Republic of Korea
Tel: 82-62-970-2211
Email: hoyo@gist.ac.kr, kjoh81@gist.ac.kr, leecheon@gist.ac.kr, sblee@gist.ac.kr, stna@gist.ac.kr

*Byeongho Choi and *Ji Ho Park
*Korea Electronics and Technology Institute (KETI)
#68 Yatap-dong, Bundang-gu, Seongnam-si, Gyeonggi-do, 463-816, Republic of Korea
bhchoi@keti.re.kr, scottie@keti.re.kr

Source: GIST and KETI

Abstract

This document describes the depth map generation and depth map coding for multi-view video coding (MVC). At the San Jose meeting, CE participants agreed to study the depth-based view interpolation scheme. In addition, CE participants agreed to encode the depth map as side information to avoid the recalculating the depth map at the decoder side [1]. In this document, we propose the segment-based multi-view depth generation scheme and its coding method. We apply the synthesized image by using the proposed schemes to the previous 'VIP P-picture' coding method [2]. By experiments, we show that the proposed depth generation and its coding scheme are good enough to synthesize the virtual image and to contribute to the improvement of coding efficiency.

1. Introduction

In MVC, the view synthesis prediction is a coding method that generates a virtual image for a certain view and uses the synthesized image for MVC. Since the synthesized image is quite similar to the image to be coded, we can expect the improvement of coding gain. In this document, we propose the multi-view depth generation scheme using segment-based disparity estimation scheme and its coding method. The depth map generation scheme consists of image segmentation, segment-based disparity estimation, and disparity to depth conversion. The multi-view depth video is encoded by H.264/AVC (4:0:0 color format). The virtual image is synthesized by using 3D warping technique. Finally, we apply the synthesized image to the 'VIP P-picture' coding method.

2. Depth Map Generation

2.1 Image Segmentation

Recently, the most powerful stereo matching algorithms use segment-based approach. The segment-based stereo matching algorithms assume that all pixels in one segment have the same disparity value. Therefore, the better segmentation results guarantee the better disparity estimation. In this document, we employ 'Graph-based Image Segmentation' scheme [3]. Fig. 1 and Fig. 2 show the segmented images for 'Exit' and 'Akko&Kayo'.

Fig. 1 Segmented Image for 'Exit' (view 1)

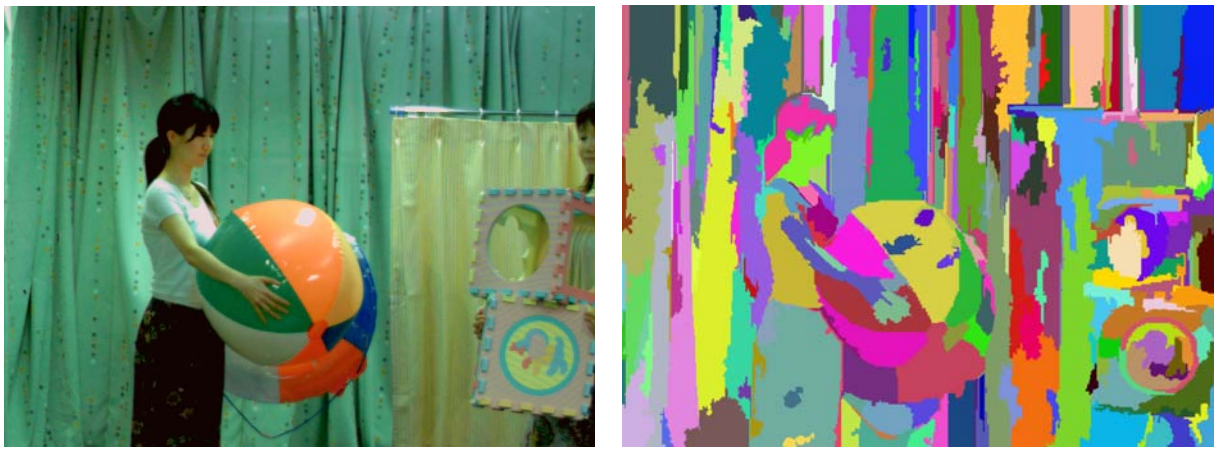

Fig. 2 Segmented Image for 'Akko&Kayo' (view 27)

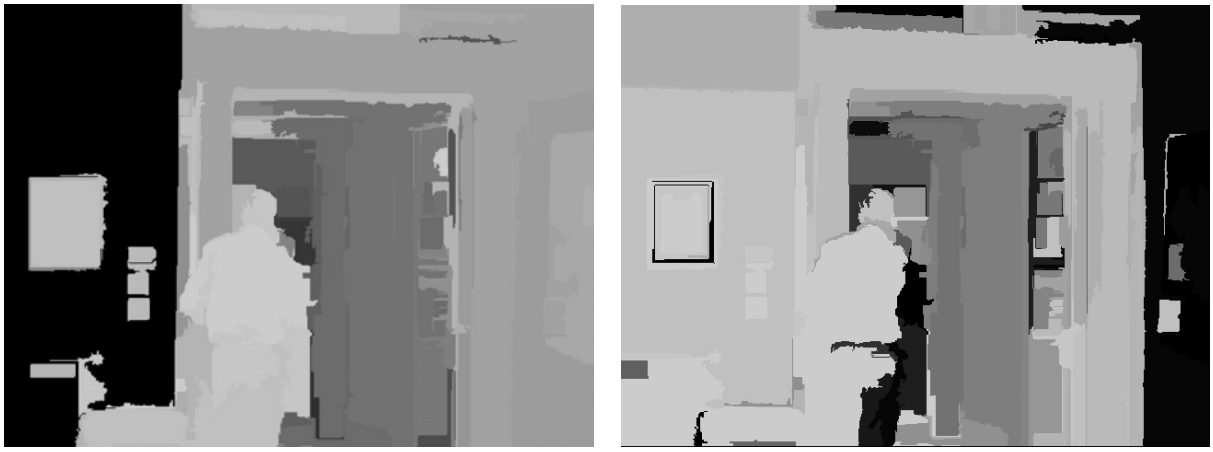

Fig. 3 Disparity Images for 'Exit' (left: right-based, right: left-based)

2.2 Segment-based Disparity Estimation

After the image segmentation, we conduct disparity estimation for each segment. To generate the disparity image for the center view, we conduct disparity estimations twice that use the left and right image respectively. Fig. 3 shows the disparity images that are obtained by using the left and right images respectively. As you can see, the left-image-based disparity image is unstable at the right part and the right-image-based disparity image is unstable at the left part of the image. We make a final disparity map for center view by combining these two images. Fig. 4 shows the final disparity images. For visualization, the disparity values are scaled to be within from 0 to 255.

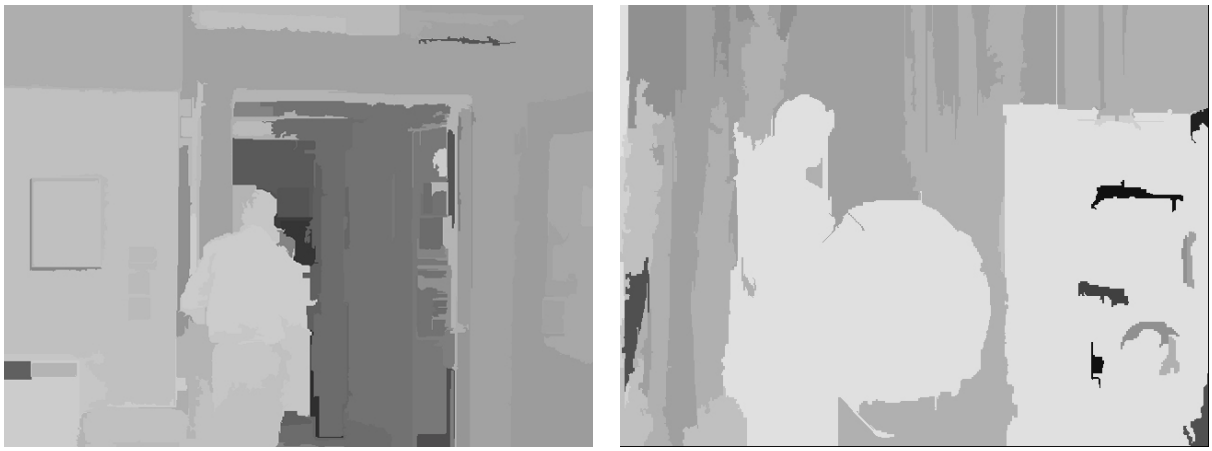

Fig. 4 Final Disparity Images (left: 'Exit' view1, right: 'Akko&Kayo' view 27)

2.4 Disparity to Depth Conversion

To synthesize a virtual image by using a 3D warping technique, we need depth information instead of disparity. In this process, we convert disparity into depth by Eq. (1).

$$Z = \frac{fB}{d} \tag{1}$$

where f means the focal length and B is the distance between baseline. d is the disparity value. Fig. 6 shows the depth images. For visualization, the depth values are scaled within from 0 to 255.

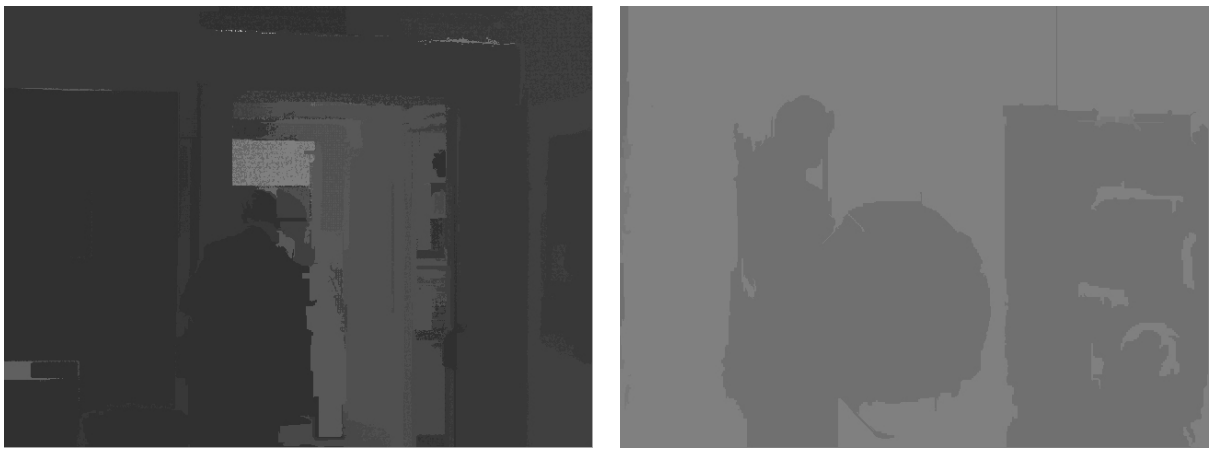

Fig. 6 Depth Images (left: 'Exit' view1, right: 'Akko&Kayo' view 27)

3. Depth Map Coding and View Synthesis

So far, we tried to reconstruct a depth map more precisely. As I mentioned in abstract, since the depth map is a kind of side information we have to encode the depth map to use the view synthesis prediction. In this section, we propose a depth map coding method and introduce a view synthesis scheme. We employ the 4:0:0 color format video coding in H.264/AVC codec. We check the possibility of the down sampling and lossy coding combined with the view synthesis. We exploit the 3D warping technique to synthesize the virtual image. By experiments, we confirmed that the quality of the synthesized image is not seriously degraded even if the depth map is down sampled up to 1/4 pel and encoded with QP value 30. Fig. 7 shows the synthesized images for 'Exit' and 'Akko&Kayo'.

Fig. 7 Synthesized Images (left: 'Exit' view1, right: 'Akko&Kayo' view 27)

4. Experimental Results

According to the description in [4], four QPs (22, 27, 32, and 37) have been used. At this time, we prepared the preliminary result for 'Exit' and 'Akko&Kayo' that are captured under the parallel camera arrangement. The proposed geometrical compensation scheme is implemented on JMVM 1.0 software on CE description [1]. By the experiments, we noticed that the depth map is coded efficiently in terms of the bit rate and quality of the synthesized image when it is down sampled by 1/4pel and coded with QP 30. The coding results of the depth maps are described in Table 1. The bitrate of the depth coding is added to the bitrate for each view. We also show the results by using 'VIP P-picture' coding [2]. These results are preliminary results that are obtained from only three views (first, second, and third views) and their GOP lengths are 12 and 15. Table 2 to 3 show the experimental results compared to the JMVM 4.0 and Fig. 7 shows the RD curves for the experimental results. The proposed 'VIP P-picture' coding scheme achieves better results compared to JMVM 1.0. The proposed view synthesis prediction scheme shows -0.01dB and 0.06 dB of PSNR gain for 'Exit' and 'Akko&Kayo' respectively.

Table 1. Coding Results of Depth Maps

Sequences	Cam. Num.	QP	Avg. PSNR (dB)	Avg. Bitrate (kbps)
Akko&Kayo	27	30	51.30	2.11
	28	30	51.00	2.14
Exit	1	30	49.28	2.14
	2	30	47.08	2.74

Table 2. Performance Evaluation for 'Exit' Sequence

Basic QP	Avg. PSNR (dB)		Avg. Bitrate (kbps)		Gain	Bit saving
	JMVM	Proposed	JMVM	Proposed		
22	40.34	40.35	923.26	924.39	-0.01dB	-0.46%
27	39.00	39.01	409.95	411.41		
32	37.27	37.27	219.81	222.15		
37	35.09	35.15	133.27	135.40		

Table 3. Performance Evaluation for ‘Akko&Kayo’ Sequence

Basic QP	Avg. PSNR (dB)		Avg. Bitrate (kbps)		Gain	Bit saving
	JMVM	Proposed	JMVM	Proposed		
22	42.37	42.38	1305.54	1305.66	0.06dB	1.00%
27	39.66	39.65	710.71	710.87		
32	36.60	36.63	405.75	408.45		
37	33.68	33.68	245.45	244.14		

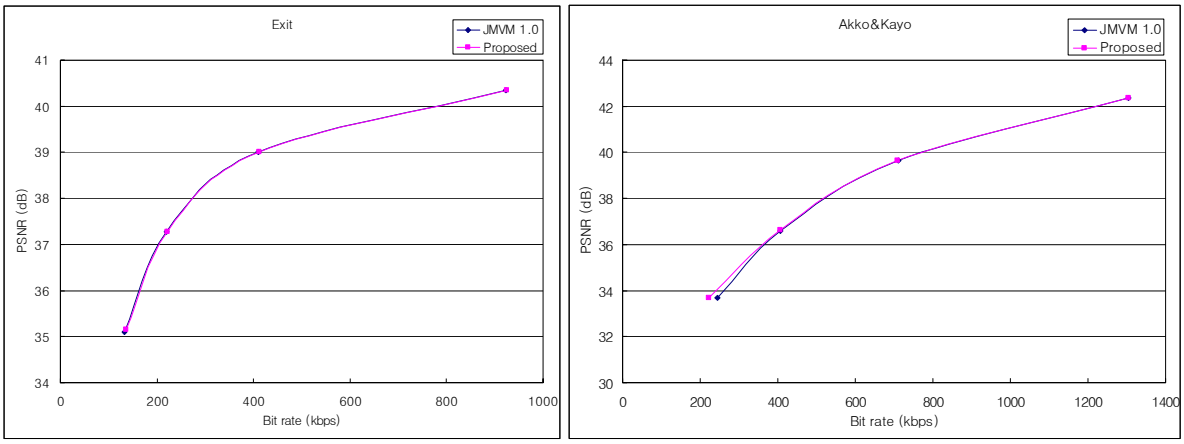

Fig. 7 Rate-Distortion Curves

5. Conclusion

In this document, the depth map generation and depth map coding scheme were proposed to be utilized in the view synthesis prediction. The proposed method reconstructed depth maps which are useful for synthesizing virtual images. By the preliminary experiments, we demonstrated that the proposed view synthesis prediction scheme improved the coding efficiency up to 0.06 dB.

6. Acknowledgements

This work was supported in part by the Information Technology Research Center (ITRC) through the Realistic Broadcasting Research Center (RBRC) at Gwangju Institute of Science and Technology (GIST), and in part by the Ministry of Education (MOE) through the Brain Korea 21 (BK21) project.

7. References

[1] H. Kimata, “CE3: View Interpolation Prediction for MVC,” ITU-T and ISO/IEC JTC1, JVT-W303, San Jose, California, USA, April 2007.

[2] Y.S. Ho, C. Lee, K.J. Oh, B.H. Choi, and J. H. Park, “CE6: View Interpolation Prediction for Multi-view Video Coding,” ITU-T and ISO/IEC JTC1, JVT-W083, San Jose, California, USA, April 2007.

[3] P.F. Felzenszwalb and D.P. Huttenlocher, “Efficient Graph-Based Image Segmentation,” International Journal of Computer Vision, vol. 59(2), pp. 167–181, 2004.

[4] Y. Su, A. Vetro, and A. Smolic, “Common Test Conditions for Multiview Video Coding,” ITU-T and ISO/IEC JTC1, JVT-U211, Hangzhou, China, October 2006.

(Append for Proposal Documents)

JVT Patent Disclosure Form

International Telecommunication Union
Telecommunication Standardization Sector

International Organization for Standardization

International Electrotechnical Commission

Joint Video Team - *Patent Disclosure Form*

(Typically one per contribution and one per Standard | Recommendation)

Please send to:

JVT Rapporteur Gary Sullivan, Microsoft Corp., One Microsoft Way, Bldg. 9, Redmond WA 98052-6399, USA
Email (preferred): Gary.Sullivan@itu.int Fax: +1 425 706 7329 (+1 425 70MSFAX)

This form provides the ITU-T | ISO/IEC Joint Video Team (JVT) with information about the patent status of techniques used in or proposed for incorporation in a Recommendation | Standard. The JVT requires that all technical contributions be accompanied with this form. *Anyone* with knowledge of any patent affecting the use of JVT work, of their own or of any other entity ("third parties"), is strongly encouraged to submit this form as well.

This information will be maintained in a "living list" by the JVT during the progress of their work, on a best effort basis. If a given technical proposal is not incorporated in a Recommendation | Standard, the relevant patent information will be removed from the "living list". The intent is that the JVT experts should know in advance of any patent issues with particular proposals or techniques, so that these may be addressed well before final approval.

This is not a binding legal document; it is provided to the JVT for information only, on a best effort, good faith basis. Please submit corrected or updated forms if your knowledge or situation changes.

This form is *not* a substitute for the formal *Patent Statement and Licensing Declaration form (see http://www.itu.int/ITU-T/ipr/index.html)*, which should be submitted by Patent Holders to the ITU TSB Director, ISO Secretary General, and IEC General Secretary at the time the patent holder believes that the patent is essential to the implementation of a draft or approved Recommendation | International Standard (in addition to the less formal reporting in the earlier proposal/contribution stages of work within the JVT).

Submitting Organization or Person:	
Organization name	Gwangju Institute of Science and Technology (GIST) Korea Electronics and Technology Institute (KETI)
Mailing address	C-404, Department of Information and Communications 1 Oryong-dong, Buk-gu, Gwangju 500-712
Country	Republic of Korea
Contact person	Yo-Sung Ho
Telephone	+82-62-970-2211
Fax	+82-62-970-2247
Email	hoyo@gist.ac.kr
Place and date of submission	Geneva, CH, 29 June – 5 July, 2007

Relevant Recommendation | Standard and, if applicable, Contribution:	
Name (ex: "JVT")	JVT
Title	CE3: Depth Map Generation and Depth Map Coding for MVC
Contribution number	JVT-X048

(Form continues on next page)

Disclosure information – Submitting Organization/Person (choose one box)

☐ 2.0 The submitter is not aware of having any granted, pending, or planned patents associated with the technical content of the Recommendation | Standard or Contribution.

or,

The submitter (Patent Holder) has granted, pending, or planned patents associated with the technical content of the Recommendation | Standard or Contribution. In which case,

☐ 2.1 The Patent Holder is prepared to grant – on the basis of reciprocity for the above Recommendation | Standard – a free license to an unrestricted number of applicants on a worldwide, non-discriminatory basis to manufacture, use and/or sell implementations of the above Recommendation | Standard.

X 2.2 The Patent Holder is prepared to grant – on the basis of reciprocity for the above Recommendation | Standard – a license to an unrestricted number of applicants on a worldwide, non-discriminatory basis and on reasonable terms and conditions to manufacture, use and/ or sell implementations of the above Recommendation | Standard.

Such negotiations are left to the parties concerned and are performed outside the ITU | ISO/IEC.

☐ 2.2.1 The same as box 2.2 above, but in addition the Patent Holder is prepared to grant a "royalty-free" license to anyone on condition that all other patent holders do the same.

☐ 2.3 The Patent Holder is unwilling to grant licenses according to the provisions of either 2.1, 2.2, or 2.2.1 above. In this case, the following information must be provided as part of this declaration:

- patent registration/application number;
- an indication of which portions of the Recommendation | Standard are affected.
- a description of the patent claims covering the Recommendation | Standard;

*In the case of any box **other than 2.0** above, please provide the following:*

Patent number(s)/status _____

Inventor(s)/Assignee(s) _____

Relevance to JVT _____

Any other remarks: _____

(please provide attachments if more space is needed)

(form continues on next page)

Third party patent information – fill in based on your best knowledge of relevant patents granted, pending, or planned by other people or by organizations other than your own.

Disclosure information – Third Party Patents (choose one box)

X 3.1 The submitter is not aware of any granted, pending, or planned patents *held by third parties* associated with the technical content of the Recommendation | Standard or Contribution.

☐ 3.2 The submitter believes third parties may have granted, pending, or planned patents associated with the technical content of the Recommendation | Standard or Contribution.

For box 3.2, please provide as much information as is known (provide attachments if more space needed) – The JVT will attempt to contact third parties to obtain more information:

3rd party name(s) _____

Mailing address _____
Country _____
Contact person _____
Telephone _____
Fax _____
Email _____
Patent number/status _____
Inventor/Assignee _____
Relevance to JVT _____

Any other comments or remarks: