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/M16670
April 2009, London, England

Source: GIST (Gwangju Institute of Science and Technology)[1]
KETI (Korea Electronics Technology Institute)[2]
Status: Report
Title: Results of EE1 on 'Pantomime' Sequence
Author: Cheon Lee[1], Jiho Park[2] and Yo-Sung Ho[1]

1. Introduction

This document reports experimental results of the depth estimation software on 'Pantomime' sequence as a response to EE1 in 3D video coding [1]. We tested the performance of the DERS 4.0 especially about the newly integrated semi-automatic depth estimation techniques.

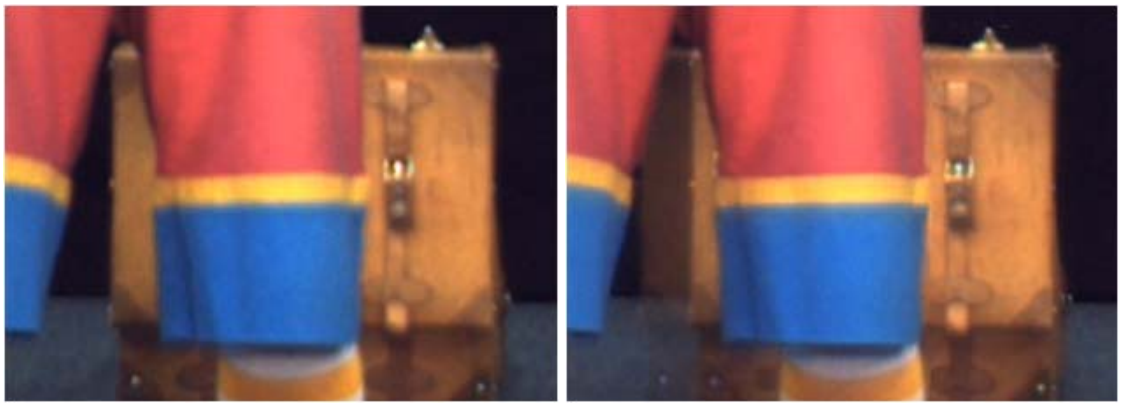
2. Results for EE1: Semi-automatic Depth Estimation

During the Maui meeting, experts agreed that the performance of semi-automatic is considerably improved comparing with that of the conventional automatic method. In this exploration experiments, we tested two semi-automatic methods, which are proposed by Nagoya University [2] and ETRI [3]. At first, we generated the depth video using Nagoya's method with provided auxiliary data since there is no available keyframe of 'Pantomime' sequence. Thereafter, we sampled the depth video for the use of key-frame, and then we ran the semi-automatic depth estimation using ETRI's method with the sampled depth maps. As described in the EE document [1], we obtained two neighboring views, view_39 and view_41, to synthesize the image of view_40 using the optimized parameter combination as described in Table 1.

Table. 1. Optimal parameter set for automatic depth estimation

Left View	Right View	Precision	Smoothing Coefficient	Temporal Enhancement	Matching Block	Image Segmentation
39	41	Half-pel	1.0	On	3x3	Off

The obtained depth results by using the semi-automatic method of Nagoya University showed improvements. The major artifact on 'Pantomime' sequence is the ghosting effects around the bag at the backside as shown in Fig 1(b), which have been reduced in the results of the semi-automatic method. The figure 1(c) and 1(d) show the reduced visual artifacts.



(a) Original image (b) Resultant image of automatic method

(c) Nagoya's semi-automatic (d) ETRI's semi-automatic

Fig. 1. Visual artifacts in the 130th frame at view_40

Table. 2. Synthesis results of semi-automatic method of Nagoya University

(unit: dB)		Automatic	Semi-automatic method					
			Temporal Weight					
			0.5	0.6	0.7	0.8	0.9	1.0
Avg. PSNR		35.24	35.84	**35.86**	35.86	35.85	35.84	35.84
PSPNR	AvPSNR_Y	35.67	35.67	**35.69**	35.69	35.68	35.67	35.67
	Av_T_PSNR_Y	54.51	54.45	**55.13**	55.06	54.66	54.45	54.43
	Av_S_PSNR_Y	40.86	40.87	**40.90**	40.90	40.88	40.87	40.86

The semi-automatic method of Nagoya University employs a parameter of 'TemporalWeight'. To find the optimum setting, we tested six values such as {0.5, 0.6, 0.7, 0.8, 0.9, 1.0}. After obtaining the depth videos, we generated intermediate view images using VSRS 4.1 and evaluated the quality of them using conventional PSNR calculation and PSPNR tool distributed by Zhejiang University on 31st May 2009. Table 2 shows the synthesis results for 100 frames using Nagoya's semi-automatic method.

The best PSNR value is 35.86 which is the result by using the temporal weight 0.6. Noticing the Av_T_PSNR_Y in the table, the temporal quality has been improved considerably.

In order to run the ETRI's semi-automatic method, we extracted one frame in every 10 frames from the Nagoya's results. Since there are two parameters in the method, we set six combinations as described in Table 3. 'O' represents the *MovingObjectsBSize* and 'M' represents the *MotionSearchBSize* in the configuration file respectively. As a result, the best PSNR value is 35.94 dB as shown in the Table 3; it is higher than that of Nagoya's results.

Table. 3. Synthesis results of semi-automatic method of ETRI

(unit: dB)		Automatic	Semi-automatic method					
			Parameters					
			O0, M0	O1, M0	O0, M1	O1, M1	O0, M2	O1, M2
Avg. PSNR		35.24	**35.94**	35.94	35.93	35.93	35.93	35.93
PSPNR	AvPSNR_Y	35.67	**35.77**	35.77	35.76	35.76	35.76	35.76
	Av_T_PSNR_Y	54.51	**55.94**	55.87	55.97	55.92	55.92	55.92
	Av_S_PSNR_Y	40.86	**41.02**	41.02	41.01	41.01	41.00	41.01

Consequently, the ETRI's semi-automatic method showed the best results if we use the sampled keyframes from the Nagoya's semi-automatic method

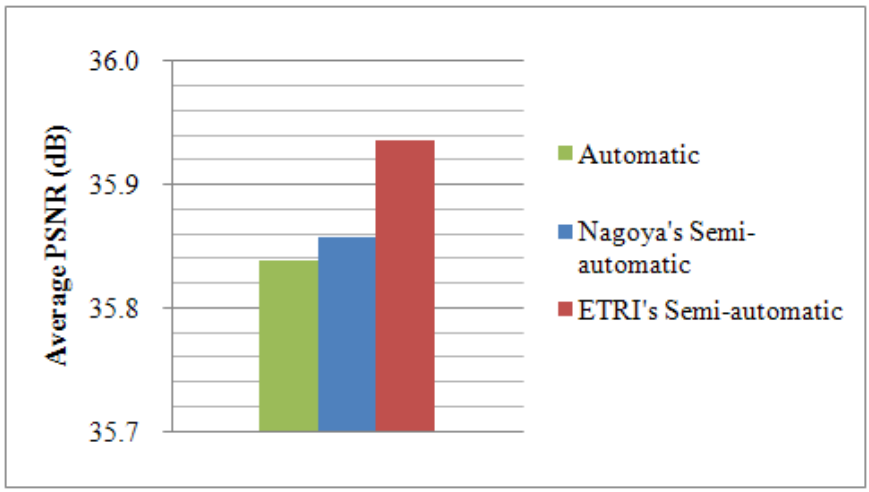


Fig. 2. Objective quality of the synthesized view for EE1

2.1. Observations of EE1

- Nagoya's semi-automatic method considerably reduces the ghosting effect in the synthesized image.

- ETRI's semi-automatic method increases the quality of depth map when the key-frame extracted from the Nagoya's result is used.

3. Conclusion

In this document, we reported that the performance of two semi-automatic depth estimation reference methods is improved significantly. In conclusion on EE1, the depth data on 'pantomime' sequence generated by semi-automatic method is acceptable for reference.

4. Acknowledgements

This work was supported in part by ITRC through RBRC at GIST (IITA-2009-C1090-0901-0017).

5. References

[1] ISO/IEC JTC1/SC29/WG11 "Description of Exploration Experiments in 3D Video Coding", N10360, February 2009.
[2] ISO/IEC JTC1/SC29/WG11 "Semi-automatic Depth Estimation for FTV", m16391, April 2009.
[3] ISO/IEC JTC1/SC29/WG11 "Depth Estimation Algorithm in SADERS 1.0", m16411, April 2009.