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/M16950
October 2009, Xian, Chnia

Source: GIST (Gwangju Institute of Science and Technology)
Status: Proposal
Title: Test Sequence for 3-D Video Coding
Author: Eun-Kyung Lee and Yo-Sung Ho

1. Introduction

This document describes a test sequence for 3-D video coding, provided by Gwangju Institute of Science and Technology (GIST). In the 88th Maui MPEG meeting, we presented a test sequence 'Delivery' [1] as a response of 'Call for Contribution of Test Material' [2]. As the exploration experiment (EE) progressed, we have noticed that the 'Delivery' sequence has some problems, such as depth inaccuracy along object boundaries and not acceptable subjective quality in synthesized images using depth maps. In this contribution, we provide an enhanced 3-D test sequence 'Delivery'. In the following, we describe the specification of the test sequence and its camera parameters.

2. Specification of the Test Sequence

The sequence was recorded by using our hybrid camera system, which is shown in Fig. 1. The system consists of five multi-view cameras and one depth camera. There is one sync-generator, sending a synchronization signal; this signal is distributed to all PCs. We captured one sequence with the 1-D parallel camera arrangement, where the camera interval is 20cm.

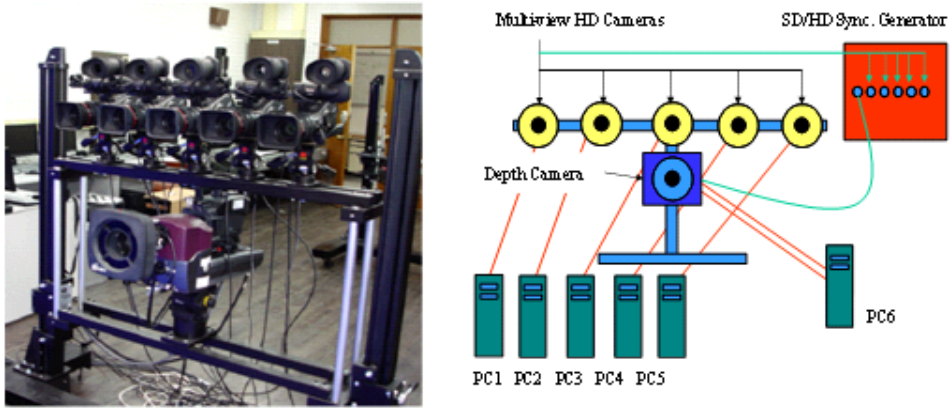

Figure 1: Hybrid camera system

Table 1: Test sequence by GIST

Sequence	Image Property	Camera Arrangement	Depth Specification
Delivery	1920x1080, 30fps (rectified)	5 cameras with 20cm spacing; 1D parallel	Z_near: 2220mm Z_far: 4200mm

Table 1 shows the specification of the test data. The original picture size is 1920(H) x 1080(V), and the frame rate is 30 frames/sec. The duration of the original sequence is 5 seconds. For the sequence, multi-view rectification and color correction were applied. Then, the images were converted to YUV 4:2:0 format. The Z_near is 2220mm and Z_far is 4200mm for correct interpretation of depth values.

3. Foreground and Background Separation

As mentioned above, since the depth sequence is inaccurate in the boundary of depth map, we add a foreground and background separation function to generate an improved depth map. Figure 2 shows the result of foreground and background separation.

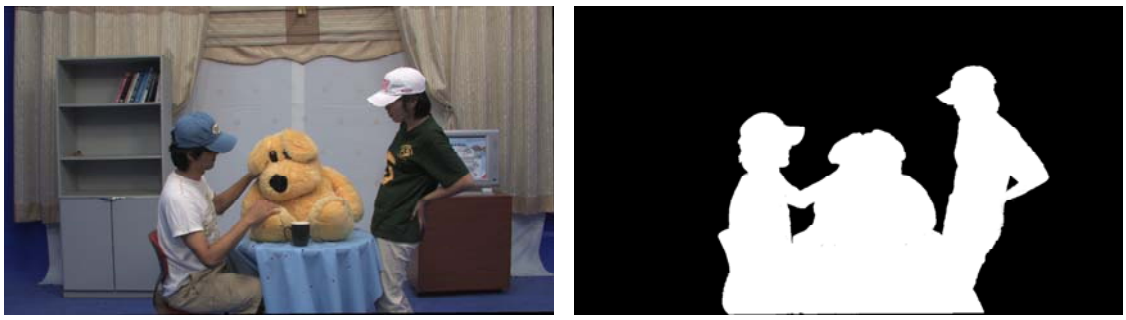

Figure 2: Foreground and background separation in 'Delivery'.

4. Occlusion and Disocclusion Detection

In general, mismatching problems occur in occlusion and disocclusion regions. This problem is caused especially by the sparse camera configuration. The occlusion and disocclusion regions are unreliable regions with respect to the stereo matching operation. Since some regions exist in the current camera position, but not in the other camera positions, mismatching problems occur in these areas. To reduce the mismatched regions, we detect the occlusion and disocclusion regions using the initial depth information. Figure 3 show the results of occlusion and dicocclusion detection.

Occlusion regions Disocclusion regions

Figure 3: Occlusion and disocclusion detection in 'Delivery'.

5. Camera Parameters

The definition, format, and validation of camera parameters of the sequence is described by referring the document, ISO/IEC JTC1/SC29/WG11 N9595, "Call for Contributions on 3D Video Test Material" [2].

5.1. Intrinsic Parameters

Table 2 shows the rectified intrinsic parameters of the sequence.

Table 2: Rectified intrinsic matrix

2059.383013	0.0	928.543706
0.0	2057.422796	503.546216
0.0	0.0	1.0

5.2. Extrinsic Parameters

Table 3 and Table 4 show the rectified extrinsic parameters of the sequence.

Table 3: Rectified rotation matrix

0.9997	0.0014	0.0251
-0.0004	0.9992	-0.0396
-0.0252	0.0396	0.9989

Table 4: Rectified translation vector for each camera

	translation		
Camera	t_x	t_y	t_z
0	50.26184	32.58658	23.03765
1	30.42681	32.58658	23.03765
2	10.59178	32.58658	23.03765
3	-9.24325	32.58658	23.03765
4	-29.07829	32.58658	23.03765

6. Experimental Results

To check the validity of this sequence, we tested the quality of synthesized image using a reference software. In the following subchapters, we demonstrate the experimental results.

6.1. Color Segmentation

To obtain depth maps in the multiview image, we first segment the multiview image by a mean-shift color segmentation algorithm [3]. Figure 4 shows the segmented image for 'Delivery' on view 3.

Figure 4: Enhanced test sequence 'Delivery'.

6.2. Depth Generation

The proposed sequence consists of 5 view videos. We generated 3-D video for five views. Figure 5 shows a snapshot of the generated 3-D video on view 3.

Figure 5: Test sequence 'Delivery'.

6.3. View Synthesis

Using the generated 3-D video, we synthesized two intermediate views. Figure 6 is a demonstration of synthesized image. The left image is the original image for viewpoint 3 and the right image is its synthesized image using 'VSRS' software. As you can see, the quality of the synthesized image is good. The average PSNR value of the synthesized images is higher than 30 dB.

Figure 6: Original image of view 3(left) and synthesized image (right)

7. Conclusion

We explained the test sequence 'Delivery' for 3-D video which consists of five views. Using the foreground/background separation and occlusion/disocclusion detection, we presented the improved sequence and the corresponding depth map. To evaluate the quality of the sequence, we conducted experiments using the released software. The results were acceptable. You can download them at our web site:
The ftp address, User ID and Password will be announced at the meeting.

8. 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-2009-C1090-0902-0017).

9. References

[1] ISO/IEC JTC1/SC29/WG11 "3-D Test Sequence – Multi-view Video and Depth Map," M16396, April 2009.

[2] ISO/IEC JTC1/SC29/WG11 N9595, "Call for Contributions on 3D Video Test Material," January 2008.

[3] D. Comaniciu and P. Meer, "Mean shift: a robust approach toward feature space analysis," *IEEE Trans. on Pattern Analysis and Machine Intelligence*, vol. 24, no. 4, pp. 603–619, 2002.