Shape Reconstruction of Human Foot from Multi-camera Images Based on PCA of Human Shape Database

Jiahui Wang, Hideo Saito
Department of Information and Computer Science, Keio University, Japan
wjh, saito@ozawa.ics.keio.ac.jp

Makoto Kimura, Masaaki Mochimaru,Takeo Kanade
Digital Human Research Center, National Institute of Advance Industrial Science and Technology, Japan
makoto.kimura, m-mochimaru, t.kanade@aist.go.jp

Abstract

Recently, researches and developments for measuring and modeling of human body are taking much attention. Our aim is to capture accurate shape of human foot, using 2D images acquired by multiple cameras, which can capture dynamic behavior of the object. In this paper, 3D active shape models is used for accurate reconstruction of surface shape of human foot. We apply Principal Component Analysis (PCA) of human shape database, so that we can represent human's foot shape by approximately 12 principal component shapes. Because of the reduction of dimensions for representing the object shape, we can efficiently recover the object shape from multi-camera images, even though the object shape is partially occluded in some of input views. To demonstrate the proposed method, two kinds of experiments are presented: high accuracy reconstruction of human foot in a virtual reality environment with CG multi-camera images and in real world with eight CCD cameras. In those experiments, the recovered shape error with our method is around 2mm, while the error is around 4mm with volume intersection method.

1. Introduction

In recent years, anthropometry has been widely used in criminological, medical applications or selective trial of people[9]. In industry design, anthropometry also acts an important part, e.g. in the design of shoes, which need to be fit to the human body very much. For this purpose accurate measuring and modeling of human foot is necessary. Some 3D foot scanners have been commercially available. Although almost all this kind of scanners can generate high accuracy 3D foot model, the motion analysis of foot has not been solved sufficiently. This is because of the measurement space of these systems is always fixed, and the position constraint of cameras is strict. Thus the movement of foot is limited. However, foot is our motor organ, the measurement of its dynamic behavior is very important for various purposes. Thus, the goal of our research is acquisition of dynamic behavior of foot in a relative free space and high accuracy modeling of foot shape. This goal will be achieved step by step: firstly, capturing information of foot surface and modeling the 3D foot model from single image frame; secondly, extending the method to cope with motion image sequence and generate a dynamic foot model. In this paper, we concentrate on the first step.

Three types of object surface measurement techniques have been widely used: laser scanning, structured light projection and multiple images-based approaches. The precision of the measurement has made laser scanning[3][7] and structured light projection[11][14] the most popular systems for surface measurement. However, depending on the size and resolution of the surface to measure, the acquisition time can range from seconds to half minute. This fact requires the object human body should be stable without motion while the measurement, so it is difficult to measure the foot shape in dynamic situation.

The multiple images-based methods make use of multiple cameras, single moving camera or single camera combined with a rotating platform to acquire a set of images around object. The images are processed to extract the silhouettes[13] or find correspondence[2], which are then combined to result in a 3D model. The system based on multiple cameras can measure dynamic events, however, because of difficulties in processing concave surface or finding corresponding points between image pairs, high accuracy cannot be achieved easily.

In our research, multiple cameras are also used to acquire images around the human foot. In order to acquire high accuracy, we applied Active Shape Model (ASM)[4] based method. In this method, PCA is implemented on human shape database, so that we can represent human's foot shape by approximately 12 variation modes. Because of the reduction of dimensions for representing the object shape, we can efficiently recover the object shape from multi-camera

1550-6185/05 $20.00 © 2005 **IEEE**

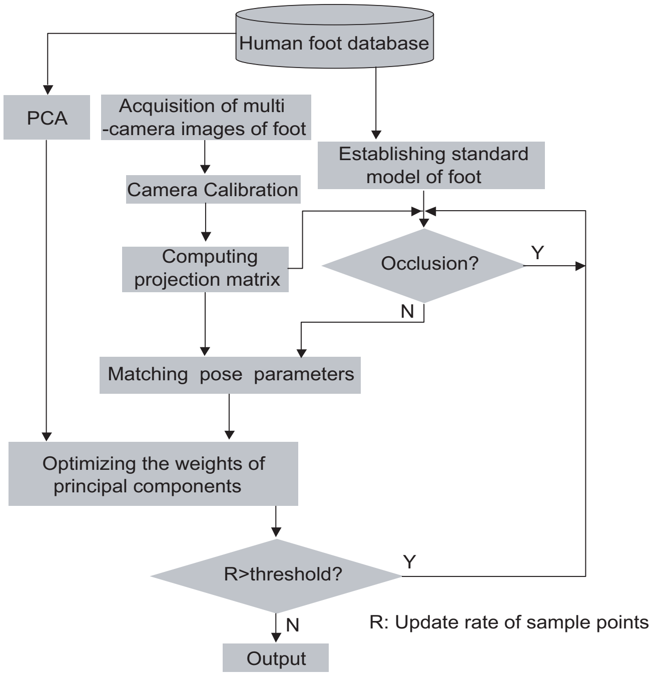
Human foot database
PCA
Acquisition of multi -camera images of foot
Establishing standard model of foot
Camera Calibration
Computing projection matrix
Occlusion?
Y
N
Matching pose parameters
Optimizing the weights of principal components
R>threshold?
Y
N
R: Update rate of sample points
Output

Figure 1. The flow chart of proposed method

images.

There are some advantages in using such human database. First advantage is avoiding occlusion. In capturing multiple view images of foot, we cannot avoid occlusion because of legs. Such occlusion caused by the legs will also be a problem even in the case for projecting a structured pattern to the object as [14]. The human database strongly contributes to recover the object shape from multiple view images with occlusion. The second advantage is reducing the ambiguities of recovered shape. Since we apply an Active Shape Models (ASM) based method for recovering the object shape, reducing the dimensions of estimated parameters is very important for stable recovery of object shape. The third advantage is making the system convenient. For existing 3D foot scanners, human anatomical landmarks need to be added for the measurement data to be handled as human data. The landmarks are generally measured manually. This is a complicated work and makes error easily. In our research work, the significant anatomical information is labeled by the statistical knowledge of database, in which the foot examples are aligned by a group of sample points. By using this sample points, the 3D foot shape will be estimated effectively.

While ASM is a powerful method in many image processing applications, some weakness is also discovered. Firstly, the surface of initial model is only allowed to adjust along the perpendicular directions. Secondly, the ASM uses only few part of the image information (mostly the edge profiles). Moreover, the 3D version ASM[5] is always implemented with a volumetric data of the object. However it is difficult in many practical cases without particular instruments (e.g. X-ray or MRI).

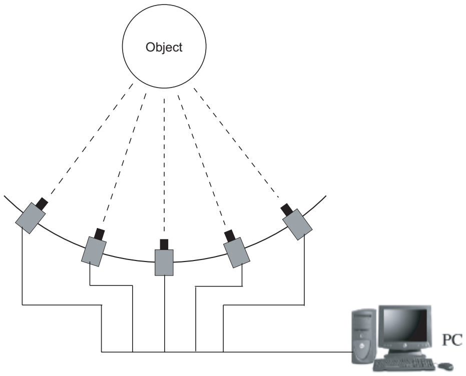

Figure 2. Multi-camera system, the IEEE1394 CCD cameras are synchronized together and connected to PC, which digitizes the images.

In this work, we used 3D ASM to search the fitting surface of human foot from its multi-camera images and reconstruct the 3D model. In order to improve the reliability of our method, we modified the 3D ASM in some points. Firstly, we used more information of the multi-camera images, e.g. intensity comparability, edge feature and silhouette information. Furthermore, the establishing of volumetric data is avoided. The fitness of the 3D model is only verified by the multi-camera images. We combine the 3D model and 2D multi-camera images by projecting the 3D model to image plane, so that the complicate process of establishing volumetric data can be skipped. The outline of the proposed method is shown in Fig.1.

2. Method

2.1. Multi-camera System

In this work, the term “multi-camera images” refers to images acquired from different positions in the space describing the same scene(Fig.2).

For our later processing, all the cameras involved in the acquisition system should be calibrated. To calibrate the camera system, various methods can be used. In this work a convenient self-calibration method is adopted. This method only requires the cameras to observe a planar pattern shown at a few different orientations and a cubic reference object. Then tracking a light maker in the calibration space from each viewing point synchronously[6]. The procedure is as follows:

- Estimating the intrinsic parameters by Zhengyou Zhang's flexible calibration method[15].

- Computing the initial estimate of the extrinsic parameters with Direct Linear Transform method[1] by using a cubic reference object.

- Tracking a distinct marker simultaneously from all the viewing points.

- By using the epipolar constraint provided by the marker positions in all the input views. The initial estimation of the extrinsic parameters is refined with the down-hill simplex optimization algorithm.

Since describing this method is not the primary purpose of this paper, for more details see the references.

2.2. Implementation of Active Shape Models

In this section we explain the ASM-based 3D foot shape modeling method.

ASM makes use of a prior model of what is expected in the images. The model is described by pose parameters and shape parameters. The pose parameters include scale, rotation and translate. On the other hand, the shape parameters are derived by combining statistical knowledge of object shape and shape variation from a training set of instances of the object by PCA. We aim at adjusting the pose and shape parameters to refine the model from its current location to the new location that as close to the object surface as it can be.

2.2.1 Initial Model

Our method starts from an initial model. The initial model is refined to fit the surface of object. The initial model can be generated from a human foot database (training set) including m foot models.

The training set is usually aligned manually when it was established. Each foot model is composed of l polygons with n sample points. The sample points of each foot model describe the corresponding characteristic positions of foot. Thus, we can establish a standard foot model $\bar{\boldsymbol{v}}$ by computing the average position of each sample point (Eq.(1)).

$$\bar{\boldsymbol{v}}_i = 1/m \sum_{j=1}^{m} \boldsymbol{v}_j \quad (i = 1 \cdots n) \tag{1}$$

where $\boldsymbol{v}_i$ is the coordinate vector of sample point. i is the index of sample points, j is the index of foot models. Then standard model is also composed of l polygons and n sample points. In our research we use the standard model as initial model.

2.2.2 Pose Matching of the Model and Object

In ASM the position and orientation of the initial model should be registered with image data.

The pose of 3D model is described by scale matrix $\boldsymbol{A} = diag(s, s, s)$, rotation matrix $\boldsymbol{R}$ by $\theta = (\theta_x, \theta_y, \theta_z)$ and translate vector $\boldsymbol{T} = (T_x, T_y, T_z)^T$. The initial model is refined to fit the object in position and orientation with optimal pose parameters.

The silhouette images of multi-camera images are used to evaluate pose parameters. Because all of the camera parameters have been estimated by thorough calibration (section 2.1), the projection matrix $\boldsymbol{P}$ can be established. Then the sample points of 3D model can be projected onto the silhouette images i.e., the world coordinates of sample points are transformed to 2D coordinate system of image plane.

In this way the 3D model and 2D multi-camera images are related. Because the intensity K_{ij} of foot image is always higher than the background in binary silhouette image, updating the pose parameters to maximize the cost function E_{pose} (Eq.(2)) can make the 3D model approach to object.

$$E_{pose}(\boldsymbol{A}, \theta, \boldsymbol{T}) = \sum_{i=1}^{f} \sum_{j=1}^{n} K_{ij} \tag{2}$$

where, i is the index of binary silhouette multi-camera images, j is the index of sample points.

2.2.3 Principal Component Analysis Based Shape Parameters

PCA is an effective approach to describe the statistical relationship within a training set of objects. It can reduce the dimensionality of the data to something more manageable.

In the training set, each foot example is represented in a $3n$ dimensional space. Every example in this space gives a set of sample points whose shape is broadly similar to that of those in the training set. Thus by moving about the sample points we can generate new shapes in a systematic way. However the n is more the cost of computation will go up distinctly. So we are anxious a good efficiency approach in low dimensional space. In order to achieve this goal, we apply PCA to the training set. For each shape in the training set we calculate its deviation from the standard model $\bar{\boldsymbol{v}}$ (Eq.(1)) as:

$$d\boldsymbol{v}_i = \boldsymbol{v}_i - \bar{\boldsymbol{v}} \tag{3}$$

Then the $3n \times 3n$ covariance matrix $\boldsymbol{S}$ of database is calculated by

$$\boldsymbol{S} = 1/m \sum_{i}^{m} d\boldsymbol{v}_i d\boldsymbol{v}_i^T \tag{4}$$

The principal components give the modes of variation of the shape, are described by $\boldsymbol{p}_k$ (k=1, $\cdots$, $3n$), which the eigenvectors of covariance matrix $\boldsymbol{S}$ such that

$$\boldsymbol{S}\boldsymbol{p}_k = \lambda_k \boldsymbol{p}_k \tag{5}$$

where λ_k is the corresponding eigenvalue of $\boldsymbol{p}_k$, $\lambda_k > \lambda_{k+1}$. It can be shown that the eigenvectors of the covariance matrix corresponding the largest eigenvalues describe the most significant modes of variations. Then the variations can be explained by a small number of modes, q. This means that the $3n$ dimensional space is approximated by a q dimensional space, where q is chosen so that the smallest number of the modes such that the sum of their variances explains a sufficiently large proportion of the total variance of all the variables.

Any samples of object can be reached by adding a linear combination of the eigenvectors to the standard model $\bar{\boldsymbol{v}}$

$$\boldsymbol{v} = \bar{\boldsymbol{v}} + \boldsymbol{P}\boldsymbol{B} \tag{6}$$

where $\boldsymbol{P} = (\boldsymbol{p}_1, \boldsymbol{p}_2, \cdots \boldsymbol{p}_q)$ is the matrix of the first q eigenvectors, and $\boldsymbol{B} = (b_1, b_2, \cdots, b_q)$ is a vector of weights. Because the foot shape is controlled by the weights vector, $\boldsymbol{B}$ is so called shape parameters.

2.2.4 Using the Method for Search Images

Optimal pose parameters and shape parameters are estimated with an iterative scheme. While these parameters are updated one by one during the iterative process, if the difference of translate is too big, the refinement of other parameters will be invalid. Consequently, we assume the optimization of rotation, scale parameter and shape parameters is started when the superposition area of the sample points' projection and silhouettes in multi-camera images is more than 80%. Then the iterative will stop until the update ratio of the sample points between the last iterative and the current iterative is under a threshold. In our experiment, the threshold is always set less than 1%

As described in section 2.2.2, the optimal pose parameters were searched by maximizing the cost function Eq.(2). Thus we also need a cost function for estimating the optimal shape parameters. We define the cost function as:

$$E_{shape}(\boldsymbol{B}) = \sum_{i=1}^{f}\sum_{j=1}^{n}(w_{ij}\left|I_{ij} - \bar{I}_j\right| + w_{ij}^{'}D_{ij}) \tag{7}$$

where i is the index of camera, j is the index of sample points. I_{ij} is the intensity of the projection pixel of sample point, D_{ij} is the distance between the sample point whose projection on the contour of the projection area of all the sample points (we call them "contour sample points") and the closest edge of foot in multi-camera images. w_{ij} and $w_{ij}^{'}$ are weight factors. $\bar{I}_j = \frac{1}{f}\sum_{i=1}^{f} I_{ij}$ is the average of I_{ij}. Thus, while the first term of the cost function is getting smaller, the sample point's position is more credible. However, this constraint will lead the projection of sample points to fall into the background field, in some extreme case. Consequently, we set a distinctive condition to the cost function: the euclidean distance between the contour sample points and the edge of object in the multi-camera images. By minimizing the cost function, we can obtain the optimal approach of the object surface. The optimization of the cost function is implemented with "Rosenbrock" algorithm[10]. We will explain how to seek the contour sample points in the next section.

2.3. Occlusion

Because the intensity of sample points' projection on multi-camera images is used to evaluate the parameters of ASM, we must think over the problem of occlusion.

In order to solve the occlusion issue, we investigate the spatial relationships of sample points' projection with polygons' projection in image plane. The relationships are arranged in two classes: Inside and Outside (Fig.3).

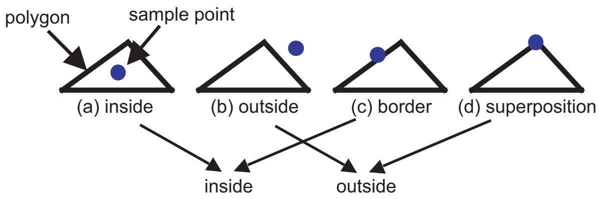

Figure 3. relationships of sample points with polygons

For "Outside" sample point will not be occluded by polygon obviously, but for "Inside", the problem is complex. In Fig.4, the projection of occluded sample point F is inside a polygon's projection. On the other hand, although D's projection is also inside a polygon's projection, D is not occluded. Thus, we should do further investigation. For instance, in order to estimate if W is occluded by $\triangle XYZ$, the world coordinates of W and the vertices of $\triangle XYZ$ are transformed to camera coordinate, as W', $\triangle X'Y'Z'$ respectively. Then we can compute the signed distance d between W' and $\triangle X'Y'Z'$.

Because in our experiments the object is always set near the center of the space, the sign of distance are considered to be the criterion of occlusion. If d is positive the sample point is between camera and a polygon, the sample point is not occluded, otherwise the sample point is considered to be occlusion.

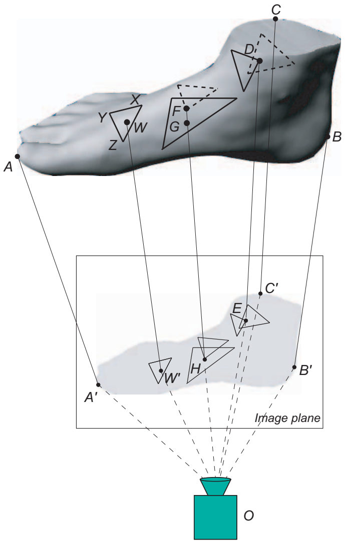

Figure 4. The solid line triangles represent the polygons that faces the camera O, broken line triangles represent the polygons on the reverse surface to the camera. The polygons on the reverse surface are invisible.

For attentively, if the sample points' projections are on the contour of all the sample points' projection area, its projection may inside no polygon'S projection. We call this kind of sample points "contour sample points" (CSP). In Fig.4 sample points A, B, and C are so-called CSP.

In order to reflect the occlusion's effect on the process, we redefine the shape parameters cost function as

$$E_{shape}(\boldsymbol{B}) = \sum_{i=1}^{m}\sum_{j=1}^{n}(c_{ij}w_{ij}\left|I_{ij} - \bar{I}_j\right| + c_{ij}^{'}w_{ij}^{'}D_{ij}) \quad (8)$$

where $c_{ij} = \begin{cases} 0 & occlusion \\ 1 & otherwise \end{cases} \quad c_{ij}^{'} = \begin{cases} 1 & CSP \\ 0 & otherwise \end{cases}$

3. Experiments

3.1. Computer Graphics Data

We apply the proposed method in a computer simulative experiment firstly. Multi-camera system includes 32 virtual cameras is created by Povray[8]. Multi-camera images of a premade foot model with random pattern texture were taken by this system. Some examples of the multi-camera images are shown in Fig.5. The resolution is 640×480 pixels.

The shape parameters are derived from PCA of a training set. In this experiment a human foot database that is composed of 397 adults' right foot is adopted. Table 1 shows the distribution of size in the database. Each foot model is composed of 740 polygons with 372 sample points.

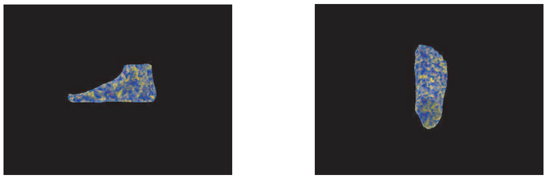

Figure 5. CG multi-camera images of foot model

Table 1. Human foot database

Size (U.S. size)	5	6	7	8	9	10
Male	-	17	56	68	48	23
Female	11	20	78	42	24	10

The cumulative contribution ratio $cum_i = \sum_{j=1}^{i} c_j$ is computed for deciding shape parameters. Where $c_i = \lambda_i/\lambda_T$ is called contribution ratio of covariance matrix's eigenvector λ_i, and $\lambda_T = \sum_{i=1} \lambda_i$. According to the cumulative contribution ratio, more than 90% of the variance is explained by the first 12 modes of variation. On the other hand, from the $13th$ mode the contribution ratios are less than 0.5%, thus the influence of them can be ignored. The weights of the first 12 modes are considered to be the shape parameters in our experiments.

The result of our proposed method is shown in Fig.6. The points group in the images is the projection of sample points. For display the result distinctly, the sample points are projected to the silhouette images that are extracted by background subtraction. It is obviously that during the iterative (from left to right) the projections of sample points are getting fitted the foot image well. Fig.7 is the foot model displayed in 3-dimensional.

Since volume intersection is a very popular surface reconstruction technique, we are very interested in the comparison with it. Fig.8 is the result of volume intersection in 3-dimensional. According to this comparison, volume intersection's result gives a little incondite impress, particularly in the concave part of the surface. On the other hand, our proposed method's result is smoother and more similar to the real human foot than volume intersection.

Because the CG foot model is known previously, the euclidean distance between the sample points and the surface of CG foot model would be a reasonable criterion. The root mean square error of the proposed method is 2.21mm, and

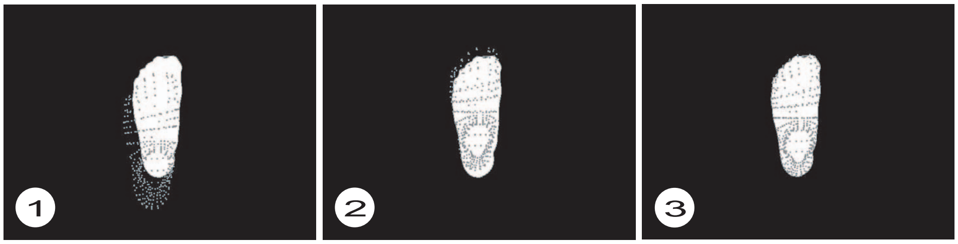

Figure 6. Silhouette images of CG foot model with the projection of sample points superimposed, during iterative process (from left to right)

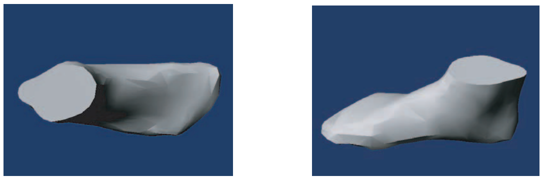

Figure 7. Reconstructed 3D model of CG foot model by proposed method

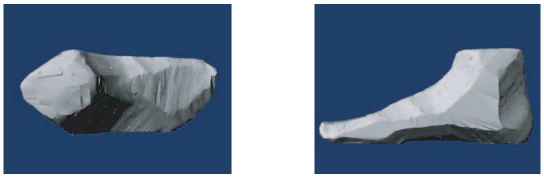

Figure 8. Reconstructed 3D model of CG foot model by volume intersection

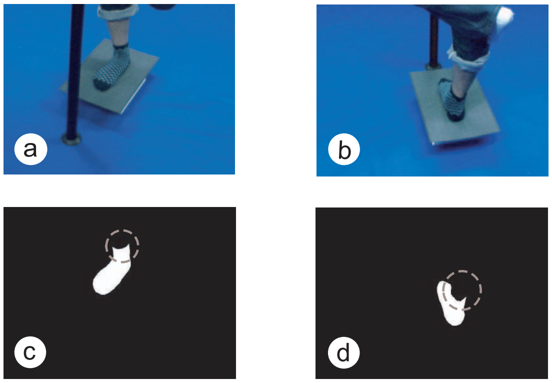

Figure 9. Multi-camera images of real human foot((a),(b)), and their silhouette images((c),(d)), in which the broken circles show the "leg occlusion"

the error of volume intersection is 4.44mm. According to this comparison, we can say our proposed method provides more precise result than volume intersection.

Moreover, in our previous research, the 3D Active Contour Models (ACM) is used for reconstruction of foot model[12]. More than 3000 control points are aligned on the surface. While smooth foot model can be reconstructed, the shape features of foot are not sufficient. This is occurred by the smooth effect of ACM is difficult to be controlled. Particularly, the root mean square error is about 3.50mm that is inferior to our proposed method.

3.2. Real Data

The proposed method is also implemented in real data experiments. The multi-camera system includes 8 CCD cameras whose frame rate is 7.5fps. The resolution is also 640 × 480 pixels (Fig.9(a)(b)). For making the surface texture strong, the object puts on socks. The silhouette images are shown in Fig.9(c)(d). In CG experiment, silhouette images are obtained by background subtraction. For real human foot, we not only take the background subtraction, but also remove the unwanted parts e.g. the leg, manually. Although the manual "leg remove" work is unefficient, because there are only 8 multi-camera images, the workload is still acceptable. However, in our future work, we will exploit automatic "leg remove" processing.

Fig.10 shows the result in 2-dimensional. The 3-dimensional result is shown in Fig.11. We also compared the result with volume intersection (Fig.12). However, in this case volume intersection's result is too rough to understand its shape. This principally happened by "leg occlusion" phenomenon, which is occurred because part of foot in images is occluded by leg. The volume intersection back-projects the voxels to multi-camera silhouette image in a cone space, then intersects all the cones to be the approach of the object. Consequently, if there are big gaps on the silhouette images, the reconstructed surface must be affected

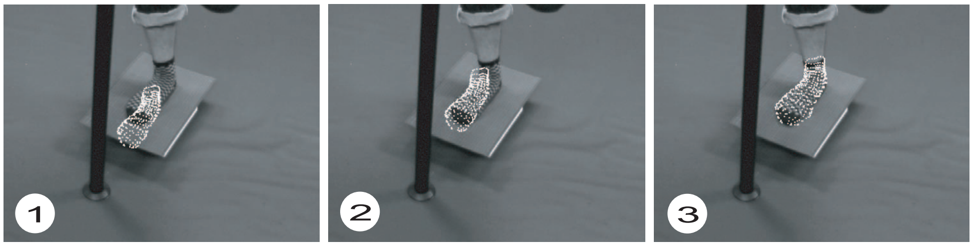

Figure 10. Multi-camera images of real human foot with the projection of sample points superimposed, during iterative process (from left to right)

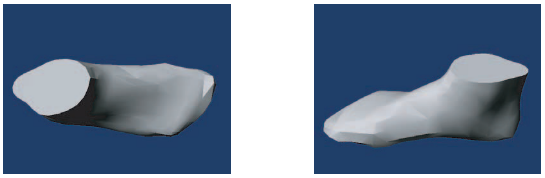

Figure 11. Reconstructed 3D model of real human foot by proposed method

Figure 13. Multi-camera images of a plastic foot model

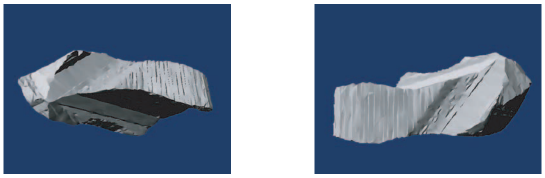

Figure 12. Reconstructed 3D model of real human foot by volume intersection

largely. In the proposed method foot database is adopted to establish an intact foot model as the initial model and statistical knowledge of human foot is applied to amend the missing data, hence the effect of leg occlusion can be covered. According to Fig.11 the same problem of volume intersection has not been occurred. Therefore, the proposed method is attested to be robust.

For real camera data, we don't know the answer previously. Thus, how to evaluate the result is becoming a difficult issue. In this work, an experiment of a plastic foot model whose position information was measured previously, was done. The multi-camera images are shown in Fig.13. The result of our iterative method is shown in Fig.14. Then the similar evaluation of CG data is carried out. The root mean square error is about 2.46mm. The error is in the same level to the CG experiment.

Since there is no effect like "leg occlusion" for plastic foot model, the result of volume intersection method (Fig.15) is not as bad as Fig.12. However, 8 cameras are too less for volume intersection algorithm, the reconstructed 3D model is still too rough to be considered an acceptable approach of foot. On the other hand, under the same condition (using 8 CCD cameras) a very satisfied 3D model is obtained by our proposed method (Fig.16).

4. Conclusions

In this paper we proposed a foot surface modeling method. The images of foot surface are acquired by multi-camera system in a free space. Because PCA of human shape database reduced the dimensions for representing the object shape, we can efficiently recover the object shape from multi-camera images by more effective and general parameters. The stable object shape can be obtained, even though the object shape is partially occluded in some of input views. However, there are still some remained issues. Firstly, the sole of foot cannot be recovered by our multi-camera system. Although we use the sole of the standard foot shape from database as substitute, the accuracy was affected. Secondly, because in our experiment we put on sock to add texture of foot, the reliability was also decreased. These disadvantages will be improved in our future work. Moreover, in our future work, the proposed method of this paper should be extended to cope with motion data. We assume the foot is in still condition on floor in the first frame of image sequence. The initial foot model is refined to fit the first frame of images sequence by our proposed method.

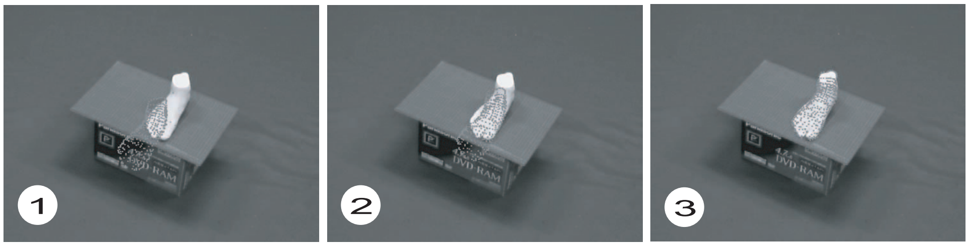

Figure 14. Multi-camera images of plastic foot with the projection of sample points superimposed, during the iterative process (from left to right)

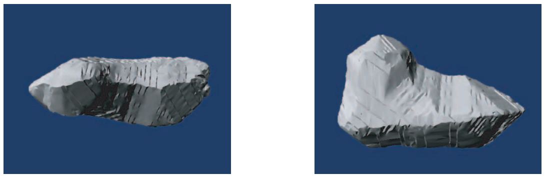

Figure 15. Reconstructed 3D model of plastic foot model by volume intersection

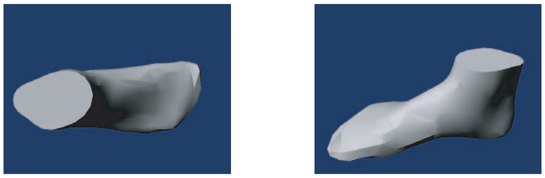

Figure 16. Reconstructed 3D model of plastic foot model by proposed method

The sample points of foot model are tracked in rest frames to generate a dynamic 3D foot model.

Acknowledgments

This work is supported in part by a Grant in Aid for the 21st century Center of Excellence for Optical and Electronic Device Technology for Access Network from the Ministry of Education, Culture, Sport, Science, and Technology in Japan.

References

[1] Y.l.Abdel-Aziz, H.M.Karara, Direct Linear Transformation from Comparator Coordinates into Object Space Coordinates in Close Range Photogrammetry, In *ASP/UI Symposium on Close-Range Photogrammetry Proc.*, pp.1-18, 1971.

[2] N.D'Apuzzo, Surface measurement and tracking of human body parts from multi-image video sequences. *Journal of Photogrammetry and Remote Sensing*, 56(5-6):360-375, 2002.

[3] M.A.Brunsman, H.Daanen, K. M. Robinette. Optimal postures and positioning for human body scanning. In *3DIM'97 Proc.*, pages 266-273, 1997.

[4] T.F.Cootes, C.J.Taylor, D.H.Cooper and J.Graham. Active Shape Models-Their Training and Application. *Computer Vision and Image Understanding*, 61(1):38-59, 1995.

[5] A.Hill, A.Thornham, and C.J.Taylor. Model-based interpretation of 3D medical images. In *BMVC'93 Proc.*, pages 339-348, 1993.

[6] http://www.library.co.jp/Eng/index.html

[7] N. Nishida, S. Fukushima and M. Minoh. A Method of Estimating Human Shape by Fitting the Standard Human Model to Partial Measured Data. In *ACCV'00 Proc.*, pages 276-281, 2000.

[8] http://www.povray.org/

[9] J.A.Roebuck,Jr.. Anthropometric Method: Designing to fit the human body. Human Factors and Ergonomics Society, 1995.

[10] H.H.Rosenbrock. A Automatic Method for Finding the Greatest or Least Value of a Function. *The Computer Journal*, 3(3):174-184, 1960.

[11] J.Siebert, S.Marshall. Human body 3D imaging by speckle texture projection photogrammetry.*Sensor Review*, 20(3):218-226, 2000.

[12] J.Wang, H.Saito, M.Kimura, M.Mochimaru and T.Kanade. Reconstruction of Human Foot from Multiple Camera Images with 3D Active Contour Models. In *AISM'04 Proc.*, pages 390-395, 2004.

[13] S.Weik, A Passive Full Body Scanner Using Shape from Silhouettes, In *Proc. ICPR'00 Proc.* pages 1750-1753, 2000.

[14] L.Zhang, N.Snavely, B.Curless, and S.M.Seitz. Spacetime Faces: High-resolution capture for modeling and animation. In *ACM SIGGRAPH'04 Proc.*, pages 548-558, 2004,

[15] Z.Zhang. Flexible Camera Calibration By Viewing a Plane From Unknown Orientations. In *ICCV'99 Proc.*, pages 666-673, 1999.