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/M16674
June 2009, London, UK

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

1. Introduction

This document represents the test sequence provided by Gwangju Institute of Science and Technology (GIST). In 88th Maui meeting, we presented a test sequence 'Delivery' [1] as a response of 'Call for Contribution of Test Material' [2]. As EE progressed, we noticed that 'Delivery' sequence has challenging properties such as depth inconsistency of boundary and not acceptable for subjective quality. Therefore, we propose temporally 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, as 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.

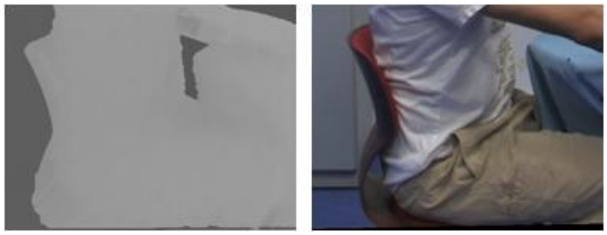
Table 1: Test sequence by GIST

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 length 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. Temporal Consistency

As mentioned before, since the depth sequence is temporally inconsistent, we add a temporal function to generate the improved depth map. Figure 3 shows the result of temporally enhanced depth map with the previous depth map. As shown in Figure 3(a), we noticed that the depth sequences have inconsistent depth near the chair, whereas the depth sequence in Figure 3(b) has consistent depth. We reduced the flickering artifacts of virtual views without any degradation of visual quality from the experimental results.



(a) Original depth map and virtual view

(b) Enhanced depth map and virtual view

Figure 3: Enhanced depth map and virtual view in 'Delivery'.

4. 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].

4.1. 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

4.2. 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

5. Experimental Results

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

5.1. 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 2 shows segmented image for 'Delivery' on view 3.

Figure 2: Enhanced test sequence 'Delivery'.

5.2. Depth Generation

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

Figure 3: Enhanced test sequence 'Delivery'.

5.3. View Synthesis

Using the generated 3-D video, we synthesized two intermediate views. Figure 4 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 synthesized images is higher than 30 dB.

Figure 4: Original image of view 3(left) and synthesized image (right)

6. Conclusion

We have explained the test sequence 'Delivery' for 3-D video which consists of 5-views. Using the temporal consistency enhancement, we have presented the improved sequence and the corresponding depth map. To evaluate the quality of the sequence, we conducted the 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.

7. 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 IITA(Institute for Information Technology Advancement) (IITA-2009-C1090-0902-0017).

8. 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.