ORIGINAL RESEARCH

Validation of Stereophotogrammetry of the Human Torso

Juhun Lee[1], Manas Kawale[2], Fatima A. Merchant[3], June Weston[4], Michelle C. Fingeret[4], Dianne Ladewig[5], Gregory P. Reece[5], Melissa A. Crosby[5], Elisabeth K. Beahm[5] and Mia K. Markey[6,7]

[1]Department of Electrical and Computer Engineering, The University of Texas at Austin, Austin, Texas, USA. [2]Department of Computer Science, University of Houston, Houston, Texas, USA. [3]Department of Engineering Technology, University of Houston, Houston, Texas, USA. [4]Department of Behavioral Science, The University of Texas M. D. Anderson Cancer Center, Houston, Texas, USA. [5]Department of Plastic Surgery, The University of Texas M. D. Anderson Cancer Center, Houston, Texas, USA. [6]Department of Biomedical Engineering, The University of Texas at Austin, Austin, Texas, USA. [7]Department of Imaging Physics, The University of Texas M. D. Anderson Cancer Center, Houston, Texas, USA. Corresponding author email: juhunlee@mail.utexas.edu

Abstract: The objective of this study was to determine if measurements of breast morphology computed from three-dimensional (3D) stereophotogrammetry are equivalent to traditional anthropometric measurements obtained directly on a subject using a tape measure. 3D torso images of 23 women ranged in age from 36 to 63 who underwent or were scheduled for breast reconstruction surgery were obtained using a 3dMD torso system (3Q Technologies Inc., Atlanta, GA). Two different types (contoured and line-of-sight distances) of a total of nine distances were computed from 3D images of each participant. Each participant was photographed twice, first without fiducial points marked (referred to as unmarked image) and second with fiducial points marked prior to imaging (referred to as marked image). Stereophotogrammetry was compared to traditional direct anthropometry, in which measurements were taken with a tape measure on participants. Three statistical analyses were used to evaluate the agreement between stereophotogrammetry and direct anthropometry. Seven out of nine distances showed excellent agreement between stereophotogrammetry and direct anthropometry (both marked and unmarked images). In addition, stereophotogrammetry from the unmarked image was equivalent to that of the marked image (both line-of-sight and contoured distances). A lower level of agreement was observed for some measures because of difficulty in localizing more vaguely defined fiducial points, such as lowest visible point of breast mound, and inability of the imaging system in capturing areas obscured by the breast, such as the inframammary fold. Stereophotogrammetry from 3D images obtained from the 3dMD torso system is effective for quantifying breast morphology. Tools for surgical planning and evaluation based on stereophotogrammetry have the potential to improve breast surgery outcomes.

Keywords: three-dimensional, anthropometry, validation, breast, photogrammetry, stereophotogrammetry, surgical planning

Introduction

Breast cancer is the most common type of cancer affecting women in the US, with approximately 200,000 new cases of invasive breast cancer diagnosed each year.[1] Current breast cancer treatment encompasses not only surgical removal of the tumor and medical adjuvant and neo-adjuvant therapies to control the cancer, but an increasing emphasis on restoration of a woman's quality of life. The goal of breast reconstruction is to recreate a breast form that is satisfying to the patient, facilitating her psychosocial adjustment to living as a breast cancer survivor.

Until recently, reconstruction outcomes such as the shape of the breast were evaluated via direct anthropometry, or occasionally by photogrammetry.[2–10] However, these techniques have several limitations. For instance, direct anthropometry is time intensive and intrusive for patients. As an example, a tape measure must be placed on the subject's body directly, which the subject may consider embarrassing or uncomfortable. Moreover, random error due to soft tissue deformation and inconsistent posture can be introduced[11] and therefore, highly trained nurses are usually necessary for the direct anthropometry. Photogrammetry on standard clinical photography (two-dimensional photographs) is unable to capture the full three-dimensional (3D) structure and properties (e.g., volume) of the breast since photography is a 2D projection of a 3D object. In addition, two-dimensional photographs contain errors due to perspective distortion.[12] Lastly, neither direct anthropometry nor photogrammetry can measure breast volume accurately.

Stereophotogrammetry has attracted substantial interest in plastic surgery as an alternative method for evaluating surgical outcomes. Stereophotogrammetry creates a 3D model from x, y, and z positions of the object in space using the triangulation method via multiple cameras with known geometries. Modern 3D imaging systems for stereophotogrammetry are non-invasive and fast, and permit the evaluation of differences in breast volume, surface area, shape, size, and contour quantitatively.[10,13–16] In previous studies, asymmetry in breast volume was evaluated for various breast sizes[15,16] and different types of breast surgery[14] (e.g., breast reconstruction). Some limitations on acquisition of the 3D model from the imaging system were reported in obese subjects or those with very ptotic breasts. Surprisingly, however, previous studies paid little attention to distance measurements. Previous analysis of non-volume stereophotogrammetric measures has been limited to normal or pre-operative 3D images of participants and a single surface distance (sternal notch to nipple), which is insufficient for evaluating breast morphology.[15,16]

In this study, we validated stereophotogrammetry in terms of its ability to accurately assess distances between several fiducial points (landmarks) on pre- and post-operative 3D images of participants. Specifically, the hypotheses evaluated in this study are: 1) stereophotogrammetry is equivalent to direct anthropometry and 2) stereophotogrammetric measurements computed from fiducial points identified from viewing a 3D image alone are equivalent to those computed from fiducial points identified from direct viewing of the subject. In order to address these questions, we utilized two sets of images for each participant. In addition to the standard image (without any markings on the participant), we collected an additional image in which fiducial points were first marked on the participant prior to the 3D image being taken. Two types of distance measurements: contoured (surface) and line-of-sight (straight) distance, were computed from each set of images. The line-of-sight distance is equivalent to the straight-line distance that can be obtained from traditional photogrammetry. It was included in this study to provide a reference for comparison with the contoured distances, which can be obtained from stereophotogrammetry or direct anthropometry, but not photogrammetry. Moreover, the line-of-sight distance can be combined with its corresponding contoured distance to provide a curvature measure.[17]

Methods and Materials
Study population

Women who underwent or were scheduled for breast reconstruction surgery in the Center for Reconstruction at The University of Texas M. D. Anderson Cancer Center were recruited for this study. Informed consent was obtained from participants following institutional review board (IRB) approval. 3D torso images of 23 participants were considered in this study. These participants ranged in age from

36 to 63. There were 17 white, non-Hispanic/Latino; 3 white, Hispanic/Latino; 2 Black/African American, non-Hispanic/Latino; and 1 Asian, non-Hispanic/Latino among 23 study participants. Of the 23 participants, 14 had not yet undergone breast reconstruction surgery, 7 underwent tissue expander (TE) reconstruction (4 unilateral, 3 bilateral), 1 had bilateral latissimus dorsi flap reconstructed breasts, and 1 had bilateral breast reconstruction utilizing a TE for one breast and a transverse rectus abdominis myocutaneous (TRAM) flap for the other breast.

3D surface imaging system

The 3dMD torso system (DSP800) (3Q Technologies Inc., Atlanta, GA) was used to obtain surface images of each participant's torso. The 3dMD system projects a random light pattern on the subject and obtains 3D geometry using precisely synchronized digital cameras at varying angles, and texture information from a color photograph. Each reconstructed surface image consists of 3D positions (x, y, and z coordinates) and their corresponding color and texture. The surface images were analyzed using a custom program described in the following section.

3D torso analysis software

Our research team (M.K. and F.A.M. at the University of Houston) has developed software for visualization, manipulation, and analysis of 3D torso data. The analysis modules allow computation of both distance (line-of-sight and contoured) and volume measurements. The software can render, display, and rotate 3D mesh images. It also supports cropping and auto/manual alignment (i.e., normalization to a predefined spatial orientation) of 3D surface images. The user can annotate fiducial points such as the nipples and sternal notch, and the software records and displays the spatial coordinates of the marked points. The user can save and retrieve the fiducial points marked; he/she can also view and save a complete report of different kinds of measurements consisting of surface path measurements (line-of-sight, Dijkstra's, and contoured distances), volume measurements, symmetry measures, linear and curvilinear ptosis. Figure 1 displays the user interface of the software and its functionality to measure line-of-sight and contoured distances from the marked fiducial points.
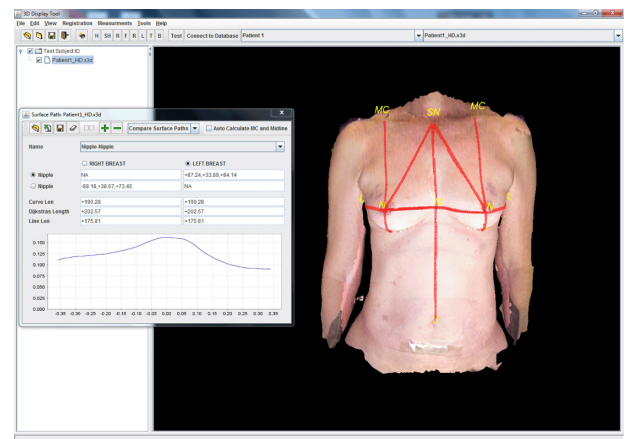

Figure 1. 3D torso image visualization software displaying a 3D surface image of one participant. Its functionality, which is line-of-sight and contoured distance computation from the marked fiducial points, is shown.

Study design and data collection

In this study, we assessed the accuracy of stereophotogrammetry computed from 3D images acquired using the 3dMD Torso system. Specifically, two hypotheses were evaluated: that stereophotogrammetry is equivalent to direct anthropometry, and that stereophotogrammetric measurements are equivalent, irrespective of whether fiducial points are marked on the participant before taking a 3D image or marked on the image itself.

In order to address the hypotheses, we used two sets of 3D images from each participant. With the participant standing and facing the 3D cameras, the first image was obtained without the fiducial points marked on the participant ("unmarked image") and the second image was obtained with the fiducial points marked on the participant ("marked image"). The stereophotogrammetric measurements computed from the marked image were compared to the corresponding direct anthropometry measurements to evaluate the first hypothesis. To assess the second hypothesis, the stereophotogrammetric measurements from the unmarked image and those from the marked image were compared.

To obtain the data necessary for this study, participants were placed in a standing position and 9 distances (listed in Table 1) between fiducial points (Fig. 2) were measured both directly on the participants using a tape measure and indirectly using a computer program, from both marked and unmarked images. For each participant, three sets of measurements were collected: direct anthropometry

Table 1. List of physical measurements.

Abbreviation	Description
SN-N	Distance from sternal notch (SN) to nipple (N).
MC-N	Distance between mid-clavicle (MC) to nipple.
MC-TP	Distance between mid-clavicle and transition point(TP), i.e., where breast mound first starts to leave the chest wall.
M-N	Horizontal distance from midline (M) to nipple.
L-N	Horizontal distance from lateral breast-chest wall junction (L) to nipple.
IMF-N	Vertical distance from inframammary fold (IMF) to nipple.
IMF-LV	Vertical distance from inframammary fold to lowest point (LV) on the breast.
N-N	Distance between the nipples.
SN-U	Distance between the sternal notch and the umbilicus (U).

using a tape measure, stereophotogrammetry on the unmarked 3D image, and stereophotogrammetry on the marked 3D image.

Direct anthropometry using tape measure

With the participant placed in a standing position, a research nurse, experienced with determining the location of the fiducial points, measured 9 distances between fiducial points directly on the participants using a tape measure. For each distance measurement listed in Table 1, the contoured distance and the

Figure 2. List of physical measurements evaluated in this study. **A, B)** The vertical distance from inframammary fold to the lowest visible point (IMF-LV) is not shown in the image. **C)** 3D image of participant #10 is shown. She has severe ptosis of the left breast and she does not have a right breast, and her umbilicus is not visible in the image. Therefore, only four distances (MC-N, MC-TP, SN-N, and L-N) were measured.

line-of-sight distance were obtained and recorded. The contoured distance was recorded by holding the tape measure to follow the contours along the body surface, whereas the line-of-sight distance was recorded by holding the tape measure taut between two fiducial points.

Stereophotogrammetry on unmarked 3D image

The breasts and torso of each participant was photographed in a standing position and facing forward using a 3dMD Torso system, prior to any marks being made on the participant. A single observer (J.L.) annotated the locations of the fiducial points on the 3D images. The contoured (geodesic) distance was recorded as the shortest path along the surface of participant's torso, whereas the line-of-sight (Euclidean) distance was computed to be the straight line distance between the fiducial points. The contoured distance and the line-of-sight distance between the annotated fiducial points on the image were computed and recorded by the previously mentioned custom program from the University of Houston.

Stereophotogrammetry on marked 3D image

With the participant placed in a standing position, each fiducial point was carefully examined by the research nurse and marked on participants with a colored marker before participants were photographed. The first author (J.L.) re-annotated the fiducial points that are shown in the 3D image (Fig. 2A–B) based on points which were initially marked by the nurse. The contoured distance and the line-of-sight distance were computed and recorded as described above.

Because the participants had a variable number of visible fiducial points depending on the stage of
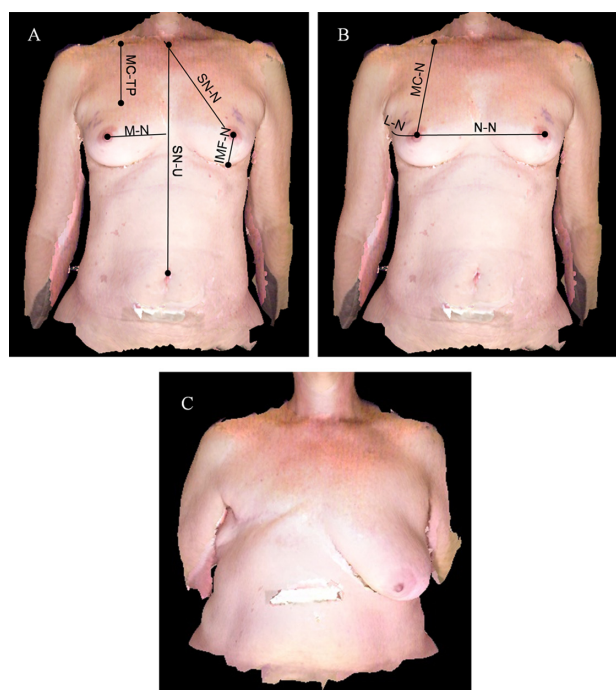
their surgical treatments, all measurements in each participant could not be obtained. For the vertical distance from the inframammary fold to nipple (IMF-N) and the vertical distance from the inframammary fold to the lowest visible point of breast (IMF-LV), images were excluded if the inframammary fold was not visible due to ptosis or pseudo-ptosis of the breast (Fig. 2C).

In addition, one professional model was recruited for assessing the intra-observer variability of direct anthropometry. At three different time points (at least one week apart), one research nurse measured 9 distances (Table 1, excluding the nipple to nipple distance) from the model using a tape measure with the model in a standing position. The study points were chosen to provide a time gap, such that it would be unlikely for the research nurse to recall any of the previous measurements obtained from the model.

Statistical analyses

Three different statistical analyses were performed: Bland-Altman analysis, Hypothesis test for equivalence, and Intraclass correlation coefficient. All analyses were performed using MATLAB® (The Mathworks, Natick, MA) statistics toolbox.

A Bland-Altman plot[18,19] provides a visual and qualitative assessment of the agreement between two assays. The assays have excellent agreement if all of the pairwise differences between them are within 1.96 standard deviations from the mean difference. A Bland-Altman plot was used to compare: 1) the stereophotogrammetric measurements computed from the marked images to the corresponding direct anthropometric measurements using a tape measure, 2) the stereophotogrammetric measurements computed from the unmarked images to those from the marked images.

A hypothesis test for equivalence[20] is another way to evaluate the agreement between two measurements. To test equivalence for the first hypothesis of this study, the null hypothesis is that the stereophotogrammetric measurements from 3D imaging are not equivalent to those from direct anthropometry, and the alternative hypothesis is that they are equivalent. Similarly, the null hypothesis for the second hypothesis of this study is that the stereophotogrammetric measurements from the unmarked images are not equivalent to those from the marked images. This test is different from a typical paired t-test in which the null hypothesis is

that the means of the two populations are equal and the alternative is that the population means are not equal. The test statistic for assessing equivalence is $t = \sqrt{n}(x' \pm \delta)/s$, where x' and s are sample mean and standard deviation of the differences in the measurements from stereophotogrammetry and direct anthropometry, and n is the sample size. The δ corresponds to an acceptable level of variability in the measurements made by the two methods. A smaller value of δ is a stricter criterion for establishing equivalence of the two measurements. The value of δ is calculated by multiplying a user-specified factor by the mean of the direct anthropometric measurements. In this paper, we used 0.1 as the factor, which means that we consider measurements to be equivalent if they are within 10% of the average value obtained from direct anthropometry.

The intraclass correlation coefficient (ICC)[21] is a measure of the reproducibility of replicate measures from the same subject. In this study, a two way random model was used since a set of 3D images is randomly sampled from the population of possible images and the set of observers, who measure the distances, is also a random sample of the observer population. The interpretation of intraclass correlation coefficient used for this study is as follows; $ICC < 0.4$ indicates poor reproducibility, $0.4 \leq ICC < 0.75$ indicates fair to good reproducibility, and $ICC \geq 0.75$ indicates excellent reproducibility.[22]

Results
Intra-observer variability of direct anthropometry

The anthropometric measurements made by the research nurse from the professional model at three different time points are shown in Table 2. If we set the measurements which have high intra-observer variability based on standard deviation values higher than 10% of the mean, 4 direct anthropometric measurements demonstrated higher intra-observer variability than others. These measurements are: the distance between mid-clavicle and the point where the breast mound leaves the chest wall (MC-TP), the horizontal distance from the lateral breast-chest wall junction to the nipple (L-N), the vertical distance from the inframammary fold to the nipple (IMF-N), and the vertical distance from the inframammary fold to lowest point on the breast (IMF-LV).

Table 2. Repeated 9 distance measurements at three different time points and their mean and standard deviation values. Intra-observer variability of direct anthropometry was evaluated by comparing the standard deviation and the mean values. High intra-observer variations were observed for the MC-TP, L-N, IMF-N, and IMF-LV distance measurements (both line-of-sight and contoured distance). For those measurements, their standard deviation values were higher than 10% of their mean values.

	Line-of-sight					Contoured				
	First (cm)	Second (cm)	Third (cm)	Mean (cm)	Std (cm)	First (cm)	Second (cm)	Third (cm)	Mean (cm)	Std (cm)
SN-N (right)	23.5	23.6	24.4	23.8	0.5	24.3	24.2	24.8	24.4	0.3
MC-N (right)	23.5	23.4	24.5	23.8	0.6	24.0	24.0	25.5	24.5	0.9
MC-TP (right)	11.0	10.7	11.3	11.0	0.3	11.2	10.8	11.4	11.1	0.3
MP-N (right)	9.0	8.6	9.6	9.1	0.5	10.7	12.0	11.5	11.4	0.7
L-N (right)	10.5	13.3	10.5	11.4	1.6*	11.7	14.4	10.6	12.2	2.0*
IMF-N (right)	5.5	5.8	6.0	5.8	0.3	7.0	5.9	8.2	7.0	1.2*
IMF-LV (right)	3.8	3.4	3.0	3.4	0.4*	3.8	3.5	3.2	3.5	0.3
SN-N (left)	24.0	24.2	24.3	24.2	0.2	25.0	25.5	25.1	25.2	0.3
MC-N (left)	26.0	25.7	25.2	25.6	0.4	26.4	26.0	25.8	26.1	0.3
MC-TP (left)	12.7	12.5	10.4	11.9	1.3*	12.8	12.8	11.0	12.2	1.0
M-N (left)	8.5	8.5	8.5	8.5	0	11.0	12.3	12.8	12.0	0.9
L-N (left)	12.0	13.6	14.6	13.4	1.3	14.2	15.7	17.1	15.7	1.5
IMF-N (left)	6.2	5.8	6.6	6.2	0.4	7.8	5.6	7.4	6.9	1.2*
IMF-LV (left)	4.5	4.0	3.4	4.0	0.6*	4.6	4.0	3.8	4.1	0.4
SN-U	32.6	32.0	32.4	32.3	0.3	33.5	33.0	32.5	33.0	0.5
NN	18.2	18.5	N/A	18.4	0.2	23.7	23.3	N/A	23.5	0.3

Note: *Measurements that their standard deviation value is larger than 10% of their mean value.

Accuracy of stereophotogrammetry

For most distance measurements (both contoured and line-of-sight distances), the difference between the mean value of stereophotogrammetric measurements from the marked images and those of direct anthropometric measurements was typically less than 1 cm (Table 3). Although the mean of the differences of a few measures (the sternal notch to nipple line-of-sight distance, nipple to nipple contoured distance) are more than 1 cm, these differences are still less than 10% of their mean values. This can be accounted for by the inherent variability associated with direct anthropometry in attempting to hold the tape measure taut between the sternal notch and the nipple which are not in the same plane, and with the nipple to nipple distance where the soft tissue of the breast mounds can have variable depression when the tape measure is pressed along the body surface to measure the contoured distance.

In general, Bland-Altman plots for most of the stereophotogrammetric measurements from marked 3D images (7 out of 9 measurements, both contoured and line-of-sight distances) showed excellent agreement with those obtained from direct anthropometry on participants. The differences between the two types of measurements (stereophotogrammetry and direct anthropometry) were frequently (except 1 or 2 measurements) within the 1.96 times standard deviation bounds (Fig. 3A–B). Due to space constraints we do not include Bland-Altman plots for all of the measures evaluated in this study, but as an example Figure 3 depicts the Bland-Altman plots for the sternal notch to nipple distances (SN-N) (Fig. 3A) and the vertical distances from the inframammary fold to nipple (IMF-N) (Fig. 3B). The intraclass correlation coefficient values for most measures (7 out of 9 measurements, both contoured and line-of-sight distances) were above 0.75, which indicates excellent agreement between stereophotogrammetry and direct anthropometry (Table 4). The hypothesis tests for equivalence showed that stereophotogrammetry on marked 3D images and direct anthropometry were in agreement for 7 out of 9 measurements (Table 4).

The vertical distance from the inframammary fold to nipple (IMF-N) (line-of-sight distance only) and the vertical distance from the inframammary

Table 3. Descriptive statistics for the stereophotogrammetric measurements (from marked and unmarked images) and the direct anthropometric measurements. For most distance cases, the difference between the mean value of stereophotogrammetric measurements from the marked images and those of direct anthropometric measurements was less than 1 cm. Moreover, the difference between the mean value of the measurements from the unmarked images and those from the marked images was less than 1 cm. Although the mean value differences of few cases (SN-U line-of-sight distance, N-N contoured distance, L-N both line-of-sight and contoured distance) are more than 1 cm, those differences are less than 10% of their mean values.

Distance (number of samples)	Contoured distance						Line-of-sight distance					
	3D photo				Tape measure		3D photo				Tape measure	
	Marked		Unmarked				Marked		Unmarked			
	Mean (cm)	Std (cm)	Mean (cm)	Std (cm)	Mean (cm)	Std (cm)	Mean (cm)	Std (cm)	Mean (cm)	Std (cm)	Mean (cm)	Std (cm)
SN-N (24)	22.75	4.16	22.85	4.22	23.4	4.1	22.48	4.01	22.57	4.07	23.0	4.1
MC-N (25)	22.49	4.14	22.9	4.15	23.3	4.1	22.20	4.01	22.60	4.07	22.9	4.2
MC-TP (33)	10.89	1.95	11.64	1.97	11.3	2.0	10.82	1.94	11.56	1.96	11.1	1.9
SN-U (17)	38.60	3.09	38.92	3.81	39.4	3.4	37.39[†]	2.79	37.34	3.07	38.4	3.2
L-N (27)	14.62	3.55	13.59*	3.07	15.1	4.1	14.16	3.56	12.93*	3.02	13.7	3.6
N-N (9)	21.82[†]	2.25	21.82	2.28	23.2	3.0	19.57	2.33	19.61	2.32	19.7	2.2
M-N (21)	11.41	1.52	10.86	1.38	11.1	1.5	10.16	1.21	10.17	1.15	9.7	1.2
IMF-N (10)	6.98	1.21	6.86	1.14	7.3	1.1	6.29	0.94	6.22	0.94	5.3	0.5
IMF-LV (15)	1.02	0.78	0.92	0.46	0.5	0.6	0.81	0.58	0.81	0.39	0.5	0.5

Notes: *Measurements from unmarked images that their mean value difference from those from marked images is more than 1 cm; [†]Measurements from marked images that their mean value difference from the direct anthropometry is more than 1 cm.

fold to the lowest visible point (IMF-LV) were problematic. For those measurements, the intraclass correlation coefficient values were below 0.45, which denotes poor agreement. Moreover, we failed to reject the null hypothesis that the measurements were not equivalent (P-values were above 0.05). Even though we excluded 3D images of women with ptosis or pseudo-ptosis, the 3D imaging system failed to capture the inframammary fold area and resulted in holes in the 3D images. These holes made the inframammary fold and the lowest visible point marked by the research nurse invisible for some participants as shown in Figure 4. Moreover, the direct anthropometry of the vertical distance from the inframammary fold to nipple (IMF-N) and the vertical distance from the inframammary fold to the lowest visible point (IMF-LV) also showed high intra-observer variability (refer to previous section and Table 2). This suggests that there is inherent practical difficulty in making these two direct anthropometric measurements and it can further explain poor agreement between the direct anthropometry and the stereophotogrammetry for these measurements.

Equivalence between stereophotogrammetry on unmarked and marked 3D images

For most measurements, the difference between the mean value of stereophotogrammetric measurements from the unmarked images and those from marked images was typically less than 1 cm (Table 3). Although the mean value difference of one distance (the lateral breast-chest wall junction to nipple, both contoured and line-of-sight distances) is more than 1 cm, the difference is still less than 10% of its mean value. In Bland-Altman plots of all measurements, the differences between measurements from the unmarked images and those from the marked images were frequently within the 1.96 times standard deviation bounds (Fig. 3C–D). The mean difference between the two types of measurements was less than 1 cm for all except the horizontal distance from lateral breast-chest wall junction to nipple (L-N). Both the intraclass correlation coefficient values and the hypothesis tests for equivalence showed that there was excellent agreement between stereophotogrammetry from the unmarked images and that from the marked images for most measures (7 out of 9 measurements) (Table 4).
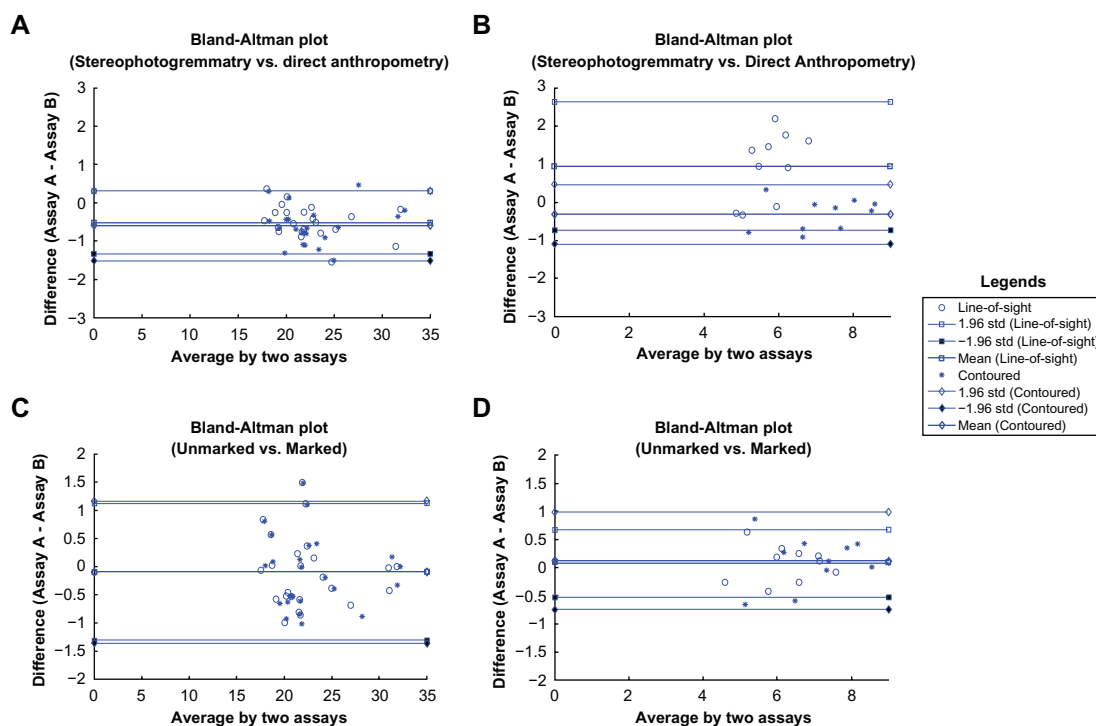
Figure 3. Bland-Altman plots for sternal notch to nipple distance (SN-N) and the vertical distance from inframammary fold to nipple (IMF-N). **A**) Comparison between stereophotogrammetry and direct anthropometry for SN-N case. **B**) Comparison between stereophotogrammetry and direct anthropometry for IMF-N case. **C**) Comparison between stereophotogrammetry from the unmarked image and that from the marked image for SN-N case. **D**) Comparison between stereophotogrammetry from the unmarked image and that from the marked image for IMF-N case. The differences between two measurements (in figure A and B, both line-of-sight and contoured distances) are within the bounds, which means the excellent agreement between the stereophotogrammetry and the direct anthropometry. However, higher mean value offset and looser bounds are shown from the line-of-sight distance case than those from the contoured distance case (figure B). The differences between two measurements (in figure C and D, both line-of-sight and contoured distances) are within the bounds, which means that the measurements from unmark images are equivalent to those from marked images.

Statistical analyses for the horizontal distance from lateral breast-chest wall junction to nipple (L-N) (Hypothesis test for equivalence only) and the vertical distance from the inframammary fold to the lowest visible point (IMF-LV) showed poor to fair agreement between stereophotogrammetry from the unmarked images and that from the marked images. Since the vertical distance from the inframammary fold to nipple (IMF-N) showed excellent agreement between the measurements from both marked and unmarked images, these results suggest that the lowest visible point and the lateral breast-chest wall junction point (L) may have caused the poor to fair agreement. This is not surprising since those two fiducial points are more vaguely defined than other points such as the nipples and the umbilicus.

Discussion

In this study, we validated stereophotogrammetry on the human torso in terms of the agreement between stereophotogrammetric and anthropometric measurements. In addition to the contoured distances, we validated the line-of-sight distances computed from the 3D torso image since these correspond to traditional photogrammetric measures on clinical photography (2D). Two hypotheses were evaluated in this study. To assess the accuracy of stereophotogrammetry, the stereophotogrammetric measurements from marked 3D images were compared to the direct anthropometric measurements. We also evaluated the equivalence between the stereophotogrammetric measurements from unmarked and marked 3D images.

In general, there was excellent agreement (ICC > 0.75 and the *P*-value of equivalence test < 0.0001) between the stereophotogrammetric measurements (both contoured and line-of-sight distances) made on 3D images and those obtained by direct anthropometry on participants. This is consistent with what has been reported in prior studies of 3D stereophotography in plastic surgery.[16,23–25] However, for some measurements such as the vertical distance from the inframammary fold to nipple (IMF-N)

Table 4. Intraclass correlation coefficient (ICC) and Hypothesis test for equivalence. Seven out of 9 distance measurements showed excellent agreement between the stereophotogrammetry and the direct anthropometry. Both ICC values and the *P*-values of Hypothesis test of equivalence for most cases except for two measurements (for IMF-N; line-of-sight distance only, for IMF-LV; both line-of-sight and contoured distance) showed poor agreement between two measurements. Seven out of 9 distance measurements made on the unmarked images showed the equivalence to those made on the marked images. ICC values of one measurement (IMF-LV) and the *P*-values of the Hypothesis test of equivalence of two measurements (L-N and IMF-LV) showed poor to fair agreement between two measurements.

Distance	Intraclass correlation coefficient (ICC)				Hypothesis test for equivalence (*P*-value)			
	Stereophotogrammetry vs. Direct anthropometry		Unmarked vs. marked		Stereophotogrammetry vs. direct anthropometry		Unmarked vs. marked	
	Line-of-sight	Contoured	Line-of-sight	Contoured	Line-of-sight	Contoured	Line-of-sight	Contoured
SN-N	0.99	0.97	0.99	0.99	<0.0001	<0.0001	<0.0001	<0.0001
MC-N	0.98	0.98	0.97	0.97	<0.0001	<0.0001	<0.0001	<0.0001
MC-TP	0.97	0.96	0.82	0.82	<0.0001	<0.0001	0.0093	0.0096
SN-U	0.85	0.89	0.95	0.91	<0.0001	<0.0001	<0.0001	<0.0001
L-N	0.98	0.99	0.81	0.84	<0.0001	<0.0001	0.4262†	0.1554†
N-N	0.9	0.79	0.97	0.98	<0.001	0.0268	<0.0001	<0.0001
M-N	0.83	0.92	0.94	0.77	<0.001	<0.0001	<0.0001	<0.0001
IMF-N	0.15*	0.9	0.94	0.92	0.9401†	0.0067	0.0002	0.0022
IMF-LV	0.24*	0.23*	0.65‡	0.57‡	0.9438†	0.9812†	0.6504†	0.5763†

Notes: *ICC values showing poor agreement between two measurements; †Statistically insignificant results for Hypothesis test for equivalence; ‡ICC values showing good agreement between two measurements.

and the vertical distance from the inframammary fold to the lowest visible point (IMF-LV), there was poor agreement between the assessment made directly on the participant and that made using the 3D image. Our results suggest that this may be due to the failure of the 3D imaging system to capture the inframammary fold area (Fig. 4) and the practical difficulty in making the direct anthropometry for these distances. Moreover, since the inframammary fold and the lowest visible point of breast are very close to each other for non-ptotic breasts, the IMF-LV distances are usually close to zero in magnitude. Therefore, it is hard to validate IMF-LV distances.
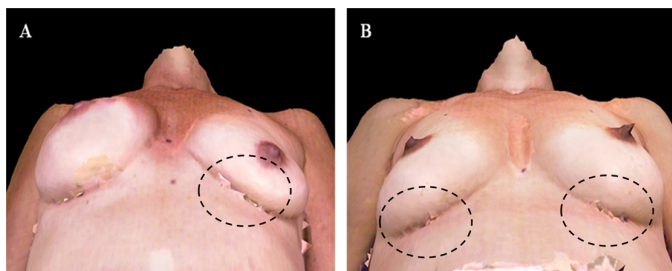

Figure 4. 3D torso images showing holes around the inframammary fold area. The dashed ellipses indicate holes from the 3D imaging system due to the failure of capturing the manually marked fiducial points (the inframammary fold and the lowest visible point).

For most measures, the stereophotogrammetric measurements made on unmarked 3D images were equivalent to those made on marked 3D images. Thus, in general, fiducial points can be effectively identified by visual inspection of 3D images. However, there were a few distances (e.g., the lateral breast-chest wall junction to nipple distance and the vertical distance from the inframammary fold to the lowest visible point) for which the agreement was only poor to fair due to the difficulty of localizing vaguely defined fiducial points such as the lowest visible point and the lateral breast-chest wall junction point. This suggests that distance measures based on well-defined fiducial points will be most reliable in future stereophotogrammetric studies.

Our analysis builds upon a prior study by Losken et al[16] in which the authors evaluated the accuracy and the reproducibility of stereophotogrammetry using the 3dMD Torso system from 3D images of 14 subjects prior to the breast reconstruction surgery and up to 20 untreated breasts. Two sets of measurements of breast volumes and one contoured distance (the sternal notch to nipple) were obtained from 3D images by two independent observers and those measurements were compared to the measurements obtained from participants directly using water replacement

(volume) and a tape measure (distance). Although this group was one of the first to assess the accuracy of stereophotogrammetry on the human torso, particularly for distance measurements, their study had limitations. They considered a single surface distance measure (the sternal notch to nipple) only, which is insufficient for evaluating breast morphology. Moreover, their analysis was limited to subjects prior to the breast reconstruction surgery.

Stereophotogrammetry has multiple advantages over traditional methods for assessing the shape of breast. Since there are less bony structures in the breast area than other areas of the body (e.g., face), there can be measurement error due to the soft tissue deformation during direct anthropometry. In contrast, stereophotogrammetry avoids these measurement errors since the measurements are made on the images, not on subjects directly. Moreover, the 3D torso image can be rotated 360 degrees, which allows evaluation of the breast morphology from any angle. This helps retrospective assessment of the 3D geometry of the breast as well as the 3D morphological changes due to breast reconstruction surgery. In addition, stereophotogrammetry can provide a measurement of breast volume using a fast mathematical model, while traditional methods are either unable to compute volume or cannot do so accurately.[10,13–16] Moreover, unlike traditional methods, stereophotogrammetry makes it easy to introduce new measurements for quantifying breast morphology. For instance, direct anthropometry is strictly limited since any new measurement must be made on the subjects's torso directly. In case of photogrammetry, a new measurement may also require new photographs of the subject's torso.

In conclusion, this study suggests that 3D torso images from stereophotography are sufficiently accurate for assessing breast morphology. Stereophotogrammetry has the potential to replace direct anthropometry and photogrammetry for the evaluation of breast surgery outcomes. Plastic surgeons may find stereophotogrammetry to be useful in their preoperative surgical planning and the postoperative evaluation of the breast aesthetic and/or reconstruction surgery outcomes.

Acknowledgements

Disclosures

This manuscript has been read and approved by all authors. This paper is unique and not under consideration by any other publication and has not been published elsewhere. The authors and peer reviewers report no conflicts of interest. The authors confirm that they have permission to reproduce any copyrighted material.

References

1. *Cancer Facts and Figures* 2009: American Cancer Society; 2009.
2. Westreich M. Anthropomorphic breast measurement: protocol and results in 50 women with aesthetically perfect breasts and clinical application. *Plastic and Reconstructive Surgery*. 1997;100(2):468–79.
3. Brown TRC, Hyland RE, Cole AA, Brotherston TM. A method of assessing female breast morphometry and its clinical application. *British Journal of Plastic Surgery*. 1999;52(5):355–9.
4. Pezner RD, Patterson MP, Hill LR, et al. Breast retraction assessment: an objective evaluation of cosmetic results of patients treated conservatively for breast cancer. *Int J Rad Oncol Bio Phy*. 1985;11(3):575–8.
5. Van Limbergen E, van der Schueren E, van Tongelen K. Cosmetic evaluation of breast conserving treatment for mammary cancer. 1. Proposal of a quantitative scoring system. *Radiother Oncol*. 1989;16(3):159–67.
6. Kim MS, Burgess A, Reece GP, et al. Understanding Surgeons' Assessments of the Aesthetic Outcome of Breast Cancer Treatment using Eye-Tracking. *Medical Image Perception Conference XII*. Iowa City, Iowa2007.
7. Kim MS, Reece GP, Beahm EK, Miller MJ, Atkinson EN, Markey MK. Objective assessment of aesthetic outcomes of breast cancer treatment: measuring ptosis from clinical photographs. *Computers in Biology and Medicine*. 2007;37(1):49–59.
8. Kim MS, Rodney WN, Cooper T, Kite C, Reece GP, Markey MK. Towards quantifying the aesthetic outcomes of breast cancer treatment: comparison of clinical photography and colorimetry. *Journal of Evaluation in Clinical Practice*. 2009 Feb;15(1):20–31.
9. Kim MS, Rodney WN, Reece GP, Beahm EK, Crosby MA, Markey MK. Quantifying the aesthetic outcomes of breast cancer treatment: assessment of surgical scars from clinical photographs. *Journal of Evaluation in Clinical Practice*. In press.
10. Kim MS, Sbalchiero JC, Reece GP, Miller MJ, Beahm EK, Markey MK. Assessment of Breast Aesthetics. *Plastic and Reconstructive Surgery*. 2008;121(4):186e–94.
11. Kouchi M, Mochimaru M, Tsuzuki K, Yokoi T. Random errors in anthropometry. *Journal of Human Ergology*. 1996;25(2):155–66.
12. Kovacs LYA, Zimmermann A, Brockmann G, et al. Optimization of 3-dimensional imaging of the breast region with 3-dimensional laser scanners. *Annals of Plastic Surgery*. 2006;56(3):229–36.
13. Nahabedian MY, Galdino G. Symmetrical breast reconstruction: is there a role for three-dimensional digital photography? *Plastic and Reconstructive Surgery*. 2003;112(6):1582–90.
14. Galdino GM, Nahabedian M, Chiaramonte M, Geng JZ, Klatsky S, Manson P. Clinical applications of three-dimensional photography in breast surgery. *Plastic and Reconstructive Surgery*. 2002;110(1):58–70.
15. Losken A, Fishman I, Denson D, Moyer H, Carlson G. An Objective Evaluation of Breast Symmetry and Shape Differences Using 3-Dimensional Images. *Ann Plast Surg*. 2005;55(6):571–5.

16. Losken A, Seify H, Denson DD, Paredes AAJ, Carlson GW. Validating three-dimensional imaging of the breast. *Ann Plast Surg.* 2005;54(5): 471–6.

17. Gupta S, Markey MK, Aggarwal JK, Bovik A. Three dimensional face recognition based on geodesic and Euclidean distances Paper presented at: IS&T/SPIE Symposium on Electronic Imaging: Vision Geometry XV2007.

18. Bland JM, Altman DG. Statistical methods for assessing agreement between two methods of clinical measurement. *Lancet.* 1986 Feb 8;1(8476): 307–10.

19. Bland JM, Altman DG. Applying the right statistics: analyses of measurement studies. *Ultrasound in obstetrics and gynecology.* 2003 Jul;22(1):85–93.

20. Wellek S. *Testing statistical hypothesis of equivalence.* Boca Raton, FL: CRC Press LLC; 2003.

21. McGraw KO, Wong SP. Forming inferences about some intraclass correlation coefficients. *Psychological Methods.* 1996;1(1):30–46.

22. Rosner B. *Fundamentals of Biostatistics.* 6th ed. Belmont, CA: Duxbury Press; 2005.

23. Plooij JM, Swennen GRJ, Rangel FA, et al. Evaluation of reproducibility and reliability of 3D soft tissue analysis using 3D stereophotogrammetry. *International Journal of Oral and Maxillofacial Surgery.* 2009;38(3): 267–73.

24. Wong JY, Oh AK, Ohta E, et al. Validity and reliability of craniofacial anthropometric measurement of 3D digital photogrammetric images. *The Cleft Palate-Craniofacial Journal: Official Publication of the American Cleft Palate-Craniofacial Association.* 2008;45(3):232–9.

25. Weinberg SM, Naidoo S, Govier DP, Martin RA, Kane AA, Marazita ML. Anthropometric precision and accuracy of digital three-dimensional photogrammetry: comparing the Genex and 3dMD imaging systems with one another and with direct anthropometry. *The Journal of Craniofacial Surgery.* 2006;17(3):477–83.