INTERNATIONAL ORGANISATION FOR STANDARDISATION
ORGANISATION INTERNATIONALE DE NORMALISATION
ISO/IEC JTC1/SC29/WG11
CODING OF MOVING PICTURES AND AUDIO

ISO/IEC JTC1/SC29/WG11
MPEG2009/M17488
April 2010, Dresden, Germany

Source: GIST (Gwangju Institute of Science and Technology)
Status: Report
Title: Results of EE1 on 'Cafe' Sequence
Author: Sang-Beom Lee, Cheon Lee, and Yo-Sung Ho

1. Introduction

This document reports experimental results of the depth estimation on 'Cafe' sequence in response to EE1 of 3D video coding [1]. After we performed the depth estimation by DERS 5.0 assisted by semi-automatic data, we manually refined the erroneous parts of depth videos. The comparison was done by original views and synthesized views.

2. Results of Depth Estimation

2.1. Description of Depth Estimation

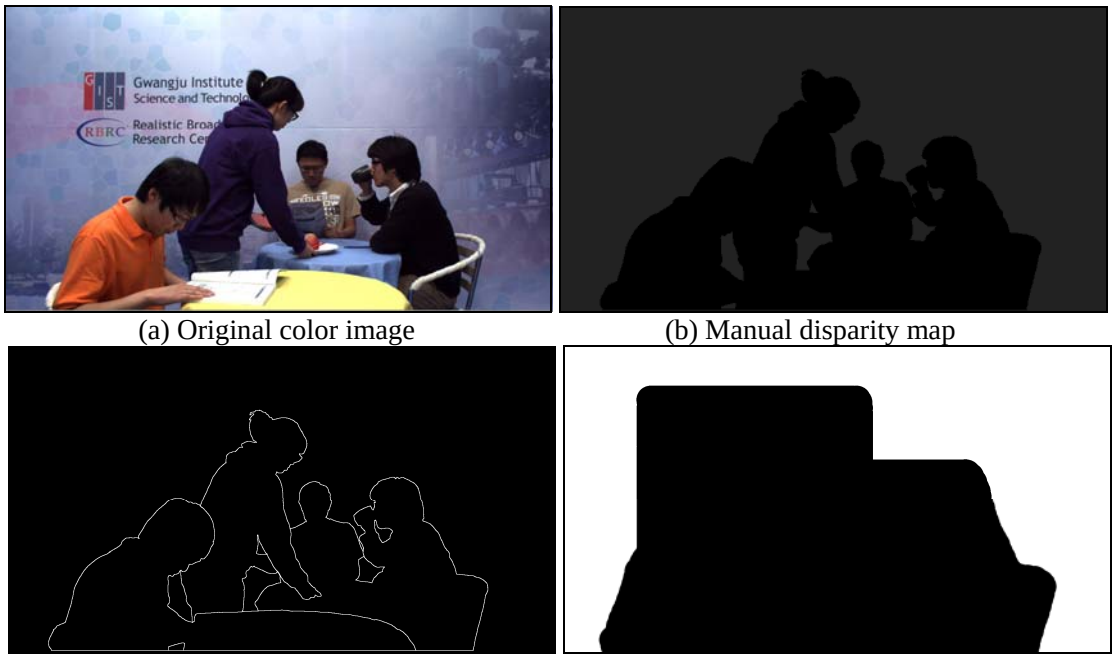
We performed the depth estimation experiment following the description of EE in the document N10720 [1]. The depth estimation was performed by using the semi-automatic mode implemented by Nagoya University. Table 1 shows parameters for depth estimation used in these experiments.

Table 1. Parameters for depth estimation

Search range	30 ~ 90
Precision	Half-pel
Smoothing coefficient	1.0
Matching block	3x3
Depth estimation mode	Semi-automatic (Nagoya)
Temporal weight	0.6
View number	2, 3, 4

2.2. Semi-automatic Mode

As mentioned before, we exploited the semi-automatic mode implemented by Nagoya University. Since supplementary data, manual disparity map, manual edge map, and manual static map, are required for the semi-automatic depth estimation, we made them by Adobe Photoshop CS4. Figure 1 shows the semi-automatic inputs for the 49th frame. For more accurate depth estimation, we added as many inputs as possible; 1st, 49th, 99th, 149th, 199th, 249th, 270th frames in each viewpoint.



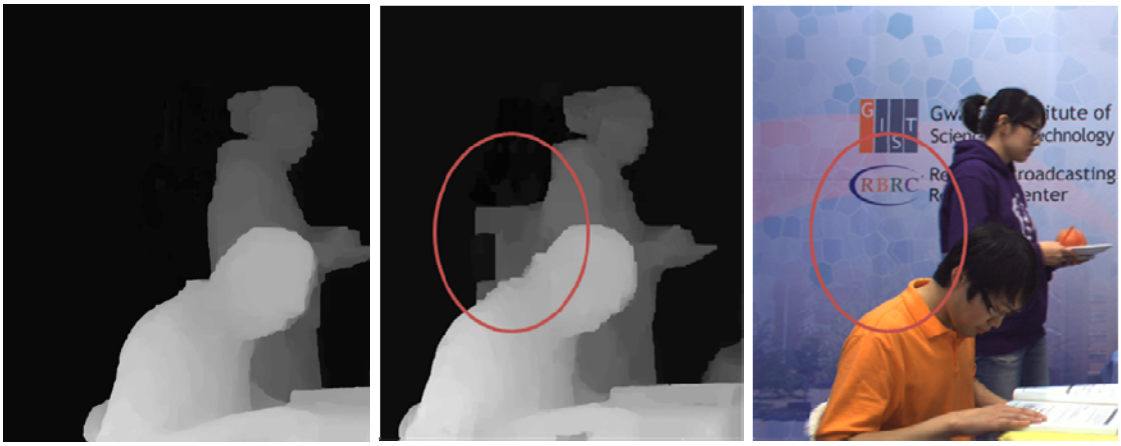
(a) Original color image (b) Manual disparity map

(c) Manual edge map (d) Manual static map

Figure 1. Semi-automatic inputs for the 49th frame of view 3

2.3. Depth Video Refinement

From the depth estimation experiments, we obtained relatively reliable quality of depth videos. However, we found some erroneous regions degrading rendering quality. Figure 2 shows results of view synthesis. As shown in Fig. 2(b) and Fig. 2(c), erroneous regions on background cause rendering errors. Therefore, we manually removed them and refined the depths on background by using the same method as reported in M17490 [2]. Figure 2(d) shows the refined depth map. As a result, the rendering quality of synthesized image as shown in Fig. 2(e) was increased.



(a) Depth map (view 2) (b) Depth map (view 4) (c) Synthesized view (view 3)

(d) Refined depth map (view 4) (e) Synthesized image (view 3)

Figure 2. Results of view synthesis

2.4. Results of View Synthesis

The final depth map and the synthesized image are shown in Fig. 3. The average PSNR between the original color video and the synthesized video for the view 3 was 34.1635 dB. From the result, we noticed that these depth videos guaranteed the good quality of synthesized video.

(a) Depth map (b) Synthesized image

Figure 3. Results of view synthesis

3. Conclusion

In this document, we reported the depth estimation and view synthesis results for ‘Cafe’ sequence. In conclusion on EE1, the depth videos guarantee the good quality for the view synthesis in terms of average PSNR of the synthesized views and the subjective quality. We are ready to demonstrate these results for the viewing test in this meeting.

4. Acknowledgements

This research was supported by the MKE(The Ministry of Knowledge Economy), Korea, under the ITRC(Information Technology Research Center) support program supervised by the NIPA(National IT Industry Promotion Agency)" (NIPA-2010-(C1090-1011-0003)).

5. References

[1] ISO/IEC JTC1/SC29/WG11 “Description of Exploration Experiments in 3D Video Coding,” N11095, Jan. 2010.

[2] ISO/IEC JTC1/SC29/WG11 “Modified Depth Data of ‘Newspaper’ Sequence,” M17490, April 2010.