Automatic fiducial localization in ultrasound images for a thermal ablation validation platform

Laura Bartha[a], Andras Lasso[a] , Thomas Kuiran Chen[a], Gabor Fichtinger[a]
[a]School of Computing, Queen's University at Kingston, Canada

ABSTRACT

PURPOSE: Development of ultrasound-based tumor ablation monitoring systems requires extensive validation. Validation is based on the comparison of ablated regions, computed from ultrasound images, to the ground truth region observed on histopathology images. Registration of ultrasound and histopathology images can be efficiently implemented by localizing fiducial lines embedded in the test phantom. Manual fiducial localization is time consuming and may be inaccurate. Current automatic localization algorithms were designed for use on images containing easily detectable fiducials in clear water, while the images produced by the ablation monitoring platform contain fiducials and ablated tissue embedded in tissue-mimicking gel. Our goal was to develop an automatic fiducial localization algorithm for the ablation monitoring platform. METHOD: A previously existing algorithm for detecting fishing line in water for ultrasound probe calibration, created by Chen *et al.*, was tested on ultrasound images of an ablation phantom. Fiducial and line point detection parameters were determined by running the algorithm multiple times with different parameter sets and searching for the set that results in the best detection success rate. The fiducial intensity scoring method was modified to use intensities from an unaltered image; this greatly reduced the number of incorrectly identified fiducials. Line finding was modified to suit the ablation phantom geometry. RESULTS: The new algorithm was tested by comparing the automatic localization results to manually identified fiducial positions. Using the optimized parameters, it was found to have a 94.1 % success rate on the tested images. Fiducial localization error was defined as the difference between the manually segmented positions and the positions found by the algorithm. Fiducial localization error was -0.04±0.18mm along the x-axis, and -0.09±0.14mm along the y-axis. CONCLUSION: We have developed an automatic algorithm that detects line fiducials at a high success rate in complex phantoms containing a tissue sample embedded in tissue-mimicking gel.

Keywords: ultrasound, pathology, tumor ablation, validation platform, automatic segmentation

1. INTRODUCTION

Thermal ablation with ultrasound guidance is an emerging focal tumor management method. However, dose must be monitored to avoid damage of healthy tissue while still preventing recurrence of the cancer. Currently, there is a lack of an affordable real-time dose monitoring solution[2]. Ultrasound is a widely available, inexpensive imaging modality that could be potentially used for monitoring dose[6], but no established ultrasound-based monitoring method exists yet. To develop, optimize, and validate such a method, extensive testing is essential. Previously, a prototype of a validation platform for ablation monitoring was constructed and presented[2].While this provided proof of the viability of a thermal ablation monitoring platform, the described system has a few limitations that make it impractical for use in a realistic, large-scale validation study. A critical problem is the lengthy manual work required to process the acquired ultrasound images. Processing the data set of one ablation takes about one hour and most of that time is spent on manually picking the fiducial marker positions on the ultrasound images. We propose an algorithm that automates this time consuming process. Previous algorithms exist[3,4], however they are designed to be used only on simple phantoms[5] where the fiducial markers are clearly visible and there are no other image features that could be easily mistaken as fiducials. Our algorithm operates on a phantom that contains a tissue sample embedded in tissue-mimicking gel, an environment in which earlier algorithms produce too many false positives to retain their functionality. Together with other improvements presented by Peikari *et al*[1] our work results in a next generation validation test bed that is efficiently usable for thorough testing of real-time ablation monitoring methods.

Medical Imaging 2011: Visualization, Image-Guided Procedures, and Modeling,
edited by Kenneth H. Wong, David R. Holmes III, Proc. of SPIE Vol. 7964, 796421

2. METHODS

The current validation test bed is comprised of ultrasound and histopathology images taken of a phantom. The phantom contains tissue that undergoes US monitored ablation. The phantom is subsequently sliced and photographed with a flatbed scanner. Both ultrasound and histopathology images contain the same fiducials. Using the graphical user interface of the software the fiducials are localized manually in the ultrasound and histopathology images, and tissue contours are marked in the histopathology images. The software then reconstructs the three-dimensional tissue surface from both ultrasound and histopathology images and compares the two[2]. The work being presented automates the fiducial localization in the ultrasound images. There is a greater need for the automation of fiducial localization in ultrasound images than in histopathology images because this test bed creates several hundred ultrasound images and just 10-12 histopathology images.

2.1 Phantom

The phantom is composed of a box with 3 N-wire[5] fiducials along the bottom and two opposing sides. The box is filled with tissue-mimicking gel containing chicken breast (Fig. 1). It is created and imaged in the manner proposed by Pompeu-Robinson *et al.*[2]. The ultrasound probe is placed on top of a phantom with N fiducials arranged as seen in Fig. 1 and imaged while moving the probe either from the front of the phantom to the back, or from the back to the front. With this arrangement, there are seven fiducials along the edge of the phantom.

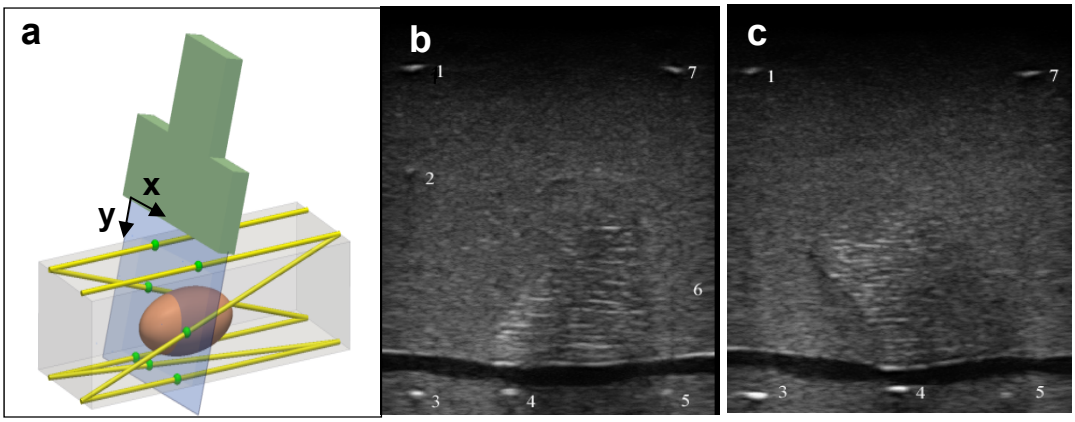

Figure 1. Ultrasound imaging setup (a) and sample images (b, c). The left ultrasound image (b) is typical of the data we used. The right image (c) is an example of an infrequent case where the side fiducials are visible. The leftmost image is a more typical case in which the side fiducials are not visible.

In most of the ultrasound images used, the two middle fiducials along each side (fiducial 2 and 6) are hardly visible because they are at an angle to the ultrasound probe. Changing the fiducial geometry so that these diagonal lines become more horizontal would make the lines more visible, but in this case they would not contribute much to the slice orientation and position detection. Using simulations, Peikari *et al.*[1] showed that slice position and orientation can be recovered even if these unreliably detectable lines are not present, so we chose not to localize them.

A first set of phantoms were created by using exactly the same recipe that Pompeu-Robinson *et al.* proposed (combining 80 parts water, 2 parts agar and 4 parts gelatine).A second set of phantoms was created for further analysis from 100 parts water, 3.5 parts agar, 8 parts glycerol. We found that without gelatine the gel is more rigid at body temperature (the starting temperature for ablation). Glycerol was needed to account for the differences in sound speed between the chicken tissue and the gel. In previous phantoms created for gel testing purposes, it was found that fiducials placed behind the chicken tissue appeared to be closer to the chicken. When glycerol was added, the fiducial was not noticeably displaced.

2.2 Fiducial Detection

The algorithm by Chen *et al.*[3] first performs a series of morphological opening operations with a 1x18 pixel bar structural element at four orientations and with a circle structural element with a radius of 2 pixels (on images with approximately 0.1mm/pixel resolution). This is done to remove speckle, reflections, and other image artifacts. Next, clustering is performed on this processed image by identifying each set of connected pixels that have intensity above the clustering threshold (*a*) value. The center of each pixel cluster is determined by computing the center of gravity of all the above-threshold pixels on the preprocessed image.

Chen *et al.* used a clustering threshold that is 10% of the maximum pixel intensity in the image. This value was determined for water phantoms, while the basis of our phantom is tissue-mimicking gel containing ablated tissue. It was found that with the same threshold value we could not detect all of the fiducials on any of the ablation phantom images. The optimal threshold was found by running the algorithm with varying threshold parameters on a large number of images.

Chen *et al.* determined an intensity score for each fiducial by summing its pixel values obtained after the opening operations[3]. Fiducials with higher scores were given priority in the following processing steps. It was found that this scoring method did not work well for our phantom, because the gel and the tissue in the phantom produced many artifacts that received higher score than real fiducials. We found that computing the score by summing above-threshold pixel intensities of the unaltered image provides a better metric, resulting mostly in higher scores for real fiducials than for artifacts.

2.3 Line Detection

Chen's algorithm first detects all collinear sets of three fiducials from the set of fiducial candidates found using the fiducial detection algorithm described above. This is implemented by an exhaustive search among all the fiducials, while applying constraints[3]. The constraints include the difference between normal directions of two-point line segments *(b)* and maximum distance of any third point from the line defined by the two other points (*c*). These limits were increased to account for fiducial line deformation in the phantom. This deformation arose during the creation of the phantom, possibly due to the insertion of the tissue sample. The minimum length of the 3-point line segment (*d*) was also increased to correspond with ablation phantom geometry.

The next step in the line detection process finds the detected lines that best match the known fiducial geometry. Chen's phantom contains two parallel lines, each consisting of 3 points[3,5]. Our ablation phantom geometry is different from Chen's geometry, and therefore this step required complete reworking. Candidate fiducials needed to be orthogonal to the base line to be considered. Those candidates that formed a line with one of the endpoints of the base line and were roughly perpendicular to the base line, that is, within the allotted angle tolerance (*e*) were stored, provided that they were within the minimum and maximum length specifications (*f, g*) as well. The stored candidate left and right side lines were then paired. Vectors connecting the candidate points to the corresponding base line endpoints were calculated. Pairs were accepted if the magnitudes of these vectors were within a specific tolerance value (*h*). The tolerance values were determined by estimating typical values (based on the phantom geometry and expected probe position) and verified by running the segmentation on numerous test images.

2.4 Validation

The algorithm performance was validated by comparing the detected fiducial point positions to ground truth fiducial positions. Ground fiducial positions were determined through manual segmentation of the ultrasound images in the open-source 3D Slicer application[7]. We marked the center point of the bright fiducial spot and used it as ground truth for the validation.

3. RESULTS

To find the optimal threshold, we recorded the percentage of frames in which all five fiducials were correctly identified for different thresholds ranging from 1 to 50% of the maximum image itensity. The results on Fig. 2.a show that any value between 1% and 5% results in 97.9% of fiducials detected as fiducial candidates, while using a threshold of 6% or

more, the percentage of fiducials found declines rapidly. We also recorded the total fiducial candidates found with each threshold (Fig. 2.b). As it is expected, the number of candidates quickly increases as we lower the threshold. We used a threshold of 5% for all further processing, as this is the highest possible threshold that still maintains the highest percentage of found fiducials but keeps the number of false positive fiducial candidates at a minimum. The dataset from the first phantom was used to generate these results.

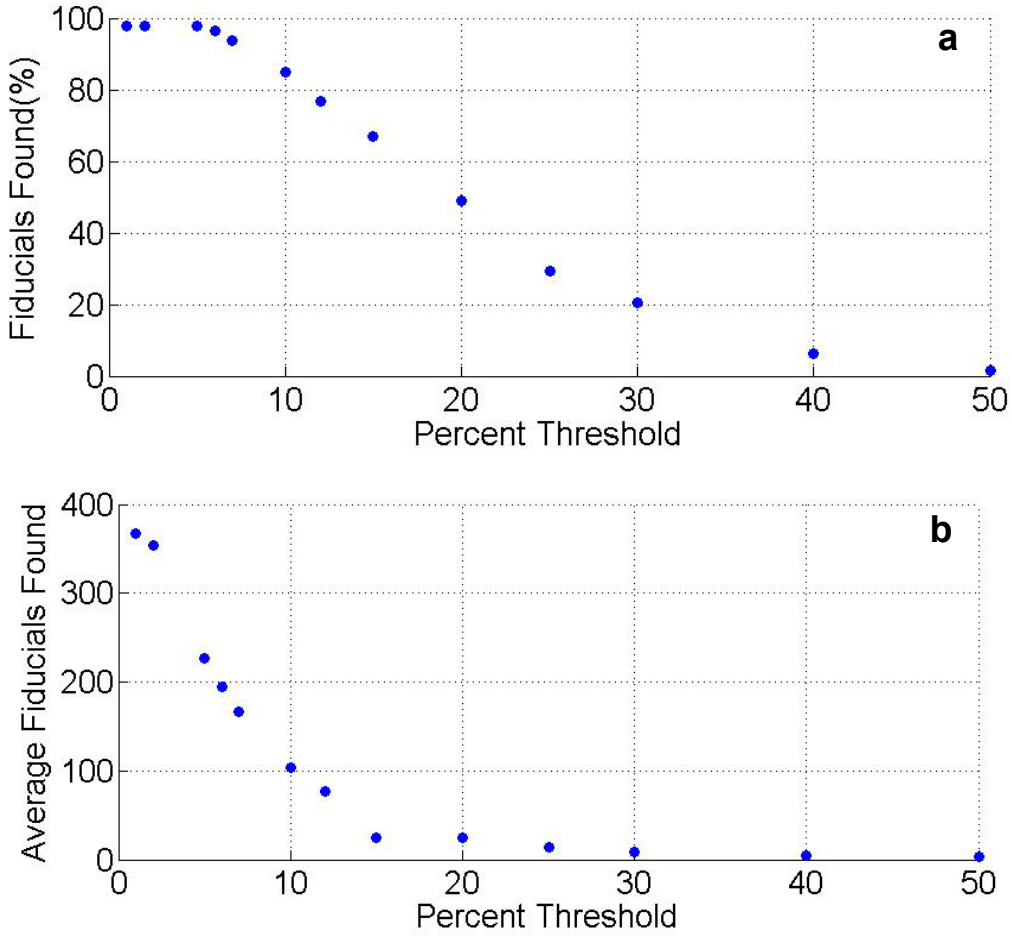

Figure 2. (a) Fiducials Found vs Threshold, and (b) Fiducial Candidates vs Threshold

The result of the clustering threshold optimization and selection of other tolerance parameter values are the following (for appr. 0.14mm pixel spacing): a = 5%, b = 20 degrees, c = 210 pixels, d = *30 pixels*, e = 10 degrees, f = 320 pixels, g = 350 pixels, h = 30 pixels. These values were used for the generation of all the results.

Tests were conducted on four scans with 147, 92,112, and 123 ultrasound frames respectively. When compared to the ground truth positions, all five fiducials were found on 94.5%, 94.6%, 94.6% and 92.7% of the frames with a tolerance of five pixels.

The mean fiducial localization errors for each of the four scans were -0.04, 0.07, -0.10, -0.08 mm along the x axis and -0.09, 0.02, -0.02, -0.02 mm along the y. The x axis increases to the right, horizontally across the image, and the y axis increases vertically towards the bottom of the image. The standard deviations were 0.18,0.27,0.30,0.24 mm and 0.14,0.11,0.12,0.12 mm for x and y respectively. The distribution can be seen for both axes in Fig. 3.

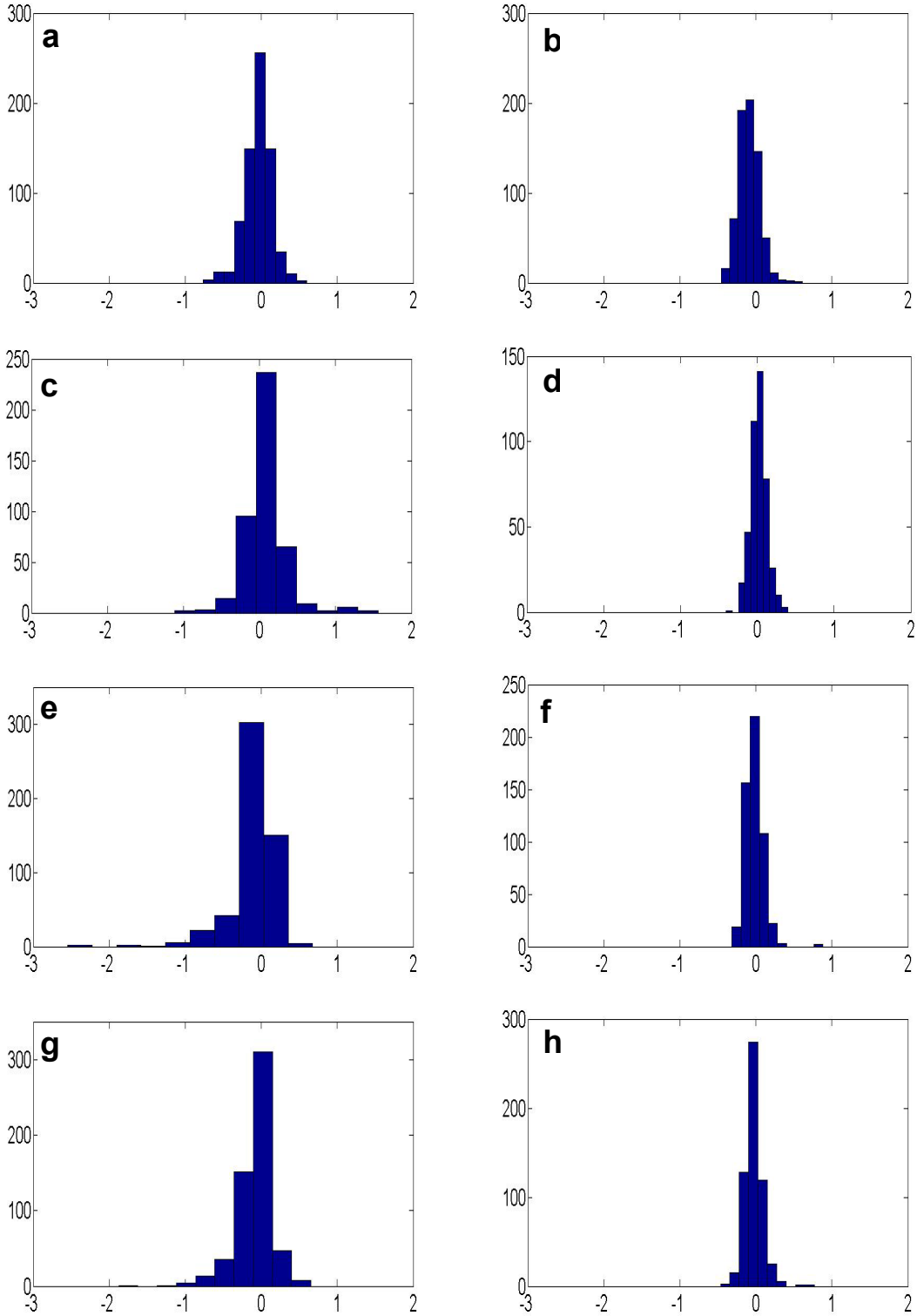

Figure 3. Fiducial localization error distribution. (a)(c)(e)(g) along X axis. (b)(d)(f)(h) along Y axis. (a)(b) are representative of a scan from the first (gelatine) phantom, while (c)(d), (e)(f), and (g)(h) are all different scans of the same phantom (glycerol) from the second set

The error on the y-axis is not exactly centered around zero, however the value is still well below the estimated accuracy of the ground truth (2-3 pixels).

The execution time of the algorithm on a 2.4GHz Intel Core2 processor (on a single thread) takes an average 90 seconds to run. Reviewing the frames where the fiducial detection was not successful may take just a few minutes, thus, the total fiducial segmentation time of an average sequence is decreased from about one hour to a few minutes.

4. DISCUSSION

On average the algorithm successfully locates all five fiducials on 94.1 % of the frames tested. Our goal was to achieve a 90% success rate or better. This percentage is enough to greatly reduce the amount of time spent manually segmenting. The second set of phantoms proved the repeatability of our results; we still achieved a success rate above 90% despite different experimental settings such as gel composition.

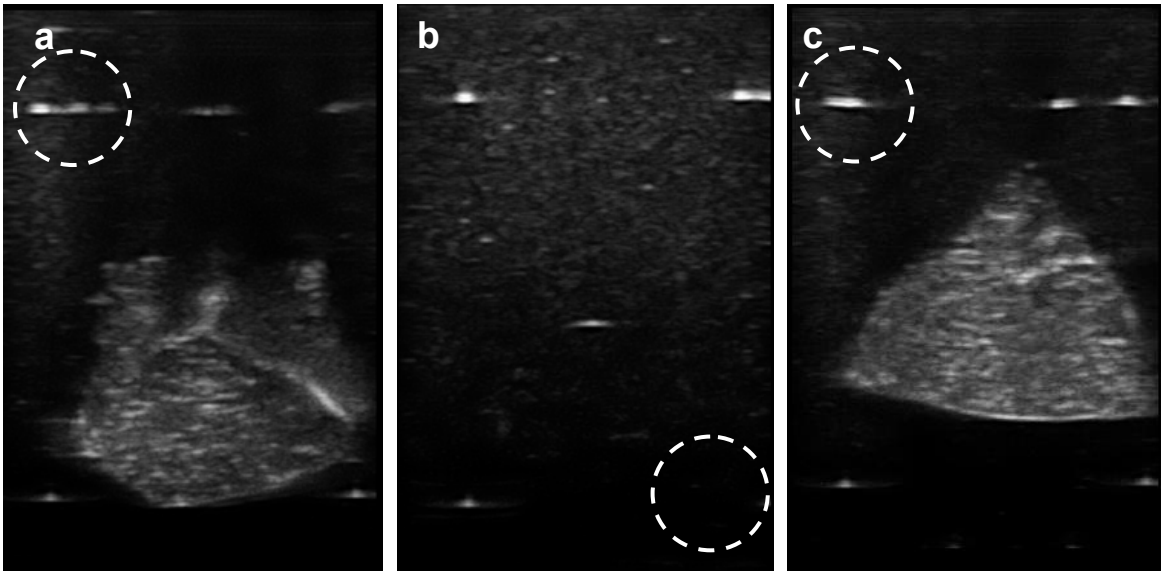

Figure 4. Failure Cases (a) artifact mistaken for real fiducial (b) fiducial not visible (c) inability to detect fiducial

Most failure cases are caused by duplication of fiducials in the image (Fig. 4.a). This occurs when the tissue boundary is nearly parallel to the y axis of the ultrasound image and vertically aligned with a fiducial position, causing the fiducial to appear as an elongated artifact with multiple intensity maxima. The duplicated 'phantom' fiducials are sometimes brighter than the real fiducials, causing the algorithm to mistake them for real fiducials. Because the algorithm ranks potential lines based on intensities, if any phantom fiducial between the two endpoints of the three point line is brighter than the real fiducial between the endpoints, then the algorithm will select the phantom fiducial as a real fiducial.

Another failure case occurs when only a fraction of the fiducials is visible in the acquired image (Fig. 4.b). This is due to an error during the scanning process in which the ultrasound probe was not centered around the fiducial lines, therefore these cases where excluded from the reported results.

A rarer cause of failure is the inability to identify fiducials among the list of candidate fiducials (Fig. 4.c). This is due to the thresholding parameters. By using more permissive thresholding parameters, these fiducials could be found, but the occurrence of the previous failure cases would increase.

During manual ground truth segmentation the locations selected may slightly differ from the real center point position, however we estimate this to be less than 2-3 pixels as the fiducials are small (average fiducial size is about 15x5 pixels).The coordinate values representative of a fiducial are sensitive to the relationship between the width and intensity

of the fiducial. More elongated fiducials with lower intensities tend to be represented by coordinates close to or on their boundaries.

The value of our algorithm is proved by its ability to correctly segment datasets it was not optimized for. The only adjustments made for the new phantoms were for line lengths, which are parameters that can vary with the dimensions of the phantom box. Thresholding was not altered and still provided accurate results. This confirms the portability of our algorithm and verifies the repeatability of our results.

The presented algorithm is easily extendable for other fiducial configurations that contain parallel lines and points with known expected position along a line, thus implementation of any of the fiducial configurations studied by Peikari *et al.*[1] (having well-visible fiducials) is straightforward.

5. CONCLUSIONS

Our contribution was automating the most time-consuming manual part of the ultrasound image processing step of the ablation monitoring validation framework. We began with an existing algorithm that worked for fishing line in water and made it usable with a complex tissue and tissue mimicking gel phantom. Specifically, thresholding was optimized to work with lower quality images. As well, the fiducial clustering was replaced and now uses intensities from the original image. Line finding geometry was redeveloped and optimized to comply with the ablation phantom geometry. The algorithm was validated with manually obtained points and was found to have a 94.1% success rate. Our purpose has been met, and fiducials can be detected with a sufficient success rate and accuracy. Thus lengthy manual post-processing can be replaced by the proposed automatic method.

6. ACKNOWLEDGEMENTS

This work was funded by the Natural Sciences and Engineering Research Council of Canada under the Idea to Innovation program. Gabor Fichtinger was supported as Cancer Care Ontario Research Chair.

7. REFERENCES

[1] Peikari, H., ,Lasso, A., and Fichtinger, G. "Improved Validation platform for ultrasound-based monitoring of thermal ablation",*Proc. SPIE Medical Imaging (2011) (Accepted*)

[2] Pompeu-Robinson, M.A., Gray, J., Marble, J., Peikari, H., Hall, J., U-Thainual, P., Aboofazeli, M., Lasso, A., Fichtinger, G. " Validation platform for ultrasound-based monitoring of thermal ablation," *Proc. SPIE Medical Imaging 7625, 76250T (2010)*

[3] Chen, T.K., Thurston, A.D., Ellis, R.E., and Abolmaesumi, P. " A real time free hand ultrasound calibration system with automatic accuracy feedback and control," *Ultrasound in Med. & Biol,* **35***(1) pp. 79–93, (2009)*

[4] Rousseau, F., Hellier, P., and Barillot, C. "Confhusius: A robust and fully automatic calibration method for 3D freehandultrasound," *Medical Image Analysis* **9***(1) 25-38(2005)*

[5] Mercier, L., Langø, T., Lindset, F. and Collins, D.L., "A review of calibration techniques for freehand 3-D ultrasound systems", *Ultrasound in Medicine and Biology,* ***31****(2), 143-165 (2005)*

[6] Rivaz, H., Fleming, I., Azzumpcao, L., Fichtinger, G., Hamper, U., Choti, M., Hager, G., and Boctor, E., " Ablation monitoring with elastography: 2d in-vivo and 3d ex-vivo studies," *Medical Image Computing and Computer-Assisted Intervention* ***11****(Pt 2), 458-466(2008)*

[7] Pieper, S., Lorensen, B., Schroeder, W., Kikinis, R. The NA-MIC Kit: ITK, VTK, Pipelines, Grids and 3D Slicer as an Open Platform for the Medical Image Computing Community. *Proceedings of the 3rd IEEE International Symposium on Biomedical Imaging: From Nano to Macro 2006; 1:698-701.*