DYNAMIC ANALYSIS OF HUMAN MOTION USING SEQUENTIAL IMAGES OF VIDEO THEODOLITE

Hirofumi CHIKATSU, Tetsuji ANAI, Hideharu YANAGI
Department of Civil Engineering, Tokyo Denki University
Hatoyama, Saitama, 350-03, Japan.
Shunji MURAI
Institute of Industrial Science, University of Tokyo
22-1, Roppongi 7 Chome, Minato-ku,Tokyo, 106, Japan.

Commission V, Working Group 5

KEY WORDS; Video theodolite, Image, Calibration, Matching, Dynamic analysis, Human motion

ABSTRACT

Video image sequences often give important information about dynamics of human motion in the field of sports training or rehabilitation. For understanding of the dynamics of human motion from video image sequences, there are two complicated subjects which have to be resolved. One is how to estimate the photogrammetric camera orientation parameters on a moving camera. The second is image processing, for example, automated recognition of human feature points such as the head, elbows or knees.
This paper describes an application of a video theodolite system and image processing procedure with regard to the above two issues for dynamic analysis of human motion.

1. INTRODUCTION

Generally, a dynamic analysis of human motion has been performed under a condition that camera position and rotation are fixed and some markers are fitted on the body.
Therefore, it is possible to calibrate the camera parameters in advance. Also, automated recognition of some human feature points such as the head, elbow or knees is possible.
In order to understand a dynamic analysis of the most natural human motion, limitation of the camera and any marker on the body should be removed. For this ideal dynamic analysis, however, camera orientation parameters should be acquired in real time while recording a moving object. Furthermore, automated recognition of some human feature points should be performed.
The effectiveness of the video theodolite system for dynamic analysis of human motion has been indicated by the authors(Chikatsu and Murai,1995). This paper describes the dynamic analysis of human motion using sequential images which are taken by video theodolite. Also, image processing techniques are described.

2. VIDEO THEODOLITE SYSTEM

In order to acquire the camera rotation parameters in real time, a CCD camera was mounted on top of a theodolite as camera axis x,y,z coincide with theodolite axis X(horizontal axis),Y(vertical axis) and Z(collimation axis). Then, ω and ϕ were defined as the vertical and horizontal angles respectively. κ is assumed to be 0 degrees as the theodolite is levelled. This concept was first published in (Huang and Harley,1989).

With this motive, the authors have been concentrating on developing the 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 were then continuously superimposed on the image frames and thus recorded as a part of the image data (Chikatsu and Murai,1994).
Last year, SOKKIA corporation developed a motorized video theodolite (MET2NV, Figure 1) in association with the authors. MET2NV were developed based on MONMOS(the first order total station), and main feature points consist of two CCD color cameras. The current values of the rotation angles(vertical and horizontal) and distance were then continuously superimposed on the image frames. CCD 1 is the upper one in figure 1 and is used for precise pointing to the target through the monitor, and CCD 2 is the lower one and is used as a finder.

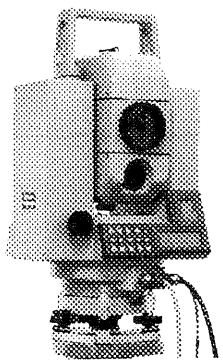


Figure 1. Video theodolite (MET2NV)

The detail of the CCD camera and the components of the video theodolite system are shown in table 1. Figure 2 shows the configuration of this system.

Table 1. Video theodolite system

Theodlite	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 ~47.2mm
A/D Convert er	FRM-512 (Photron,12.28Mhz)	
Video Recorder	PVM-1454Q (Sony)	
Monitor	HR-SC1000 (Victor)	
PC	PC-9801BX(NEC)	

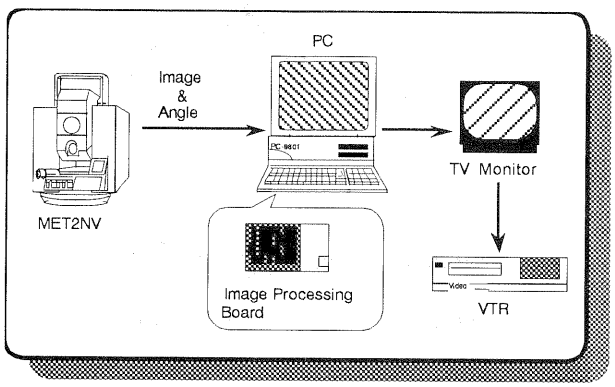


Figure 2. System configuration

3. CAMERA CALIBRATION

In order to confirm the relationship between the rotation parameters (ω , ϕ) and the reading angles (zenith, horizontal) of the video theodolite, the following experiments were performed using a target field (Figure 3). These targets were prepared to check the accuracy. Test field A means orientation image for camera calibration and B and C mean sequential image when video theodolite is rotated.

Camera calibration was performed with 9 control points which were produced in the following steps in figure 4.

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 image. Next, when the video theodolite is rotated H 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 video theodolite is rotated V 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. Ground coordinates for these control points are calculated with the following equation,

$$X = -D\cos V_0 \sin(H_i - H_0)$$

$$Y = D\left\{\cos V_i \sin V_0 - \sin V_i \cos V_0 \cos(H_i - H_0)\right\} \quad (1)$$

$$Z = D\left\{\sin V_i \sin V_0 + \cos V_i \cos V_0 \cos(H_i - H_0)\right\}$$

where, these ground coordinates are opposite the telescope coordinate system in figure 5 and H_i is measured counterclockwise.

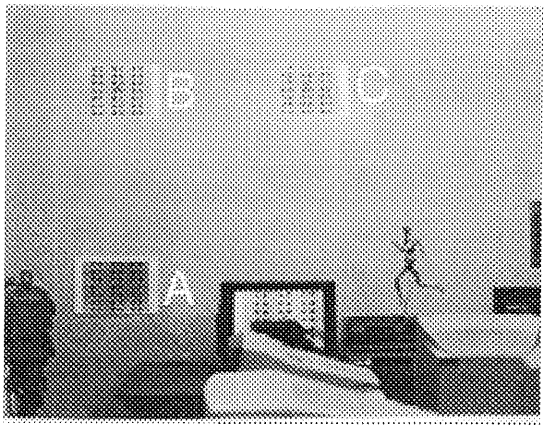


Figure 3. Test field

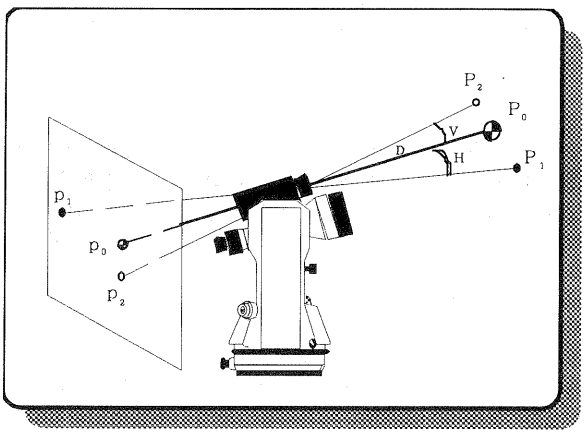


Figure 4. Production of control points

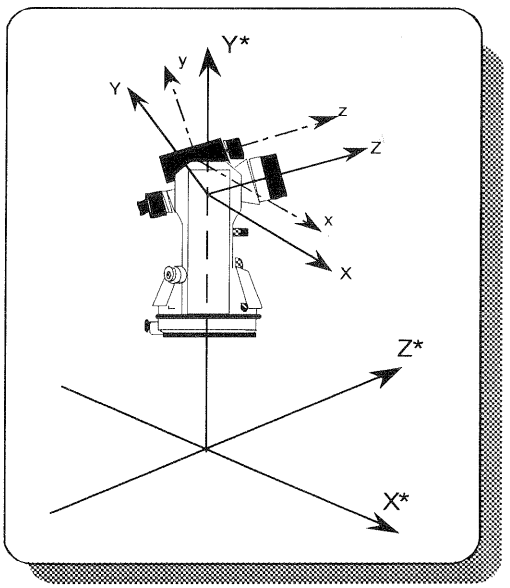


Figure 5. Coordinate system

X*,Y*,Z*; theodolite coord. system
X,Y,Z ; telescope coord. system
x,y,z ; camera coord. system

Bundle adjustment is applied for camera calibration between the camera coordinate and telescope coordinate system. Unknown parameters are as follows: exterior orientation parameters { ω, ϕ, κ (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)}.

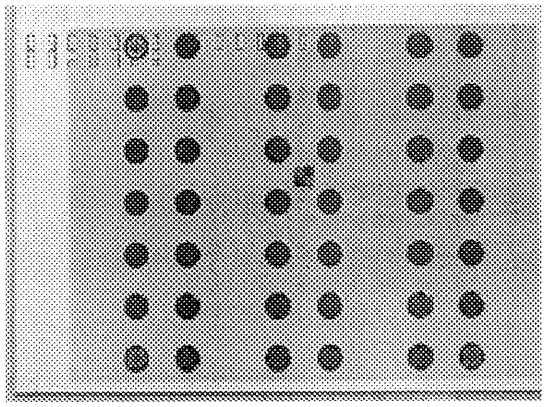


Figure 6 . Target field A

Image A (Figure 6) was taken from about 2.0m as an orientation image. The left-hand number superimposed in the upper left corner shows the zenith angle (V_0=0 ° 0' 0",V=90°-Zenith angle). The number just to the right is the horizontal angle (H_0=0 ° 0'1") and the distance (D=2.0128m).

Table 2 shows the calibration results for the target field A, using 9 control points. The rectangular target in the image A corresponds to the above P_0 point and the black circles mean check points. RMSE for 42 check points are shown in table 3. With regard to automated targeting for check points, the following image processing procedure was developed. The basic steps of this image processing procedure are shown in figure 7.

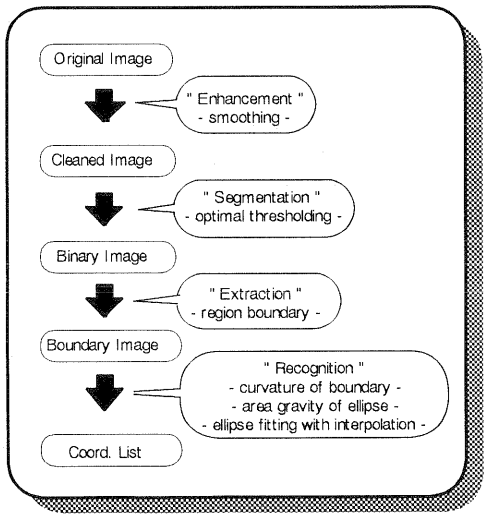


Figure 7. Flow of automated targeting

Table 3. RMSE for check points

Target Field	A	B	C
σ_{XY}	$\pm 0.19^{mm}$	$\pm 0.26^{mm}$	$\pm 0.45^{mm}$

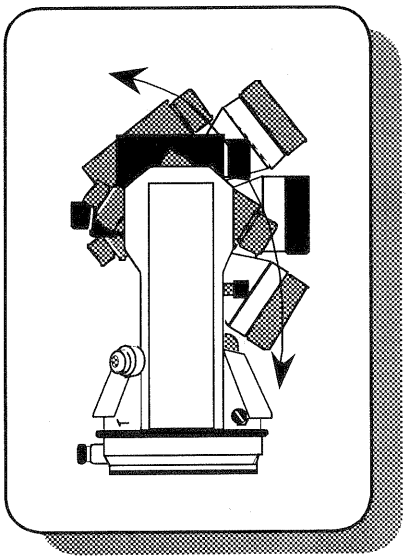


Figure 8. Multiexposure station

Next, the target field B was taken by rotating the video theodolite 0 ° 0'17" clockwise and 27 ° 46'02" above the horizon, giving ΔV_1=27°46'02" and ΔH_1=0°0'17". Finally, target field C was taken by rotating the video theodolite 25°05'20" clockwise and 27° 43'52" above the horizon, giving ΔV_2=25°05'20" and ΔH_2=27°43'52". Unknown parameters, ω and ϕ for each sequential image 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 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.

$$\omega = \omega_0 + \Delta V_i$$
$$\varphi = \varphi_0 + \Delta H_i \qquad (2)$$

where, ΔV_i =(zenith angle-90 °) and ΔH_i>0 in the clockwise case due to the relation between telescope coordinate(X,Y,Z) and the theodolite coordinate system (X*,Y*,Z*) in figure 5.

However, each sequential image was taken at a different exposure station as indicated in figure 8, due to the discord between the center of the theodolite and the lens of the CCD camera. Each camera position has to be corrected to respond to the rotation of the video theodolite by the following equation,

Table 2. Calibration results for the target field A

X_0	Y_0	Z_0	ω	ϕ	κ	x_0	y_0	f	a_1	a_2	p_1 ($\times 10^{-7}$)
mm	mm	mm				pixel	pixel	mm			
-0.026	-56.024	1957.799	−0°0'30"	−0°41'30"	−0°2'16"	256.071	256.074	18.219	131.855	-0.215	0.806

$$X_0^* = D^* \cos V^* \sin(\varphi - H_0)$$
$$Y_0^* = D^* \{ \cos\omega \sin V^* - \sin\omega \cos V^* \cos(\varphi - H_0) \} \quad (3)$$
$$Z^* = D^* \{ \sin\omega \sin V^* + \cos\omega \cos V^* \cos(\varphi - H_0) \}$$
$$Z_0^* = (D - Z^*) \cos V_0$$

where, X_0^*, Y_0^*, Z_0^*; corrected camera position,

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

Then, the X and Y coordinates for checking points in each sequential image are computed from the following equation,

$$X = X_0^* + \frac{a_{11}x + a_{21}y - a_{31}f}{a_{12}x + a_{22}y - a_{32}f}(Z - Z_0^*)$$
$$Y = Y_0^* + \frac{a_{13}x + a_{23}y - a_{33}f}{a_{12}x + a_{22}y - a_{32}f}(Z - Z_0^*) \quad (4)$$

where, X, Y, Z; object coordinate, x, y; image coordinate, f; focal length, a_{ij}; rotation matrix with three parameters, ω, ϕ, κ.

Table 3 shows the RMSE for check points on the target field B and C by using calibration parameters which were acquired as noted above.

Consequently, utilization of the video theodolite system is expected to become a useful tool in the field of sports training and rehabilitation since the camera parameters can be acquired in real time and the position for feature points of a human can be calculated.

4. APPLICATION OF VIDEO THEODOLITE

The relatively slow 4 images acquired per second is perhaps due to the ability of image processing board or MET2NV. Test for dynamic analysis of human motion were performed in a gymnasium by using MET2NV(Figure 9).

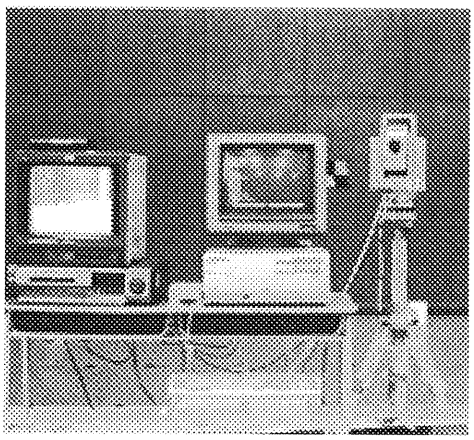


Figure 9. Test field for dynamic analysis

Sequential images were taken from 4.5 m at the above noted intervals. Calibration was performed by using one sheet reflector fitted on the wall. Table 4 shows the calibration results with 9 control points.

The authors have previously analyzed the dynamics of the sprinter Carl Lewis and that of boat rowing by using an image procedure and animation technique(Chikatsu and Murai,1992,1994). For the automated recognition of some feature points of a human, the image processing procedure and animation technique were combined in this paper from the above experience. The most remarkable point of this automated recognition is template matching. Templates have coordinates for some feature points of a human in advance. Then, only template matching between original binary image and template binary image are needed. Feature points of a human are shown in figure 10.

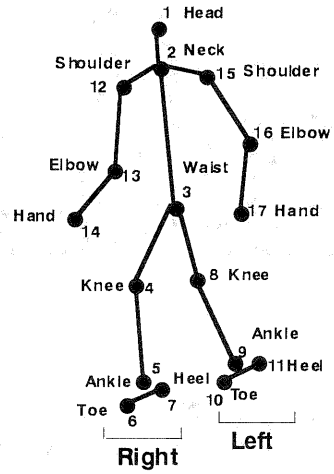


Figure 10. Feature points

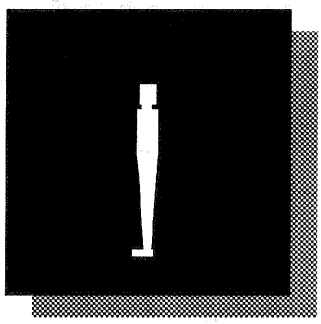


Figure 11. Base template

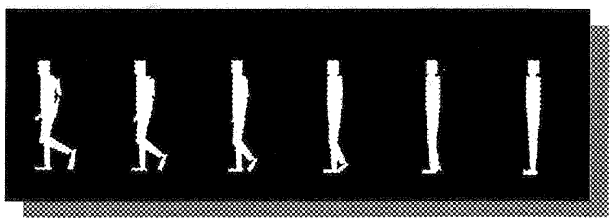
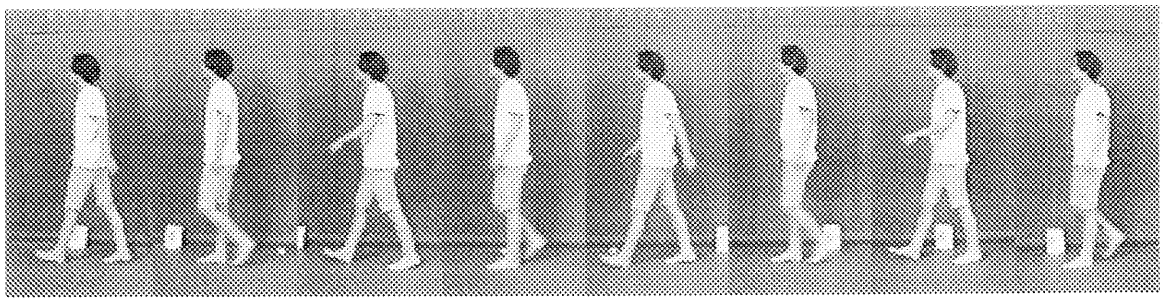


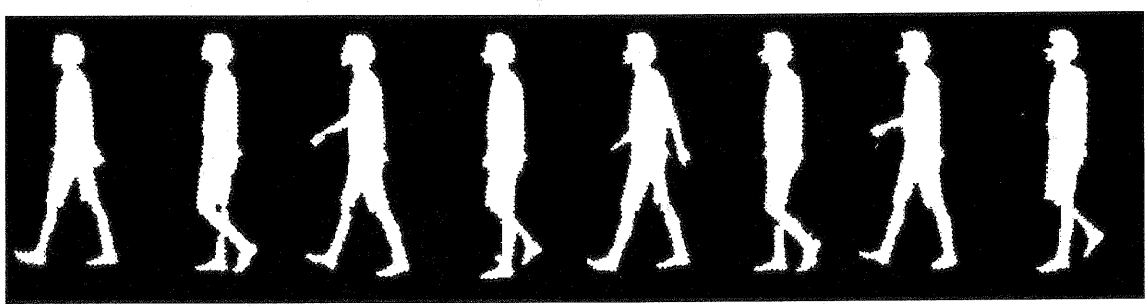
Figure 12. Animated cartoons

Table 4. Calibration results

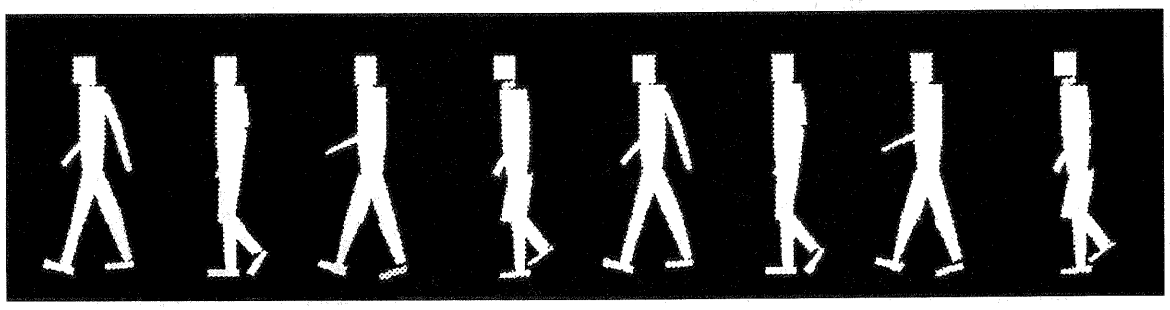
X_0	Y_0	Z_0	ω	ϕ	κ	x_0	y_0	f	a_1	a_2	$p_1 (\times 10^{-6})$
mm	mm	mm				pixel	pixel	mm			
-0.358	-55.991	4496.293	−0°5′50″	−0°22′43″	−0°9′18″	277.998	257.249	5.938	131.792	-0.070	0.412



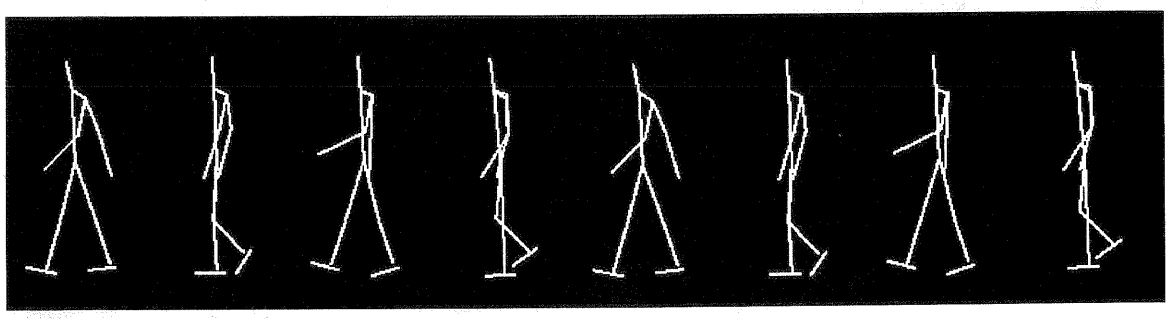
(a) Original image



(b) Binary image of original



(c) Animated image of original



(d) Skeleton image

Figure 13. Flow of automated extraction of feature points

A basic template image is produced from personal data such as height and chest. A series of animated cartoons are then produced using animation techniques. Figure 11 shows a basic template for a model. A series of animated cartoons are shown in Figure 12. Because the model is imaged in profile the right shoulder, elbow and hand are almost hidden from view. These feature points are interpolated by using the respective left points.

An automated feature points extraction is achieved from this template matching procedure. Figure 13 (a) shows sequential images as the model is walking along the wall perpendicular to the camera. The camera height for each sequential image is the same height which was computed from camera calibration. (b) shows a binary image of the original image and (c) shows the template image corresponding to each original image. (d) are the skeleton images which were only lined coordinates for each feature point. X and Y coordinates for the human feature points in each sequential image were calculated from equation (2).

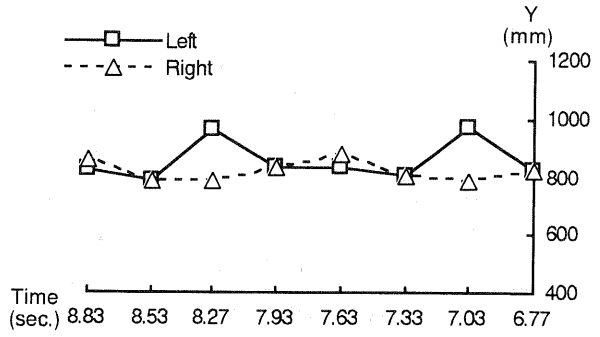
Figures 14 and 15 show the vertical and horizontal displacement for the hand and ankle. Figure 16 shows a cycle trace for these feature points at 6.77 to 8.83 second interval. The center of body gravity of a human is one of the most important elements for dynamic analysis of human motion. In order to calculate the center of body gravity, a human body was segmented into 12 sections using feature points (Figure 10). Therefore, the center of body gravity can be obtained as the sum of product weight ratio which are shown in figure 17 and gravity coordinates for each section.

Figure 18 shows the displacement for the center of body gravity. The horizontal displacement in these figures means the displacement from the head. The values below zero mean that each feature point is located behind the head.

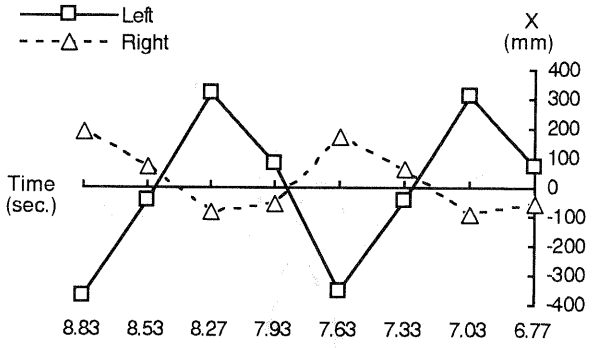
It may be seen from these figures that,

1. Displacement for the circle motion of the left hand is larger than the right, while there are no noticeably differences between left and right ankle.
2. The left ankle is located the farthest behind the head at 7.03 and 8.27 sec. and the left hand is located at the highest position and the farthest in front of the head at the same moment. This cycle means one stride of walking and was analyzed as 1.24 seconds.
3. Horizontal displacement for the center of body gravity is always located behind the head.
4. When the left ankle is located the farthest behind (7.03 and 8.27 sec.), the vertical displacement for the center of body gravity becomes the lowest values and the horizontal displacement between the body gravity and the head become shortest.
5. It can be said that the changing of the center of body gravity is performed at 7.03, 7.63 and 8.27 sec. where both ankles are most extended.
6. The model has large impulsive force from 6.77 to 7.03sec. and from 7.93 to 8.27sec. since the large displacement of the body gravity means that kinetic energy and impulsive force are also large.
7. All feature points except the right hand have large displacement at the same interval.

Generalizing these results, it can be said that that the motion of the model in walking is controlled by the left hand.

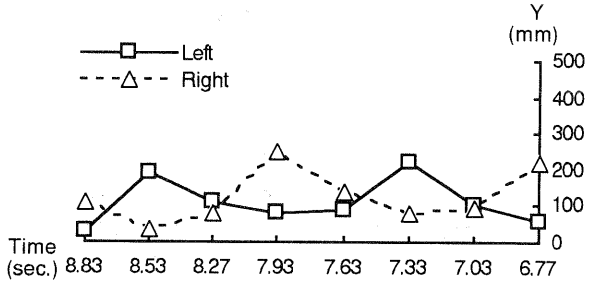


(a) Vertical displacement

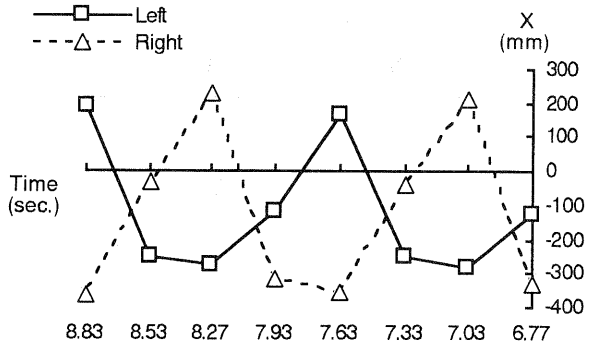


(b) Horizontal displacement

Figure 14. Displacement of the hand



(a) Vertical displacement



(b) Horizontal displacement

Figure 15. Displacement of the ankle

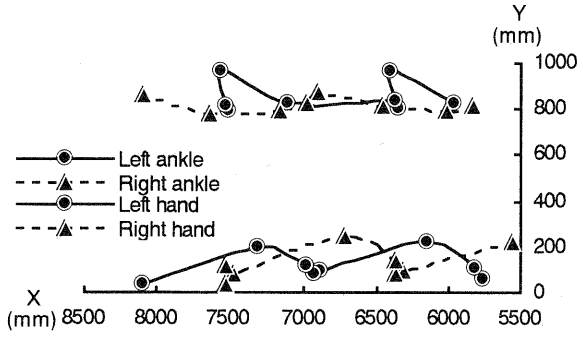


Figure 16. Locus of feature points

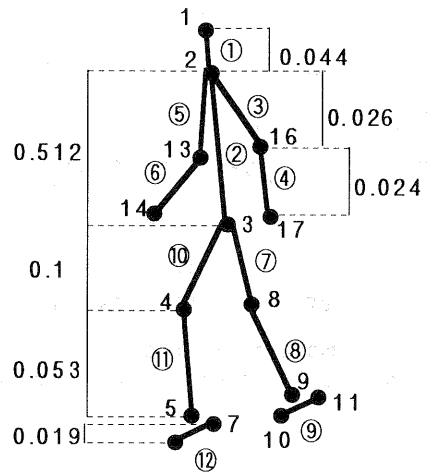
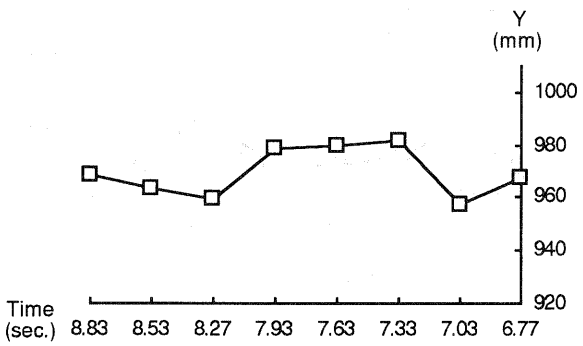
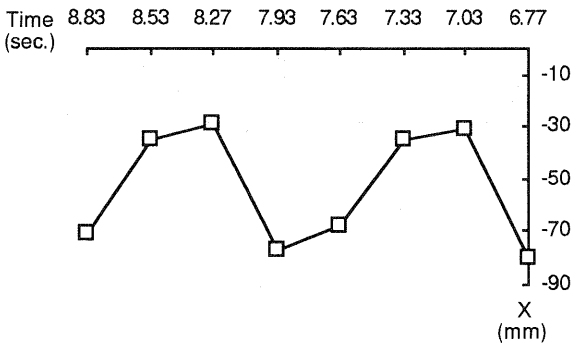


Figure 17 . Weight ratio



(a) Vertical displacement



(b) Horizontal displacement

Figure 18. Displacement for center of body gravity

5. CONCLUSION AND FURTHER WORK

Two issues on dynamic analysis of human motion using sequential images have been described in this paper. One was data acquisition of camera rotation parameters for sequential image using video theodolite system.

It has been shown that the unknown rotation parameters, ω and ϕ for sequential images can be obtained as the sum of changing vertical and horizontal values resulting in ω_0 and ϕ_0 respectively. Furthermore, it has been demonstrated that two dimensional coordinates for human feature points in sequential images can be calculated using the above parameters (ω , ϕ) and parameters other than ω and ϕ are considered as the same values as the calibration results for the orientation image.

With regard to the second issue, automated extraction for some feature points of a human have been achieved by using template matching which is combined image processing procedure and animation technique.

Dynamic analysis for human motion in walking has been demonstrated. There are , however, some issues which need to be resolved before video theodolite system becomes operational. These problems include, increased speed for data acquisition, 3D template model and efficient production of templates.

However, it is concluded that video theodolite systems are a useful tool not only for sports training and rehabilitation but also for various real-time photogrammetric fields since the rotation parameter can be acquired in real time while recording a moving object.

References

H. CHIKATSU, S. MURAI, 1992. Sports Dynamics of Carl Lewis through 100m Race using Video Imagery, 17th International Archives of Photogrammetry and Remote Sensing, Vol.29, Part 5, pp.875-879.

H. CHIKATSU, S. MURAI, 1994. Application of Image Analysis to Rowing Dynamics using Video Camera, Proceedings of the Commission V Symposium, Vol.30, Part 5, pp.35-40.

H.CHIKATSU, S.MURAI, 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.

H.CHIKATSU, S.MURAI,1995. Application of a Video Theodolite System for Sports Dynamics, International Archives of Photogrammetry and Remote Sensing, Vol.30, Part 5W1,pp.110-115.

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