Joint Video Team (JVT) of ISO/IEC MPEG & ITU-T VCEG
(ISO/IEC JTC1/SC29/WG11 and ITU-T SG16 Q.6)
25th Meeting: Shenzhen, CN, 21 – 26 October, 2007

Document: JVT- Y065
Filename: JVT-Y065.doc

Title: **CE3: Depth Map Generation and Virtual View Synthesis**

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

The multi-view depth can be utilized in virtual view synthesis for free viewpoint video (FVV)/free viewpoint television (FTV) and view synthesis prediction (VSP) for multi-view video coding. However, only 'Breakdancers' has a depth among the current test sequences. This document introduces the depth generation scheme and virtual view synthesis using depth data. An analysis of virtual view synthesis and its relationship to depth coding and preprocessing is also reported.

1. Introduction

Recently, the depth is in the spotlight as alternatives for 3DTV and FTV/FTV with multi-view video [1]-[3]. Nevertheless, the importance of study for depth is not recognized as much as multi-view video. The multi-view video is directly obtained from multiple cameras in general, while the multi-view depth artificially calculated from multi-view video. Surely, the depth camera already exists but its price is too expensive and it still has some limitations to represent real depth. Thus, artificial multi-view depth generation is an importation issue. This document introduces multi-view depth generation and rendering results under various conditions. The depth generation is a segment-based approach and uses the 3D warping technique. However, there are remaining problems such as inaccurate object segment and low temporal correlation. This works are still on going. The experiments on depth demonstrate that down-sampling is not good and median filtering can give a benefit to depth coding.

2. Depth Map Generation

2.1 Image Segmentation

In this proposal, we utilize the segment-based depth estimation scheme. The segment-based approaches assume that all pixels in one segment have the same depth value. Therefore, the better segmentation results guarantee the better depth results. In this document, we employ 'Mean Shift Image Segmentation' scheme in [4]. Fig. 1 and Fig. 2 show the segmented images for 'Akko&Kayo' and 'Rena'. 'Mean Shift' algorithm shows the more stable results compared to 'Graph-based mage Segmentation' scheme [5].

Fig. 1 Segmented Image for 'Akko&Kayo' (view 27)

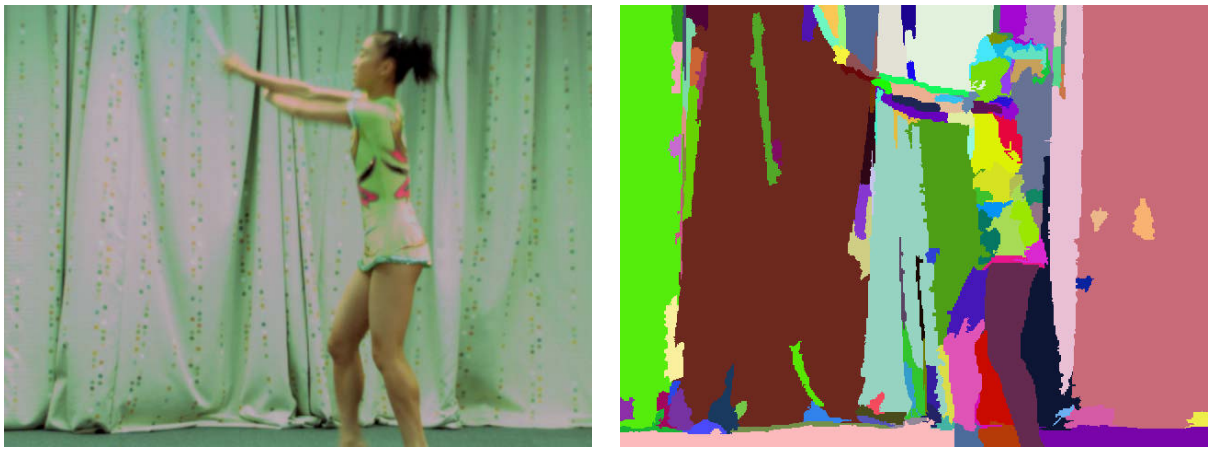

Fig. 2 Segmented Image for 'Rena' (view 50)

2.2 Depth Estimation

After the image segmentation, we conduct depth estimation for each segment. We directly obtain the depth data using 3D warping technique. To generate the depth image for the center view, both left and right views are considered simultaneously. The most well-known matching functions are SD (Squared intensity Difference) and AD (Absolute intensity Difference). Since these functions are not robust to illumination/color changes, we add the absolute gradient difference term to existing AD. Fig. 3 shows the final depth images for 'Akko&Kayo' view 27 and 'Rena' view 50.

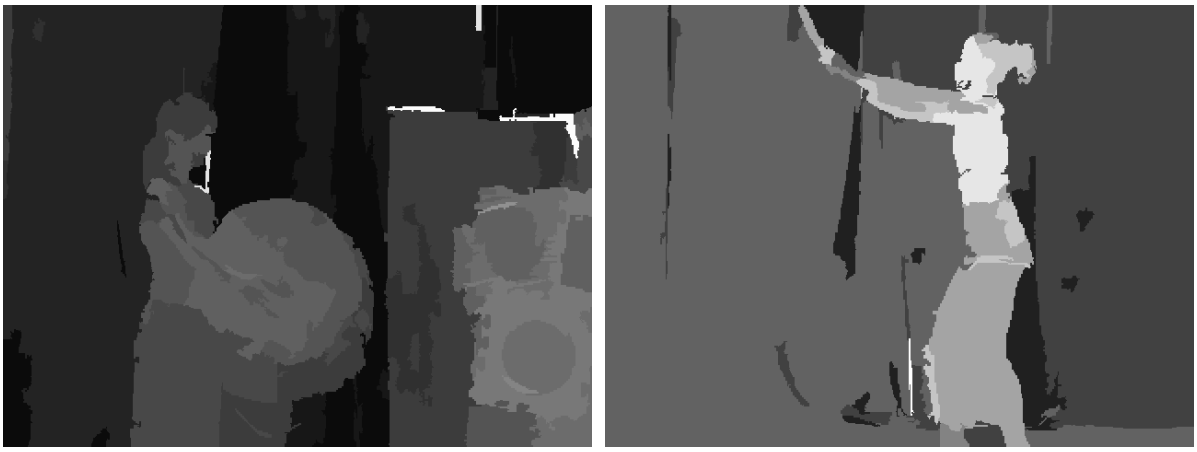

Fig. 3 Final Depth Images (left: 'Akko&Kayo' view 27, right: 'Rena' view 50)

3. Virtual View Synthesis

We can generate the virtual view by using multi-view video and its depth. The depth video is simple but its value is more important compared to conventional video data. The followings are two experiments for depth map coding and virtual view synthesis.

3.1 Down-sampling and View Synthesis

We make experiments on down-sampling for depth. We compare the quality of the synthesized image using the encoded original depth and down-sampled/up-sampled depth with the same bit-rate. For experiments, first ten frames of 'Breakdancers' view 0 and view 2 are used and we use JM12.4 and IPPP structure. Table 1 shows the bit-rate comparison. We intentionally adjust the bit-rate for down-sampled depth with the bit-rate for encoded original depth. Fig. 4 and Fig. 5 show the synthesized images at high bit-rate and low bit-rate, respectively. As shown in figures, the synthesized image using down-sampled/up-sampled depth is worse than the synthesized image using encoded original depth at the same bit-rate. Since down-sampling affects the edges which are important for view synthesis, down-sampling is not a good method for depth coding.

Table 1. Bit-rate Comparison (Breakdancers: view 0 and view 2) [bits, (QP)]

View 0		View 1	
Original	Down-sampling	Original	Down-sampling
1,089,564 (22)	1,011,084 (13)	969,444 (22)	925,392 (13)
590,460 (27)	580,164 (18)	530,928 (27)	521,832 (18)
306,132 (32)	291,576 (24)	280,872 (32)	260,952 (24)
161,640 (37)	158,940 (29)	144,720 (37)	142,560 (29)

(a) encoded original depth (b) original depth (c) down-sampled/up-sampled depth

Fig. 4 Synthesized Images at high bit-rate

(a) encoded original depth (b) original depth (c) down-sampled/up-sampled depth

Fig. 5 Synthesized Images at low bit-rate

3.2 Median Filtering and View Synthesis

We make experiments on median filtering for depth. To check the benefit of median filtering, we compare the quality of the synthesized image using original depth and median-filtered depth in Fig. 6. Fig. 7 shows the original depth and median-filtered depth. In addition, we compare the bit-rates in Table 2. The test conditions are the same as above. The quality of the synthesized image using median-filtered depth is quite similar to the original result, while we can save the bit-rate using median filtering.

Table 2. Bit-rate Comparison (Breakdancers: view 0) [bits, (QP)]

QP	Original	Median filter (7x7)	Bit Saving (%)
22	969,444	776,592	19.89
27	530,928	447,012	15.81
32	280,872	246,888	12.10
37	144,720	136,044	5.60

Fig. 6 Synthesized Images (left: from original depth, right: from 7x7 median-filtered depth)

Fig. 7 Depth Comparison (left: original depth, right: 7x7 median-filtered depth)

4. Conclusion

In this document, multi-view depth generation and rendering results under various conditions were reported. The image segmentation and 3D warping techniques were used for depth generation. However, there are remaining problems such as inaccurate object segment and low temporal correlation. This works are still on going. The experiments on depth map demonstrated that down-sampling is not good and median filtering can give a benefit to depth coding.

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] A. Smolic, K. Müller, P. Merkle, C. Fehn, P. Kauff, P. Eisert, and Thomas Wiegand, “3D Video and Free Viewpoint Video – Technologies, Applications and MPEG Standards”, ICME 2006, IEEE International Conference on Multimedia and Expo, Toronto, Ontario, Canada, July 2006.

[2] A. Smolic, and P. Kauff, “Interactive 3D Video Representation and Coding Technologies”, Proc. of the IEEE, Special Issue on Advances in Video Coding and Delivery, vol. 93, no. 1, Jan. 2005.

[3] P. Merkle, A. Smolic, K. Mueller, and T. Wiegand, “MVC: Experiments on Coding of Multi-view Video plus Depth,” ITU-T and ISO/IEC JTC1, JVT-X064, Geneva, CH, July 2007.

[4] D. Comaniciu and P. Meer, “Mean Shift: A Robust Approach toward Feature Space Analysis,” IEEE: PAMI, vol. 24(5):603–619, May 2002.

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

(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	Shenzhen, CN, 21 – 26 October, 2007
Relevant Recommendation | Standard and, if applicable, Contribution:	
Name (ex: "JVT")	JVT
Title	CE3: Depth Map Generation and Virtual View Synthesis
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: