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/m15419

Archamps, France, April 2008

Title: **Multiview Video Test Sequence and Camera Parameters**

Source: **Gwangju Institute of Science and Technology (GIST)**

Author: **Yo-Sung Ho, Eun-Kyung Lee, and Cheon Lee**

1. Introduction

This document describes the test sequence provided by Gwangju Institute of Science and Technology (GIST). We present the multiview test sequence, and its intrinsic and extrinsic camera parameters. We have constructed a nine-view camera system and captured the test sequence with 1-D parallel camera arrangement. 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 nine-view camera system, as shown in Fig. 1. The system consists of one master PC and eight slave PCs. 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 5cm. The viewing zone is 40 cm in length.

Figure 1: Nine-view camera system

Table 1: Test sequence by GIST

Sequence	Image Property	Camera Arrangement	Depth Specification
Newspaper	1024x768, 30fps (rectified)	9 cameras with 5cm spacing; 1D parallel	Z_near:3600mm Z_far:7800mm

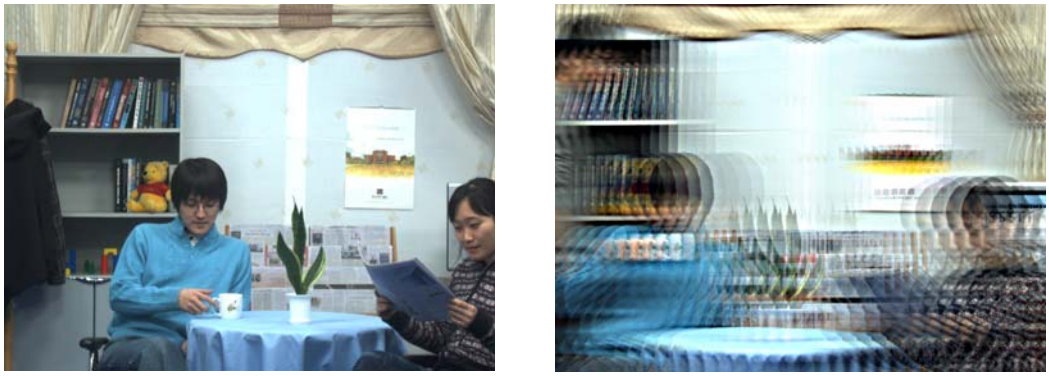
Figure 2: Original test sequence "Newspaper"

The cameras are Point Grey Research Flea with 1/3-inch Sony CCD IEEE-1394 camera. Table 1 shows the specification of the test data. The original picture size is 1024(H)x768(V), and the frame rate is 30 frames/sec. The length of the original sequence is 10 seconds. For the sequence, rectification was applied. Then, the images were converted to YUV 4:2:0 format. The Z_near is 3600mm and Z_far is 7800mm for correct interpretation of depth values.

3. Image Rectification

Image rectification is a process that makes epipolar lines of two images captured at different position parallel each other. Then, the vertical coordinates of all image points of two images become identical and there remain only the horizontal disparities.

In order to rectify multi-view images, we calculate the baseline that minimizes the sum of squared distances to the camera centers. Then we assume ideal virtual cameras on the baseline with the same interval. Ideal virtual cameras have the same intrinsic parameters. The horizontal axis of each camera is parallel to the baseline and the principal axis of each camera is perpendicular to the baseline. We can obtain rectified multi-view images by transformations from unrectified images of original cameras to rectified images of virtual cameras. Figure 3(a) is the result of the rectified input image and Fig. 3(b) shows the overlapped images with identical vertical coordinates and the equal interval.



(a) **(b)**

Figure 3: Rectified test sequence "Newspaper".

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" [1].

4.1. Intrinsic Parameters

The intrinsic parameters for each camera are represented as follows:

Format of Intrinsic Parameter

focal_length_x	skew_parameter	principal_point_x
0	focal_length_y	principal_point_y
0	0	1

Table 2 is the original intrinsic parameters of the each camera. Table 3 shows the rectified intrinsic parameters of the sequence.

Table 2: Original intrinsic matrix

I	0	1	2	3	4	5	6	7	8
A_{11}	2526.23	2530.44	2529.75	2525.72	2538.89	2525.87	2521.99	2536.97	2529.59
A_{12}	0	0	0	0	0	0	0	0	0
A_{13}	616.10	596.90	636.72	592.47	594.02	591.96	613.29	619.04	609.27
A_{21}	0	0	0	0	0	0	0	0	0
A_{22}	2525.26	2528.57	2525.42	2521.91	2533.97	2518.43	2513.39	2524.38	2513.03
A_{23}	420.03	439.54	408.05	434.57	435.49	417.72	399.84	402.91	395.18
A_{31}	0	0	0	0	0	0	0	0	0

A_{32}	0	0	0	0	0	0	0	0	0
A_{33}	1	1	1	1	1	1	1	1	1

Table 3: Rectified intrinsic matrix

2869.494052	0.0	386.560376
0.0	2862.706449	422.693294
0.0	0.0	1.0

4.2. Extrinsic Parameters

Table 4 and Table 5 show extrinsic parameters of the sequence. Table 4 is the original extrinsic parameters of the sequence. Table 5 and Table 6 show the rectified extrinsic parameters. In these tables, "i" means *i*th camera, R_{mn} means (m, n) coordinate of rotation matrix for each camera for the sequence, and t_k means *k*th coordinate of translation vector for each camera for the sequence. Rotation matrix and translation vector is 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

To verify that camera parameters and world to camera equation are consistent a 3D-point $\mathbf{M}_w$ is projected on all camera planes. The following 3D-Point $\mathbf{M}_w$ is given (see below on how to get this) in world coordinates. The 3-dimensional world point is now mapped into a 2-dimensional camera point by:

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

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 4: Original rotation matrix and translation vector for each camera

i	0	1	2	3	4	5	6	7	8
R_{11}	0.99376	0.99395	0.99227	0.99366	0.99364	0.99317	0.99293	0.99272	0.99247
R_{12}	0.00802	0.00401	0.00548	-0.00428	-0.00374	0.00365	0.00083	-0.00089	0.01121
R_{13}	0.11122	0.10968	0.12392	0.11230	0.11249	0.11659	0.11864	0.12036	0.12189
R_{21}	0.01388	0.00896	0.01209	0.00153	0.00194	0.01001	0.00830	0.00583	0.01792
R_{22}	-0.99855	-0.99896	-0.99854	-0.99866	-0.99872	-0.99849	-0.99800	-0.99843	-0.99837
R_{23}	-0.05199	-0.04467	-0.05263	-0.05163	-0.05041	-0.05397	-0.06253	-0.05559	-0.05403
R_{31}	0.11065	0.10939	0.12345	0.11237	0.11254	0.11622	0.11836	0.12022	0.12108
R_{32}	0.11065	0.04538	0.05372	0.05147	0.05031	0.05476	0.06307	0.05589	0.05581
R_{33}	-0.99243	-0.99296	-0.99089	-0.99233	-0.99237	-0.99171	-0.99096	-0.99117	-0.99107
t_1	-110.91	-146.86	-234.07	-246.01	-287.89	-343.12	-407.12	-460.73	-521.30
t_2	151.02	132.68	152.49	145.23	143.48	154.36	184.20	168.70	168.92
t_3	2725.407	2724.24	2715.36	2715.71	2710.23	2706.49	2700.72	2695.62	2688.38

Table 5: Rectified rotation matrix

0.999399	0.0134989	0.031922
0.015158	-0.998514	-0.052334
0.031168	0.052787	-0.998119

Table 6: Rectified translation vector for each camera

i	0	1	2	3	4	5	6	7	8
t_1	108.75	64.16	19.56	-25.02	-69.61	-114.21	-158.80	-203.39	-247.98
t_2	149.53	149.53	149.53	149.53	149.53	149.53	149.53	149.53	149.53
t_3	2725.61	2725.61	2725.61	2725.61	2725.61	2725.61	2725.61	2725.61	2725.61

5. **Conclusion**

We have explained the test sequence provided by GIST. Using our nine-view camera system, we have obtained the sequence named “Newspaper”. You can download them at our web site:

The ftp address, User ID and Password will be announced at the meeting.

Acknowledgements

This research was supported by the Ministry of Knowledge Economy, Korea, under the Information Technology Research Center support program supervised by the Institute of Information Technology Advancement.

References

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