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

**ISO/IEC JTC1/SC29/WG11**
**MPEG2011/M19277**
**January 2011, Daegu, Korea**

**Source: GIST (Gwangju Institute of Science and Technology)**
**Status: Report**
**Title: MPEG 3DV EE1 Results on Café**
**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]. In order to obtain depth videos, we performed several steps; depth estimation using DERS 5.0 assisted by semi-automatic data, bilateral depth filtering, manual refinements. We confirmed that the resultant depth data have improved subjective quality.

## 2. Results of Depth Estimation

In the last meeting, we had presented the refined depth videos and synthesized views, however, those data still generated visual artifacts [2]. As a result, the evaluation on Café sequence was ‘Slight artifacts’. In order to solve this problem, we modified them with additional steps, post-processing with a bilateral filter as described in the last document [3]. The overall process is depicted in Fig. 1.

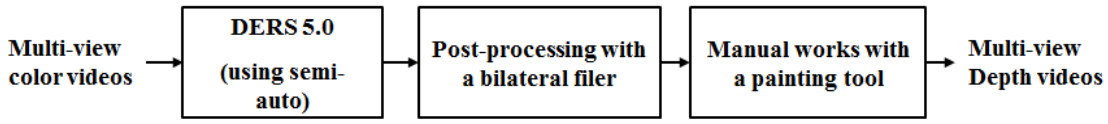
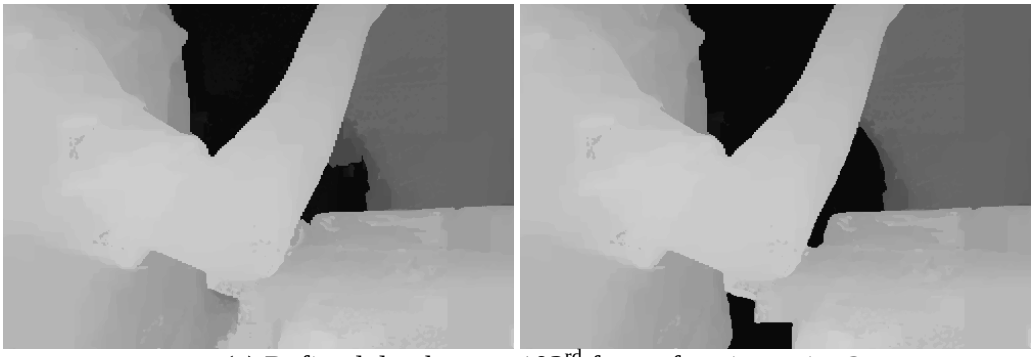



Fig. 1. Depth generation for Café sequence

### 2.1. Depth Video Refinement

Although the depth videos were refined by the post-processing step as described above, erroneous depth values are still remained. In order to remove them, we manually refined depth map again using Adobe Photoshop CS4. The most erroneous regions are found near the foremost man and woman. Therefore, we modified them from $50^{th}$ frame to $199^{th}$ frame for three viewpoints. Figure 2 shows results of manual depth refinement. As shown in Fig. 2, erroneous depth values are cleaned out.



(a) Refined depth map: 103rd frame for viewpoint 2

(b) Refined depth map: 148th frame for viewpoint 2

(c) Refined depth map: 145th frame for viewpoint 4

Fig. 2. Manually refined depth data

## 2.2. Results of View Synthesis

We performed view synthesis for the 3rd view using two color and depth videos for the 2nd and 4th viewpoint. Then, we compared the original view to the synthesized view in terms of PSNR. In order to avoid the common hole problem, we used the modified hole filling method with the boundary noise removal method. Figure 3 shows the rendering quality comparison. The average value of the previous depth data reported in 93rd meeting is 33.14 dB and the average value of the refined depth data is 33.13 dB. Notice that although the PSNR values are almost the same, the rendering quality was improved.

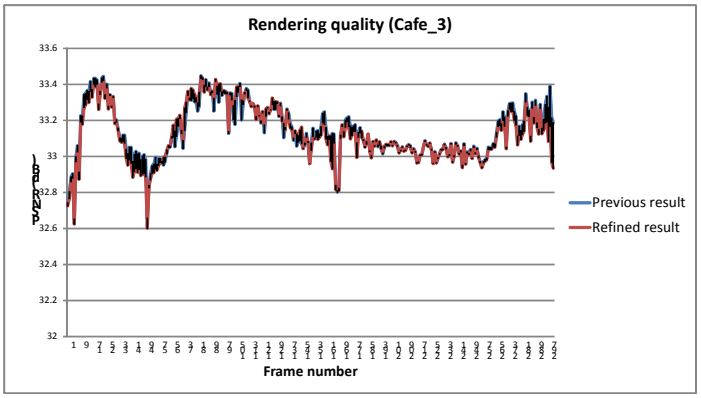



Fig. 3. Rendering quality of the synthesized images

The final depth map and the synthesized image for the 3rd view are shown in Fig. 4. From the results, we noticed that the refined depth videos guaranteed the good rendering quality of synthesized views.

(a) Depth map (91st frame)

(b) Synthesized image (91st frame)

(c) Depth map (251st frame)

(d) Synthesized image

Fig. 4. Results of view synthesis

## 3. Conclusion

In this document, we have reported the depth estimation and view synthesis results for ‘Cafe’ sequence. In conclusion on EE1, we have confirmed that the depth videos guaranteed good quality in terms of average PSNR of the synthesized views and subjective quality. In this meeting, we are ready to demonstrate our experimental results for the viewing session.

## 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-2011-(C1090-1011-0003)).

## 5. References

[1] ISO/IEC JTC1/SC29/WG11 “Description of Exploration Experiments in 3D Video Coding,” N11630, Oct. 2010.

[2] ISO/IEC JTC1/SC29/WG11, “Results of EE1 on ‘Cafe’ Sequence,” M18512, Oct. 2010.

[3] ISO/IEC JTC1/SC29/WG11, “Common-hole Filling for Boundary Noise Removal in VSRS,” m18514, Oct. 2010.