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

ISO/IEC JTC1/SC29/WG11
MPEG2008/M15847
October 2008, Busan, Korea

Source: GIST (Gwangju Institute of Science and Technology)
Status: Report
Title: EE1: Results on 'Pantomime' Sequence using Nagoya SW
Author: Yun-Suk Kang, Cheon Lee and Yo-Sung Ho

1. Introduction

This document reports the results of the exploration experiment on depth estimation and view synthesis of 'Pantomime' sequence [1][2]. We obtained depth maps and intermediate view images by using Nagoya SW [3]. From the experiments, we found a fine parameter set that confirms the quality of the synthesized view image. By using this parameter set, we finally obtained whole frames of reference view's depth maps and synthesized views for the narrow and medium baseline cases.

2. Description of Experiment

To obtain the fine parameter set, we experimented on ten frames of the sequence with changing six parameters: smoothing coefficient, baseline basis, precision, search level, filter, and matching method. The other factors related to the disparity are fixed during the experiment as shown in Table 1. We generated depth maps of view 38 and view 41 and then synthesized images at view 39 and view 40 for ten frames. Finally, we obtained a fine parameter set by calculating and comparing average PSNR values of ten frames of each view.

Table 1. Input parameters for depth estimation SW

Parameters	Value for 10 frames	Value for whole frames
Minimum Value of Disparity Search Range	2	2
Maximum Value of Disparity Search Range	10	20
Minimum Value of Disparity Range	0	0
Maximum Value of Disparity Range	10	25

After determining the parameters, we increased the disparity search range and the disparity range. Then, we generated whole frames of depth maps of view 38 and view 41

for the narrow baseline case and view 37 and view 42 for the medium baseline case. With these depth maps, we synthesized whole frames of images at view 39 and view 40 for both of two baseline cases.

2.1. Depth Map Estimation

To find the fine parameter set, we changed the first parameter and fixed the others to their default values. After finding the best with respect to PSNR of the synthesized image for the first parameter, we changed the second parameter and fixed the others. By performing this process sequentially, we finally determined the fine parameter set for depth estimation. We experimented on the narrow baseline case for parameter determining and the default values are indicated in Table 2.

Table 2. Default values of input parameters for depth estimation

Parameters	Default value	Detail
Smoothing Coefficient	0	Smoothing coefficient to compute depth map
Baseline Basis	0	Minimum baseline
Precision	1	Integer-pel
Search Level	1	Integer-pel
Filter	0	Bi-linear
Matching Method	0	Conventional

2.2. View Synthesis

We used Nagoya view synthesis SW of version 'VS2_3', and it requires Z_{near} and Z_{far} values. For ten frames, we used small disparity search range and disparity range, and obtained that Z_{near} and Z_{far} values were 5297.544529 and 8221.650623, respectively. However, we increased the disparity search range and the disparity range for whole frames, and only Z_{near} decreased to 3454.569912.

With the obtained depth values, we synthesized images at view 39 and view 40 by using the estimated depth maps for both of the narrow baseline case and the medium baseline case. We used the same parameters for the precision and the filter which are used in depth estimation.

3. Experimental Results

From the experimental results, we noticed that a few parameters have effects on depth maps, and these depth maps also influences in the quality of the synthesized images. The following results show the process to determine the fine parameter set.

3.1. Smoothing Coefficient

We generated depth maps with variant smoothing coefficients. Then, we obtained synthesized views by using these depth maps. Figure 1 shows that depth maps of view 38 and synthesized images of view 39 of each first frame when coefficients are 0, 1.0, 2.0,

3.0, 4.0 and 5.0. When the coefficient is zero, depth values of both foreground and background are noisy. However, as the coefficient increases, depth maps are smoothed and have less noise. Depth of the background also has the unique value when the coefficient is increasing.

Therefore qualities of synthesized views are also increased as the coefficient is increased. Table 3 shows average PSNR values of ten frames of synthesized views in accordance with the smoothing coefficient. Since we obtained the highest PSNR value with the coefficient 5.0, this value is selected for the fine parameter set.

(a) Smoothing coefficient 0 (b) Smoothing coefficient 1.0

(c) Smoothing coefficient 2.0 (d) Smoothing coefficient 3.0

(e) Smoothing coefficient 4.0 (f) Smoothing coefficient 5.0

Fig. 1. Depth maps of view 38 and synthesized images at view 39 according to various smoothing coefficients

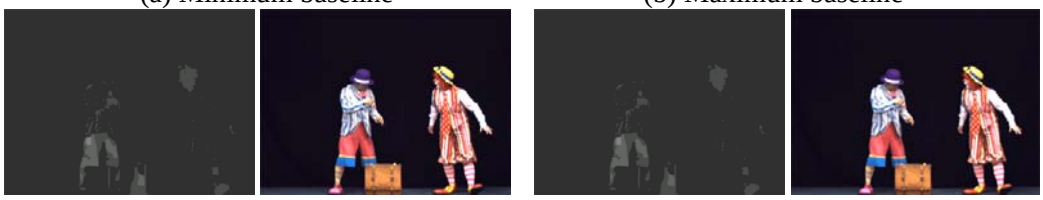
Table 3. Average PSNR according to the smoothing coefficient

Synthesized View	Smoothing Coefficient					
	0	1.0	2.0	3.0	4.0	5.0
View 39	22.8868	34.3030	36.0001	36.4500	36.6037	36.6727
Veiw 40	23.8568	33.6268	35.0018	35.4072	35.6044	35.6995

3.2. Baseline Basis

There are four types of baseline bases: minimum baseline, maximum baseline, left baseline, and right baseline. We generated depth maps of two reference views for each baseline and synthesized images with fixing the first parameter that is the smoothing coefficient. However, we noticed that this baseline basis has no effect on the quality of view synthesis. Thus, we just selected minimum baseline for the fine parameter set. Figure 2 and Table 4 show the results according to baseline bases.

(a) Minimum baseline (b) Maximum baseline



(c) Left baseline (d) Right baseline

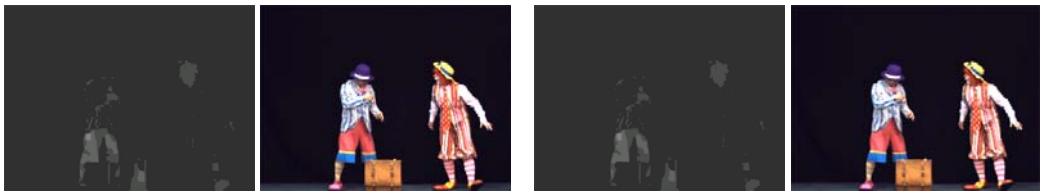
Fig. 2. Depth maps of view 38 and synthesized images at view 39 according to various baseline bases

Table 4. Average PSNR according to the baseline basis

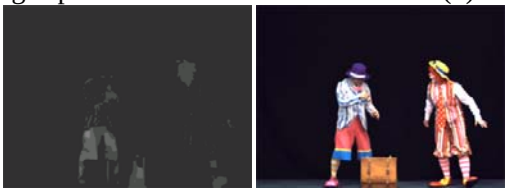
Synthesized View	Baseline Basis			
	Minimum	Maximum	Left	Right
View 39	36.6727	36.6727	36.6727	36.6727
View 40	35.6995	35.6995	35.6995	35.6995

3.3. Precision

We experimented on three pixel precisions which are integer-pel, half-pel, and quarter-pel with the previously selected two parameters. However, this precision also has no effect on quality of view synthesis. Thus, we just selected the integer-pel precision for the fine parameter set. The results are shown in Fig. 3 and Table 5.



(a) Integer-pel (b) Half-pel



(c) Quarter-pel

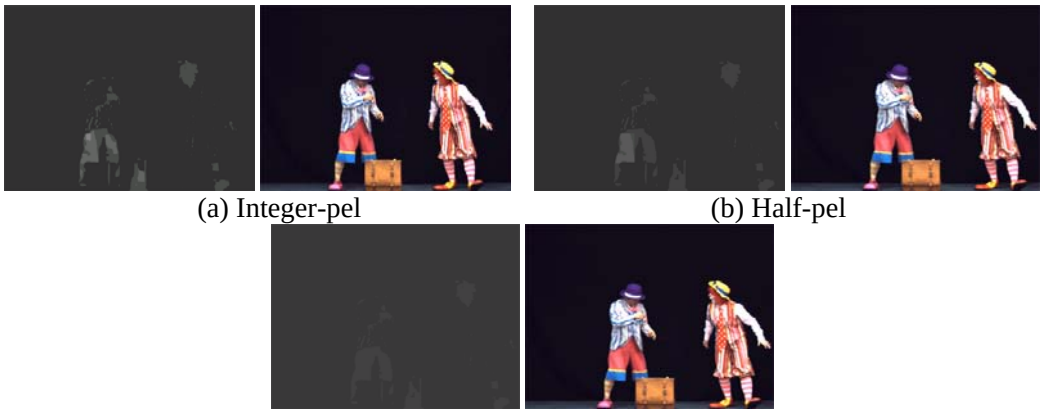
Fig. 3. Depth maps of view 38 and synthesized images at view 39 according to various precisions

Table 5. Average PSNR according to the precision

Synthesized View	Precision		
	Integer-pel	Half-pel	Quarter-pel
View 39	36.6727	36.6727	36.6727
View 40	35.6995	35.6995	35.6995

3.4. Search Level

Three search levels influence on the view synthesis quality. When we used the quarter-pel search level for depth estimation with the selected parameters, average PSNR values was slightly increased than the other cases in view 40. While the integer-pel search level had the highest average PSNR in view 39. However, the PSNR differences among three search levels are quite small. Therefore, we selected the quarter-pel search level because the smaller search level usually means more accuracy and we can evaluate the effect of filters. Figure 4 and Table 6 show the results according to search levels.



(a) Integer-pel (b) Half-pel

(c) Quarter-pel

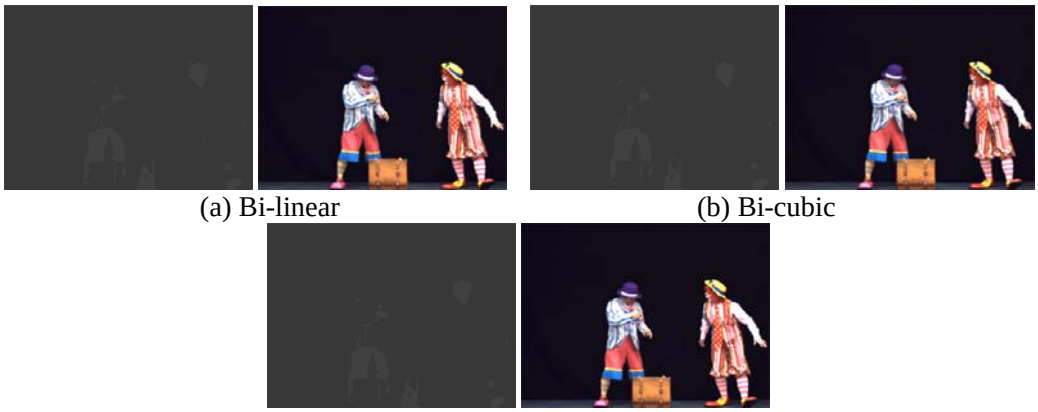
Fig. 4. Depth maps of view 38 and synthesized images at view 39 according to various search levels

Table 6. Average PSNR according to the search level

Synthesized View	Search Level		
	Integer-pel	Half-pel	Quarter-pel
View 39	36.6727	36.4453	36.4965
View 40	35.6995	35.7809	35.8313

3.5. Filter

Three filters did not have effects on the results. As shown in Fig. 5 and Table 7, we could not find any differences among the results according to the filter. Therefore, we selected the bi-linear filter as the fifth parameter with previously selected ones.



(a) Bi-linear (b) Bi-cubic

(c) MPEG-4 AVC 6-tap

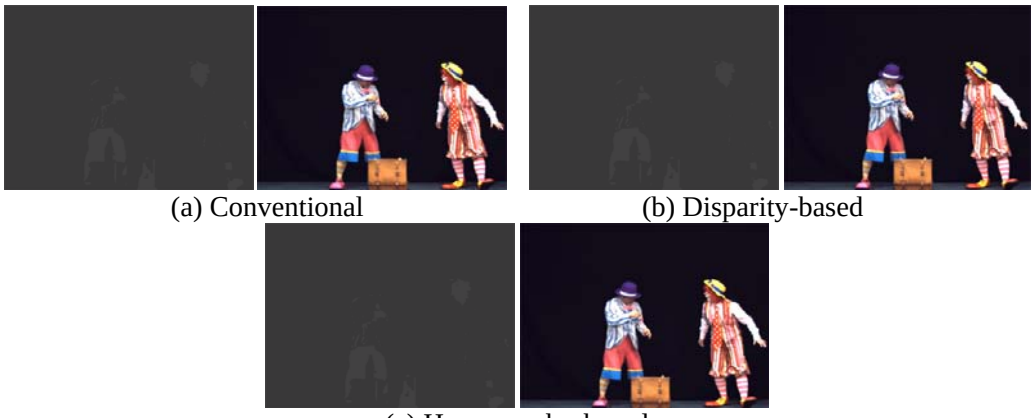
Fig. 5. Depth maps of view 38 and synthesized images at view 39 according to various filters

Table 7. Average PSNR according to the search level

Synthesized View	Filter		
	Bi-linear	Bi-cubic	MPEG-4 AVC 6-tap
View 39	36.4965	36.4965	36.4965
View 40	35.8313	35.8313	35.8313

3.6. Matching Method

The last parameter is the matching method. There are conventional method, disparity-based method, and homography-based method. However, we noticed that this matching method has little effect on the quality of the view synthesis with the previously selected five parameters. In addition, there were no difference between the disparity-based and the homography-based methods as shown in Fig. 6 and Table 8. Thus, we just selected the conventional method for the matching.



(a) Conventional (b) Disparity-based

(c) Homography-based

Fig. 6. Depth maps of view 38 and synthesized images at view 39 according to various matching methods

Table 8. Average PSNR according to the matching method

Synthesized View	Matching Method		
	Conventional	Disparity-based	Homography-based
View 39	36.4965	36.5052	36.5052
View 40	35.8313	35.8330	35.8330

3.7. Fine Parameter Set

Finally, we obtained the fine parameter set shown in Table 9. By using this parameter set, we generated whole frames of depth maps and synthesized view images for both of the narrow and medium baseline cases. The average PSNR values of the synthesized images at view 39 and view 40 are shown in Table 10.

Table 9. Fine parameter set for Nagoya depth estimation SW

Parameters	Determined value	Detail
Smoothing Coefficient	5.0	Smoothing coefficient to compute depth map
Baseline Basis	0	Minimum baseline
Precision	1	Integer-pel
Search Level	4	Quarter-pel
Filter	0	Bi-linear
Matching Method	0	Conventional

Table 10. Average PSNR of the synthesized views

Baseline	View	Average PSNR
Narrow	View 39	32.6099
	View 40	32.3945
Medium	View 39	30.2485
	View 40	31.2323

4. Conclusion

We experimented on depth estimation and view synthesis of ‘Pantomime’ sequence by using Nagoya SW. We generated depth maps of two reference views and synthesized two images for both of the narrow baseline and the medium baseline cases with changing parameters. By comparing average PSNR values, we selected parameters that confirm the high quality and defined them as the fine parameter set. As the results, smoothing factor 5.0, minimum baseline, integer-pel precision, integer-pel search level, bi-linear filter, and conventional matching method are selected. When we experimented on depth estimation and view synthesis of whole frames of ‘Pantomime’ sequence with the fine parameter set, we obtained average PSNR values about 32.5 dB for narrow baseline and 30.75 dB for medium baseline.

5. Acknowledgements

This work was supported in part by ITRC through RBRC at GIST (IITA-2008-C1090-0801-0017).

6. References

[1] ISO/IEC JTC1/SC29/WG11 “Description of Exploration Experiments in 3D Video Coding,” N9991, July 2008.

[2] ISO/IEC JTC1/SC29/WG11 “1D Parallel Test Sequences for MPEG-FTV,” M15378, April 2008.

[3] ISO/IEC JTC1/SC29/WG11 “Reference Software for Depth Estimation and View Synthesis,” M15377, May 2008.