A system for automatic monitoring of surgical instruments and dynamic, non-rigid surface deformations in breast cancer surgery

Winona L. Richey1, Ma Luo1, Sarah E. Goodale1, Logan W. Clements1, Ingrid M. Meszoely2, Michael I. Miga1,3,4,5

1 Vanderbilt University, Department of Biomedical Engineering, Nashville, TN USA
2 Vanderbilt University Medical Center, Division of Surgical Oncology, Nashville, TN USA
3 Vanderbilt University Department of Radiology and Radiological Sciences, Nashville, TN USA
4 Vanderbilt Institute for Surgery and Engineering, Nashville, TN USA
5 Vanderbilt University Medical Center, Department of Neurological Surgery, Nashville, TN USA

ABSTRACT

When negative tumor margins are achieved at the time of resection, breast conserving therapy (lumpectomy followed with radiation therapy) offers patients improved cosmetic outcomes and quality of life with equivalent survival outcomes to mastectomy. However, high reoperation rates ranging 10-59% continue to challenge adoption and suggest that improved intraoperative tumor localization is a pressing need. We propose to couple an optical tracker and stereo camera system for automated monitoring of surgical instruments and non-rigid breast surface deformations. A bracket was designed to rigidly pair an optical tracker with a stereo camera, optimizing overlap volume. Utilizing both devices allowed for precise instrument tracking of multiple objects with reliable, workflow friendly tracking of dynamic breast movements. Computer vision techniques were employed to automatically track fiducials, requiring one-time initialization with bounding boxes in stereo camera images. Point based rigid registration was performed between fiducial locations triangulated from stereo camera images and fiducial locations recorded with an optically tracked stylus. We measured fiducial registration error (FRE) and target registration error (TRE) with two different stereo devices using a phantom breast with five fiducials. Average FREs of 2.7 ± 0.4 mm and 2.4 ± 0.6 mm with each stereo-camera device demonstrate considerable promise for this approach in monitoring the surgical field. Automated tracking was shown to reduce error when compared to manually selected fiducial locations in stereo camera image-based localization. The proposed instrumentation framework demonstrated potential for the continuous measurement of surgical instruments in relation to the dynamic deformations of a breast during lumpectomy.

Keywords: breast cancer, lumpectomy, surgical guidance, registration, computer vision, breast conservation therapy, image guided surgery

1. INTRODUCTION

Breast cancer is the most common cancer among women, second only to lung cancer in mortality1. With advancements in detection, an increasing number of cases are being diagnosed at early stages. These early stage diagnoses result in more operations on smaller tumors, and a rising number of candidates eligible for breast conserving therapy, which includes surgical resection and radiation therapy2,3. For patients who are eligible, breast conserving therapy with negative margins has been shown to be as effective as mastectomy, the complete removal of breast tissue2,5. With respect to the surgical goals of lumpectomy, success is determined by the presence of pathologically negative margins occurring when there are no cancer cells at the edges of the resected tissue. Breast conserving therapy also offers patients favorable cosmetic outcomes and equivalent survival outcomes in the case of negative margins. However, the re-operation rates for lumpectomy are currently estimated at 10-59%9-14 and returning to the operating room negatively impacts patient satisfaction, cosmetic outcomes, and reintroduces risk of surgical complications. It is commonly accepted that improved localization reduces reoperation rates, as localization procedures are often measured by their effect on reoperation rates15,16. Therefore, the high re-resection rates of BCT indicate the need for improved intraoperative tumor localization.

Breast cancer tumor localization proves difficult for two main reasons: the huge differences between diagnostic and intraoperative presentations and the limited extent of geometric and spatial cues during surgery to guide resection. Magnetic resonance is generally considered the most sensitive imaging modality for breast cancer17,18. While not a
standard diagnostic, diagnostic magnetic resonance images are often taken and in the prone position, with the breast pendant. Additionally, a patient’s arms are either at their sides or extended above their head in order to enter the imaging bore. For surgery, the patient is in a supine position with the arm extended laterally. Tumor displacements between prone and supine breasts are reported between 18-60 mm\(^1\text{9,21}\). In addition, mammographic examination is another common breast cancer image screening method. In this case, a standing patient will place the breast between two compressive plates. This process also introduces large tissue displacements and shape change from its preoperative counterpart due to tissue compression.

With respect to intraoperative localization methods, the most common form of surgical guidance is the use of guide wires. A guide wire is placed in the center of the tumor by a radiologist during diagnostic imaging, and the patient waits with a wire protruding from the breast until surgery. There is discomfort during preoperative wait time and guide wires can shift during that period. In addition, the implanted wire path to the center of the tumor is constrained by the diagnostic presentation and may not be the best surgical path. Guide wires also only provide localization of tumor center and do not provide information about tumor margins. Intraoperative ultrasound visualizes tumor margins in the operating room but takes time, can require a radiologist, and most importantly, approximately half of breast tumors are not echogenic\(^22\). Intraoperative supine MRI offers clear location information, but is costly, cumbersome, time inefficient, and only offers initial tumor location, i.e. the location and deformations are not tracked or updated throughout surgery. Alternatively, imprint cytology is the practice of pressing a tumor to a glass slide following removal; cancerous cells stick to the slide and non-cancerous cells come off clean. This method is good for evaluating tumor margins, but interferes with workflow and, most importantly, is not always effective.

Non-rigid, dynamic deformation of breast tissue during surgery makes tumor localization difficult. Further, these deformations are challenging to compensate for, especially with the large variance in material properties across patients due to age, genetics and hormonal changes\(^23,24\). Existing modeling and prediction is often driven by surface deformations. The relationship between surface and subsurface measurements has been previously explored in a trial done with 12 patient volunteers. In that study, surface and subsurface feature deformations were quantified and compared, showing that surface deformations are predictive of subsurface deformations\(^25\). Initial registrations offer surgeons improved localization at the start of surgery, but if registrations are not updated, tumor margins remain difficult to ascertain and the guidance becomes less valuable upon incision. Continuous monitoring of the breast surface is needed for registration and modeling updates to maintain accuracy after the initial data collection and throughout the procedure.

In this work, we couple optical instrument tracking with a stereo camera’s visual surface tracking. Conventional tracking can be used for surgical instruments and the deforming breast surface can be more readily monitored with a stereo camera. Stereo camera point localization has been previously shown to be reliable\(^26\), and we extend this work to hypothesize that a computer vision tracking algorithm can be applied to automatically and continuously monitor the breast surface. The surface information provided throughout surgery would allow for ongoing non-rigid registration with a computational model. Registration between preoperative and intraoperative data would give the surgeon access to continuously updating spatial information. This work is an important and exciting first step towards automatic and continuous surgical guidance in real-time for breast cancer surgery.

2. METHODS

A bracket was designed to maximize the overlap between the volume of a Polaris Vicra optical tracker (Northern Digital, Waterloo, ON, Canada) and the volumetric field of view of the Bumblebee XB3 (FLIR, formerly Point Grey Research, Richmond, BC, Canada). The stereo cameras offer point reconstruction capabilities from triangulation of stereo camera images and the automated tracking algorithms of OpenTLD software\(^27\). The stereo camera and optical tracker were rigidly paired and registration was performed on a phantom breast (Breast Probe, SIMULAB Corporation, Seattle, WA) with five fiducials. To evaluate the performance of this system with the Bumblebee XB3 stereo camera, three trials were conducted, each consisting of a deformed and undeformed state. To further evaluate the capabilities of optical tracking paired with stereo camera automated tracking, the Bumblebee XB3 was replaced with two Grasshopper cameras (FLIR, formerly Point Grey Research, Richmond, BC, Canada). The grasshoppers, shown attached to the bracket in Figure 1A and 1B, offer improved resolution with 1600 x 1200 pixels, compared to the 1280 x 960 pixels of Bumblebee XB3 images. Three trials were conducted with the two Grasshopper cameras, each trial consisted of five deformation states.
2.1 Bracket Design

The Polaris Vicra was angled towards the Bumblebee XB3’s field of view. The angle that maximizes volume overlap was calculated to be 3 degrees, and a bracket was designed and machined to rigidly couple the two devices. With the y-axis as the range, or depth, of view, and x-axis as the plane perpendicular to the cameras, the Z dimension was parallel to the cameras and was optimized by centering the two devices (i.e. the center of the Polaris Vicra is aligned with the center of the set of stereo cameras). The volume was therefore optimized as an area in the XY plane with most overlap. The volume overlap in the XY plane is shown in Figure 1C with the Grasshopper cameras.

2.2 Stereo Camera Points

Bumblebee XB3 stereo camera images were collected through a modified version of Point Grey software, displaying output from all three cameras of the Bumblebee. Example output of the Bumblebee can be seen in Figure 2, with superimposed arrows to describe locations where forces were applied. Stereo camera images from two grasshopper cameras were collected with a custom-built image capture software. The coupled devices were centered over the phantom and each Grasshopper was tilted until the right camera’s field of view was about 1 cm to the right of the left field of view, and both views were approximately centered on the deformation apparatus. Calibration for both sets of cameras was performed with stereo camera calibration algorithms from Matlab’s Computer Vision System Toolbox. Sixteen images were used for Bumblebee XB3 calibration resulting in a mean reprojection error of 0.29 pixels, and twelve images were used for Grasshopper images resulting in a mean reprojection error of 0.46 pixels.
tracking algorithm. Three methods were used to determine the fiducial centers within each bounding box: two automatic methods, and manual selection. The automatic fiducial localization methods were bounding box averaging and region growth. For the bounding box averaging method, the top left and bottom right corners of each bounding box were averaged to calculate the center, and these centers were used in the triangulation process. For the region growth technique, a fast marching algorithm was used to produce a fiducial region segmented from the image. This segmentation process uses gray scale intensity difference weights calculated from the intensity values at the seed locations via Matlab’s Image Processing toolbox (utilizing graydiffweight with a roll-off factor of 0.2). Because bright fiducial markers were used, this method was automated by using the brightest pixel as a seed point for the algorithm. The centroid of the fiducial of interest can then be computed by determining the centroid of this region. To evaluate the accuracy of our automated methods, manually selected fiducial centroids were also collected at each point for each deformation.

2.3 Optical Tracker Points

The Polaris Vicra and an optical tracking stylus shown in Figure 3A and 3D were used to record optical tracker fiducial locations. The optical tracking system detects the location of retro-reflective markers by emitting infrared light and detecting the reflections from the markers on a specific instrument. Each instrument has a specific geometry of reflective markers to allow for simultaneous tracking of multiple instruments. At each fiducial point, the optically tracked fiducial location was computed by averaging 8 rapid acquisitions via the optical tracking system.

2.4 Experimental Setup

Five yellow fiducials with 6 mm diameters were used as shown in Figure 3. With the Bumblebee XB3, three tests were performed, each with two deformation states. Tests 1 and 2 induced deformation to produce the second states, while Test 3 was a relaxation from a deformed state to a second, undeformed state. With the two Grasshopper cameras, three tests were performed, each with five deformation states. All three tests consisted of an undeformed state, and then a deformation at each of the arrows shown in Figure 2. The experimental setup is shown in Figure 3.

![Figure 3: Coupled device prototype (A) coupled stereo camera and optical tracker attached to a surgical arm (B) monitor displaying the location of the instrument tip within the optical tracker’s volume (C) monitor displaying stereo camera images for surface tracking (D) tracked surgical instrument (E) breast phantom (F) fiducial markers](image)

2.5 Registration Technique

All three sets of triangulated fiducial locations were registered to optically tracked points via a point based rigid registration. Stereo camera triangulated points from the initial left and right frames were registered to the optically tracked NDI points from the first deformation state. A point set results from each method of fiducial localization in stereo images: bounding box averaging, region growth, and manual selection. The transformation that results from registering each of these three point sets to optically tracked NDI points, each define the relationship between the stereo
camera space and the optical tracker instrumentation space. However, given that each method has disparate sources of error, slight differences in these transformations are to be expected. For this reason, these point sets were, separately, each registered to the optically tracked NDI points. The registration computed with the 5 fiducial points in the first frame, served as the calibration transformation between Vicra and stereo-pair and was applied to all successive frames in the trial. This registration defines a transformation $T$ consisting of a rotation and transformation to relate the two sets of points: $x$, the stereo camera point set, and $y$, the optically tracked NDI point set. $T(x)$ represents the transformed stereo camera points, i.e. the stereo camera points represented in the optically tracked instrumentation space. Fidelity was reported as a fiducial registration error (FRE) a measure of overall landmark misalignment, and is shown below, with $N$ representing the number of fiducials:

$$FRE = \frac{1}{N} \sum_{i=1}^{N} (T(x_i) - y_i)^2$$  \hspace{1cm} (1)

To evaluate target localization, target registration error (TRE), a measure of target misalignment, was computed using a leave-one-out method. Coordinates from four fiducials across all deformation states were used to register, and the calculated transformation, $T_i$, was applied to the fifth fiducial with $i$ ranging from 1 to $N$, the number of fiducials. This means that the TRE for a target point is the magnitude of the vector between the transformed stereo camera point, $T_i(p)$, and the optically tracked NDI point, $q$. The equation for TRE is shown below:

$$TRE = ||T_i(p) - q||$$  \hspace{1cm} (2)

### 2.6 Optical Tracker Point Localization Evaluation

In these experiments we use optically tracked NDI points to evaluate the accuracy of our computer vision tracking system, considering the optically tracked NDI points to be the true, or “gold standard”, point locations. Therefore, it is useful to evaluate the same five points localized multiple times with the optically tracked NDI stylus to provide some sense of the fidelity of the “gold standard”. Five fiducial locations were collected following the same procedure as outlined in Section 2.3. These same, undeformed points were recollected four additional times for a total of five point sets each with five fiducials. All combinations of point sets were registered systematically to establish an estimate of FRE. In addition, a leave-one-out test was performed to establish an estimate of TRE.

### 3. RESULTS

Comparing NDI tracked points and triangulated stereo camera points can be reliably registered with an FRE of $2.7 \pm 0.4$ mm for the Grasshopper Cameras and $2.4 \pm 0.6$ mm for the Bumblebee, using the region growth method to calculate fiducial centroids on stereo camera images. FREs for each camera were averaged across all deformations and are presented in Figure 4A. TRE averages were of $2.8 \pm 1.1$ mm for the Grasshopper Cameras and $2.7 \pm 1.4$ mm for the Bumblebee using the region growth method to calculate fiducial centroids on stereo camera images.

Figure 4: Errors reported averaged across all deformation states, in millimeters with error bars reflecting standard deviation; Fiducial Registration Error (Left) and Target Registration error, (Right).
Our “gold standard” validation of the optically tracked NDI point accuracy yielded an average FRE of 0.8 ± 0.3 mm, with a leave-one out TRE of 1.0 ± 0.3 mm.

The first state contains the stereo-camera points that were localized from user-defined bounding boxes for the automated fiducial localization methods. To evaluate the accuracy of our tracking method alone, average FREs and TREs were also computed without the point sets from the first state. Registration was still performed using the first state. For the Bumblebee camera, FRE averaged 2.7 ± 0.9 mm and TRE averaged 2.7 ± 1.0 mm. For the Grasshopper cameras, FRE averaged 2.9 ± 0.9 mm and TRE averaged 2.8 ± 1.1 mm.

A two-sample t-test with unequal variances was performed to compare the two stereo camera systems. The camera systems were compared using the results from the automatic region growth fiducial localization. The Bumblebee FRE was found to be significantly less than that of the Grasshoppers (p < 0.01), while TRE was not found to be significantly different between the two camera systems.

4. DISCUSSION

As was somewhat expected, the region growth algorithm outperformed the bounding box algorithm. As a more precise alternative, it suffers less subjective error with user-defined bounding box. Even for bounding boxes initialized with the fiducial center exactly in the center, as the tracking algorithm progresses there is the potential for the bounding box to shift with subtle shape and location changes that could create error in the bounding box method. More surprisingly perhaps, is that the region growth method outperforms the manual selection process. Upon further investigation, the difference between manually selected points and automatically selected points is generally minimal in the plane of the camera image. However, the manually selected method suffered higher error in the triangulated depth than the automatic region growth method. Even with great care taken to select the centroids of fiducials, we hypothesize this error stems from discrepancies in correspondences. That is, the same point must be selected in the left and right images for triangulation to be correct; automatic methods lend themselves to more reliably selecting the same pixel at the center, provided that the bounding boxes are reasonably similar. A two-sample t-test with unequal variances showed the Bumblebee XB3 had statistically significantly lower average error (2.4 mm average FRE) than the two Grasshoppers (2.7 mm average FRE) for the region growth method. However, overall, the Grasshoppers produced more reliable results with lower standard deviations, and more consistent results with less average error across all three methods. Because the grasshopper cameras can be angled toward each other to create more overlap in their field of view, we have found in results not reported here that the point cloud reconstructions have been more satisfying, which will prove valuable as we move forward. For future work, we will likely adopt the Grasshopper configuration.

While care was taken to minimize deformation from stylus contact with the phantom breast, it should be noted that the optical tracker stylus is a contact method and potentially introduces error due to deformations. A stylus was used to collect points as opposed to fixing near infrared-visible markers to the breast phantom which could have improved localization and consequently calibration. The registration of the repeated optically tracked points (Section 2.3) allows us to determine the baseline for accuracy of our validation method. Our ground truth has an average error of 0.8 ± 0.3 mm.

It is well known that target error is minimized when targets are surrounded by localized fiducials. Because registration accuracy decreases as the target moves further from the centroid of the fiducials, the leave one-out method for TRE presented here has an inherent error, i.e. the target in our experiments was always completely distant from the centroid of the fiducials. To evaluate the target error without this bias, an additional data point was collected at the nipple of our mock breast phantom setup. With the nipple as a target, the bounding box method was used to locate the center of the fiducial resulting in an average TRE of 2.0 mm (note region growth could not be used due to the phantom coloring in this case). When using the “gold standard” registration, TRE of the nipple was 0.6 mm.

5. CONCLUSION

The proposed approach is an important step towards real-time, automated monitoring of dynamic intraoperative breast deformation. The tracking method is robust to occlusions of fiducials and multiple fiducials of similar size and shape. Visual fiducials adhered to the breast and tracked with computer vision are extremely practical for intraoperative breast tracking, allowing for easy demarcation of fiducial points and reliable tracking with computer vision. Using passive optical tracking, multiple tools can be simultaneously monitored with minimized surface tracking interference. Additionally, because the instrumentation tracking volume extends outside the stereo camera’s field of view (in the Z
direction, or left-right axis of images), the tip of surgical instruments can continue to be tracked even if their reflective markers fall outside the view of the stereo camera.

Computer vision applications in image guided breast surgery have not been well investigated. The system presented is a realization of a novel pairing of computer vision technologies and instrumentation tracking, each focused at different aspects of the surgical problem. By monitoring instrumentation and surgical field separately, each modality takes advantage of the different tracking strengths. Conventional optical tracking gives precise instrument locations of multiple objects. The stereo camera allows automation of the surface tracking process for real time intraoperative image guidance in dynamic, non-rigid breast deformation. The functionality to track the instrument tip with precision is not new, but its incorporation into an automatically deforming image guidance system is novel.

This system is workflow friendly; it does not require expensive imaging equipment and personnel to operate (as in iMR), sits away from the clinician’s surgical field, and requires less than one minute of initialization and then no further interaction. This technique could be easily managed by a surgeon or surgical assistant, taking less than a minute for the one-time initialization of surface markers. The results here show that points can be tracked and localized reliably with an average FRE below 3.0 mm. Both vision approaches (Grasshopper or Bumblebee XB3) are acceptable and the region growth method is the best for fiducial localization in stereo camera images. This work shows promise toward computer vision surface tracking in breast cancer lumpectomy and breast cancer therapy. This system shows the potential to provide the surgeon with tumor localization even as the tissue moves due to gravity and surgical instruments. If this system is fully realized and validated, surgeons could be presented with real-time changes to surface and subsurface breast features in relation to their surgical instruments.

ACKNOWLEDGEMENTS

This work was supported in part by the NIH-NCI grant R01CA162477 and NIH-NIBIB grant R21EB022380.

REFERENCES