CONSTRUCTION OF STEREO VISION SYSTEM FOR 3D OBJECTS MODELING

Tsutomu KAKIUCHI*, Hirofumi CHIKATSU**
* AERO ASAHI CORPORATION
Minamidai 3-1-1,Kawagoe, Saitama, 350-1165, Japan
kaki@g.dendai.ac.jp

** Professor, Department of Civil Engineering
Tokyo Denki University
Hatoyama, Saitama, 350-0394, JAPAN
chikatsu@g.dendai.ac.jp

Commission V, Working Group V/3

KEY WORDS: Video Theodolite, Stereo Adapter, Camera Calibration, Stereo Image Sequence. 3D Modeling.

ABSTRACT

Recently, spatial data acquisition for virtual environment applications, 3D objects modeling, 3D-GIS and so on have been extensively carrying out using laser-range finder or photogrammetric method. However, there are some issues for rapid reconstruction of objects and visualization of 3D special data. These problems include texture mapping in the case of laser-range finder and camera calibration aspects in photogrammetry. In order to resolve the twin issues, a stereo vision system which a stereo adapter was fitted to a lens of a CCD camera of a video theodolite was constructed so that stereo image can be acquired with one CCD camera. The most remarkable points for this stereo vision system are follows;

1) Camera calibration can be performed with one control point since the camera rotation parameters and the synchronized stereo image Sequences can be acquired simultaneously.

2) Stereo image can be acquired with one CCD camera since right and left image are taken as the odd field and even field by a liquid crystal shutter.

A test was performed in the laboratory using a stereo vision system, and system evaluation for 3D object modeling was investigated.

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 has been demonstrated (Chikatsu and Murai, 1995, Chikatsu and et al., 1996). The application of the video theodolite system to the ski jump also has been demonstrated (Chikatsu and et al., 1997), The application of the video theodolite system to auto-tracking and real-time positioning for a moving object (Kakiuchi and Chikatsu, 1998) and automated panoramic imaging using video theodolite also has been demonstrated (Nakano and Chikatsu, 1999).

The most remarkable points of the video theodolite system are its ability to obtain the camera rotation parameters and to calculate the object position in real-time.

In order to investigate innovative application fields, e.g. virtual environment applications, 3D objects modeling, 3D-GIS and so on, the authors constructed the stereo vision system based on the video theodolite system fitting a stereo adapter to a lens of a CCD camera of the video theodolite. Fitting a stereo adapter, 3D data for wide area can be obtained by the stereo matching because the camera rotation parameters and the synchronized stereo image sequences can be acquired. Furthermore, textures for object are simultaneously taken.

With these circumstances, in order to evaluate the stereo vision system, pre-experiment for the accuracy of 3D measurement of this system was demonstrated. Furthermore, indoor experiment for reconstructing and 3D modeling for indoor space in the Tokyo Denki University was demonstrated.

This paper describes the details of system architecture, pre-experiment results and 3D object modeling for indoor space using a stereo vision system.

2 STEREO VISION SYSTEM

2.1 Video Theodolite

Figure 1 shows the motorized video theodolite (SOKKIA, MET2NV) used in construction of a stereo vision system. The video theodolite has two CCD cameras. CCD1, the upper one in the figure 1, is used for precise pointing to the target through the monitor, and CCD2, the upper one in the figure 1, is used as a finder. The lens of the CCD2 camera can zoom from 5.9 to 47.2 mm focal length. In order to take wide scene, the CCD2 camera was mainly used in this paper, and fitting a stereo adapter to the lens of the CCD2 camera of the video theodolite.
The horizontal and vertical angles and the focal length of the CCD2 camera can be controlled by the personal computer via a RS232C port. At the same time, horizontal and vertical angles can be received by the personal computer. The current values of the rotation parameters and distance are continuously superimposed on image frames.

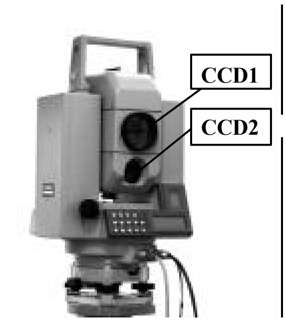

Figure 1. Video Theodolite

2.2 Stereo Adapter

Figure 2 shows the stereo adapter (3-D VIDEO INC., Nu-View) used in this investigation. Right and left image are taken as the odd field and even field on the same image by a liquid crystal shutter. The even lines on the odd field are interpolated using a just upper odd line respectively; similar the odd lines on the even field are interpolated in this paper. Figure 3 shows an interpolation concept. Thus, stereo image can be acquired with one camera by fitting the stereo adapter to the lens of the CCD 2 camera.


Figure 2. Stereo Adapter

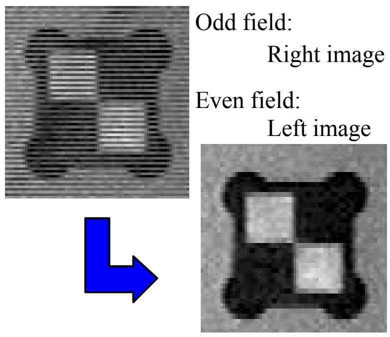

Figure 3. Interpolation concept

2.3 System Configuration

The stereo vision system consisting of the video theodolite, the stereo adapter, a personal computer, an image processing board, a TV monitor and a video recorder was constructed in this paper.
Fitting the stereo adapter to the lens of the CCD2 camera of the video theodolite, the synchronized stereo image sequences for wide area can be acquired with one camera rotating the video theodolite. Rotation parameters of the CCD2 camera for each stereo image sequences are acquired in real-time as the angles of the video theodolite.
Camera calibration can be performed with one control point since the rotation parameters and stereo image can be acquired simultaneously. Furthermore, 3D object position and textures for object are simultaneously taken.
3D spatial data acquisition procedures, i.e. the acquisition of stereo images for the right and left image, extraction for the image coordinates of feature points, stereo matching, 3D measurement, panoramic imaging, wire frame model and texture mapping for 3D modeling are performed on the personal computer. Live image can be confirmed via the TV monitor. The current values of the rotation parameters are continuously superimposed on image frames and thus recorded as a part of the image data via the video recorder.
The detail of the components of the stereo vision system is shown in Table 1. Figure 4 shows the configuration of the system.

Theodolite	MET2NV (Sokkia, accuracy 2")	
CCD Camera	CCD 1	CCD 2
	CCB-GC5 (Sony, 510H 492V)	EVI-310 (Sony, 768H 494V)
Lens	f = 300mm	f = 5.9 47.2 mm
A/D Converter	MTAT-CL (Microtehcnica. 15.734KHz)	
Video Recorder	HR-SC1000 (Victor)	
Monitor	PVM-1454Q (Sony)	
PC	Optiplex GXM 5200 (DELL)	
Stereo Adapter	Nu-View (3-D VIDEO INC.)	

Table 1. System components

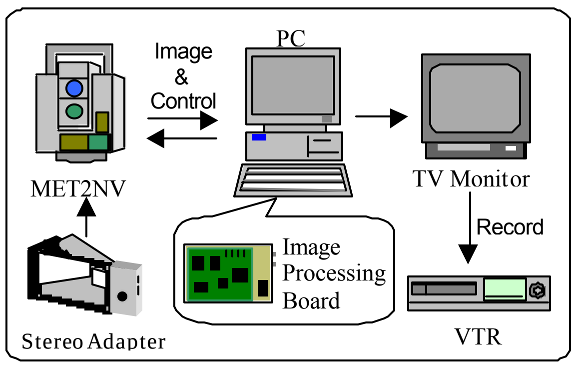

Figure 4. System configuration

3 CAMERA CALIBRATION USING VIDEO THEODOLITE

For camera calibration, control points should be generated and its image coordinates also should be acquired. Camera calibration procedures are as follows: When the telescope is pointing to P_0 point under the condition that the distance is D from the center of the video theodolite 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 right and left image respectively. Next, when the video theodolite is rotated H_1 angle, the image point to P_0 is taken as p_1 on the right and left image, and the object position to image point p_1 becomes P_1. Similarly, when the video theodolite is rotated V_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 stereo image and are produced in the space respectively based on Figure 5.

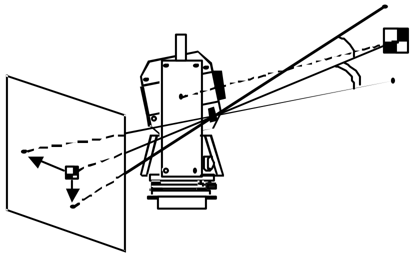

Figure 5. Camera calibration

Ground coordinates for these control points are calculated following equation.

$$
\begin{aligned}
X_i &= D\cos V_0 \sin(H_i - H_0) \\
Y_i &= D\{\cos V_i \sin V_0 - \sin V_i \cos V_0 \cos(H_i - H_0)\} \\
Z_i &= D\{\sin V_i \sin V_0 + \cos V_i \cos V_0 \cos(H_i - H_0)\}
\end{aligned}
\qquad (1)
$$

Camera calibration for the right and left image was performed using 9 control points, which were produced in the above procedure.

4 PRE EXPERIMENT

4.1 Experiment Method

In order to estimate the accuracy of 3D measurement of the stereo vision system and the relationship between the rotation parameters (ω, ϕ) and the reading angles (zenith, horizontal) of the video theodolite, the following pre-experiments were performed using the test field (Figure 6).

Figure 7 shows the three test models (Z=0mm, 20mm, 50mm) were used in this paper. The black white rectangular target is a control point corresponds the above P_0, and other 38 black circle points are check points for checking 3D accuracy.

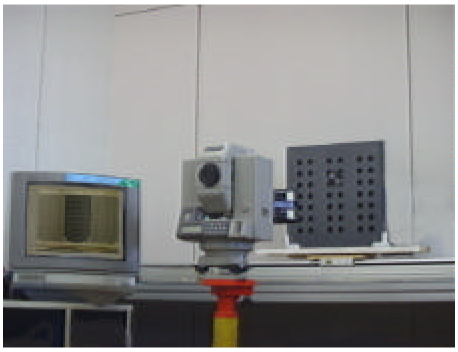

Figure 6. Test site

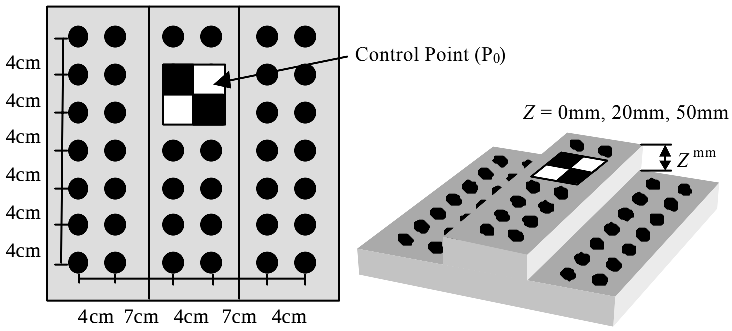

Figure 7. Test model

The following image A1 (Figure 8) was taken from 2.4980m as the original orientation image. The left-hand number superimposed in the upper left corner shows the horizontal angle (H_0=0°0'0"). The number just to the right is the zenith (V_0=0°0'0", V=90°-Zenith angle) and the distance (D=2.4980m).

Figure 9 shows the right and left image for camera calibration which was interpolated by above concept.

Camera calibration for the right and left image was performed by the same procedure noted above using one control point on the test model.

For the unknown parameters, exterior orientation parameters{ ω_0, ϕ_0, κ_0 (rotation parameters), X_0, Y_0, Z_0 (camera positions)} and interior orientation parameters {f, x_0, y_0 (principal points), a_1, a_2 (scale factor), p_1(lens distortion)}.

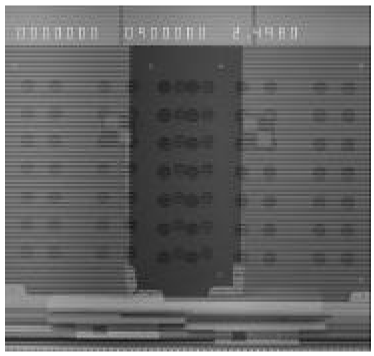

Figure 8. Orientation image

After the camera calibration for the right and left image, the video theodolite tracked the test model was slowly moving to left side. Stereo image A2 was taken at +500mm. The changing values of the test model on the actuator were then controlled through a personal computer. The rotation angles of the video theodolite were 20°12'47" clockwise, 0°00'14" under the horizon, thus giving ΔV_2 = 0°00'14" and ΔH_2 = 20°12'47". Stereo image A3 was taken at -500mm and ΔV_3 = -0°00'14", ΔH_3 = -27°57'31".

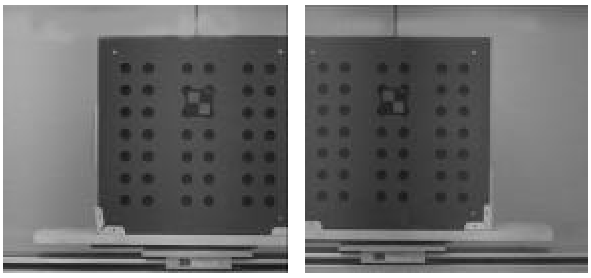

Figure 9. Right and Left image

The unknown parameters, ω and ϕ for each stereo images should be estimated as the sum of changing vertical and horizontal values resulting in ω_0 and ϕ_0 respectively. Here, ω_0 and ϕ_0 are the calibration results of the both orientation image. Consequently, ω and ϕ are calculated as follows using the changing values in vertical (ΔV_i), horizontal direction(ΔH_i) and ω_0, ϕ_0.

$$\omega = \omega_0 + \Delta V_i \qquad \phi = \phi_0 + \Delta H_i \tag{2}$$

However, each stereo image was taken at a different exposure station, due to the discord between the center of the theodolite and the lens of the CCD 2 camera. Each camera position has to be corrected to respond to the rotation of the video theodolite by the following equation,

$$\begin{aligned} X_0^* &= D^* \cos V^* \sin(\phi - H_0) \\ Y_0^* &= D^* \{\cos\omega \sin V^* - \sin\omega \cos V^* \cos(\phi - H_0)\} \\ Z^* &= D^* \{\sin\omega \sin V^* + \cos\omega \cos V^* \cos(\phi - H_0)\} \\ Z_0^* &= (D - Z^*)\cos V_0 \end{aligned} \tag{3}$$

Where, X_0^*, Y_0^*, Z_0^*, corrected camera position, $V^* = \tan^{-1}\{Y_0/(D - Z_0)\}$, $D^* = \sqrt{X_0^2 + Y_0^2 + (D - Z_0)^2}$

4.2 Pre-Experiment Results

Table 3 shows the R.M.S.E. for check points in each stereo image.

Stereo Image	A3			A1			A2		
Changing Values	-500mm			0mm			+500mm		
Height Difference	0mm	20mm	50mm	0mm	20mm	50mm	0mm	20mm	50mm
$_{xy}$ (mm)	±0.462	±0.421	±0.462	±0.381	±0.401	±0.413	±0.461	±0.481	±0.523
$_{z}$ (mm)	±4.402	±4.552	±4.824	±3.962	±4.381	±4.121	±4.765	±4.827	±4.923

Theoretical accuracy; $_{xy}$ = ±0.416 mm, $_{z}$ = ±4.167 mm

Table 2. R.M.S.E. for check points

Almost the same low accuracy for the Z-coordinate can be found, probably due to the lack of horizontal line, i.e. odd and even field was utilized in this paper, and the short of baseline of the stereo adapter, but it is concluded that the stereo vision system is a useful tool in a small inner space.

5 3D SPATIAL DATA ACQUISITION

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. Furthermore, textures for object are simultaneously taken.

In order to evaluate the utilization of the stereo vision system, indoor experiment for reconstructing and 3D modeling was demonstrated.

Figure 10 shows the elevator hall in the Tokyo Denki University (TDU) used in this investigation.

Reconstructing and 3D modeling procedure are as follows.

Figure 10. Elevator hall in the TDU

5.1 3D Data Acquisition

In order to perform camera calibration for the right and left image, one control point was set the stairs at the side of the elevator in the figure 10. Then, camera calibration was performed by the same procedure noted above.
After the camera calibration, stereo images for wide area of the indoor space are acquired rotating the video theodolite. The unknown parameters, ω and ϕ for each stereo images can be estimated as the sum of changing vertical and horizontal values resulting in ω_0 and ϕ_0 for the orientation image respectively and parameters other than ω and $\phi$are considered as the same values as the calibration results for the both orientation image. Left and Right image are acquired by above concept. Feature points such as the corner of stairs, pillar and so on are extracted on the right image just by the mouse click. Consequently, 3D data for each feature points in stereo images are calculated by the area based stereo matching.

However, due to the each stereo images were taken at a different exposure station, a few different color tone are found. Then, tone correction for each stereo images were performed with the following equation

$$L_{out} = \frac{\sigma_{obj}}{\sigma_{ref}}\left(L_{out} \quad m_{ref}\right) + m_{obj} \tag{4}$$

Where, L_{out}; output image
L_{in}; input image
σ_{obj}; standard deviation of output image
m_{obj}; mean value of output image
σ_{ref}; standard deviation of input image
m_{ref}; mean value of input image

3D data for each feature points in stereo image can be calculated using the calibration parameters of the stereo image, the rotation angles of the video theodolite and the image coordinates for each future points.

5.2 Wire Frame model

3D data for each future points are continuously displayed on the computer screen as a wire frame. Wire frame model can be achieved using 3D data for every feature points by perspective projection techniques. Wire frame data can be converted as a DXF file format.
Figuare11 shows the wire frame model of the elevator hall. Wire frame model can be viewed form different angles just by mouse click.

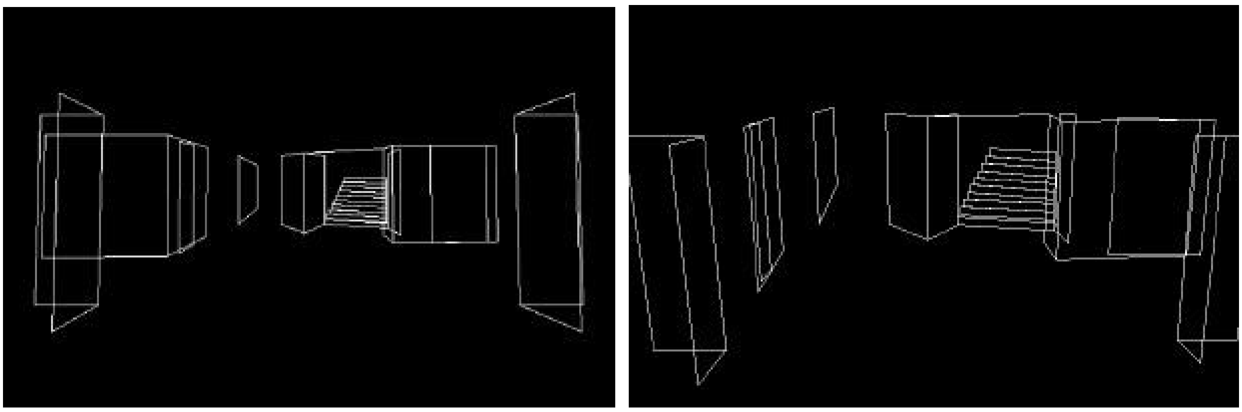

Figure 11. Wire frame model

5.3 Panoramic Imaging

In order to acquire texture images to mapping for the wire frame model, panoramic images were obtained rotating the video theodolite in this paper. Each sequential image as a central projection image can be transformed into ortho projection image using spatial data. Panoramic imaging for the right image can be achieved automatically since the rotation angles of the video theodolite are controlled by a personal computer (Nakano, K., and Chikatsu, H., 1999).

For automated panora mic imaging, vertical and horizontal angles for panoramic area should be previously obtained, and rotation angles for each sequential right image can be calculated using the camera angle, which is calculated using the relationship between the focal length and effective sensor area. Then, camera angle is calculated with the following equation.

$$\theta = 2\tan^{-1}\left(n \bullet s / 2f\right) \tag{5}$$

Where, θ;camera angle, f; focal length, n; pixel number, s; pixel size

On the other hand, each sequential image should be transformed into ortho projection image since the sequential right images are central projection image, and central projection image were resampled into ortho projection image. Here, automated panoramic imaging can be achieved overlapping with corresponding coordinate since the ortho images have common coordinate system.
Consequently, spatial data for the indoor space are calculated from the following equation on the assumption that distance from the video theodolite to walls are constant.

$$\begin{aligned} X &= X_0^* + \frac{a_{11}x + a_{21}y + a_{31}f}{a_{13}x + a_{23}y + a_{33}f}\left(Z^* - Z_0^*\right) \\ Y &= Y_0^* + \frac{a_{12}x + a_{22}y + a_{32}f}{a_{13}x + a_{23}y + a_{33}f}\left(Z^* - Z_0^*\right) \end{aligned} \tag{6}$$

where, X_0, Y_0, Z_0; corrected camera position X, Y, Z; object coordinate, x ,y; image coordinate f; focal length, a_{ij}; rotation matrix with three parameters, ω, ϕ, κ

The detail procedures for automated panoramic imaging are as follows:
+ Camera calibration
+ Calculation of camera angle
+ Establishment of panoramic area
+ Calculation of rotation angle for each sequential image
+ Calculation of spatial data for all pixels
+ Resampling
+ Mosaicking
+ Tone correction
Figure 5 shows the panoramic image was consist of 52 sequential images.

Figure 12. Panoramic image (Elevator hall in the Tokyo Denki University, Japan)

5.4 Modeling

Reconstruction and 3D modeling of indoor space are achieved since the 3D spatial data can be calculated by stereo images. Figure 13 shows the textured model of the elevator hall in the Tokyo Denki University, Japan. This texture model was achieved by graphic software using panoramic images and wire frame model. Due to the relatively slow manual texture mapping, the texture model was not achieved automatically

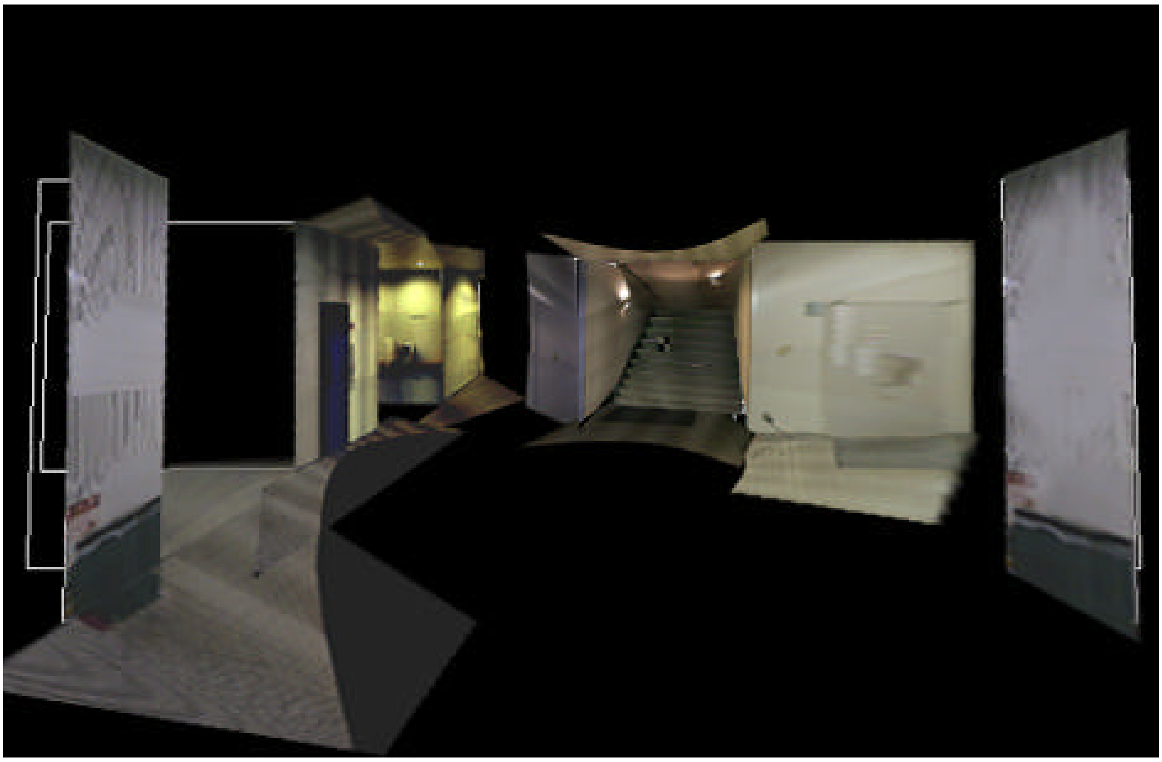

Figure 13. Textured model of the elevator hall in the Tokyo Denki University, Japan

5.5 3D Spatial Data Acquisition Procedures

The detail procedures for the 3D spatial data acquisition are as follows:
+ Camera calibration for the right and left image is performed using one control point
+ Stereo images for the indoor space are acquired rotating the video theodolite, then right and left image are simultaneously taken.
+ Tone correction is performed for right and left image.
+ Every feature points are extracted on the right image just by mouse click.
+ Area based stereo matching is performed for each stereo image.
+ 3D data of the future points is calculated using the calibration parameters, the camera rotation angles and the image coordinate of each future points.
+ Wire frame model is acquired by perspective projection techniques using 3D data.
+ Panoramic images are automatically obtained using resampling, mosaicing and tone correction procedure.
+ 3D modeling is achieved by texture mapping using graphic software.
These procedures are shown in Figure 14.

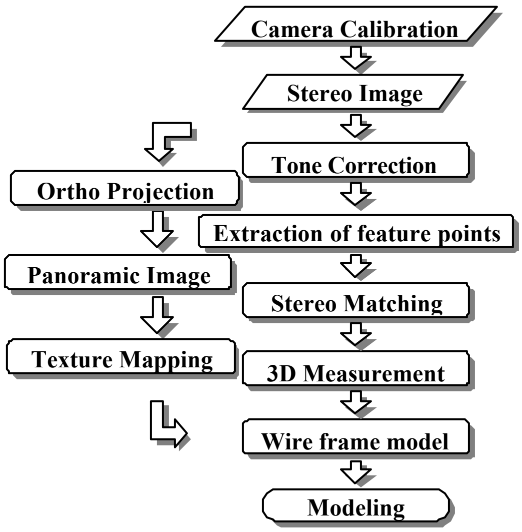

Figure 14. 3D Spatial Data Acquisition Procedures

6 CONCLUSIONS AND FURTHER WORK

Camera calibration and acquisition of 3D spatial data using the stereo vision system have been investigated in this paper. Fitting the stereo adapter to the lens of the CCD 2 camera, stereo image can be obtained with one camera, and stereo image sequences for wide area also obtained rotating the video theodolite. Furthermore, 3D spatial data can be obtained simultaneously. And camera calibration for the stereo image can be performed with one control point since the rotation parameters and stereo image can be acquired simultaneously.
The effectiveness of the stereo vision system not only human motion, auto-tracking and real-time positioning of a moving objects but also acquisition of 3D spatial data were demonstrated. There are issues, however, for further work. These problems are automatic modeling, texture mapping. And from the test, positioning errors can be found due to the short of baseline of the stereo adapter, but it is concluded that the stereo vision system is a useful tool in a small inner space.
With this circumstance, a stereo vision system is expected to become a useful tool in various photogrammetric fields since the stereo image sequences and the camera rotation parameters can be acquired in real-time.

ACKNOWLEDGMENTS

Finally, this paper has been made as a part of investigation of Chikatsu laboratory in Tokyo Denki University.

REFERNENCES

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

Heck, U., 1993. 3D-motion of an object determined by an image sequence of a video theodolite. Optical 3-D Measurement Techniques II, pp.538-545.

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 et al., 1996. Dynamic Analysis of Human Motion using Sequential Images of Video Theodolite. International Archives of Photogrammetry and Remote Sensing, Vol.XXXI, PartB5, pp.82-87.

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

Kakiuchi, T. and Chikatsu, H., 1998. Auto-tracking and Positioning for Moving Object using Video Theodolite, International Archives of Photogrammetry and Remote Sensing, Vol. XXXII, Part5, pp.735-738.

Nakano, K., and Chikatsu, H., 1999. Automated Panoramic Imaging Using Video Theodolite. International Workshop on Vision-Based Techniques in Visualization and Animation, Vol.XXXII, pp.145-148.