A Pilot Study on Using Eye Tracking to Understand Assessment of Surgical Outcomes from Clinical Photography

Min Soon Kim · Angela Burgess · Andrew J. Waters · Gregory P. Reece ·
Elisabeth K. Beahm · Melissa A. Crosby · Karen M. Basen-Engquist · Mia K. Markey

Published online: 18 September 2010
© Society for Imaging Informatics in Medicine 2010

Abstract Appearance changes resulting from breast cancer treatment impact the quality of life of breast cancer survivors, but current approaches to evaluating breast characteristics are very limited. It is challenging, even for experienced plastic surgeons, to describe how different aspects of breast morphology impact overall assessment of esthetics. Moreover, it is difficult to describe what they are looking for in a manner that facilitates quantification. The goal of this study is to assess the potential of using eye-tracking technology to understand how plastic surgeons assess breast morphology by recording their gaze path

while they rate physical characteristics of the breasts, e.g., symmetry, based on clinical photographs. In this study, dwell time, transition frequency, dwell sequence conditional probabilities, and dwell sequence joint probabilities were analyzed across photographic poses and three observers. Dwell-time analysis showed that all three surgeons spent the majority of their time on the anterior–posterior (AP) views. Similarly, transition frequency analysis between regions showed that there were substantially more transitions between the breast regions in the AP view, relative to the number of transitions between other views. The results of both the conditional and joint probability analyses between the breast regions showed that the highest probabilities of transitions were observed between the breast regions in the AP view (APRB, APLB) followed by the oblique views and the lateral views to complete evaluation of breast surgical outcomes.

M. S. Kim
Department of Health Management and Informatics,
University of Missouri School of Medicine,
Columbia, MO, USA

A. Burgess
The Psychology Department, Wichita State University,
Wichita, KS, USA

A. J. Waters
Uniformed Services University of the Health Sciences,
Bethesda, MD, USA

G. P. Reece · E. K. Beahm · M. A. Crosby
Department of Plastic Surgery, The University of Texas M. D.
Anderson Cancer Center,
Houston, TX, USA

K. M. Basen-Engquist
Department of Behavioral Science, The University of Texas M. D.
Anderson Cancer Center,
Houston, TX, USA

M. K. Markey (✉)
The University of Texas Department of Biomedical Engineering,
1 University Station C0800,
Austin, TX 78712, USA
e-mail: mia.markey@mail.utexas.edu

Keywords Breast neoplasm · Eye movements · Biomedical image analysis · Decision support · Evaluation research

Introduction

There are approximately 250,000 new cases of breast cancer diagnosed each year in the US [1]. Improved screening and treatment methods have increased the breast cancer survival rate such that the majority of women with localized, early stage disease can expect to live many years after a diagnosis of breast cancer. Thus, in addition to continuing to strive to improve detection and survival rates, more work is needed on issues that influence the quality of breast cancer survivors' lives. One important factor is appearance change as a result of surgical treatment for breast cancer.

Breast reconstructive surgery is an important component to improve the patient's quality of life. Approximately 57,000 breast reconstructive surgeries are performed annually in the US, according to the American Society of Plastic Surgeons [2]. 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. Currently, physicians, patients, or other observers evaluate characteristics of the reconstructed breasts, such as symmetry and proportion, in a subjective, qualitative manner [3]. However, such assessments are typically based on vaguely defined rating scales that have low intra- and inter-observer agreement. Their qualitative nature also restricts the analyses that can be performed. Current quantitative approaches to breast esthetic measurements include: measurements on the patient's body (anthropometry) [4, 5], measurements on 2D [6–9] or 3D [10, 11] imaging. Anthropometry can be a useful tool for quantifying esthetic outcomes. However, this method can be complicated and time intensive, making it impractical for routine use. Using photographs has advantages over anthropometry. A photograph is more efficient and less intrusive for the patient because it serves a stable record for a variety of measurements. However, this method also suffered from substantial intra- and inter-observer variability, due to the lack of consistency in the manual identification of anatomic landmarks [12]. Three-dimensional imaging has tremendous potential for analysis of breast appearance. However, it is not routinely collected in clinics due to the high cost.

Quantitative, objective measures are needed to enable outcomes research. Understanding how different kinds of observers perceive reconstruction and developing of tools for quantifying breast morphology could impact breast cancer care in a variety of ways. In the future, such measures may also be used directly in surgical planning. Another possible application is in surgical education. In addition to laying the groundwork for assisting surgeons in performing better reconstructions, such advances would enable the development of patient decision aids. Another use is in setting appropriate reimbursement rates for different procedures. To the best of our knowledge, there are no prior studies that substantively investigate the relationships between specific surgical variables reliably and psychosocial adjustment. In fact, this lack of detailed understanding of the psychological consequences of reconstructive procedures motivates our overall research agenda on relating changes in breast morphology to the perceptions of cancer survivors and caregivers. However, in developing quantitative, objective measures that capture critical aspects of breast appearance, it would be helpful to understand how people subjectively evaluate breast morphology.

The goal of this investigation is to explore the potential of eye-tracking technology for elucidating the process of subjective assessment of breast reconstruction surgery outcomes on 2D images. In this pilot study, we record the gaze path of three plastic surgeons while he or she rates breast characteristics based on clinical photographs. Eye tracking is a valuable tool for understanding medical image perception. For example, eye tracking has been used in numerous studies on the visual search processes of radiologists. Eye tracking has enabled scientists to study factors such as the effect of lesion conspicuity on the visual search strategy of radiologists' in mammogram reading [13], the relationship between duration of gaze at lesion and correct diagnosis of it [14], the relationship between lesion subtlety and detection of the lesion [15], and to predict radiologists' diagnosis using spatial frequency representation of regions [16]. Eye tracking has also been used in some studies to investigate how body-image assessment is related to preferential attention toward the body part. For example, eye tracking was used to investigate how age and fatigue judgments are made using the facial cues as participants viewed full-face digital photographs [17] and to investigate the visual assessments by men of female attractiveness who viewed computer-morphed anterior–posterior (AP) photographs of the woman with same face, but differ for waist to hip ratio [18]. In comparison, there have been few studies employing eye tracking in the field of plastic and reconstructive surgery. For example, Ishii et al. reported statistically significant differences in the patterns of scan paths of eight naive observers when gazing at images of faces with vs. without surgical deformities [19]. Moreover, we are unaware of any previous studies in breast reconstructive surgery that employed eye-tracking technology. Thus, there is a clear opportunity to introduce this powerful methodology to a new clinical research area where we can evaluate and reengineer the current methods used to assess breast reconstructive surgery.

Materials and Methods

Eye Tracker

An Applied Science Laboratories (Bedford, MA) Eye-Tracking System (Model 504) was used to track the gaze of plastic surgeons in the Behavioral Research and Treatment Center at The University of Texas M. D. Anderson Cancer Center (UT MDACC). This machine is a desktop eye-tracking system which can unobtrusively collect eye position data while a subject examines an image presented on the computer screen. A height-adjustable chin rest was used to have observer's head to be stable. A research coordinator (AB) based at UT MDACC and the first author

(MSK) conducted the eye tracking experiments at UT MDACC. The Model 504 eye tracker is an optical magnetic head tracker where the eye is illuminated by the beam from near infrared LEDs on the pan/tilt optics module. An auto-focusing lens system in the pan/tilt module focuses a telephoto image of the eye onto a video sensor (eye-camera). The pan/tilt mechanism can rotate the illumination source in order to follow the eye as the subject moves his/her gaze. According to the specifications provided by the eye-tracker manufacturer, the measurable field of view is about 25° visual angle to either side of the optics, about 25° above the optics, and about 10° below the optics. Precision refers to how closely individual measurements agree with each other. The precision of the Model 504 eye tracker is better than 0.5° of the visual angle, which corresponds to a circle of 0.55 cm in radius on the monitor which displays the image at a viewing distance of 63.5 cm. Accuracy refers to how closely a measured value agrees with the correct value. The accuracy between the true eye position and the computed measurement is less than 1° of the visual angle 1.11 cm, which corresponds to a circle of 1.11 cm in radius on the monitor under the viewing conditions defined above.

Calibration

The eye-tracking system was connected to a computer that ran the calibration and interface software (ASL) and saved the data, while another computer was used to display the images in a darkened room. In this setup, the subject was seated in the eye-tracker room equipped with monitoring facilities that allows communication with monitoring room. Images were displayed on a standard LCD monitor with 1,024×768 pixels resolution using the Internet Explorer browser (Microsoft, Redmond, WA). The raw data measured by the eye tracker are the separation between the pupil center and the corneal reflection. The relationship between these raw values and the eye line of gaze differs for each subject and for different optical units. The purpose of the calibration process is to provide data that allow the eye tracker to account for individual subject differences. The objective is to have the subject look at (fixate on) each of the nine calibration points, which are at known locations. The points are numbered from left to right; 1–3 for the top row, 4–6 for the middle row, and 7–9 for the bottom row. The points cover about 80% of the monitor screen area and are separated by 15–20° visual angle horizontally and 10–15° vertically. These are ideal specifications. The actual distribution of the nine points is taken from the scene monitor and entered into memory with the eye-tracker “set target points” function. The system was calibrated at the beginning of the session and repeated before each case for

every subject. The observers were given a short break after each case because they felt that the chin rest arrangement was not comfortable enough for longer viewing sessions.

Data Collection

Eye data were collected from three expert plastic surgeons (EKB (Observer 1), GPR (Observer 2), MAC (Observer 3)) and their experience with breast reconstructive surgery extends from 5 to 24 years with approximately 150 cases a year. They viewed color clinical photographs of eight women aged 21 years or older who underwent breast reconstruction from January 1, 1990 to June 1, 2003 at UT MDACC. The data pool consisted of digitized conventional 35-mm photograph, slide film and digital photographs. The digital cameras used were a Nikon 990 Coolpix (Nikon, Tokyo, Japan) or Cannon T90 35 mm SLR with 50-mm lens (Canon, Tokyo, Japan). No artificial calibration was made because doing so would influence the assessment of the outcomes. All photographs were taken against a sky blue background, following standard guidelines for clinical photography in plastic surgery [20]. A forward facing flash unit, but no studio umbrella flash, was used in a room lit by overhead fluorescent lights. The images were selected by an experienced plastic surgeon (GPR) to demonstrate a large variety of breast esthetic characteristics, such as size, symmetry, ptosis, and projection. This study was approved by the Institutional Review Boards (IRB) of The University of Texas at Austin and The University of Texas M. D. Anderson Cancer Center (UT MDACC). A web-based interface was used to display the clinical images and observer ratings of breast characteristics, e.g., symmetry, during the eye-tracker experiments. Five images of each patient are displayed: an anterior–posterior view and lateral and oblique views of both the right and left breasts (Fig. 1). The oblique views are taken at an angle of about 45° from the true lateral. The vertical extent of the views is from just below the chin down to the top of the pubic bone. Three plastic surgeons rated breast properties, e.g., symmetry. The rating scale used in the manuscript is a part of a preliminary breast esthetic rating instrument developed under the leadership of an experienced behavioral scientist at UT MDACC that is currently being validated. The draft instrument has 25 items, each of which uses a 10-point scale. Each rating scale in the lexicon includes a thorough description to help standardize the terminology used to report the subjective assessment of esthetic outcomes in reconstructive surgery and improve communication. The draft instrument consists of 11 symmetry ratings items, 14 individual breast ratings items, and a global rating on overall appearance before and after the entire rating items. Out of these 25 items, six rating items that most impact and/or reflect the surgical outcomes were selected based on

Fig. 1 The visual scan path recorded by eye-tracking equipment while Observer 2 examined a case to rate the “Symmetry of size of breast mounds”. Regions of interest were identified around the breast areas in each view. The scan path demonstrates that the observer’s eyes were more drawn to the breast regions where surgical scars were present. Five images of each subject are displayed: an anterior–posterior (AP) view and lateral and oblique views of both the right and

left breasts. The rating item text is displayed only before and after the images are displayed to prevent the observer from being distracted by the item text. *Rectangle*, (□): regions of interest. *Dots* (•): fixations of an observer with size proportional to the amount of time. *Line* (—): gaze path *dots* on the nipple and the *dots* (•), reference points for coordinate transformation from eye tracker to displayed image

preliminary statistical analyses and discussion among the developers (Table 1). Unfortunately, we are not prepared to present the validity and reliability at this point because the scale is still being evaluated; therefore, statistical analysis between the magnitude (score) of ratings and eye-tracker data were not presented in this study. The rating items used were: (1) initial impression of overall appearance of the breasts, (2) symmetry of size of breast mounds, (3) symmetry of shape of breast mounds, (4) esthetic shape, (5) natural shape, and (6) final impression of overall appearance of the breasts. Vertical and horizontal coordinates of eye position and pupil diameter were saved to an “Eyedat file” on the Interface PC hard disk. In addition, experimenters marked the events to separate the recordings for each item in the control room. A field of data, consisting of the elements including eye tracking data, event marks and item ratings, were recorded every 60th of a second (60 Hz update rate).

Data Analysis

EyeNal and Fixplot (ASL, Bedford, MA) were used to plot and analyze the raw eye data on the images. Fixation is defined as the mean eye position over a minimum time period (0.1 s) during which the gaze stays within one visual angle. We are unaware of a precise, universally accepted definition of a fixation and its minimum/maximum amplitude. A 1° visual angle and a minimum time of 100 ms was considered reasonable based on discussions with behavioral scientists on our team who had previously applied eye

tracker in many human behavior studies. Higher fixation frequency on a particular area can be indicative of greater interest in the area under consideration. Dwell is defined as the amount of time during which a contiguous series of one or more fixations remains within a region of interest (ROI). In this study, median dwell times across rating items, patients, and observers were compared.

Two sets of regions were created to further stratify the data, one for regions for the areas of five photographic views (five ROIs): (1) AP view, (2) left oblique view, (3) right oblique view, (4) left lateral view, and (5) right lateral view, and another set for the areas of breast regions within the five photographic views (six ROIs): (1) left breast in AP view, (2) right breast in AP view, (3) both breasts in the left oblique view, (4) both breasts in the right oblique view, (5) left breast in the left lateral view, and (6) right breast in the right lateral view. Fixations were then plotted for each individual item within the regions of interest. Dwell-time tables and transition tables were analyzed.

Results

Dwell-Time Analysis

Median dwell times across the photographic views (an anterior–posterior view and lateral and oblique views of both the right and left breasts), rating items, patient cases, and observers were compared. The average time spent

Table 1 Descriptions of the breast esthetic observer rating instrument

Initial impression of overall appearance of the breasts	This scale rates your initial impression of the overall appearance of the breasts from an esthetic point of view
Symmetry of size of breast mounds	This scale compares the symmetry of the size of the breast mounds, irrespective of the composition of the skin of the breast, i.e., native breast skin, flap skin, or some combination of the two types of skin.
Symmetry of shape of breast mounds	This scale compares the symmetry of the shape of the breast mounds, irrespective of the composition of the skin of the breast, i.e., native breast skin, flap skin, or some combination of the two types of skin.
Esthetic shape	This scale rates observer's impression of the overall esthetic shape of each of the breasts. A highly esthetic breast should be pleasing to eye by being proportional to the chest wall for volume (size and projection) and contour and by being symmetrical with the opposite breast (for size, shape, color, contour and position of the breast mound, medial take off of the breast, nipple, and areola). The esthetic breast should not be pseudoptotic or ptotic (at most—minor ptosis).
Natural shape	This scale rates your impression of the overall natural appearance of the shape of each breast. A natural-shaped breast should be compatible with the native breast. The breast should not look surgically altered nor have the appearance of a breast augmentation, reduction, or mastopexy. Some characteristics of an unnaturally shaped breast are superior pole fullness, elements of a distorted contour, and excessively round contours.
Final impression of overall appearance of the breasts	This scale rates your initial impression of the overall appearance of the breasts from an esthetic point of view.

This rating scale is part of observer rating scale that consists of 11 symmetry ratings items, 14 individual breast ratings items, and a global rating on overall appearance before and after the entire rating items

Out of these 25 items, six rating items that most impact and/or reflect the breast properties were selected based on the consensus from the discussions of plastic surgeons

Each rating scale is explained with the lexicon prepared to help standardize the terminology used to report the subjective assessment of esthetic outcomes in reconstructive surgery and improve communication between reconstructive surgeons and radiation oncologists

collecting data for rating all six rating items for a patient case was: 14.0 s for Observer 1, 19.9 s for Observer 2, and 95.5 s for Observer 3, suggesting that there is some inter-observer variability in how plastic surgeons approach this task. The amount of dwell time across rating items varied by observer. For example, Observer 1 spent more time on the items pertaining to shape ((3) symmetry of shape of breast mounds, (4) esthetic shape, and (5) natural shape) than to the other items, whereas Observer 2 and Observer 3 spent the most time on the first item (initial impression of overall appearance) and less time on subsequent items (Fig. 2). Another key trend was noted when photographic views were compared. All three of the surgeons spent more time looking at the AP views than at the lateral and oblique views (Fig. 2).

In a similar manner, median dwell times across the breast regions (separate regions in the anterior–posterior view and lateral and oblique views of both the right and left breasts), rating items, cases, and observers were compared. All three surgeons spent the most time on the breast regions in AP views; specifically, more time on left breast regions than right breast region. Please refer to Fig. 1 for an example of how the images were presented for viewing.

Dwell Transitions Frequency Analysis

A transition from region *i* to region *j* is defined as a dwell period in area *i* followed immediately by a dwell period in area *j*. The frequency of transitions between the breast

regions in each view was analyzed to investigate the way that the surgeons compare the reconstruction outcomes of the two breasts (Table 2). The regions of interest used were: (APRB) patient's right breast area in AP view, (APLB) patient's left breast area in AP view, (ROB) patient breast area in right oblique view, (LOB) patient's breast area in left oblique view, (RLB) patient's right breast in right lateral view, (LLB) patient's left breast in left lateral view (Fig. 1). The average number of transitions between regions across the eight cases was computed for each observer for each item. For brevity, we present the results for Item 1 only because similar results were obtained across the rating items. In the transition matrices in Table 2, the region given in the row label is the starting position and the region named in the column label is the ending position. When transitions between the regions were analyzed, some consistent patterns were seen across observers and rating items. The results showed that there were many transitions between the breast regions in the AP view (APRB, APLB) across all three observers and across all six items. The largest number of transitions was between region APLB and region APRB for Observer 1 and Observer 2 across the rating items, while Observer 3 had a large number of transitions between regions in the oblique and lateral views as well as for regions in the AP views. Few transitions were made between the breast regions of the oblique and lateral views, in comparison to the many transitions observed between the breast regions in the anterior–posterior view.

Fig. 2 Median dwell time of five photographic views spent by three experienced plastic surgeons of assessing the outcomes of eight patients who underwent reconstruction. When dwell time was compared across the rating items, the results showed that Observer 1 (EKB, *Top*) spent more time on the items pertaining to shape than on the other items (*Item 4*), while Observer 2 (GPR, *Center*) and Observer 3 (MAC, *Bottom*) spent the most time on the items on initial impression of overall impression (*Item 1*) and spent less time on the subsequent items. Another key trend was all three of the surgeons spent more time looking at the AP views than at the lateral and oblique views. The rating items are: (*Item 1*) initial impression of overall appearance of the breasts, (*Item 2*) symmetry of size of breast mounds, (*Item 3*) symmetry of shape of breast mounds, (*Item 4*) esthetic shape, (*Item 5*) natural shape, and (*Item 6*) final impression of overall appearance of the breasts

This result is consistent that surgeons primarily utilize the AP views of the breast regions in order to assess surgical outcomes and only move on to either oblique or lateral views to complete the assessment.

Dwell Sequence Conditional Probabilities

Average conditional probabilities of transitions between the breast regions in each view were computed and analyzed to identify possible navigation patterns among views as surgeons compare the reconstruction outcomes of the breasts (Table 3). Conditional probability is the likelihood that given a dwell on region i , the next dwell will be on another specific region j . Dwell sequence conditional probability $P(i | j)$ is the number of dwell transitions from region A_i to A_j divided by the number of dwells on A_i Eq. 1. In the transition matrices in Table 3, the region given in the row label is the starting position and the region named in the column label is the ending position. For brevity, we present the results for Item 1 only because similar results were obtained across the rating items. This table reads the same way as the transition table except that it is a matrix representing the conditional probability of transitional activity between regions of interest (ROI's).

$$P(A_j | A_i) = \frac{\text{Number of dwell transitions from } A_i \text{ to } A_j}{\text{Number of dwells on } A_i} \quad (1)$$

When conditional probabilities between the regions were analyzed, some consistent patterns were seen across observers and rating items. Overall, the highest conditional probabilities were between the breast regions in the AP view (APRB and APLB): $P(APLB|APRB)=40\%$ and $P(APRB|APLB)=34\%$. Consistent with the dwell-time analysis, this result suggests that surgeons primarily evaluate morphology based on what is seen in the breast regions in AP views. The second highest conditional probabilities were observed between the oblique views and between the lateral views: $P(LOB|ROB)=19\%$ and $P(LLB|RLB)=20\%$. This result may indicate that the observers tend to first cross-check the morphology within the images taken at the same angle then against the images taken at different angles to complete the assessment. The smallest conditional probabilities ($P \leq 10\%$) varied depending on rating item, observer, and view.

Dwell Sequence Joint Probabilities

The average of joint probabilities between regions across the eight cases was computed for each observer for each item (Table 4). Joint probability analysis takes account of the dependence between regions of interest, as well as the

Table 2 Mean transition activity across six regions of interest recorded while three observers examined eight patients who underwent breast reconstruction case to rate the initial impression of overall appearance of the breasts (Item 1)

	APRB O1/O2/O3	APLB O1/O2/O3	ROB O1/O2/O3	LOB O1/O2/O3	RLB O1/O2/O3	LLB O1/O2/O3
APRB	-/-	0.9/0.9/1.9 ^a	0.1/0.1/0.1	-/-	-/0.1/-	-/-/0.1
APLB	1.5/1.3/2.1 ^a	-/-	0.1/-/0.3	-/-	0.1/-/0.4	-/-
ROB	0.3/-/-	0.1/-/-	-/-	-/0.3/0.8	-/0.1/0.4	-/-
LOB	-/-	-/-	-/-/0.4	-/-	-/0.1/-	-/0.1/0.4
RLB	-/-/0.3	0.3/-/0.1	-/0.1/0.3	-/-	-/-	-/0.3/1.0
LLB	-/-	-/-/0.3	-/-	-/-	-/0.1/0.5	-/-

When the number of transitions was compared between the regions, the results showed that there were a notably high number of transitions between the breast regions in AP views (APRB, APLB)

Transitions either from or to areas outside of the breast regions were not considered in the analyses because we were mainly interested in transitions among breast regions

A dash (-) was used to decrease the number of zero's and to allow numbers with importance to stand out

^a The highest transitional activity was from region APLB to region APRB for the all three surgeons, Observer 1 (1.5), Observer 2 (1.3), and Observer 3 (2.1) followed by from region APRB to region APLB for the all three surgeons, Observer 1 (0.9), Observer 2 (0.9), and Observer 3 (1.9)

distributions. Joint probability is the relative likelihood (or relative frequency) of a transition between two particular regions as compared to transitions between other pairs of regions. Dwell sequence joint probability $P(i,j)$ is the number of dwell transitions from region A_i to A_j divided by the total number of dwells transition pairs Eq. 2. In the transition matrices in Table 4, the region given in the row label is the starting position and the region named in the column label is the ending position. For brevity, we present the results for Item 1 only because similar results were obtained across the rating items. This table reads the same way as the transition table except that it is a matrix

representing the joint probability of transitional activity between regions of interest (ROI's).

$$P(A_j|A_i) = \frac{\text{Number of dwell transitions from } A_i \text{ to } A_j}{\text{Total number of dwell transition pairs}} \quad (2)$$

When joint probabilities between the regions were analyzed, some consistent patterns were seen across observers and rating items. Overall, the results showed that the highest joint probabilities were between the breast regions in the AP view (APRB, APLB): $P(\text{APRB}, \text{APLB}) = 8\%$ and

Table 3 Conditional probabilities across six regions of interest recorded while three observers examined eight patients who underwent breast reconstruction case to rate the initial impression of overall appearance of the breasts (Item 1)

	APRB O1/O2/O3	APLB O1/O2/O3	ROB O1/O2/O3	LOB O1/O2/O3	RLB O1/O2/O3	LLB O1/O2/O3
APRB	-/-	42/60/53 ^a	6/1/2	-/-	-/6/-	-/-/3
APLB	49/65/29 ^a	-/-	3/-/4	-/-	3/-/5	-/-
ROB	25/9/-	13/-/-	-/-	-/25/26	-/13/11	-/-
LOB	-/-	-/-	-/-/15	-/-	-/13/-	-/13/23
RLB	-/-/5	25/-/3	-/13/7	-/-	-/-	-/8/34
LLB	-/-	-/-/6	-/-	-/-	-/6/11	-/-

The most probable transitions are between the breast regions in the AP view (APRB, APLB)

Probabilities either from or to outside of the breast regions were not considered in the analyses because we were mainly interested in transitions among breast regions.

A dash (-) was used to decrease the number of zero's and to allow numbers with importance to stand out

^a The two highest transitional activity were from region APLB to region APRB for the all three surgeons, Observer 1 (49%), Observer 2 (65%), and Observer 3 (29%), and from region APRB to region APLB for the all three surgeons, Observer 1 (42%), Observer 2 (60%), and Observer 3 (53%)

Table 4 Joint probabilities across six regions of interest recorded while three observers examined eight patients who underwent breast reconstruction case to rate the initial impression of overall appearance of the breasts (Item 1)

	APRB O1/O2/O3	APLB O1/O2/O3	ROB O1/O2/O3	LOB O1/O2/O3	RLB O1/O2/O3	LLB O1/O2/O3
APRB	-/-	10/14/18 ^a	1/2/-	-/-	-1/-	-/-1
APLB	19/19/6 ^a	-/-	1/-/3	-/-	1/-/4	-/-
ROB	2/-	10/-	-/-	-3/15	-1/-	-/-
LOB	-/-	-/-	-/1	-/-	-2/-	-/8
RLB	-/-	-/-	-2/13	-/-	-/-	-1/3
LLB	-/-	-/13	-/-	-/-	-1/1	-/-

The most probable transitions are between the breast regions in the AP view (APRB, APLB)

Probabilities either from or to outside of the breast regions were not considered in the analyses because we were mainly interested in transitions among breast regions

A dash (-) was used to decrease the number of zero's and to allow numbers with importance to stand out

^a The two highest transitional activity were from region APLB to region APRB for the all three surgeons, Observer 1 (19%), Observer 2 (19%), and Observer 3 (6%), and from region APRB to region APLB for the all three surgeons, Observer 1 (10%), Observer 2 (14%), and Observer 3 (18%)

$P(\text{APLB}, \text{APRB}) = 10\%$. This result is consistent with the conclusions of the dwell-time analysis, dwell transition frequency, and conditional probabilities that surgeons primarily assess morphology based on the breast regions in the AP views and move on to the oblique or lateral views to complete the assessment. Joint probabilities other than for the AP views were very small ($P \leq 3\%$) and varied depending on the rating item, observer, and view (Table 4).

Discussion and Conclusion

Tracking people's eye movements can help one understand visual information processing and the factors that may influence an observer's focus. Measuring aspects of eye movements, such as fixations, suggests the level of neural activity devoted to an observed event or object. While this method is indirect, it provides a non-invasive means of detecting some cognitive activity. The goal of this study was to investigate the potential of using eye-tracking technology to understand how plastic surgeons assess breast morphology. By tracking a surgeon's gaze while he or she evaluated clinical photographs of breast morphology, we were able to document several patterns in their assessments and showed that eye tracking provides a useful method for conducting an observational study of surgical outcomes. For example, when the dwell time was compared across the views, the results showed that all three surgeons spent the most time on the AP views. Similarly, when transition activity between regions was analyzed, there were many transitions observed between the breast regions in the AP view, while few transitions were observed between other views. These results are consistent with our expect-

ations that the surgeons primarily use the AP view to assess surgical outcomes. One possible explanation for why so much time was spent on the AP views and there are numerous transitions between the breast regions in the AP view is that the surgeon is searching for asymmetry between the left and right breasts [21, 22]

Visibility, the extent to which the surgical scar detracts from the esthetic outcome, is the relevant property for assessing scarring following breast reconstructive surgery [7, 8]. Itti and Koch [23] reviewed several studies of visual attention. In recent models, visual attention is broken down into localized analysis problems. Neurons at the earliest stages of processing are tuned to visual attributes including contrast, intensity, color contrast, orientation, direction, spatial frequency, and velocity of motion. While recognition and identification of visual stimuli may or may not take place at the same time as visual attention (though in different parts of the brain), we assume that the surgeons in our experiments could quickly perceive abnormalities and focus their attention on those features instead of other features that would draw the attention of a naïve observer (contrast, intensity, etc.). Surgical scars around reconstructed breasts and donor site have high contrast or intensity. We speculate that the surgeons did not pay attention to the donor site scarring which present high contrast or intensity, whereas a naïve observer might be distracted by donor site scarring and that further research is needed on comparing the perceptions of surgeons and naïve observers.

We acknowledge certain limitations of this study. Even though care was taken to control the experimental conditions, there was some inevitable variability due to technical difficulties and the physical condition of the

subjects. The surgeons who participated in this study suffered from extreme fatigue that resulted in longer calibration time which may have impacted variability of the eye data. Also, as this was a pilot study, we employed a limited number of observers and cases.

In this study, we demonstrate that eye-movement tracking is an important, quantitative technique that can afford useful advantages for the in-depth analysis of the clinical assessment of breast morphology. As a future study, it would be interesting to compare different groups of observers, particularly breast cancer survivors and plastic surgeons. It would also be interesting to investigate the relationship between image perception as measured by eye tracking and body image as measured by survey instruments. Eye tracking would be a valuable addition to current methods for evaluating surgical outcomes, such as observer rating scales, and analysis of 2D and 3D images [3]. Growth in the use of the eye-tracking technology in clinical assessment will continue as the technology becomes increasingly more affordable, less invasive, and easier to use.

References

1. Breast Cancer Facts & Figures: Atlanta: American Cancer Society, 2009–2010
2. 2000/2006/2007 National Plastic Surgery Statistics, Arlington Heights, IL: American Society of Plastic Surgeons, 2008
3. Kim MS, Sbalchiero JC, Reece GP, Miller MJ, Beahm EK, Markey MK: Assessment of breast aesthetics. *Plast Reconstr Surg* 121:186e–194e, 2008
4. Smith DJJ, Palin WEJ, Katch V, Bennett JE: Surgical treatment of congenital breast asymmetry. *Ann Plast Surg* 17:92–101, 1986
5. Westreich M: Anthropomorphic breast measurement: Protocol and results in 50 women with aesthetically perfect breasts and clinical application. *Plast Reconstr Surg* 100:468–479, 1997
6. Kim MS, Reece GP, Beahm EK, Miller MJ, Neely Atkinson E, Markey MK: Objective assessment of aesthetic outcomes of breast cancer treatment: measuring ptosis from clinical photographs. *Comput Biol Med* 37:49–59, 2007
7. 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. *J Eval Clin Pract* 15:20–31, 2009
8. Kim MS, Rodney WN, Reece GP, Beahm EK, Crosby MA, Markey MK (2010) Quantifying the aesthetic outcomes of breast cancer treatment: assessment of surgical scars from clinical photographs. *J Eval Clin Pract* (in press)
9. Pezner RD, et al: Breast retraction assessment: an objective evaluation of cosmetic results of patients treated conservatively for breast cancer. *Int J Radiat Oncol Biol Phys* 11:575–578, 1985
10. Galdino GM, Nahabedian M, Chiaramonte M, Geng JZ, Klatsky S, Manson P: Clinical applications of three-dimensional photography in breast surgery. *Plast Reconstr Surg* 110:58–70, 2002
11. Losken A, Seify H, Denson DD, Paredes AAJ, Carlson GW: Validating three-dimensional imaging of the breast. *Ann Plast Surg* 54:471–476, 2005
12. Vrieling C, et al: Validation of the methods of cosmetic assessment after breast-conserving therapy in the EORTC “boost versus no boost” trial. *Int J Radiat Oncol Biol Phys* 45:667–676, 1999
13. Mello-Thoms C: How does the perception of a lesion influence visual search strategy in mammogram reading? *Acad Radiol* 13:275–288, 2006
14. Krupinski EA: Visual scanning patterns of radiologists searching mammograms. *Acad Radiol* 3:137–144, 1996
15. Krupinski EA: Visual search of mammographic images: influence of lesion subtlety. *Acad Radiol* 12:965–969, 2005
16. Mello-Thoms C, Dunn S, Nodine CF, Kundel HL, Weinstein SP: The perception of breast cancer: What differentiates missed from reported cancers in mammography? *Acad Radiol* 9:1004–1012, 2002
17. Nguyen HT, Isaacowitz DM, Rubin PAD: Age- and fatigue-related markers of human faces: an eye-tracking study. *Ophthalmology* 116:355–360, 2009
18. Dixon B, Grimshaw G, Linklater W, Dixon A: Eye-tracking of men's preferences for waist-to-hip ratio and breast size of women. *Arch Sex Behav.*, 2009. doi:10.1007/s10508-009-9523-5
19. Ishii L, Carey J, Byrne P, Zee DS, Ishii M: Measuring attentional bias to peripheral facial deformities. *Laryngoscope* 119:459–465, 2009
20. Yavuzer R, Smirnes S, Jackson IT: Guidelines for standard photography in plastic surgery. *Ann Plast Surg* 46:293–300, 2001
21. Herbert AM, et al: Searching for symmetry: eye movements during a difficult symmetry detection task. *J Vis* 6:24–24, 2006
22. Mappus IV, RL, Ferguson RW, Czechowski K, Corballis PM: Spotting Differences: How Qualitative Asymmetries Influence Visual Search. Stresa, Italy, 2005
23. Itti L, Koch C: Computational modelling of visual attention. *Nat Rev Neurosci* 2:194–203, 2001