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/M16065
February 2009,Lausanne,CH

Source: GIST (Gwangju Institute of Science and Technology)
Status: Proposal
Title: Additional Test Sequence for 3D Video
Author: Cheon Lee, Jae-Il Jung, Yun-Suk Kang and Yo-Sung Ho

1. Introduction

This document represents an additional test sequence provided by GIST (Gwangju Institute of Science and Technology). In 84th Antalya meeting, we provided a test sequence 'Newspaper' [1] as a response of 'Call for Contribution of Test Material' [2]. As EE progressed, we noticed that 'Newspaper' sequence has challenging properties such as larger disparity range, a lot of motion, and depth impression. However, there were several defects on the sequence. For example, due to color mismatch problem, flickering artifacts were generated in the synthesized images. In order to alternate with the sequence, we propose a new test sequence named 'Study' with 10 views.

Figure 1: 10-view camera system

2. Specification of Test Sequence

The sequence was captured by 10-view multiview camera rig as shown in Fig. 1. The camera configuration is specified in Table 1. General setup and environment are similar with 'Newspaper' sequence except for the number of cameras. Figure 2 shows a snapshot of the proposed sequence. The used cameras are 'Point Grey Research Flea' with 1/3-inch 'Sony CCD IEEE-1394' camera. The original picture size is 1024(H) x 768(V), and the frame rate is 25fps. The number of frames is 250. For the sequence, rectification [3] and color correction [4] were applied. Then, the images were converted to YUV 4:2:0.

Table 1: General Information of ‘Study’ Test

Sequence	Image Property	Camera Arrangement
Study	1024x768, 25fps (rectified, color corrected)	10 cameras with 5cm spacing; 1D parallel

Figure 2: Original test sequence ‘Study’

3. Camera Parameters

3.1. Intrinsic Parameters

The intrinsic parameters for each camera are represented as follows:

Format of Intrinsic Parameter

focal_length_x	0	principal_point_x
0	focal_length_y	principal_point_y
0	0	1

We adjusted every intrinsic parameter to one common camera in rectification. Table 2 shows the common intrinsic parameters for all cameras.

Table 2: Common Intrinsic Parameter

1400.681	0.0	492.9781
0.0	1400.043	302.5895
0.0	0.0	1.0

3.2. Extrinsic Parameters

Table 3 and Table 4 show extrinsic parameters of the sequence. Table 3 is the common rotation parameter for all cameras. Since every camera has been rectified into 1D parallel, all cameras are aligned to one direction. Table 4 shows the translation vectors. In these tables, "i" means *i*th camera, and t_k means *k*th coordinate of translation vector for each camera.

Rotation matrix and translation vector are described by the following formats. The external transformation is performed by rotation **R** and translation **t** from a 3D point M_c in the camera coordinate system to the related 3D point M_w in the world coordinate system by

$$M_w = \mathbf{R} M_c + \mathbf{t} \tag{1}$$

The external transformation from a 3D point M_w in the world coordinate system to the related 3D point M_c in the camera coordinate system is performed by

$$M_c = \mathbf{R}^{-1} M_w - \mathbf{R}^{-1}\mathbf{t} \tag{2}$$

The extrinsic parameters for each camera are represented as follows:

Format of Rotation Matrix

R_{11}	R_{12}	R_{13}
R_{21}	R_{22}	R_{23}
R_{31}	R_{32}	R_{33}

Format of Translation Vector

t_1
t_2
t_3

The following 3D-point $\mathbf{M}_w$ is given in world coordinates. The 3-dimensional world point is mapped into a 2-dimensional camera point by:

$$s * \mathbf{m} = \mathbf{A} * (\mathbf{R}^{-1}\mathbf{M}_w - \mathbf{R}^{-1}\mathbf{t}) \tag{3}$$

where *s* is an arbitrary scaling factor to make the third coordinate of **m** equal to one. Note that **R** is the rotation of the camera relative to the world coordinates and **t** is given in world coordinates.

Table 3: Common Rotation Matrix

1	0	0
0	-1	0
0	0	-1

Table 4: Rectified translation vector for each camera

	t_1	t_2	t_3
Cam 1	96.749279	0	0
Cam 1	45.325846	0	0
Cam 2	-6.097587	0	0
Cam 3	-57.521020	0	0
Cam 4	-108.944453	0	0
Cam 5	-160.367886	0	0
Cam 6	-211.791319	0	0
Cam 7	-263.214752	0	0
Cam 8	-314.638185	0	0
Cam 9	-366.061618	0	0

4. Experimental Results

In order 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.

4.1. Depth Estimation

The proposed sequence consists of 10 view videos. Among them, we estimated depth video of two views, view 4 and view 6, with following parameters in Table 5. Figure 3 shows a snapshot of an estimated depth image on view 4.

Table. 5. Parameters for depth estimation

Target View	3		6	
Reference Views	Left Ref.	Right Ref.	Left Ref.	Right Ref.
	2	4	5	7
Disparity Range	8 ~ 26			
Search Range	8 ~ 25			
Smoothing Coefficient	3.0			
Depth Type	1 (Depth from the origin of 3D space)			
Baseline Basis	1 (maximum baseline)			
Precision	4 (Quater-pel)			
Search Level	4 (Quater-pel)			
Filter	2 (MPEG-4 AVC 6-tap)			
Matching Method	1 (Disparity-based)			
Temporal Enhancement	1 (on)			
	Threshold		1.50	

Fig. 3. Example of depth image; (left) original color image, (right) depth image

4.2. View Synthesis

Using the estimated depth video, we synthesized two intermediate views. Figure 5 is a demonstration of synthesized image. The left image is the original image for viewpoint 4 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 28 dB.

Fig. 4. Original image of view 4(left) and synthesized image (right)

5. Conclusion

We have explained a new test sequence 'Study' for 3D video which consists of 10-view. All views are rectified and color corrected. In order to evaluate the quality of the sequence, we conducted the experiments using the released software. The results were acceptable for a test material. We are ready to release the sequence for a test material on 3D video. In conclusion, we suggest that MPEG 3DV group needs a discussion for the sequence.

6. Acknowledgements

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

7. References

[1] ISO/IEC JTC1/SC29/WG11 “Multiview Video Test Sequence and Camera Parameters,” M15419, April 2008.

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

[3] Yun-Suk Kang and Yo-Sung Ho, "Geometrical Compensation Algorithm of Multi-view Images for Arc Multi-Camera Array," Lecture Notes in Computer Science (LNCS), 5353, pp. 543-552, 2008.

[4] Jae-Il Jung and Yo-Sung Ho, "Color Correction Method Using Gray Gradient Bar For Multi-view Camera System," International Workshop on Advanced Image Technology (IWAIT), pp. 1-6, 2009.