

Electronic Letters on Computer Vision and Image Analysis 7(4):51-61, 2009

Distortion Correction for 3D Scan of Trunk Swaying Human Body Segments

Takuya Funatomi*

Masaaki Iiyama⁺

Koh Kakusho*

Michihiko Minoh*

* *Academic Center for Computing and Media Studies, Kyoto University, Yoshida-Honmachi, Kyoto, Japan*

⁺ *Graduate School of Economics, Kyoto University, Yoshida-Honmachi, Kyoto, Japan*

Received 9th June, 2008; accepted 11th February, 2009.

Abstract

We propose a method for acquiring a 3D shape of human body segments accurately. Using a light stripe triangulation range finder, we can acquire accurate 3D shape of a motionless object in dozens of seconds. If the object moves during the scanning, the acquired shape would be distorted. Naturally, humans move slightly for making balance while standing even if the subject makes an effort to stay still for avoiding the distortion in acquired shape. Our method corrects the distortion based on measured subject's motion during the scanning.

Experimental results show the accuracy of the proposed method. Trunk swaying degrades the accuracy of the light stripe triangulation from 1mm to 10mm. We can keep the accuracy of as good as 2mm by applying our method.

Key Words: Anthropometry, Human Body Modeling, 3D Reconstruction, Light Stripe Triangulation, Trunk Swaying

1 Introduction

For computer-aided design (CAD) of apparel, ergonomic and medical products, the 3D shape of each particular human body becomes more and more useful. In fact, several commercial products [1, 2, 3] based on light stripe triangulation rangefinders have been developed to reconstruct 3D shape of an individual.

Although the light stripe triangulation rangefinder can accurately ($< 1\text{mm}$) acquire the 3D shape of an object as a dense 3D point-cloud on the surface, it takes a certain time to scan the whole object with a laser sheet. If the object moves during the scanning, the acquired shape would be distorted and the accuracy is degraded. Naturally, humans move slightly for making balance while standing even if the subject makes an effort to stay still and stop breathing for avoiding the distortion of the shape. Such movement is called *trunk sway*. Our goal is to maintain the accuracy of reconstructing the 3D human shape by light stripe triangulation.

In order to cope with the distortion caused by the trunk sway in human body measurement by the light stripe triangulation, several methods [4, 5, 6, 7] can reduce the distortion by speeding up the measurement. On the

Correspondence to: <funatomi@media.kyoto-u.ac.jp>

Recommended for acceptance by Francisco Perales

ELCVIA ISSN:1577-5097

Published by Computer Vision Center / Universitat Autònoma de Barcelona

other hand, we propose in this paper to measure the motion of human body due to the trunk sway so that we correct the distortion of the acquired shape based on the subject's motion. To obtain the motion, we estimate the rigid motion of each segment of the human body at each moment during the measurement by assuming that the human body consists of several rigid segments. For rigid motion estimation, we put markers on the skin and measure their 3D position with cameras.

This paper is organized as follows. In section 2, we show how the 3D shape is reconstructed by the light stripe triangulation and discuss how the acquired shape is distorted by the subject's motion. We present our method in section 3. We show experimental evaluation of the accuracy of the measurements for all segments and the acquired shape of a right upper arm and breast in section 4. We conclude the paper with some discussion and future works in section 5.

2 Background

2.1 The principle of light stripe triangulation

Light stripe triangulation reconstructs the 3D shape of an object as follows:

1. Project a laser sheet on the object.
2. Observe the projected laser sheet, which is observed as a thin stripe on the surface of the object, with calibrated camera(s).
3. Calculate the 3D position of the points that are on the observed stripe in the camera image from the position of the plane of the laser sheet and the pose of the camera using the triangulation.

From camera image at a moment, we can acquire only the partial shape of the whole object where is illuminated by the laser sheet. We call each of partial shapes as *stripe shape*. In order to acquire the full shape of the object, the laser scans the whole object and the cameras observe the scanning (see also Fig.1.) In comparison with other 3D shape capture methods, this method can acquire high resolutional and accurate shape but takes much time for scanning the whole object with the laser sheet.

2.2 Problem of human body scanning

With light stripe triangulation for a stationary object (see Fig.2-a), we can acquire an accurate shape. As for moving objects (see Fig.2-b), the acquired shape would be distorted by the object's motion during the scanning.

This distortion problem will often occur when scanning the human body. Some commercial products [1, 2, 3] take about a dozen of seconds to scan the whole body with accuracy of about 0.5mm. In order to acquire the undistorted 3D shape, the subject needs to stop moving during the scanning. However, the subject moves slightly for making balance while standing (trunk sway.)

The trunk sway is observed to assess postural stability in medical, physiological and biomechanical science. Generally, they observe the movement of body center-of-mass (COM) as the trunk sway and regard the trunk sway as the oscillation of an inverted pendulum. As has been reported in [9], the average absolute amplitude and the average frequency of the oscillation are about 5mm and 1Hz, respectively. Such oscillation will distort the shape acquired by light stripe triangulation.

We discuss two approaches for avoiding the distortion due to trunk sway. One approach is speeding up the measurement for reducing the distortion. Speedy scanning will make the distortion small. The other, which we propose in section 3, is correcting the distortion based on the motion of the body.

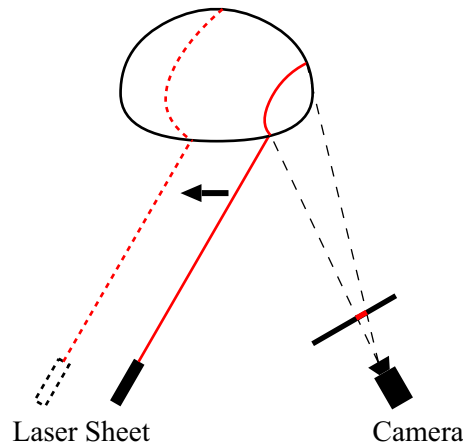


Figure 1: Principle of the light stripe triangulation.

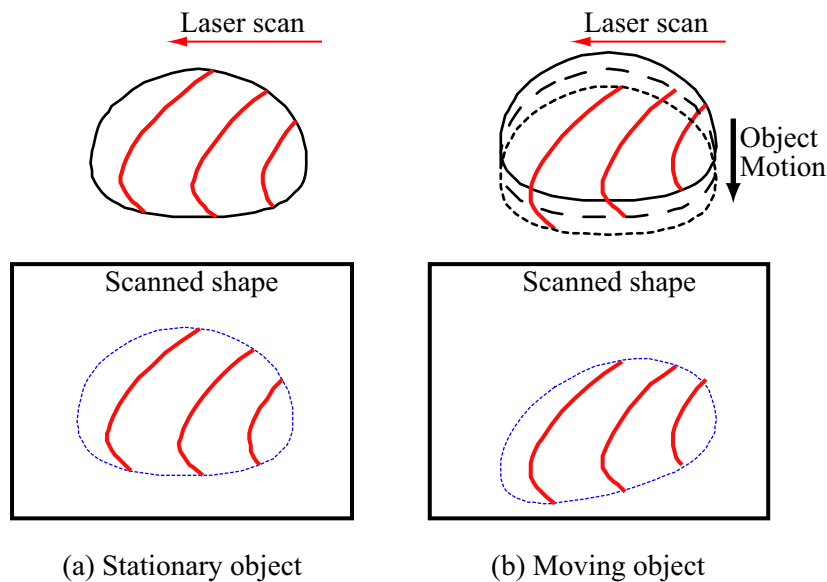


Figure 2: Shape distortion due to object motion.

2.2.1 Speeding up the measurement

Several methods [5, 6, 7] have been proposed to speed up the measurement by parallelizing the scanning with multiple lasers. Although these methods are applicable to a stationary object, a problem will occur when they are applied to a moving object. Fig.3 illustrates vertical scanning of an object which moves horizontally. With a single laser (see Fig.3-a), the acquired shape is distorted due to the object's motion. In comparison with single scanning, parallel scanning can acquire less distorted shape (see Fig.3-b). However, the acquired shape is segmented, and each segment (A, B in Fig.3-b) will not match because the time of capture is different between borders ($t = T/2$ at the bottom of A and $t = 0$ at the top of B). As mentioned above, speeding up the measurement can reduce the distortion, but cannot acquire the correct shape of the moving object.

Sotoca et al.[8] has been proposed a method of capturing the topography of the large surface with just a single image, avoiding swaying problems. They applied the method to measure the human back and presented the accuracy of as good as 3mm. Since this method is suitable to be applied to large objects, the accuracy will become degraded for measuring objects of complex geometry. Therefore, this method is not suitable for measuring the whole body, which has complex geometry.

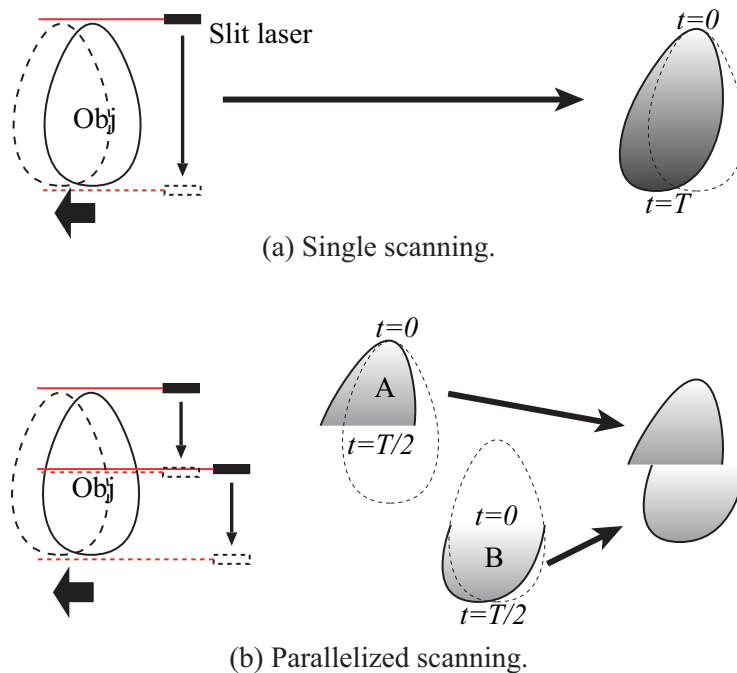


Figure 3: Mismatching problem in parallel scanning.

3 Distortion correction process

In this paper, we propose an alternative approach to human body measurement that corrects the distortion of the acquired shape based on trunk sway.

Generally, the human body will change its shape with various postures due to muscle contractions. However, the shape will not change much locally when the subject tries to maintain the posture. We assume that the human body can be divided into some rigid segments, and the trunk sway of the subject is described as the set of rigid motions of each segment. Here, we divide the whole body into 11 rigid segments ($s := \text{head, breast, waist and right and left upper arms, forearms, thighs and legs}$).

The acquired shape of a segment s by light stripe triangulation consists of 3D points on the segment's surface as mentioned in section 2.1. The position of each point would be changed by the rigid motion of s due to trunk sway. This distorts the shape acquired by light stripe triangulation. Acquiring the rigid motion during the scanning, we calculate the original position of each point and acquire the undistorted whole shape. Consequently, we have to measure shape and motion of the subject simultaneously.

3.1 Rigid motion acquisition

To acquire the trunk sway during the scanning ($t = 0, 1, \dots, T$), we put N^s (more than or equal to three) markers on the surface of each segment s , and observe them with more than 2 synchronized and calibrated cameras to calculate 3D position using the triangulation method.

To calculate the 3D position of the markers, we need the 2D position of each marker on each camera's image plane and the correspondence of each 2D position among the cameras. The markers' positions on camera images are detected with pixel values. Usually, light stripe triangulation is performed in a dark place for detecting only the light stripe on the object clearly in order to measure the shape accurately. However, normal marker will be unobservable in a dark place. If we use LED as marker, although we can observe it in a dark place, it will be obstacle to measure the shape since the LED has power line. For realizing marker detection and 3D measurement in such dark place without preventing the shape measurement, we utilize the black light

and fluorescent markers. Under the black light, only the projected laser sheet and the markers are observable with the cameras since the fluorescent markers will glow under the black light. In order to measure 3D position of markers, we use more than two synchronized and calibrated cameras. We employ Zhang's method [10] and the factorization method [11] for the camera calibration.

Let M_k^s ($k = 1, \dots, N^s$) denote the k -th marker on a segment s . At t , we express its 3D position as a 4x1 column vector in homogeneous coordinates, $M_k^s(t)$. $M_k^s(t)$ is calculated using triangulation with 2D positions of M_k^s on each camera image which acquired at time t . In our method, we track the markers on sequence of images $I(t)$ for each camera to get 2D marker positions $\mathbf{m}_k^s(t)$. We give the 2D position $\mathbf{m}_k^s(0)$ for all M_k^s of the initial frame $I(0)$ of each camera manually. For other frames $I(t)$, $\mathbf{m}_k^s(t)$ are estimated sequentially from their pixel value at $\mathbf{m}_k^s(t-1)$ on $I(t-1)$ and the pixel values around $\mathbf{m}_k^s(t-1)$ on $I(t)$. $\mathbf{m}_k^s(t)$ is estimated as the center of the area that has values similar to the pixel value of $\mathbf{m}_k^s(t-1)$ on $I(t-1)$.

We give the correspondences of the markers' 2D position among cameras manually. Then, we can calculate markers' 3D positions $M_k^s(t)$.

3.1.1 Rigid transformation estimation

From the obtained $M_k^s(t)$, we estimate a rigid motion of segment s from $t=i$ to j , $\mathbf{W}_{i,j}^s$ which is expressed as a 4x4 rigid transformation matrix. $M_k^s(j)$ is given by $M_k^s(i)$ and the rigid transformation matrix $\mathbf{W}_{i,j}^s$ as:

$$M_k^s(j) = \mathbf{W}_{i,j}^s M_k^s(i). \quad (1)$$

We define an error function $E^s(\mathbf{W})$ as follows:

$$E^s(\mathbf{W}) = \frac{1}{N^s} \sum_{k=1}^{N^s} \|M_k^s(j) - \mathbf{W} M_k^s(i)\|. \quad (2)$$

Here, N^s is the number of markers and \mathbf{W} is a rigid transformation matrix which is defined by 6 parameters. $\mathbf{W}_{i,j}^s$ is estimated as the \mathbf{W} that minimizes $E^s(\mathbf{W})$ using the Powell minimization algorithm[12].

3.2 Distortion correction

With $\mathbf{W}_{i,0}^s$ for all $t=i$ and all segments s , we can correct the distortion of the acquired shape. The distorted shape consists of $T+1$ stripe shapes which were acquired from the rangefinder at $t=0, 1, \dots, T$. For recovering the undistorted shape, we transform each stripe shape which acquired at $t=i$ into that at $t=0$ as using:

$$\mathbf{p}^s(0) = \mathbf{W}_{i,0}^s \mathbf{p}^s(i). \quad (3)$$

Here, $\mathbf{p}^s(i)$ denotes the 3D position of the point that composes the stripe shape acquired at $t=i$, and is expressed as a 4x1 column vector in homogeneous coordinates. With the above correction, we can acquire an undistorted shape as a set of $\mathbf{p}^s(0)$, which represents the subject's shape at $t=0$.

To correct the distortion for every time t and every segment s , we segment the acquired full body shape into s segments manually, and separate the shape of segment s into the $T+1$ stripe shapes which acquired at $t=0, 1, \dots, T$.

4 Results

4.1 Experiments for accuracy evaluation

The proposed method require at least only three markers for acquiring the rigid motion of each segment s . However, in order to evaluate the accuracy, we put a lot of markers $M_k^s(k=1, \dots, N^s)$ and measure their 3D position $M_k^s(t)$ with the above procedure (mentioned in 3.1) at every moment $t=0, \dots, T$. We evaluate 4 kinds of accuracy in this experiment:

- Accuracy of marker position measurement
- Non-rigid deformation of the human body segments while standing still
- Accuracy of the light stripe triangulation (with distortion)
- Accuracy of our method.

4.1.1 Non-rigidity of human body segments

To validate the assumption that the whole body can be divided into rigid segments s , we evaluate the rigidity for every segment in human body with standing still. We evaluate the non-rigidity as a residual of the eq. (2) which is used for estimating the rigid transformation $\mathbf{W}_{i,j}^s$ from a set of marker positions $M_k^s(i)$ to a set of $M_k^s(j)$ for all pairs of time (i, j) . When we suppose that the rigid transformation is adequately estimated, the residual of the eq. (2) becomes zero if the segment is completely rigid (and there is no error in marker measurement). We evaluate how largely the segment deforms non-rigidly with the residual. Let $\widehat{\mathbf{W}}_{i,j}^s$ denote the estimation of $\mathbf{W}_{i,j}^s$ from minimizing eq. (2). We evaluate the non-rigidity of segment s as an average residual $E^s(\widehat{\mathbf{W}}_{i,j}^s)$, E_W^s , which is defined as follows:

$$E_W^s = \frac{1}{T(T+1)} \sum_{\substack{i=0 \\ i \neq j}}^T \sum_{j=0}^T E^s(\widehat{\mathbf{W}}_{i,j}^s). \quad (4)$$

4.1.2 Accuracy of measuring marker position

Besides non-rigid deformation of segment s , error in measuring the marker position $M_k^s(t)$ will also influence the average residual E_W^s . We evaluate error in the measuring using the average residual of using a rigid object instead of a human body. As the rigid object, we use a mannequin which is held by a human to make the same motion as a trunk sway. We evaluate an accuracy of the marker position measurement as the average residual of $E^s(\widehat{\mathbf{W}}_{i,j}^s)$ for the rigid object, $\overline{E_W}$.

4.1.3 Accuracy of light stripe triangulation

To evaluate the influence of the trunk sway on light stripe triangulation, we simulate the distortion in acquired shape by light stripe triangulation with using the markers M_k^s .

The simulation is performed using a pair of sequences of images, $I^C(t)$ and $I^L(t)$, from two cameras C and L which observe marker positions. If a laser sheet which goes through lens center of the camera L scans the object, the stripe on the subject is observed as a straight line on camera image $I^L(t)$. Here, we suppose that the laser plane is parallel to the x-axis of the image plane of camera L and the laser is scanned along the y-axis going from $y=0$ on the image plane at $t=0$ to $y=1$ at $t=T$. The stripe at t is projected to the camera image $I^L(t)$ as the line, $y=t/T$ ($0 \leq x \leq 1, 0 \leq y \leq 1$) (see Fig.4-a). Therefore, the observation of the laser scanning from the camera L is done by integrating the lines from a sequence of the images $I^L(t)$ into a single image, which we call the *integrated scanline image* \widehat{I}^L (see also Fig.4-b).

With the integrated scanline image \widehat{I}^L , we can simulate the markers position, \widehat{M}_k^s , on the acquired shape by light stripe triangulation. When the marker M_k^s is observed on \widehat{I}^L at coordinate $\widehat{m}_k^s = (x_k^s, y_k^s)$, \widehat{M}_k^s is given by $M_k^s(\widehat{y}_k^s/T)$ since the marker would be illuminated by the laser sheet at time \widehat{y}_k^s/T . For all t , $M_k^s(t)$ is obtained using the pair of sequences of images $I^C(t)$ and $I^L(t)$, so \widehat{M}_k^s can also be acquired.

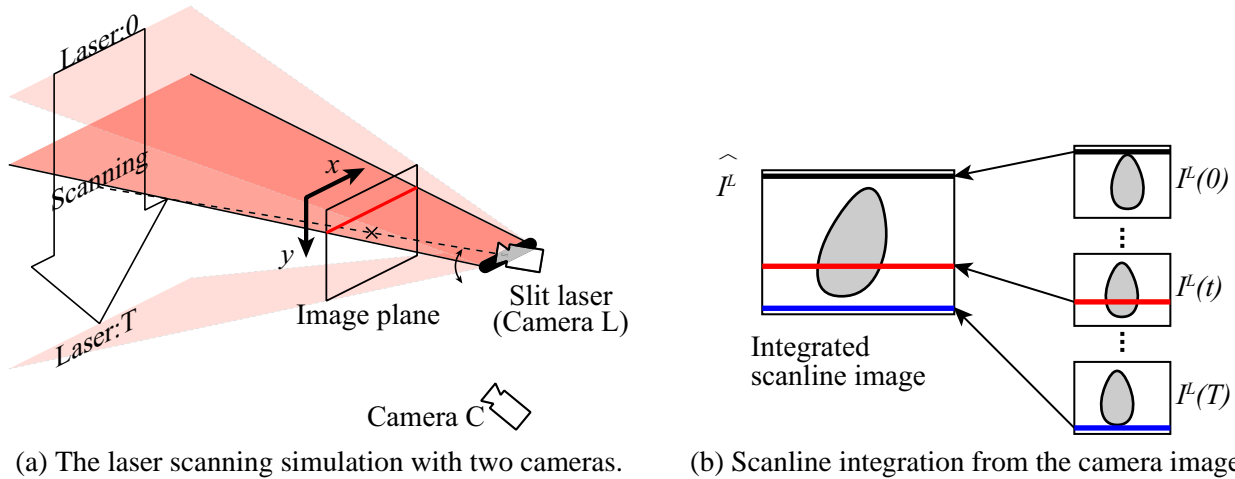


Figure 4: The simulation of the light stripe triangulation.

Using \widehat{M}_k^s and $M_k^s(t)$, we evaluate the accuracy of light stripe triangulation for a trunk swaying subject, $E_L^s(t)$, as follows for each segment s :

$$E_L^s(t) = \frac{1}{N_s} \sum_{k=1}^{N_s} \|\widehat{M}_k^s - M_k^s(t)\| = \frac{1}{N_s} \sum_{k=1}^{N_s} \|M_k^s(\widehat{y}_k^s/T) - M_k^s(t)\|. \quad (5)$$

4.1.4 Accuracy of the proposed method

To evaluate the accuracy of the proposed method, we correct the distorted shape \widehat{M}_k^s . We use only 3 of N^s markers for rigid motion estimation and use the other N^s-3 markers for accuracy evaluation. First, we calculate the estimate of the rigid transformation matrix $\widetilde{W}_{t,0}^s$ from 3 markers positions $M_k^s(t)$. Then, we transform \widehat{M}_k^s of the remaining markers into their corrected position $\widetilde{M}_k^s(0)$ with $\widetilde{W}_{t,0}^s$. Finally, we evaluate the accuracy, E_S^s , by comparing $\widetilde{M}_k^s(0)$ and $M_k^s(0)$ as follows:

$$E_S^s = \frac{1}{N_s} \sum_{k=1}^{N_s} \|\widetilde{M}_k^s(0) - M_k^s(0)\| = \frac{1}{N_s} \sum_{k=1}^{N_s} \|\mathbf{W}_{\widehat{y}_k^s/T,0}^s M_k^s(\widehat{y}_k^s/T) - M_k^s(0)\|. \quad (6)$$

4.2 Accuracy evaluation results

First, we show the result of evaluating the accuracy of the marker position measurement. We put fifty-eight 3-mm-square markers on a mannequin with about 20mm grid spacing (see Fig.5-a), and measure their 3D position for 10 seconds with 15fps cameras, that is $T = 150$. Fig.5-b illustrates the markers position with wire-frame at $t = 18, 123$. We calculated the marker measurement accuracy from all pairs of time t , \overline{E}_W was 0.73mm. Fig.5-c also illustrates markers position at $t = 123$ transformed into markers at $t = 18$ using the estimated rigid transformation matrix. This figure shows that the rigid transformation is properly estimated.

Next, we show the result of evaluating the non-rigidity of human body E_W^s , the accuracy of light stripe triangulation $E_L^s(t)$ and the accuracy of the proposed method E_S^s . As mentioned above, we put about 70 markers on the breast and waist, and about 30 markers on the head, upper arms, forearms, thighs and legs and measured the marker positions for 30 seconds, $T = 450$. Also, the subject makes an effort to keep standing still and stop breathing to avoid non-rigid deformation.

We show the camera image at $t = 0$, $I^L(0)$, and the integrated scanline image \widehat{I}^L from an upper arm observation in Fig.6. Fig.6-b shows how the trunk sway of the subject will distort the acquired shape.

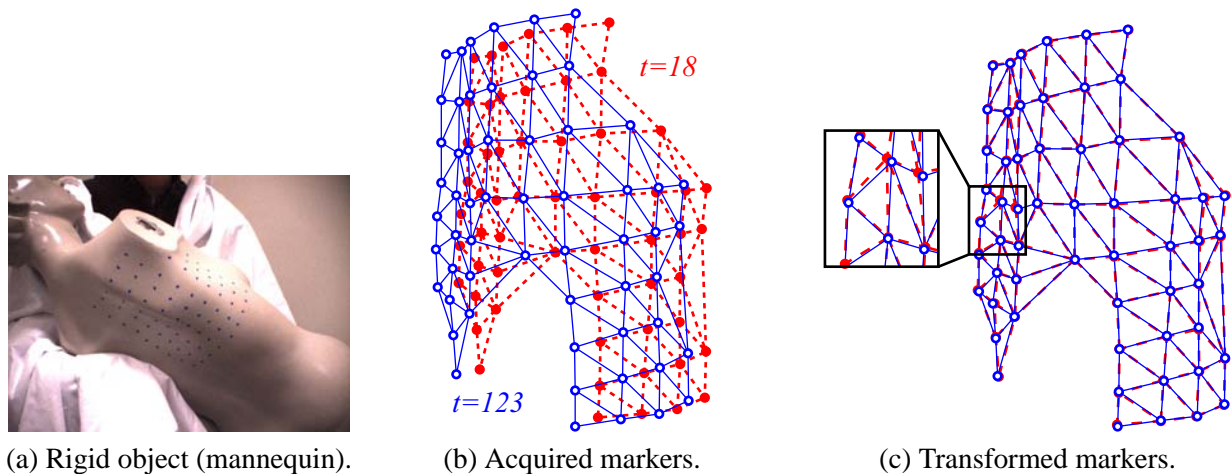


Figure 5: Experiment to evaluate the accuracy of marker position measurement.

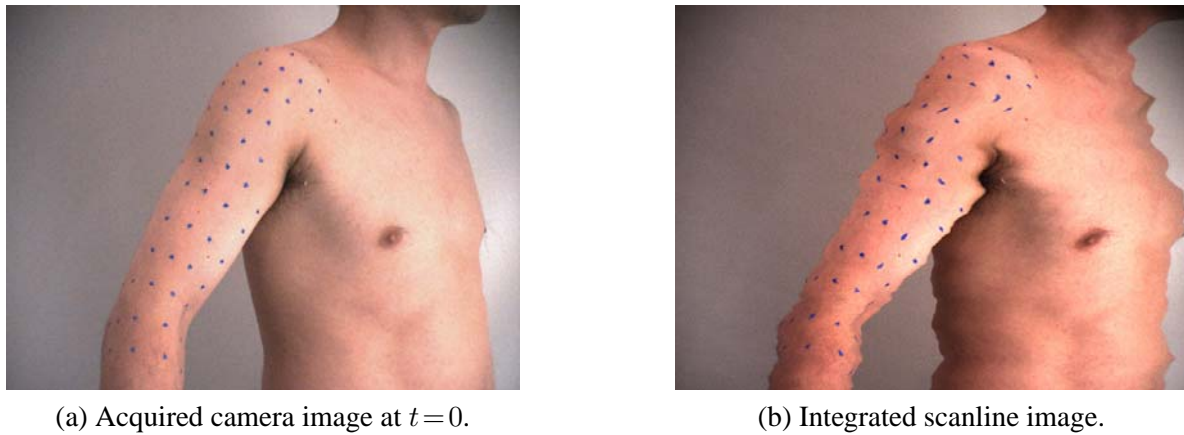


Figure 6: Experimental result of integrated scanline image.

To evaluate the accuracy, we measure the marker positions at $t = 0$, $M_k^s(0)$, the marker positions acquired with simulated the light stripe triangulation, \widehat{M}_k^s , and the positions as corrected by the proposed method, $\widetilde{M}_k^s(0)$. Fig.7 illustrates acquired marker positions $M_k^s(0)$, \widehat{M}_k^s , $\widetilde{M}_k^s(0)$ as measured from Fig.6. We chose 3 markers to estimate rigid motion (the black markers in Fig.7) and evaluate the accuracy of the correction using the remaining markers (the white markers).

We show the E_W^s , $E_L^s(0)$ and E_S^s calculated for each segment s in Table 1. From the results, we conclude that:

- In standing still, each segment of the body will deform non-rigidly less than about 1mm (E_W^s).
- Trunk sway degrades the measurement accuracy $E_L^s(0)$ to about 10mm.
- The proposed method maintains the accuracy within about 2mm (E_S^s) with distortion correction.

4.3 Measurement result

We show the acquired shape of a right upper arm and breast with two methods, traditional light stripe triangulation and the proposed method. We reconstruct the surface from acquired point-cloud by making Delaunay

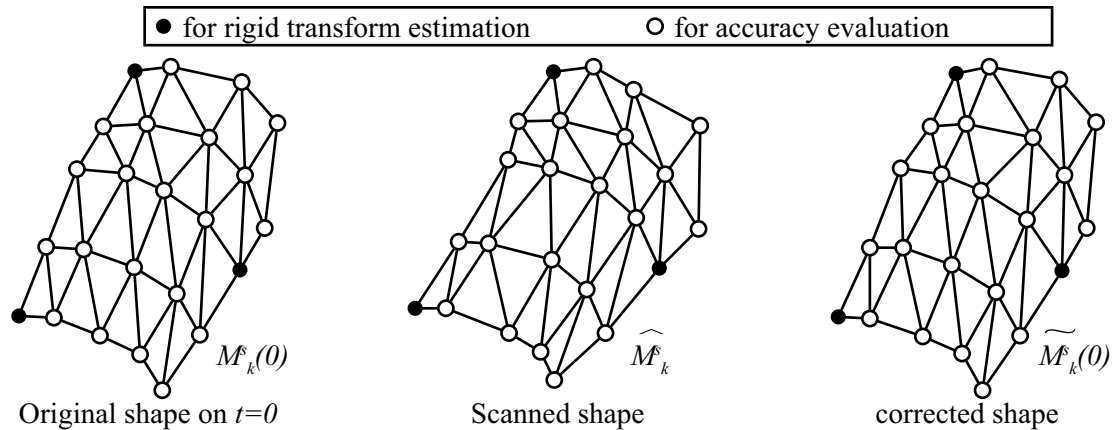


Figure 7: Experiment for markers position measurement and correction.

Table 1: Measurement accuracy evaluation for different body-parts (unit:mm).

Segment s	Non-rigidity E_W^s	Distortion $E_L^s(0)$	Correction E_S^s
head	0.69	10.32	1.09
breast	0.70	7.48	1.47
waist	0.89	6.56	1.56
upper arm	1.54	31.50	1.86
forearm	1.14	7.16	1.99
thigh	1.16	3.94	1.25
leg	1.87	3.35	2.64
Average	1.14	10.04	1.69

mesh[13]. Fig.8 illustrates the camera images for the marker position measurement and the light stripe triangulation in a dark place under black light. This figure shows that we can observe only light stripe and glowing markers. Fig.9 illustrates the shape acquired with the light stripe triangulation and the result of the proposed method. The distortion which we can see in encircled area of Fig.9-a is corrected by the proposed method, Fig.9-b.

5 Conclusion and Future works

In this paper, we discussed that the trunk sway distorts the 3D shape of a human body acquired with the light stripe triangulation and proposed a method which corrects the distortion using measured body motion. We presented how each segment of the human body undergoes less than 1mm of non-rigid deformation in standing still. Nevertheless, experimental results show that the accuracy of the light stripe triangulation is degraded from less than 1mm of error to about 10mm distortion due to the trunk sway. Our method allows for trunk swaying during the measurement with about 2mm accuracy.

We will apply our method to shape measurement of the whole body. Furthermore, all manual procedures in our method have to be automated in future.

References

- [1] Cyberware Inc., <http://www.cyberware.com/>

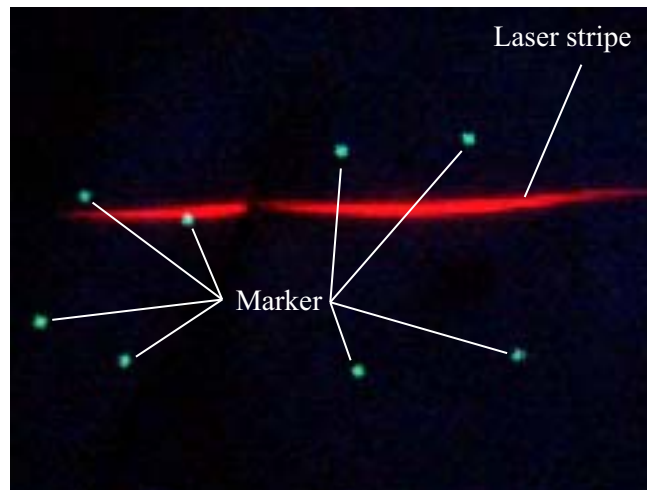
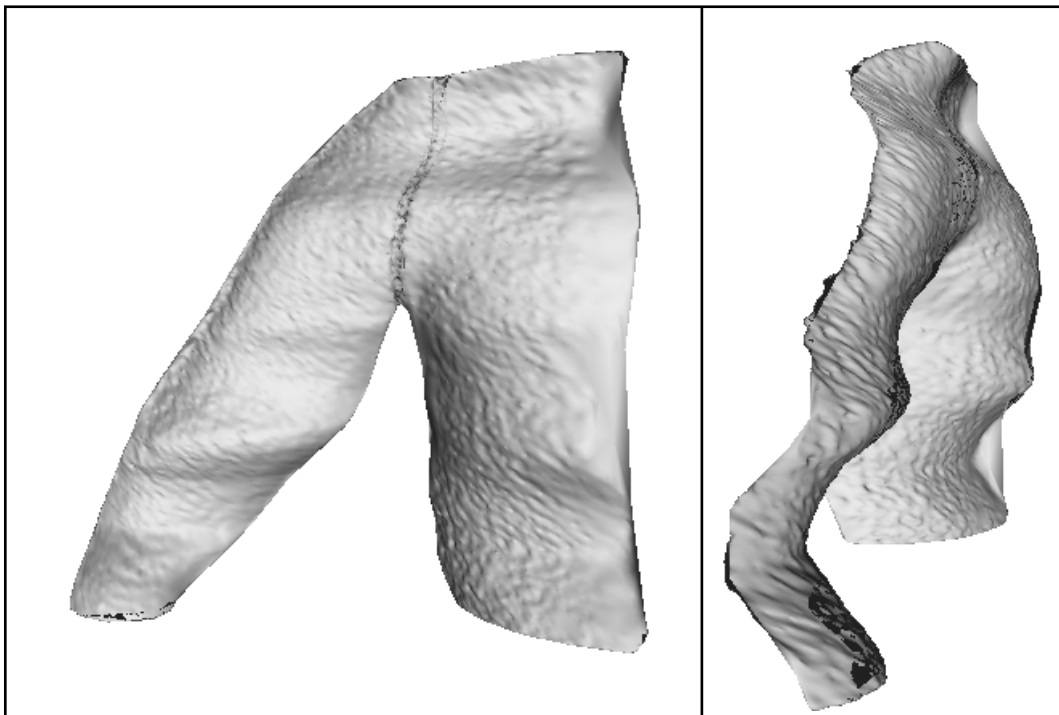
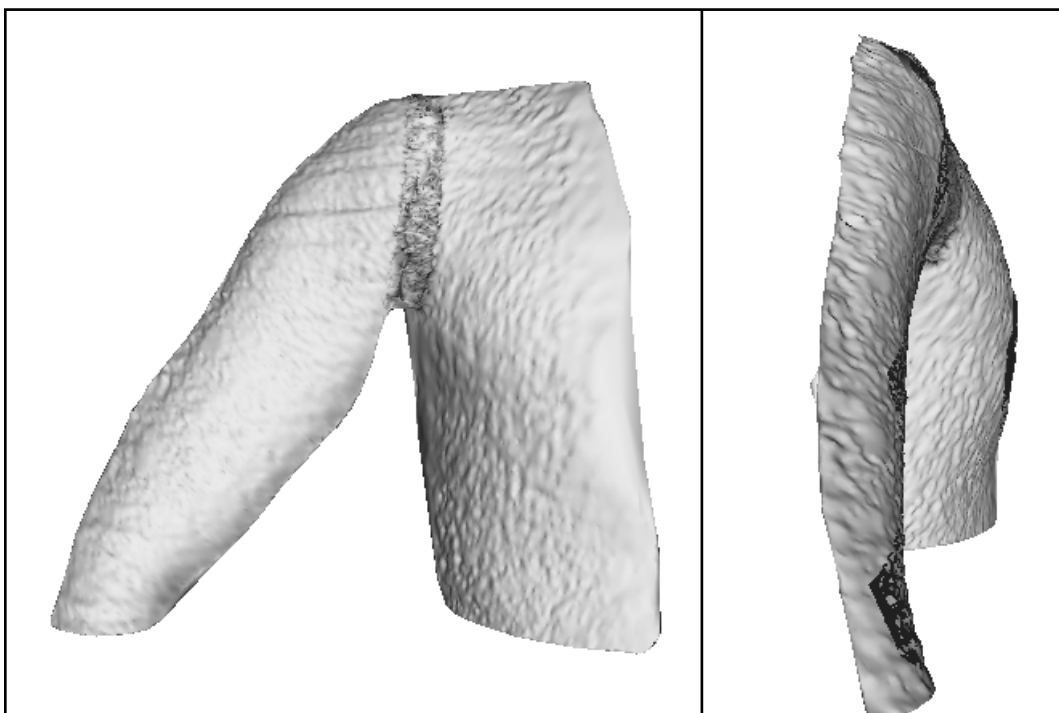


Figure 8: Synchronous observations of the markers and laser scans.

- [2] VOXELAN Inc., <http://www.voxelan.co.jp/>
- [3] Hamamatsu Inc., <http://www.hpk.co.jp/>
- [4] Y. Oike, M. Ikeda, K. Asada. "A COMS Image Sensor for High-Speed Active Range Finding Using Column-Parallel Time-Domain ADC and Position Encoder," *IEEE Trans. on electron devices*, vol.50 no.1, pp.152-853, 2003.
- [5] P. Vuylsteke, A. Oosterlinck. "Range Image Acquisition with a Single Binary-Encoded Light Pattern," *IEEE Trans. on Pattern Analysis and Machine Intelligence*, vol. 12, no. 2, pp. 148-164, 1990.
- [6] C. Sinlapeecheewa, K. Takamasu. "3D Profile Measurement by Color Pattern Projection and System Calibration," *Proc. of the 2002 IEEE International Conference on Industrial Technology*, pp.405-410, 2002.
- [7] K. Hattori, Y. Sato. "Accurate Rangefinder with Laser Pattern Shifting," *Proc. of the International Conference on Pattern Recognition*, vol.3 no.3, pp.849-853, 1996.
- [8] J.M. Sotoca, M. Buendia, J.M.Inesta. "A new Structured Light Calibration Method for Large Surface Topography," *Pattern Recognition and Applications. Frontiers in Artificial Intelligence and Applications*, pp. 261-270, 2000.
- [9] Winter DA, Patla AE, Prince F, Ishac M, Gielo-Perczak K. "Stiffness Control of Balance in Quiet Standing," *J. Neurophysiology*, vol. 80, no. 3, pp. 1211 - 1221, 1998.
- [10] Z. Zhang. "A Flexible New Technique for Camera Calibration," *Microsoft Research, MSR-TR-98-71*, 1998.
- [11] T. Ueshiba, F. Tomita. "A Factorization Method for Projective and Euclidean Reconstruction for Multiple Perspective views via iterative depth estimation," *Proc. European Conf. Computer Vision*, vol.1, pp.296-310, 1998.
- [12] W.H. Press, B.P. Flannery, S.A. Teukolsky, W.T. Vetterling. "Numerical Recipes in C, the Art of Scientific Computing, Second Edition," *Cambridge University Press*, 1992.
- [13] T. Funatomi, I. Moro, S. Mizuta, M. Minoh. "Surface reconstruction from point cloud of human body by clustering," *Systems and Computers in Japan*, Vol.37, No.11, pp.44-56, 2006.



(a) A distorted shape obtained by traditional light stripe triangulation.



(b) Undistorted shape by the proposed method.

Figure 9: Comparison of traditional and proposed method for shape reconstruction.