DEVELOPMENT OF HYBRID VIDEO THEODOLITE AND APPLICATION FOR 3D MODELING OF PANORAMIC OBJECTS

Hirofumi CHIKATSU, Tetsuji ANAI
Tokyo Denki University, Japan
Department of Civil Engineering
chikatsu@g.dendai.ac.jp

Working Group V/ 5

KEY WORDS: Hybrid Video Theodolite, Camera Calibration, Panoramic Objects, Ortho Image, 3D Modeling.

ABSTRACT

Digital archives or VR museums for structures of architectural significance and objects of importance to the World's cultural heritage have recently received more attention. However, there are some issues for effective operational of digital archives or VR museums. These problems include real-time imaging, spatial data acquisition, and modeling. In particular, efficient spatial data acquiring techniques in a site should be developed or investigated.
With this objective, and for multiple applications such as human motion analysis, auto-tracking, real-time positioning and so on, Hybrid Video Theodolite (HVT) system was developed by the authors consisting of 6 parts: sensor, pan head and tilt body, imaging, recording, control and monitor. The most remarkable points of this HVT system are its ability to obtain synchronized stereo image sequences and rotation parameters in real-time while tracking a moving object. As for further additional point of this system, automated camera calibration without target can be achieved.
This paper describes the HVT system, and investigates the effectiveness of this system for 3D modeling of panoramic objects in architecture and archeology.

1 INTRODUCTION

The authors have been concentrating on developing a video theodolite system consisting of a CCD camera, a theodolite and a video recorder where the camera rotation parameters can be determined in real-time while recording a moving object. The current values of the rotation parameters are continuously superimposed on image frames and thus recorded as a part of the image data (Chikatsu and et al., 1994). The effectiveness of the video theodolite system for dynamic analysis of human motion have been demonstrated (Chikatsu and Murai, 1995, Chikatsu and et al., 1996, Anai and Chikatsu, 1999) and the application of the video theodolite system to the ski jump (Chikatsu and et al., 1997) and panoramic imaging (Nakano and Chikatsu, 1999) also have been demonstrated.
Furthermore, the authors constructed the stereo vision system based on the video theodolite system fitting the stereo adapter to the lens of the CCD camera of the video theodolite (Kakiuchi and Chikatsu, 1998). Right and left image are taken as the odd field and even field by the liquid crystal shutter. The remarkable points of the stereo vision system are its ability to obtain the synchronized stereo image sequences and camera rotation parameters in real-time, and 3D modeling for indoor space became possible (Kakiuchi and Chikatsu, 2000).
There are still, however, some issues which need to be resolved before this system may become operational. These problems include, necessity of increased speed for tracking and long base line for stereo image for large base-depth ratios.
With these motives, the Hybrid Video Theodolite (HVT) System was developed by the authors consisting of 6 parts: sensor, pan head and tilt body, imaging, recording, control and monitor. After describing the HVT system, the effectiveness of this system for 3D modeling of panoramic objects in architecture and archeology are shown in this paper.

2 HYBRID VIDEO THEODOLITE (HVT) SYSTEM

The HVT system was developed based on the video theodolite system which have been developing by the authors, for multiple applications such as human motion analysis, auto-tracking, real-time positioning and so on (Anai and Chikatsu 2000, Yoshida and Chikatsu 2000).
The HVT system consists of 6 parts: sensor, pan head and tilt body, imaging, recording, control and monitor. The sensor part consist of 3 color CCD cameras and laser range finder mounted on pan head, and pan head mounted on

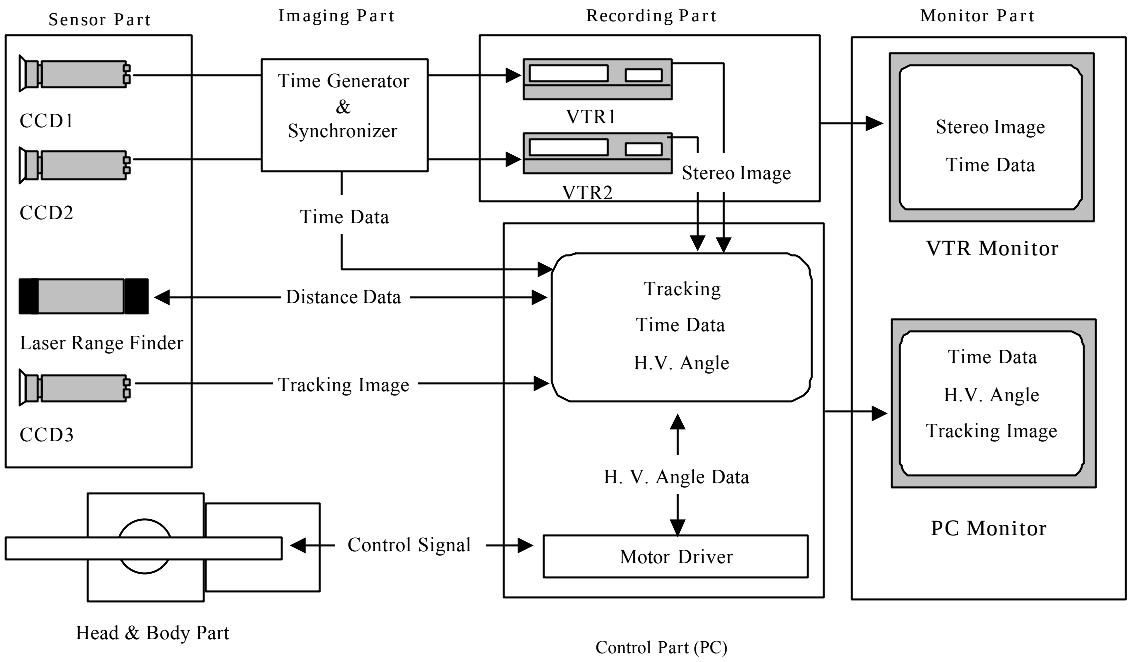


Figure 1. System configuration of HVT system

tilt body. Figure1 shows the system configuration and Figure 2 shows the appearance. The remarkable features of this system are as follows:

・ Sensor part

+ Stereo image sequences are taken by the CCD 1 and 2 camera (Sony XC-711, 768H×493V)
+ Tracking is performed by the CCD3 camera (Sony XC-75, 768H×493V).
+ Object distance above datum is measured by the laser range finder.

・ Head and body Part

+ Pan and tilt rotation angles are controlled with 10 seconds.
+ Maximum rotation speed of pan and tilt are 45°per second.
+ Maximum base line is 5m.
+ Rotation range of the head part is 360°.
+ Rotation range of the body part is 180°(-90°~90°).
+ Rotation of the cameras is synchronous to the rotation for the pan head and tilt body.

・ Imaging part

+ 30 frame images per second are taken based on NTSC format.
+ Synchronization for stereo image is performed.
+ Time for each image is generated at interval of 1/100 seconds.

・ Recording part

+ Synchronized stereo image sequences and times are continuously superimposed on image frames and thus recorded on video recorder.

・Control part

+ Pan and tilt rotation angles are controlled.
+ Time, pan and tilt rotation angles are connected, and thus recorded on PC.
+ Laser range finder is controlled.

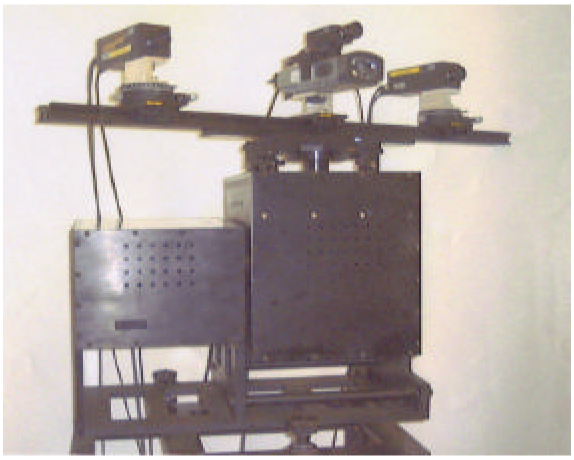


Figure 2. Appearance of HVT

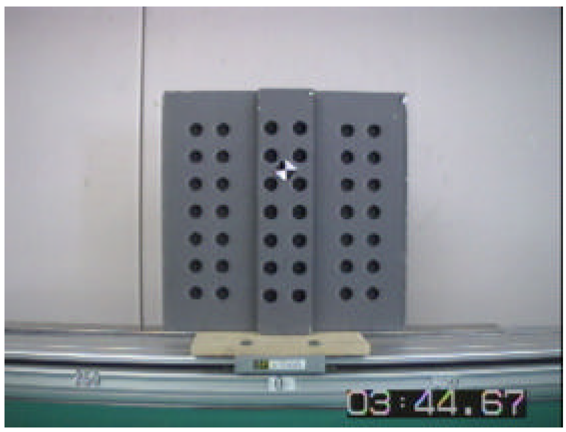


Figure 3. VTR monitor

+ Tracking is controlled.
・ Monitor part
+ Stereo image with times are displayed on VTR monitor.
+ Time data, pan and tilt rotation angles and tracking image are displayed on PC monitor

Then, the synchronized stereo image sequences, times, pan and tilt rotation angles can be obtained in real-time while tracking a moving object. Figure 3 shows the VTR monitor image. The number of right lower corners (03' 44.67") shows the superimposed time.

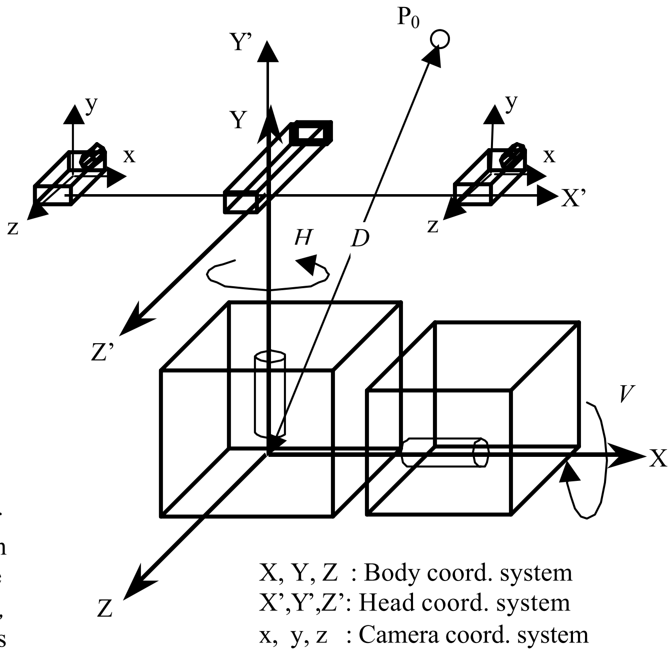


Figure 4. Coordinate system of HVT

3 CAMERA CALIBRATION

Figure 4 shows coordinate system for the HVT system. The unknown parameters for the exterior orientation parameters {$ù_0$, $ö_0$, $ê_0$ (rotation parameters), X_0, Y_0, Z_0 (camera positions)} and the interior orientation parameters {f (focal length), x_0, y_0 (principal points), a_1, a_2, (scale factor), p_1 (lens distortion)} for the CCD1 and CCD2 camera can be calculated by the same procedure in basically for the video theodolite with one target (Huang and Harley, 1989).

One reflector as a target was used to measure object distance above datum. The most remarkable point of this HVT system, however, automated camera calibration can be achieved without any reflector since object distance can be measured by the laser range finder. The brief calibration procedures are as follows:

When the CCD1 and CCD2 camera are pointing to P_0 point under the condition that the distance is D from the center of the body part of the HVT to P_0, horizontal angle H_0 and vertical angle V_0, the image point to P_0 is taken as p_0 on the image respectively. Next, when the HVT is rotated V_1 angle, the image point to P_0 is taken as p_1, and the object position to image point p_1 becomes P_1. Similarly, when the HVT is rotated H_1 angle, the image point to P_0 is taken as p_2, the object to p_2 becomes P_2. Repeating this operation, control points are taken on the image and are produced in the space respectively. 9 control points were produced in this paper, and the ground coordinates for these control points are calculated with following equation,

$$\begin{bmatrix} X_i \\ Y_i \\ Z_i \end{bmatrix} = \begin{bmatrix} cos(H_i) & sin(H_i)sin(V_i) & sin(H_i)cos(V_i) \\ 0 & cos(V_i) & -sin(Vi) \\ -sin(H_i) & cos(H_i)sin(V_i) & cos(H_i)cos(V_i) \end{bmatrix} \begin{bmatrix} cos(H_0) & 0 & -sin(H_0) \\ sin(H_0)sin(V_0) & cos(V_0) & cos(H_0)sin(V_0) \\ sin(H_0)cos(V_0) & -sin(V_0) & cos(H_0)cos(V_0) \end{bmatrix} \begin{bmatrix} 0 \\ 0 \\ D \end{bmatrix} \quad (1)$$

Where, these ground coordinates are opposite the body coordinate system in Figure 4, and H_i, V_i are measured clockwise.

Furthermore, when the stereo image are taken by rotating the body V_j and head H_j, each camera position has to be corrected to respond to the rotation of the HVT by the following equation since the stereo image are taken at a different exposure station due to the discord between the center of the body and the lens.

$$\begin{bmatrix} X_0^* \\ Y_0^* \\ Z_0^* \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & cos(V_j) & -sin(V_j) \\ 0 & sin(V_j) & cos(V_j) \end{bmatrix} \begin{bmatrix} cos(H_j) & 0 & sin(H_j) \\ 0 & 1 & 0 \\ -sin(H_j) & 0 & cos(H_j) \end{bmatrix} \begin{bmatrix} X_0 \\ Y_0 \\ Z_0 \end{bmatrix} \quad (2)$$

where, X_0^*, Y_0^* and Z_0^* are corrected camera position.

It has been shown that the unknown rotation parameters, $ù_j$、 $ö_j$ for each sequential image can be obtained as the sum of changing vertical and horizontal values resulting in $ù_0$, $ö_0$, respectively (Chikatsu and et al., 1996).
However in case of the HVT system, similarly due to the discord between the center of the body and the lens, rotation parameters, $ù_j$, $ö_j$, $ê_j$, for each stereo image has to be corrected to respond to the rotation of the body V_j and head H_j by the following equation,

$$\omega_j = -\tan^{-1}(c_{23} / c_{33}), \qquad \varphi_j = \sin^{-1}(c_{13}), \qquad \kappa_j = -\tan^{-1}(c_{12}/c_{11}) \tag{3}$$

where,

$$\begin{bmatrix} c_{11} & c_{12} & c_{13} \\ c_{21} & c_{22} & c_{23} \\ c_{31} & c_{32} & c_{33} \end{bmatrix} = \begin{bmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{bmatrix} \begin{bmatrix} \cos(H_j) & 0 & \sin(H_j) \\ 0 & 1 & 0 \\ -\sin(H_j) & 0 & \cos(H_j) \end{bmatrix} \begin{bmatrix} 1 & 0 & 0 \\ 0 & \cos(V_j) & -\sin(V_j) \\ 0 & \sin(V_j) & \cos(V_j) \end{bmatrix}$$

and

$$\begin{aligned} a_{11} &= \cos\varphi\cos\kappa, & a_{12} &= -\cos\varphi\sin\kappa, & a_{13} &= \sin\varphi \\ a_{21} &= \cos\omega\sin\kappa + \sin\omega\sin\varphi\cos\kappa, & a_{22} &= \cos\omega\cos\kappa - \sin\omega\sin\varphi\sin\kappa, & a_{23} &= -\sin\omega\cos\varphi \\ a_{31} &= \sin\omega\sin\kappa - \cos\omega\sin\varphi\cos\kappa, & a_{32} &= \sin\omega\cos\kappa + \cos\omega\sin\varphi\sin\kappa, & a_{33} &= \cos\omega\cos\varphi \end{aligned}$$

4 AQURACY EVALUATION OF HVT

In order to evaluate an accuracy of the HVT, following fundamental experiments were performed using 16mm lens. Figure 5 shows the concept of the test field and Figure 6 shows the details of the target. The target set on the actuator so that the changing values of the target can be controlled through a PC. Black and white target in Figure 5 was used as the P_0 in calibration procedure, and 42 black circle points are check points for checking accuracy.

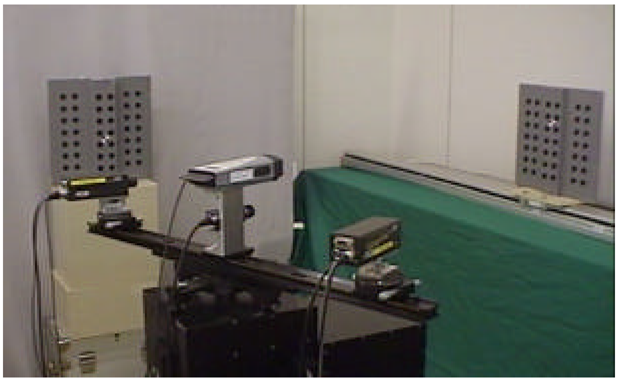


Figure 5. Test field

+Test1
Camera calibrations were performed for the target at 0mm position, and stereo image A_1 was taken. The HVT rotated to left side and stereo image A_2 was taken at -250mm, and A_3 was taken at -500 mm. Similarly, stereo image A_4 and A_5 were taken at +250 mm, +500 mm on the right side, respectively. Then, the HVT rotated to -90º, and stereo image A_0 was taken.

+Test 2.
On the contrary test 1, camera calibrations were performed for the target at -90º position and stereo image A_0 was taken. Then, the HVT rotated +90º, and stereo image A1 was taken at 0mm position. Similarly, the HVT rotated to left side and stereo image A_2 and A_3 were taken at -250mm, -500 mm position, and stereo image A_4 and A_5 were taken at +250 mm, 500 mm on the right side, respectively.

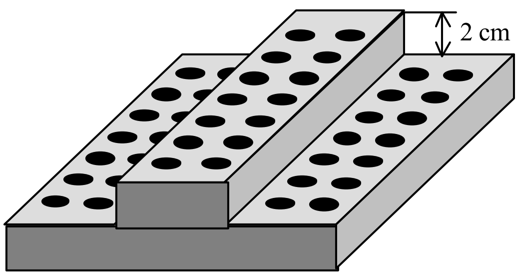


Figure 6. Configuration of target

Table 1. RMSE for check points

(a) Test1

Stereo-image Position		A_3 -500mm	A_2 -250mm	A_1^* 0mm	A_4 +250mm	A_5 +500mm
RMSE	σ_{XY} (mm)	±0.479	±0.382	±0.358	±0.297	±0.373
	σ_z (mm)	±0.837	±0.871	±0.746	±0.674	±0.665

-90 °

Stereo-image Position		A_0
RMSE	σ_{XY} (mm)	±0.337
	σ_z (mm)	±0.823

(b) Test 2

Stereo-image Position		A_3 -500mm	A_2 -250mm	A_1 0mm	A_4 +250mm	A_5 +500mm
RMSE	σ_{XY} (mm)	±0.660	±0.544	±0.469	±0.438	±0.473
	σ_z (mm)	±0.772	±0.736	±0.707	±0.682	±0.774

+90 °

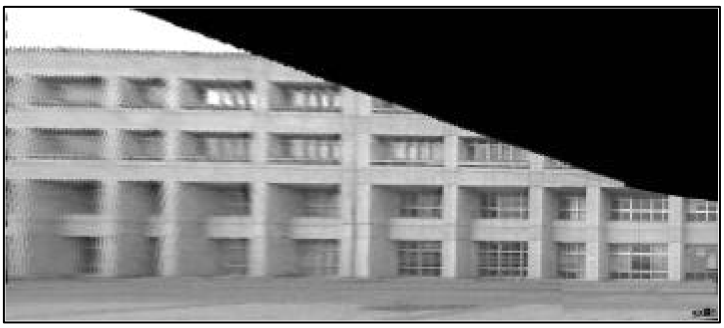
Stereo-image Position		A_0^*
RMSE	σ_{XY} (mm)	±0.348
	σ_z (mm)	±0.680

where, * means calibration position, and base-depth ratios is 0.48.

Table 1 shows the R.M.S.E. for 42 check points on each stereo image, and it is concluded from the results of this experiments that the HVT system is expected to become a useful system in the various application fields since the synchronized stereo image sequences are obtained and rotation parameters for each stereo image can be acquired in real-time.

Furthermore, the most remarkable point from the result that spatial data acquisition for 360º in horizontal become possible

(a) Front building



(b) Left side building

Figure 7. Panoramic ortho image

5 3D MODELING OF PANORAMIC OBJECTS

In general, it cannot get enough distance for taking panoramic objects in a site. Nevertheless, it often encounters the situation that there are two more panoramic objects at the same area, which should be taken. On the assumption of these situation, and in order to perform of the HVT functions, the HVT was set at 6.018m on the front of 54m long building and 10.631 m from the left side 84m long building (near the cross point of both buildings).
The heights of both buildings are 13m (54m building) and 17m (84m building) respectively, and the stereo images for both buildings are taken using 8mm lens and 3m rail (base line).
The ortho panoramic imaging and 3D modeling procedures in this paper are as follows:
+ Camera calibrations were performed using produced 9 control points.
+ Stereo images for whole buildings were taken by rotating the HVT.
+ Ortho imagings were performed using calibration parameters.
+ Extraction of feature points for each stereo image was performed using canny filter.
+ Line extractions for each stereo image were performed by image processing procedures.
+ Image coordinates for both ends of each line were automated recorded in PC.
+ Panoramic ortho imaging was performed using extracted points in overlap and sidelap area.
+ 3D modeling for panoramic objects were achieved by semi-automatic line matching.

Figure 7 shows panoramic ortho image for the two buildings, and Figure 8 shows the 3D wire frame model.

Figure 8. 3D wire frame model

6 CONCLUSION AND FURTHER WORK

The Hybrid Video Theodolite (HVT) System was developed by the authors for multiple applications such as human motion analysis, auto-tracking, real-time positioning and so on, and it is concluded from the results of the experiments that the HVT system is expected to become a useful system in the various application fields since the synchronized stereo image sequences can be obtained and rotation parameters for each stereo image can be acquired in real-time. In particular, the remarkable points as additional results are its ability to acquire spatial data for 360º in horizontal. The effectiveness as an application of the HVT to 3D modeling of the panoramic objects were demonstrated.
There are issues, however, for further work. These problems are improvement of the HVT such as mobility, function for vertical range in left and right side and image processing procedures.

REFERENCES

Anai T. and Chikatsu H., 2000. Dynamic Analysis of Human Motion using Hybrid Video Theodolite System. IAPRS, Vol. XXXIII, Amsterdam.

Anai T. and Chikatsu H., 1999. Development of Human Motion Analysis and Visualization System. IAPRS, Vol. XXXII, Part 5-3W12, pp.141-144.

Chikatsu H. and Murai S., 1994. Utilization of a Video Theodolite System for Dynamic Analysis of Human Motion. Journal of the Japan Society of Photogrammetry and Remote Sensing, Vol.33, No.3, pp.77-80.

Chikatsu H. and Murai S., 1995. Application of a Video Theodolite for Sports Dynamics. IAPRS, Vol.30-5W1, pp.110-115.

Chikatsu H. and et al., 1996. Dynamic Analysis of Human Motion using Sequential Images of Video Theodolite. IAPRS, Vol. XXXI, PartB5, pp.82-87.

Chikatsu H. and et al., 1997. On Automated Distance Measurement of Ski Jump using a Video Theodolite. Optical 3-D Measurement Techniques VI, pp.383-390.

Huang Y. D. and Harley I., 1989. A New camera calibration method needing no control field. Optical-3D Measurement Techniques, pp.49-56.

Kakiuchi T. and Chikatsu H., 1998. Auto-Tracking and 3D Measurement for Moving Object using Video Theodolite. IAPRS, Vol. XXXII, Part 2W1, 5W1, 5/3W, pp.141-144. pp.p2-4-1-p22-4-4.

Kakiuchi T. and Chikatsu H., 2000. Construction of Stereo Vision System for 3D Objects Modeling. IAPRS, Vol. XXXIII, Amsterdam.

Nakano K. and Chikatsu H., 1999. Automated Panoramic Imaging using Video Theodolite. IAPRS, Vol.XXXII, Part 5-3W12, pp.145-148.

Yoshida S. and Chikatsu H., 2000. An Integrated ERGOMA System for Human Motion Analysis. IAPRS, Vol. XXXIII, Amsterdam.