



Published in final edited form as:

*J Eval Clin Pract.* 2011 December ; 17(6): 1129–1136. doi:10.1111/j.1365-2753.2010.01477.x.

## Automated calculation of symmetry measure on clinical photographs

Mugdha Dabeer<sup>1</sup>, Edward Kim<sup>1</sup>, Gregory P. Reece<sup>2</sup>, Fatima Merchant<sup>3</sup>, Melissa A. Crosby<sup>2</sup>, Elisabeth K. Beahm<sup>2</sup>, and Mia K. Markey<sup>1</sup>

<sup>1</sup> The University of Texas Department of Biomedical Engineering

<sup>2</sup> Department of Plastic Surgery, The University of Texas M. D. Anderson Cancer Center

<sup>3</sup> The University of Houston, Department of Engineering Technology, USA

### 1 ABSTRACT

Breast cancer is one of the most prevalent forms of cancer in the world. More than 250,000 American women are diagnosed with breast cancer annually. Fortunately, the survival rate is relatively high and continually increasing due to improved detection techniques and treatment methods. The quality of life of breast cancer survivors is ameliorated by minimizing adverse effects on their physical appearance. Breast reconstruction is important for restoring the survivor's appearance. In breast reconstructive surgery, there is a need to develop technologies for quantifying surgical outcomes and understanding women's perceptions of changes in their appearance. Methods for objectively measuring breast anatomy are needed in order to help breast cancer survivors, radiation oncologists, and surgeons quantify changes in appearance that occur with different breast reconstructive surgical options. In this study, we present an automated method for computing a variant of the normalized Breast Retraction Assessment (pBRA), a common measure of symmetry, from routine clinical photographs taken to document breast cancer treatment procedures.

### Keywords

BRA; pBRA; Automated Detection; Digital Photographs; Umbilicus; Nipple Complex; Breast Cancer

### 2 INTRODUCTION

Breast cancer is one of the most prevalent forms of cancer in the world. Each year, there are approximately 250,000 cases of breast cancer diagnosed among women in the USA [1]. Fortunately, the survival rate is relatively high and continually increasing due to improved detection techniques and treatment methods. However, maintaining quality of life is a factor often poorly addressed for breast cancer survivors. Breast cancer treatments are often highly invasive and can adversely affect the breast tissue and lead to deformation of the breast. Currently, assessment of breast appearance is not objectively reported; instead subjective assessments of breast appearance by doctors, patients, and other observers are used [2]. Subjective assessments often have high inter- and intra-observer variability. As a result, there is a need for objective, quantifiable methods of assessing breast appearance that are more reliable than subjective observations in order to understand the impact of deformity on

patient quality of life, to guide selection of current treatments, and to make rational treatment advances [2,3].

Breast morphology can be objectively quantified using distances between fiducial points at common anatomical landmarks or ratios of such distances. Such approaches have been successfully implemented in the field of breast conservation therapy (e.g. [4,5]). We have demonstrated that similar techniques can be designed for use in evaluating breast reconstruction [6].

Traditionally these measurements have been made directly on the body using a tape measure. However, there are drawbacks with traditional anthropometry. For example, the process is repetitive, tedious and time consuming and may be embarrassing for the patient.

The advent of digital photography has provided a practical alternative known as photogrammetry. Photogrammetry is faster than traditional measuring techniques, since the patient need not wait until all measurements are made. Photographs can be stored and re-analyzed at a later date. Images taken at various time points can be used to track changes in the appearance of the patient over time.

In the field of radiation oncology, a widely accepted method for quantifying the symmetry of the nipples positions' is the Breast Retraction Assessment measure, or BRA, developed by Pezner et al [7]. A normalized version of the BRA measure, called pBRA, was developed by Limbergen et al. [8] in order to account for different magnifications of the patient images. The pBRA value is equal to the BRA value divided by the distance from a reference point, the sternal notch to the nipple of the untreated breast. This distance from the sternal notch to the nipple of the breast that did not undergo breast conservation therapy is always greater than the distance from the sternal notch to the treated breast [4,5]. Van Limbergen et al. found smaller intra- and inter-observer variability based on pBRA than on subjective assessments [8].

However, a significant drawback of current approaches to computing measures of breast appearance such as pBRA is that they are dependent upon manual identification of the fiducial points on a digital photograph by a human observer. There is inherent variability in localizations made by human observers, which can thus introduce variability into measures computed from the annotated fiducial points.

In a previous study in our laboratory, Kim et al. [6] studied the variability in manual localization of some fiducial points as part of an investigation of a new measure of ptosis that can be computed from digital photographs. Kim et al. demonstrated that the variability introduced by manual annotation is modest for some fiducial points, such as the nipples, and thus the impact on the measure computed from those annotations is minimal. More variation was observed with subtle points, such as the lateral terminus of the inframammary fold.

Nevertheless, manual localization of fiducial points remains a key-limiting factor in the widespread adoption of computerized analysis of breast surgery outcomes from digital photography. Manual identification of fiducial points is time consuming and disruptive to the typical clinical workflow.

Traditionally, the sternal notch has been used in the normalization of pBRA. However, the sternal notch is not always visible in digital photographs, particularly in photographs of women with a higher body-mass index (BMI). Hence, in this study, we propose the umbilicus as a reference point for normalizing our symmetry measure.

The goal of this study was to develop and implement a fully automated method of localizing the umbilicus and the nipples on anterior-posterior digitized/digital photographs of women. The locations were used to automatically calculate a variant of pBRA, which we term pBRA2, for each subject. The automatic localizations were compared to manually-obtained fiducial points' locations to assess the algorithms' accuracy. Likewise, the corresponding pBRA measurements obtained using the automatically determined vs. manually determined fiducial points were compared.

### 3 MATERIALS & METHODS

#### 3.1 Overview

Methods for the automatic localization of the umbilicus (Section 3.5) and nipples (Section 3.6) were designed based on a development set of clinical photographs (Section 3.2). A normalized measure of symmetry, pBRA2 (Section 3.7), was computed based on the locations of these fiducial points. The localization algorithms and the automatic pBRA2 calculation were evaluated relative to manual localization of the umbilicus and nipples (Section 3.8) using a separate test set of clinical photographs (Section 3.2). All of the image-processing techniques and symmetry calculations were performed using MATLAB® (The Mathworks, Natick, MA).

#### 3.2 Data Set

Two different sets of images were used for development and testing under IRB approved protocols.

The patient population for the development data set consisted of women aged 21 years or older who underwent breast reconstruction surgery at The University of Texas M. D. Anderson Cancer Center from January 1, 1990 to June 1, 2003. The development set consisted of images of women taken prior to surgical treatment for breast cancer and reconstructive surgery. Each subject is in a standard frontal pose (anterior-posterior, AP) against a standard blue background. However, other aspects of the clinical photography were highly variable. We did not have control over the quality of these images since these were *retrospectively* collected under an IRB approved protocol. Typical photographic errors included: full frame of the body was not visible (from chin to pubic bone), shadows due to use of flash, and shoulders of the patient were not leveled. We chose the 18 images demonstrating the best photographic technique out of 23 available for use as the development set in this study.

The patient population for the testing data set consisted of women aged 21 years or older who underwent breast reconstruction surgery at The University of Texas M. D. Anderson Cancer Center from February 1, 2008 to August 1, 2008. The testing dataset consisted of pre-operative and post-operative digital/digitized photographs that were prospectively collected under an IRB protocol. The testing set consisted of images of women in a standard pose (AP) against a standard blue background. While the quality of these images was better controlled than that of the development set, some photographic errors were still noted. Typical photographic errors included: subject was wearing some clothes to cover lower body or the frame exceeded the region defined by the chin and pubic bone. We chose the 10 best images out of 26 available for use as the testing set in this study.

#### 3.3 Color Representation & Preprocessing

Color is usually specified using three coordinates, or parameters, such as brightness, hue, and saturation. These parameters describe the position of the color within the color space being used. Different color spaces are better for different applications, for example some

equipment has limiting factors that dictate the size and type of color space that can be used. Some color spaces are perceptually linear, i.e., a 10 unit change in the stimulus will produce the same change in perception wherever it is applied. Many color spaces, particularly in computer graphics, are not linear in this way [9].

In this study we used color information to help locate anatomical landmarks of interest. For example, we can see color differences between the subject and the blue background and between the color of the areola and the surrounding skin.

In this study we used 2 color spaces: the RGB (Red Green Blue) color space to differentiate the background from the subject and the YIQ (Luma In-phase Quadrature) color space to detect the variation in color between the areola and the surrounding skin.

Since each photograph is taken against a standard blue background, the background region of each image can be simply removed by identifying the pixels for which the blue signal is high. Specifically, using the RGB representation, a pixel is classified as background if the intensity in the blue channel is larger than that of both the red and the green channels. This method of separating background was 100% accurate for the test and development sets.

The YIQ color space defines color by 3 different aspects. These are: (1) Brightness – The amount of light regardless of coloring (Y color space); (2) Hue – the predominant color such as red, green, yellow (I color space); (3) Saturation – the degree of purity of color (Q color space). Since a single-frequency color rarely occurs alone, saturation determines the amount of the colors of the amount of pastel shading [10]. Although brightness and hue are important in defining the color space, saturation is the parameter that determines difference in the shades of color and hence was used for evaluating the change in shade in the areolar region. To convert from the more familiar RGB to YIQ [9], a linear transform is used, as shown:

$$\begin{bmatrix} Y \\ I \\ Q \end{bmatrix} = \begin{bmatrix} 0.299 & 0.587 & 0.114 \\ 0.596 & -0.274 & -0.321 \\ 0.211 & -0.523 & 0.311 \end{bmatrix} \begin{bmatrix} R \\ G \\ B \end{bmatrix}$$

### 3.4 Overview of Automatic Localization Algorithm

The algorithm for automatic localization of fiducial points is divided into two components, which are implemented in series:

1. Umbilicus localization by using Canny edge detection algorithm to analyze linear structures in the intensity (Y) image (Section 3.5)
2. Nipple localization by using a skin-tone differentiation algorithm in the Q image (Section 3.6)

### 3.5 Umbilicus Localization

As stated in the Color Representation & Preprocessing section, the blue background is filtered out and only the region of the image containing the subject is used for further processing. After background removal, the RGB image is converted into the YIQ color space. Subsequent processing to localize the umbilicus is performed on the Y (intensity) image (Figure 1).

The Canny edge detection algorithm is used to detect the edges in the Y image (Figure 1). This edge detection technique has a low error rate and has one response to a single edge.

When photographic standards are maintained, the umbilicus almost always appears as an edge in the edge map of the AP view.

The top 50% of each image is then removed from consideration, as it is extremely unlikely that this region of the image will contain the umbilicus given the standard subject positioning used for clinical photography. Likewise, the search region is limited in the horizontal direction to the middle 20% of the non-background portion of the image, i.e., the middle 20% of the subject. The removal of these regions, which are unlikely to contain the umbilicus, reduces the likelihood of incorrectly localizing an umbilicus (Figure 1).

The umbilicus almost always appears as a horizontal line in the edge map. Hence, any circular objects in the image are filtered out to remove noise. We observe that circular objects in the edge maps are typically freckles.

A Gaussian is centered at the point that lies at approximately the middle of the subject. This location was estimated as the point that is the average height of the umbilicus from the bottom of the image in the 18 images in the development set. The semi-major and semi-minor axes of the Gaussian are the standard deviations of the location of umbilicus in the two directions, respectively, as estimated using the development set. The candidate edge that has maximum weightage as per this Gaussian is selected as the best estimate of the location of the edge corresponding to the umbilicus.

An edge consists of multiple pixels. In the development set it was found that the point with lowest intensity on this edge was the best pixel representation of the umbilicus. The algorithm marks the point with lowest intensity as the umbilicus on the edge detected in the previous step (Figure 1).

### 3.6 Nipple Localization

As stated in the Color Representation & Preprocessing section, the blue background is filtered out and only the region of the image containing the subject is used for further processing. After background removal, the RGB image is converted into the YIQ color space. Subsequent processing to localize the nipples is performed on the Q image. The contrast variations between the nipples and the surrounding skin in the Q color space are subtle. Hence, a contrast stretch is applied to the Q image to enable clearer visualization of the steps of the nipple localization algorithm. Each image is contrast stretched linearly such that the entire range of available intensities is utilized for visualization.

The bottom 30% and top 20% of each image is removed from consideration, as it is extremely unlikely that these regions of the image will contain a nipple given the standard subject positioning used for clinical photography. Likewise, the search region is restricted in the horizontal direction to the middle 10% of the non-background portion of the image, i.e., the middle 10% of the subject. The removal of these regions unlikely to contain a nipple reduces the likelihood of incorrectly localizing a nipple (Figure 2).

The remaining portion of the image that is to be searched for nipples is split into two equal regions using a vertical axis. All subsequent stages of processing are performed separately for the right and left search regions since the appearance of the right and left nipples can be different. A linear contrast stretch is used to ensure that each step of the localization algorithm can be clearly visualized. Separate contrast stretch operations are applied for the right and left search regions.

The Q channel of the image is thresholded such that the 5% of pixels with highest Q intensity are retained. The threshold level of 5% was empirically determined based on the

development set. For most images in the development set, the top 5% of pixels in Q intensity includes the nipples. If a larger region is retained, the possibility of false localization increases. If a smaller portion of the image is analyzed, then the method may fail to localize the nipples because they are not included in the search region. While nipples typically have high intensity in the Q channel, the nipples do not always have the highest intensity relative to other regions of the torso (Figure 2).

Nipples appear as circular objects in the Q image. However, the arms, shoulders and breast borders appear as lines. The Q image is eroded using a ‘line’ structural element to eliminate horizontal and vertical lines in the image due to arms, breast borders, or shoulders.

On clinical photographs, freckles and other local skin variations appear as small regions of high intensity in the Q representation. A median filter is used to eliminate this “salt and pepper” noise. The image is then dilated using a ‘disk’ structural element. Disk structuring element is used to connect disjoint pieces of white areas (to connect areolar areas) (Figure 2).

Up to this point in the algorithm, multiple candidate locations of the nipple have been retained. The purpose of this stage is to select the location from among these candidates that is most likely to contain the nipple. The inputs to this stage of the processing are the Q intensities of small regions of the image. In order to reduce each region to a single localization, an estimate of the middle of the nipple, the centroid of the Q intensities each candidate region is computed.

A Gaussian is centered at the point that lies at approximately the middle of the subject. This location was estimated as the point that is 55% of the vertical extent of the search region from the bottom of the image and at the midway point horizontally. The parameter 0.55 was selected empirically based on observation of the 18 development images. The semi-major and semi-minor axes of the Gaussian are the standard deviations of the location of nipples in the two directions, respectively, as estimated using the development set. The candidate localization that is most likely according to this Gaussian is selected as the best estimate of the location of the nipple (Figure 2).

### 3.7 Breast Retraction Assessment

The Breast Retraction Assessment (BRA) measure [7] is defined as

$$BRA = \sqrt{(L_x - R_x)^2 + (L_y - R_y)^2}$$

where  $L_x$  is the distance from the sternal notch to the left nipple in the  $x$  direction,  $L_y$  is the distance from the sternal notch to the left nipple in the  $y$  direction,  $R_x$  is the distance from the sternal notch to the right nipple in the  $x$  direction, and  $R_y$  is the distance from the sternal notch to the right nipple in the  $y$  direction.

A normalized version of the BRA measure, called pBRA, was developed by Limbergen et al. [8] in order to account for different magnifications of the patient images. The pBRA value is equal to the BRA value divided by the distance from a reference point, the sternal notch, to the nipple of the untreated breast. The development of pBRA was an important advancement, but this measure cannot be directly used for our analyses for two reasons. First, “untreated breast” is a term that has no meaning in breast reconstruction. Typically, both breasts are surgically altered in order to maximize symmetry. Second, the sternal notch is not always visible in digital clinical photographs.

Consequently, we define a similar measure, pBRA2, which is based on using the umbilicus as the reference point and which is normalized based on the maximum distance from either nipple.

$$pBRA2 = \frac{\sqrt{(L_{x2} - R_{x2})^2 + (L_{y2} - R_{y2})^2}}{\max\left(\sqrt{L_{x2}^2 + L_{y2}^2}, \sqrt{R_{x2}^2 + R_{y2}^2}\right)}$$

where,  $L_{x2}$  is the distance from the umbilicus to the left nipple in the  $x$  direction,  $L_{y2}$  is the distance from the umbilicus to the left nipple in the  $y$  direction,  $R_{x2}$  is the distance from the umbilicus to the right nipple in the  $x$  direction, and  $R_{y2}$  is the distance from the umbilicus to the right nipple in the  $y$  direction.

### 3.8 Manual Location of Fiducial Points

A group of three non-clinical observers used a MATLAB® graphical user interface to manually mark the locations of the fiducial points. The manually identified coordinates of the umbilicus were averaged across observers for each image. The manual localizations of the right and left nipples were likewise averaged, respectively. The automatically detected fiducial point locations were compared to the average locations marked by the human observers.

The values of pBRA2 for manually localized and automatically localized points were compared using the intraclass correlation coefficient (ICC). Typically, an ICC value of 0 to 0.4 is considered a sign of poor agreement between two measurements. A value between 0.4 and 0.7 is considered good. Values above 0.7 are considered to denote excellent agreement.

## 4 RESULTS

### 4.1 Umbilicus Localization

Out of the 18 development images, the umbilicus was correctly localized in 15 images. The 3 failures were due to incorrect photography techniques. Following accepted guidelines for this type of clinical photography [11], the image should be from the chin to the pubis bone. In the 3 images that failed, the full frame of the body was not included. Errors resulted when the shoulders were not included, the image frame extended lower than the pubic bone, or the image did not include pubic bone (Figure 3).

Out of the 10 test images, the umbilicus was correctly detected in 6 images. Three of the failures were due to incorrect photography techniques. In one image, the frame extended all the way down to the subject's thighs and hence the umbilicus localization failed because a portion of the subject's clothes was incorrectly identified as the umbilicus (Figure 3). In two other images, the detection of the umbilicus failed because a reference scale was taped to the subject's abdomen rather than the wall (Figure 3). The reference scale was incorrectly identified as being the umbilicus.

The sole algorithm failure that was not due to photographic error resulted from a mole on the subject's body. Every mole does not appear as a circle in the edge map. Hence, the strategy of filling and removing circular objects in order to remove noise due to moles and freckles can fail to eliminate all moles. The contrast strategy failed since this particular mole has low intensity and a high red component (Figure 3).

In summary, the umbilicus was correctly detected in 15 out of 15 (100%) images in the development set which met the photographic standards (ignoring the 3 images that were later discarded due to incorrect photography techniques). For the test set, which was not used in designing the algorithm, the umbilicus was correctly detected in 6 out of 7 (86%) images which met the photographic standards (ignoring the 3 images that were later discarded due to incorrect photography techniques).

#### 4.2 Nipple Localization

As humans have two nipples, there are two targets to localize per image. We report our algorithm results in terms of both individual nipple localizations and also the number of images for which both nipples were correctly identified. Out of the 36 nipples depicted on 18 images in the development set, our algorithm correctly located 28 of them (78%). Out of the 18 images in the development set, both nipples were correctly identified in 13 images (72%). Overall, the algorithm failed to detect one or both nipples on only 5 of the 18 images in the development set.

The effectiveness of the nipple detection algorithm was more rigorously assessed by evaluating it on a test set of images that were not used in designing the algorithm. Out of 10 test images, the nipple detection algorithm failed on 1 image due to incorrect photography technique. The image did not include the full frame of the subject (from the chin to the pubis bone). Out of the 18 nipples depicted on 9 images in the test set which met the photographic standards, our algorithm correctly located 14 of them (78%). Out of the 9 images in the test set which met the photographic standards, both nipples were correctly identified in 6 images (67%).

For both, the development and test sets, failures of the nipple localization algorithm could not be attributed to a common stage of the methodology. Rather, there are a couple of different assumptions in the algorithm that can sometimes be violated. For example, in one case the subject's right breast was considerably smaller than average, leading to failure of the stage that infers the most likely target based on prior knowledge of where the nipples typically fall in the image (Figure 4). As another example, one subject did not exhibit as much contrast as is typical between the nipple regions and the surrounding skin in the Q image. Hence, less of the areolar region was captured in the thresholded image and both the nipples were misidentified (Figure 4).

#### 4.3 pBRA2 Comparisons

The three images in the test set in which both of the nipples and the umbilicus were correctly detected were used to calculate pBRA2 values. Four pBRA2 values were calculated for each image (Table 1). Three values were computed based on manually localized fiducial points for each of three human non-clinical observers and the fourth value was computed using the automatically localized fiducial points. The intraclass correlation coefficient (ICC) of pBRA2 measurements was calculated based on fiducial points localized by all pairs of the four observers, with the automatic localization algorithm treated as an observer (Table 2). All the values were higher than 0.7. An ICC value above 0.7 indicates excellent agreement between two measurements.

### 5 DISCUSSION

Digital photography is commonly used to document surgical outcomes. Thus, it would be valuable if quantitative measures of anatomy could be automatically calculated from digital photographs. The goal of this study was to develop and implement a fully automated method of calculating a variant of the common symmetry measure pBRA, which we term pBRA2.



The method requires the automatic localization of the umbilicus and the nipples on images of anterior-posterior digitized/digital clinical photographs. The automatic localizations were compared to manually obtained fiducial point locations to assess the localization algorithms' accuracy. Likewise, the corresponding pBRA2 measurements obtained using the automatically determined vs. manually determined fiducial points were compared.

In the nipple localization algorithm, color differences between areola and skin are used. Conventionally, colorimeters or spectrophotometers are used to quantitatively assess skin color. However, we have previously shown that color measurements obtained by clinical digital photography are equivalent to those obtained using colorimetry [12]. Hence, the extraction of color information from clinical digital photographs is appropriate and accurate.

To provide "ground truth" for the locations of the nipples and umbilicus on each image, a group of three non-clinical observers manually marked the locations of the fiducial points. In a previous study, we demonstrated that non-clinical observers using a simple MATLAB® GUI are able to accurately mark the locations of some fiducial points including the nipples [6]. There is excellent agreement between the pBRA2 measurements computed based on manually detected fiducial points and the pBRA2 measurements computed based on automatically localized fiducial points. Moreover, the agreement between the pBRA2 measurements based on automatically detected fiducial points and those based on manually detected fiducial points is as good as the agreement between pBRA2 measurements based on manual localizations made by two different human observers.

While our results support the use of an algorithm for quantitative measurements of breast properties such as pBRA2, it is important to note that standardization of photographic conditions is critical. Substantial variability may be observed if the photographic conditions are not controlled. Standardization of lighting, consistent camera-to-subject distances, proper patient positioning and the standardized positioning of the camera (vertical or horizontal view) are essential features for standardized photographic documentation in plastic surgery [11]. For our algorithm, important parameters to standardize include: frame of the body photographed, distance and direction from the light to the subject, camera lighting settings, background color, and the subject's pose relative to the camera.

In conclusion, this paper presents new algorithms for automatically localizing the umbilicus and nipples on standard anterior-posterior clinical photographs, enabling the automatic calculation of measure of breast symmetry (pBRA2). The algorithms were designed using a development set of retrospectively collected clinical photographs and evaluated using a separate test set of prospectively collected clinical photographs. This study demonstrates that under standardized photographic conditions, automatic localization of fiducial points and subsequent computation of measures such as symmetry is feasible. The algorithms presented here for quantifying surgical outcomes in breast reconstructive surgery will provide a foundation for future work on assisting a cancer patient and her surgeons in selecting and planning reconstruction procedures that will maximize the woman's psychosocial adjustment to life as a breast cancer survivor.

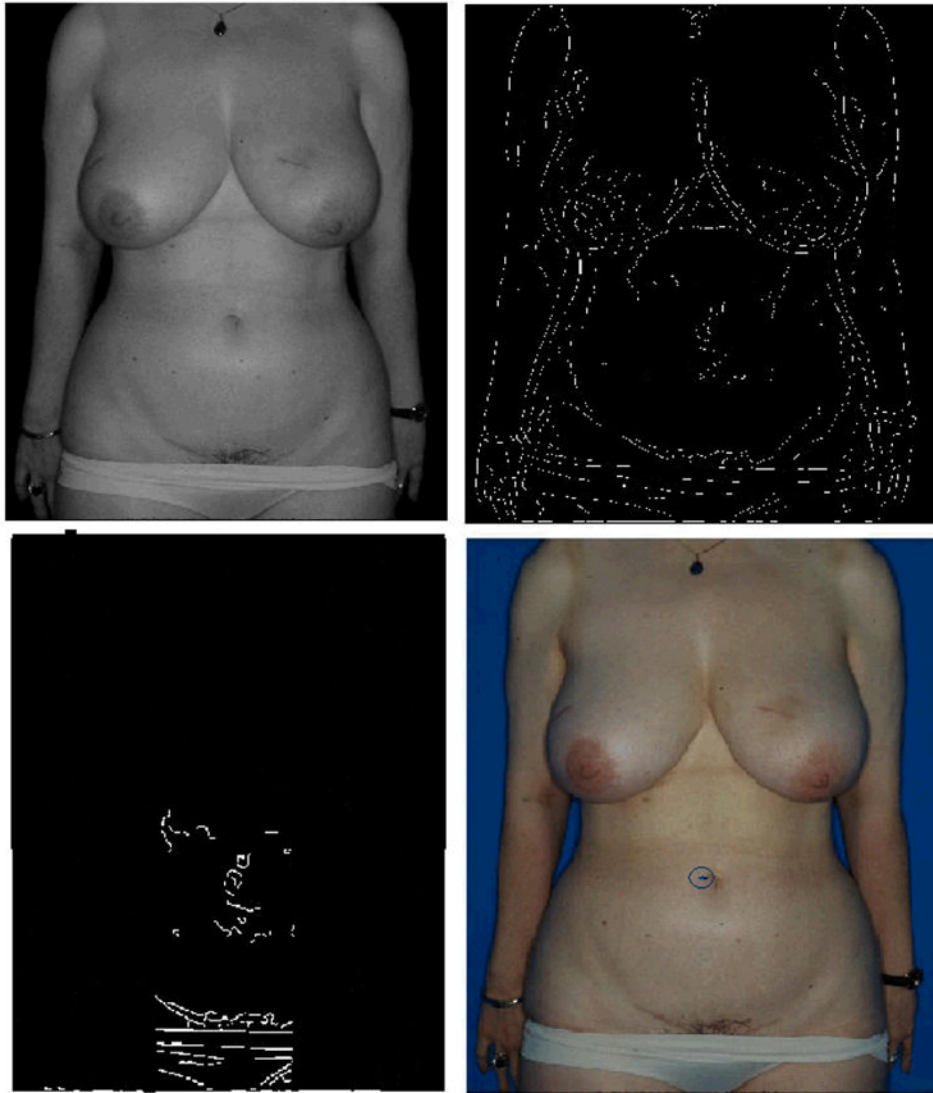
## Acknowledgments

This study was supported in part by grant 1R21 CA109040-01A1 from the National Institutes of Health and grant RSGPB-09-157-01-CPPB from the American Cancer Society.

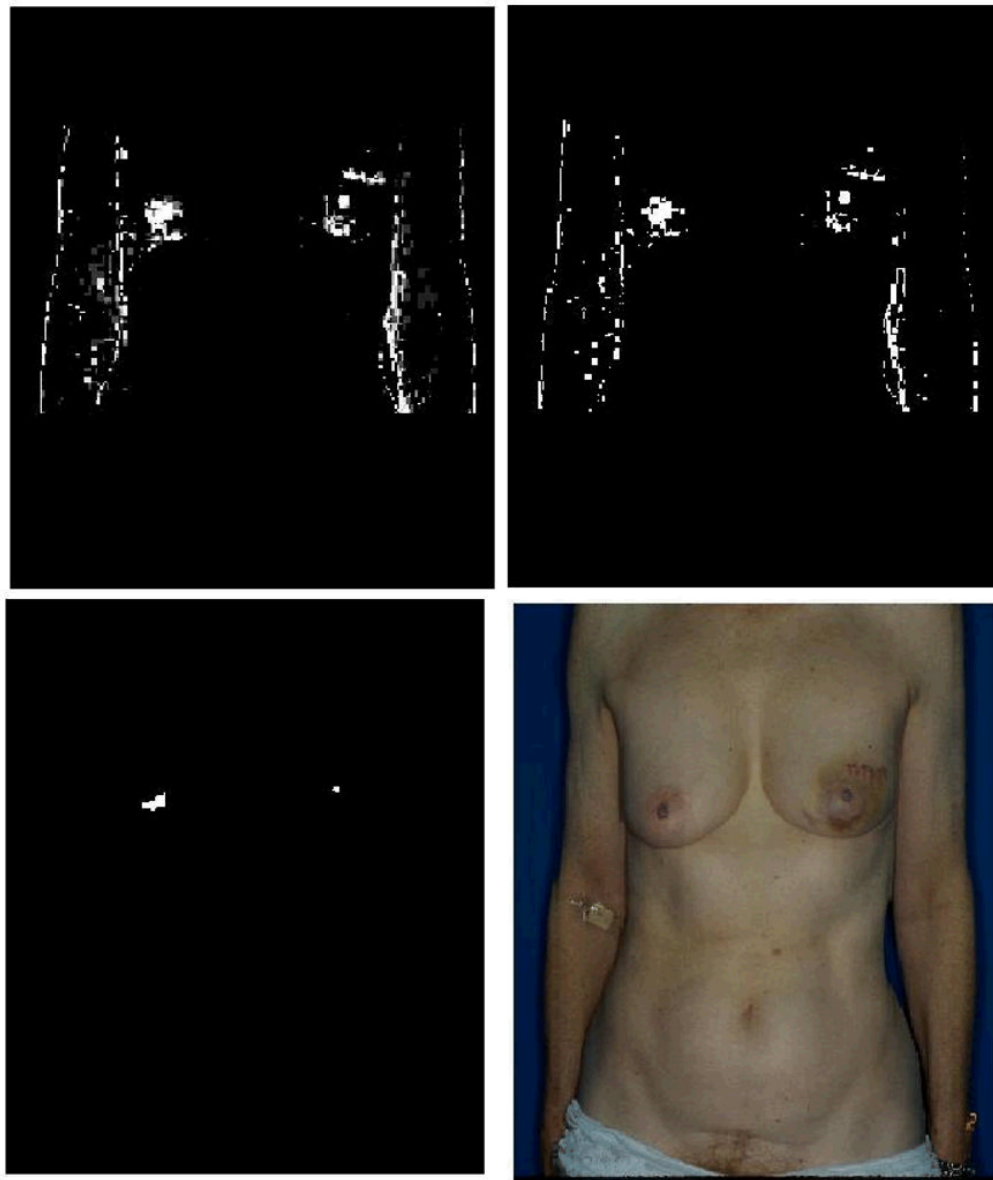
## References

1. American Cancer Society. Cancer facts & figures 2009. Atlanta: 2009.

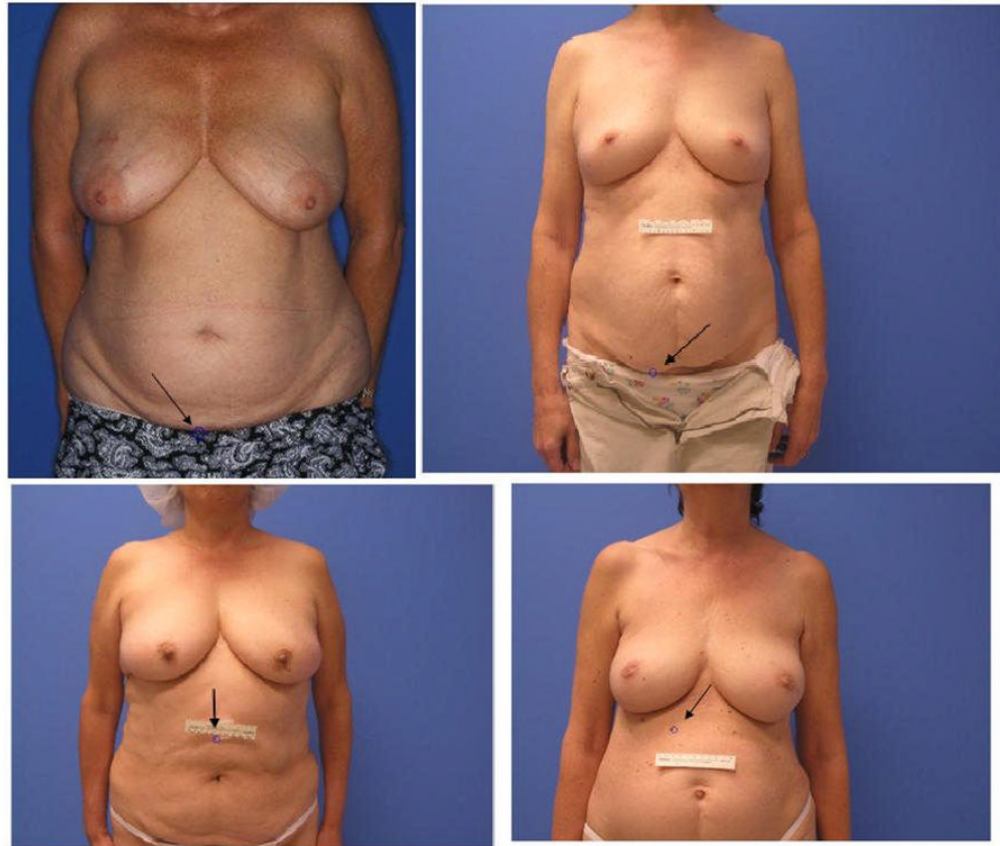
2. Kim MS, Sbalchiero JC, Reece GP, Miller MJ, Beahm EK, Markey MK. Assessment of breast aesthetics. *Plastic and Reconstructive Surgery*. 2008; 121(4):186e–194e.
3. Dabeer M, Fingeret MC, Merchant F, Reece GP, Beahm EK, Markey MK. A research agenda for appearance changes due to breast cancer treatment. *Breast Cancer: Basic and Clinical Research*. 2008; 2:1–3. [PubMed: 21655363]
4. Fabry HFJ, Zonderhuis BM, Meijer S, Berkhof J, Van Leeuwen PAM, Van der Sijp JRM. Cosmetic outcome of breast conserving therapy after sentinel node biopsy versus axillary lymph node dissection. *Breast Cancer Research and Treatment*. 2005; 92(2):157–162. [PubMed: 15986125]
5. Vrieling, C.; Collette, L.; Fourquet, A.; Hoogenraad, WJ.; Horiot, JC.; Jager, JJ.; Pierart, M.; Poortsman, PM.; Struikmans, H.; Van der Hulst, M.; Van der Schueren, E.; Bartelink, H. *Int J Radiat Oncol Biol Phys*. European Organization for Research and Treatment of Cancer; 1999. The influence of the boost in breast-conserving therapy on cosmetic outcome in the eortc “Boost versus no boost, in EORTC Radiotherapy and Breast Cancer Cooperative Groups; p. 677-85.
6. 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. [PubMed: 16438948]
7. Pezner RD, Patterson MP, Hill LR, Vora N, Desai KR, Archambeau JO, Lipsett JA. 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.
8. 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. [PubMed: 2587807]
9. Ford, ARA. *Colour space conversions*. 1998.
10. Buchsbaum, WH. *Color tv servicing*. 3. Englewood Cliffs: Prentice-Hall; 1975.
11. Ellenbogen RE, Jankauskas SJ, Collini FJ. Achieving standardized photographs in aesthetic surgery. *Plastic & Reconstructive Surgery*. 1989; 86(5):955–958. [PubMed: 2236321]
12. 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; 15(1):20–31. [PubMed: 19239578]



**Figure 1.** This figure shows the steps of the umbilicus localization algorithm. The upper left image shows the Y image after the background has been filtered out. The upper right image shows the edges detected by the Canny edge detection algorithm. The bottom left image shows the process of restricting the area of interest by cropping the regions that are unlikely to contain the umbilicus given the standard pose employed. The bottom right shows the correctly localized umbilicus marked on the original RGB image.

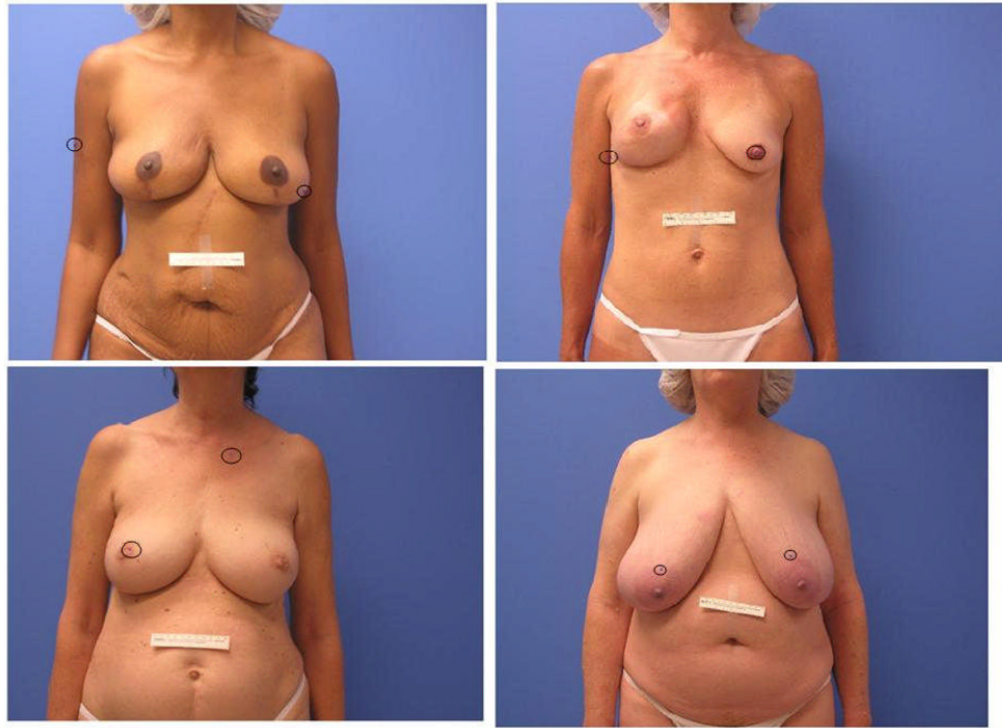


**Figure 2.** This figure shows the steps of the nipple localization algorithm. The image on the top left shows the contrast stretched Q image. The image on top right shows the image thresholded such that only the 5% highest intensity pixels from the previous image are retained. The bottom left panel shows the image after erosion, dilation, and pre-processing in the nipple detection technique. The bottom right image shows the correctly localized nipples marked on the original RGB image.



**Figure 3.**

In the top left image, there was inaccurate localization of the umbilicus due to incorrect photography techniques; the subject's shoulders should be included in the image. The image should be taken from the chin to the pubis bone. The patient should not be wearing any clothes. In the top right image, inaccurate localization of umbilicus occurred due to incorrect photography techniques; again, the image should be taken from the chin to the pubic bone only and the patient should not be wearing any clothes. In the bottom left image, there was inaccurate localization of the umbilicus due to placement of tape on the body. Tape should be placed on the wall instead of on the subject. In the bottom right image, inaccurate localization of the umbilicus occurred due to a mole on subject's body.



**Figure 4.**

This figure illustrates the variability in the errors of the nipple localization algorithm, which could not be attributed to a common stage of the methodology. Rather, there are a couple of different assumptions in the algorithm that can sometimes be violated. For example in the image in the upper right panel, the subject's right breast was considerably smaller than average, leading to failure of the stage that infers the most likely target based on prior knowledge of where the nipples typically fall in the image.

**Table 1**

pBRA2 was computed for each of the three subjects based on automatically detected fiducial points and also based on fiducial points manually localized by each of three observers.

Subject #	Automatic	Observer 1	Observer 2	Observer3
1	0.911	0.917	0.896	0.907
2	0.813	0.800	0.813	0.806
3	0.841	0.872	0.849	0.869

**Table 2**

The intra-class correlation coefficients between pairs of pBRA2 calculated using automatically localized fiducial points and manually localized fiducial points indicate excellent agreement.

Comparison of pBRA2 computed using different localizations of the fiducial points	ICC
Automatic vs. Observer 1	0.936
Automatic vs. Observer 2	0.975
Automatic vs. Observer 3	0.943
Observer 1 vs. Observer 2	0.928
Observer 1 vs. Observer 3	0.992
Observer 2 vs. Observer 3	0.955