3D Shape Reconstruction of Trunk Swaying Human Body Segments

Takuya Funatomi, Masaaki Iiyama, Koh Kakusho, and Michihiko Minoh

Kyoto University, Yoshida-Honmachi, Kyoto, 606-8501, Japan

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 the 3D shape of a motionless object in a dozen of seconds. If the object were to move during the scanning, the acquired shape would be distorted. Naturally, humans move slightly for making balance while standing even if the subject tries to stay still for avoiding the distortion of the shape. Our method corrects the distortion based on measured motion during the scanning.

Experimental results show the accuracy of our shape measurements. 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.

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 the light stripe triangulation rangefinders have been developed to reconstruct the 3D shape of an individual.

Although the light stripe triangulation rangefinder can accurately (< 1mm) acquire the 3D shape of an object as a dense 3D point-cloud on the surface, it takes time to scan the whole object with a laser sheet. If the object moves during the scanning, the acquired shape becomes distorted and the accuracy is degraded. Naturally, humans move slightly for making balance while standing even if the subject tries 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 the light stripe triangulation for reconstructing the 3D human shape.

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

The light stripe triangulation method 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 the 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 [8], the average absolute amplitude and the average frequency of the oscillation are about 5mm and 1Hz, respectively. Such oscillation will distort the shape acquired with the light stripe triangulation method.
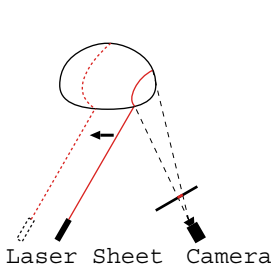
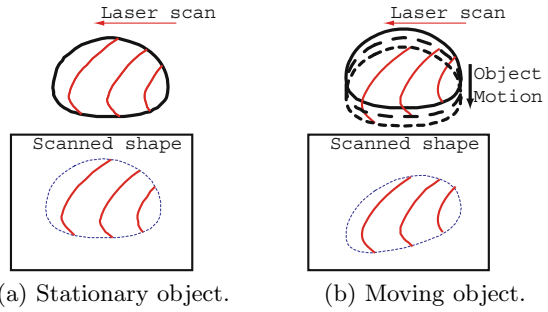Fig. 1. Principle of the light stripe triangulation

(a) Stationary object.   (b) Moving object.

Fig. 2. Shape distortion due to object motion

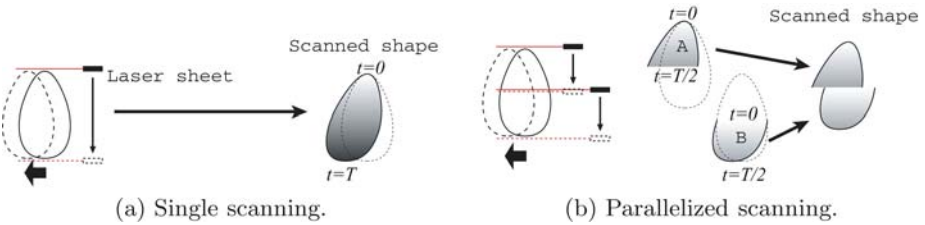(a) Single scanning.   (b) Parallelized scanning.

Fig. 3. Mismatching problem in parallel scanning

We discuss two approaches for avoiding the distortion from 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 due to trunk sway.

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 (the bottom of A and 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.

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 :=head, breast, waist and right and left upper arms, forearms, thighs and legs).

The acquired shape of a segment s using the 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 with the light stripe triangulation. Acquiring the rigid motion during the scanning, we calculate the original position of each point and acquire the undistorted 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 ($t = 0, 1, \cdots, T$), we put N^s (at least 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. To observe the markers with cameras vividly, the observation should be performed in a bright place. On the contrary, the light stripe triangulation should be performed in a dark place to detect the light stripe clearly and measure the shape accurately. To solve this conflict, we use 4 synchronized and calibrated cameras as 1 unit: 2 are adjusted in exposure for observing markers and 2 are for observing the laser stripe. We employ Zhang's method [9] and the factorization method [10] for the camera calibration.

Let M_k^s ($k = 1, \cdots, N^s$) denote the k-th marker on a segment s. At time t, we express its 3D position as a 4x1 column vector in homogeneous coordinates, $\boldsymbol{M}_k^s(t)$. $\boldsymbol{M}_k^s(t)$ 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 $\boldsymbol{m}_k^s(t)$. We give the 2D position $\boldsymbol{m}_k^s(0)$ for all M_k^s of the initial frame $I(0)$ of each camera manually. For other frames $I(t)$, $\boldsymbol{m}_k^s(t)$ are estimated sequentially from their pixel value at $\boldsymbol{m}_k^s(t-1)$ on $I(t-1)$ and the pixel values around $\boldsymbol{m}_k^s(t-1)$ on $I(t)$. $\boldsymbol{m}_k^s(t)$ is estimated as the center of the area that has values similar to the pixel value of $\boldsymbol{m}_k^s(t-1)$ on $I(t-1)$.

We give the correspondences of the markers' 2D position among cameras manually. Then, we can calculate makers' 3D positions $\boldsymbol{M}_k^s(t)$.

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

$$\boldsymbol{M}_k^s(j) = \boldsymbol{W}_{i,j}^s \boldsymbol{M}_k^s(i). \tag{1}$$

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

$$E^s(\boldsymbol{W}) = \frac{1}{N^s} \sum_{k=1}^{N^s} ||\boldsymbol{M}_k^s(j) - \boldsymbol{W}\boldsymbol{M}_k^s(i)||. \tag{2}$$

Here, N^s is the number of markers and $\boldsymbol{W}$ is a rigid transformation matrix which is defined by 6 parameters. $\boldsymbol{W}_{i,j}^s$ is estimated as the $\boldsymbol{W}$ that minimizes $E^s(\boldsymbol{W})$ using the Powell minimization algorithm[11].

3.2   Distortion Correction

With $\boldsymbol{W}_{i,0}^s$ 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 $t = 0, 1, \cdots, T$. To recover the undistorted shape, we transform each stripe shape which acquired at $t = i$ into that at $t = 0$ as using:

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

Here, $\boldsymbol{p}^s(i)$ denotes the point position which compose the stripe shape 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 $\boldsymbol{p}^s(0)$.

  To correct the distortion for each time t and each 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 $t = 0, 1, \cdots, T$.

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, we put a lot of markers $M_k^s(k=1, \cdots, N^s)$ and measure their 3D position $\boldsymbol{M}_k^s(t)$ with the above procedure (mentioned in 3.1) at all times $t = 0, \cdots, T$ for evaluating the accuracy. 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.

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 each human body segment with standing still. Non-rigidity is evaluated as a residual of the eq. (2) for estimating the rigid transformation $\boldsymbol{W}_{i,j}^s$ from a set of marker positions $\boldsymbol{M}_k^s(i)$ to a set of $\boldsymbol{M}_k^s(j)$ for all pairs of time (i, j). Let

(a) The laser scanning simulation with two cameras.

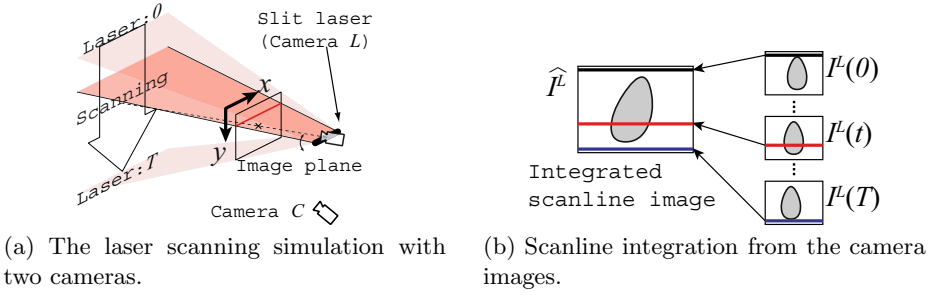(b) Scanline integration from the camera images.

Fig. 4. The simulation of the light stripe triangulation

$\widetilde{\boldsymbol{W}^s_{i,j}}$ denote the estimation of $\boldsymbol{W}^s_{i,j}$ from minimizing eq. (2). We evaluate the non-rigidity of segment s as an average residual $E^s(\widetilde{\boldsymbol{W}^s_{i,j}})$, E^s_W, which is defined as follows:

$$E^s_W = \frac{1}{T(T+1)} \sum_{\substack{i=0 \\ i \neq j}}^{T} \sum_{j=0}^{T} E^s(\widetilde{\boldsymbol{W}^s_{i,j}}). \tag{4}$$

Accuracy of the marker position measurement. Besides non-rigid deformation of segment s, error in measuring the marker position $\boldsymbol{M}^s_k(t)$ will also influence the average residual E^s_W. 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 $E^s(\widetilde{\boldsymbol{W}^s_{i,j}})$ for the rigid object, $\overline{E_W}$.

Accuracy of the light stripe triangulation. To evaluate the influence of the trunk sway on the light stripe triangulation, we simulate the measurement using the markers M^s_k and acquire the distorted shape.

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, 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, $\widehat{\boldsymbol{M}^s_k}$, on the acquired shape using the light stripe triangulation. When the marker M^s_k is observed on $\widehat{I^L}$ at coordinate $\widehat{\boldsymbol{m}^s_k} = (\hat{x}^s_k, \hat{y}^s_k)$, $\widehat{\boldsymbol{M}^s_k}$ is given by $\boldsymbol{M}^s_k(\hat{y}^s_k/T)$ since the marker would be illuminated by the laser sheet at time $\hat{y}^s_k/T$. For all

t, $\boldsymbol{M}_k^s(t)$ is obtained using the pair of sequences of images $I^C(t)$ and $I^L(t)$, so $\widehat{\boldsymbol{M}_k^s}$ can also be acquired.

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

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

Accuracy of the proposed method. To evaluate the accuracy of the proposed method, we correct the distorted shape $\widehat{\boldsymbol{M}_k^s}$. 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 $\widetilde{\boldsymbol{W}_{t,0}^s}$ from 3 markers positions $\boldsymbol{M}_k^s(t)$. Then, we transform $\widehat{\boldsymbol{M}_k^s}$ of the remaining markers into their corrected position $\widetilde{\boldsymbol{M}_k^s}(0)$ with $\widetilde{\boldsymbol{W}_{t,0}^s}$. Finally, we evaluate the accuracy, E_S^s, by comparing $\widetilde{\boldsymbol{M}_k^s}(0)$ and $\boldsymbol{M}_k^s(0)$ as follows:

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

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 matrix is estimated properly.

Next, we show the result of evaluating the non-rigidity of human body E_W^s, the accuracy of the 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.

To evaluate the accuracy, we measure the marker positions at $t = 0$, $\boldsymbol{M}_k^s(0)$, the marker positions acquired with simulated the light stripe triangulation, $\widehat{\boldsymbol{M}_k^s}$, and the positions as corrected by the proposed method, $\widetilde{\boldsymbol{M}_k^s}(0)$. Fig.7 illustrates
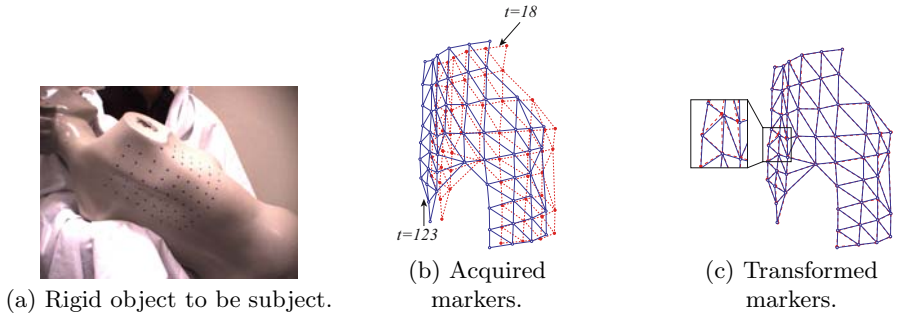
(a) Rigid object to be subject.

(b) Acquired markers.

(c) Transformed markers.

Fig. 5. Experiment to evaluate the accuracy of marker position measurement

(a) Acquired camera image at $t=0$.

(b) Integrated scanline image.
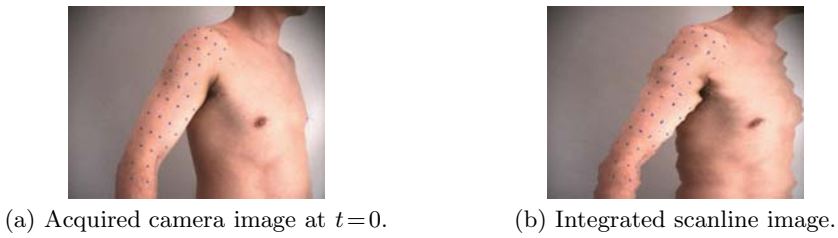
Fig. 6. Experimental result of integrated scanline image

● for rigid transform estimation ○ for accuracy evaluation

$M_k^s(0)$ $\widehat{M_k^s}$ $\widetilde{M_k^s}(0)$

Original shape on t=0   Scanned shape   corrected shape
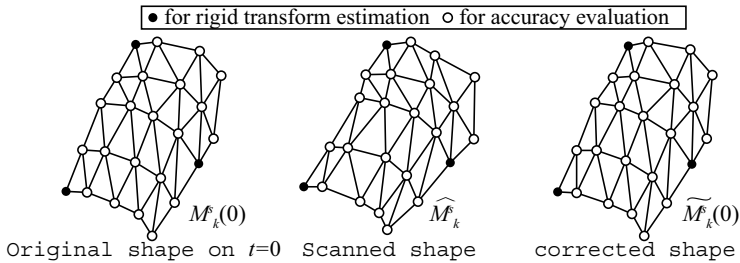
Fig. 7. Experiment for markers position measurement and correction

acquired marker positions $\boldsymbol{M}_k^s(0), \widehat{\boldsymbol{M}_k^s}, \widetilde{\boldsymbol{M}}_k^s(0)$ 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.

**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
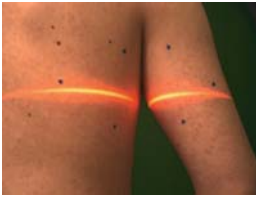
(a) Camera image for the marker
position measurement.

(b) Camera image for the light stripe
triangulation.

Fig. 8. Synchronous observations of the markers and the laser scans
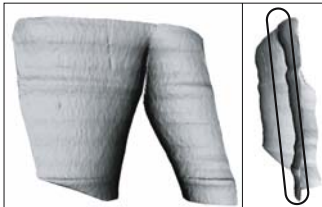
(a) A distorted shape.
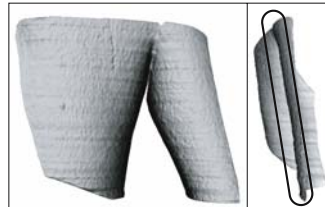
(b) Result of the proposed method.

Fig. 9. Comparison of contemporary and proposed method for shape reconstruction

4.3   Measurement Result

We show the acquired shape of a right upper arm and breast with two meth-
ods, the light stripe triangulation and the proposed method. We reconstruct the
surface from acquired point-cloud by making Delaunay mesh. Fig.8 illustrates
the camera images for the marker position measurement and the light stripe
triangulation. Fig.9 illustrates the shape acquired with the light stripe triangu-
lation 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/
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 ICIT'02, pp.405-410, 2002.
7. K. Hattori, Y. Sato. "Accurate Rangefinder with Laser Pattern Shifting," Proc.of ICPR, vol.3 no.3, pp.849-853, 1996.
8. 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.
9. Z. Zhang. "A Flexible New Technique for Camera Calibration," Microsoft Research, MSR-TR-98-71, 1998.
10. 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.
11. 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.