Journal of Medical Imaging

MedicalImaging.SPIEDigitalLibrary.org

Effects of line fiducial parameters and beamforming on ultrasound calibration

Golafsoun Ameri
John S. H. Baxter
A. Jonathan McLeod
Terry M. Peters
Elvis C. S. Chen

Effects of line fiducial parameters and beamforming on ultrasound calibration

Golafsoun Ameri,^{a,b,*} John S. H. Baxter,^{a,b} A. Jonathan McLeod,^{a,b} Terry M. Peters,^{a,b} and Elvis C. S. Chen^{a,b}

^aRobarts Research Institute, London, Ontario, Canada

^bWestern University, Biomedical Engineering Graduate Program, London, Ontario, Canada

Abstract. Ultrasound (US)-guided interventions are often enhanced via integration with an augmented reality environment, a necessary component of which is US calibration. Calibration requires the segmentation of fiducials, i.e., a phantom, in US images. Fiducial localization error (FLE) can decrease US calibration accuracy, which fundamentally affects the total accuracy of the interventional guidance system. Here, we investigate the effects of US image reconstruction techniques as well as phantom material and geometry on US calibration. It was shown that the FLE was reduced by 29% with synthetic transmit aperture imaging compared with conventional B-mode imaging in a Z-bar calibration, resulting in a 10% reduction of calibration error. In addition, an evaluation of a variety of calibration phantoms with different geometrical and material properties was performed. The phantoms included braided wire, plastic straws, and polyvinyl alcohol cryogel tubes with different diameters. It was shown that these properties have a significant effect on calibration error, which is a variable based on US beamforming techniques. These results would have important implications for calibration procedures and their feasibility in the context of image-guided procedures. © 2017 Society of Photo-Optical Instrumentation Engineers (SPIE) [DOI: 10.1117/1.JMI.4.1.015002]

Keywords: ultrasound calibration; phantoms; synthetic aperture imaging; ultrasound-guided interventions; magnetic tracking.

Paper 16075RRR received May 5, 2016; accepted for publication Feb. 8, 2017; published online Feb. 28, 2017.

1 Introduction

Ultrasound (US) imaging has become a mainstay in image-guided interventions, providing real-time information about the anatomy of interest. In the case of spinal anesthesia delivery, US is the modality of choice, particularly in obstetrics, due to counter-indications that prevent the use of modalities with ionizing radiation.¹ US has the additional benefits of being widely available, low cost, and mobile.² These advantages of US have led to its integration as an intraoperative imaging technique in image-guided interventions, often in conjunction with an augmented reality (AR) environment for navigation.³ These environments aim to complement the surgeon's conventional visualization by displaying the surgical tools, preoperative and intraoperative images, and virtual models all in one common frame of reference.⁴ The AR guidance systems are made possible with the integration of tracking technology to localize and track surgical tools and intraoperative images. A fundamental criterion for the integration of US images in AR environments is US calibration, the process whereby images from a tracked US probe are placed in the context of the tracker, i.e., a common frame of reference. More specifically, US calibration is the process of calculating the transformation (rotation, translation, and scaling) between the image plane and the origin of the sensor mounted on the US probe.

Commonly, US calibration is performed by capturing US images of a tracked, stationary calibration phantom with a known geometry at various depths and phantom poses. Then, the phantom features, i.e., fiducials, are segmented in the images. Knowing the physical location of the fiducials and

the US probe in tracker space and the location of the fiducials in the US image space, the transformation between the US image and the probe sensor can be calculated. Many different US probe calibration techniques using a variety of phantom geometries have been proposed over the years. One of the most common of these is a Z-bar calibration technique, using taut wires in a Z shape as fiducials. This technique has the advantage of low cost phantoms⁵ and a fully automatic pipeline developed and available in the open-source PLUS library.⁶ Alternative proposed US calibration methods take advantage of the properties of different geometries, which are not immediately representable as points in the US image. Wall calibration,⁷ for example, uses a single highly reflective plane, which, seen in cross-section, becomes a line. The main advantage of this approach is that segmentation of a line has properties that make it more robust. Other volume-based methods, such as the rooftop phantom,⁸ have been used to provide similar orientation information in 3-D US calibration. Phantomless calibration has also been of increasing interest as surgical tools or needles could be used in lieu of a dedicated calibration phantom.⁹ These tools are often modeled as lines or points in a cross-sectional view, similar to the wires in the Z-bar phantom.

Regardless of the calibration phantom type, the basis of US calibration is the accurate and rapid segmentation of fiducials. The imaging and segmentation steps of an US calibration pipeline traditionally rely on conventional B-mode imaging, where highly reflective materials, such as taut wires, are commonly used as fiducials. Due to ringing effects, specular reflection, beam width, and low resolution outside the focal region, such fiducials appear widespread and/or blurred in US images,¹⁰

*Address all correspondence to: Golafsoun Ameri, E-mail: gameri@robarts.ca

Fig. 1 B-mode images of two nylon Z fiducials, that are 5 mm apart, placed (a) in focus and (b) out of focus of the US image.

making their localization in the US image a challenging task. For example, Fig. 1 shows US images of braided wire fiducials of 0.3 mm in diameter placed within and out of the focal region of the image, with the foci at 7 and 7.5 cm in both cases. These fiducials, even when in focus, appear spread out and much larger than their actual size in the US image, making their accurate segmentation prone to error. Inaccuracies in fiducial localization can result in calibration error, which ultimately decreases the accuracy of image-guided interventions.

To overcome the challenges, we have investigated the effects of US image beamforming and reconstruction¹¹ as well as phantom material and geometry¹² on fiducial appearance and localization in US images and, ultimately, US calibration accuracy. To improve US image quality of the fiducials, synthetic aperture imaging (SAI) was used. SAI is a focusing technique, providing dynamic focusing and high resolution throughout the entire image, which could lead to improved fiducial localization and subsequently calibration accuracy. In addition, since hypo-echoic materials with acoustic properties similar to tissue produce fewer ringing effects in US images, it is expected that fiducials made out of such materials would be more easily segmented in these images, leading to increased US calibration accuracy. A comprehensive overview of these studies, submitted to SPIE Medical Imaging 2014 and 2015, is described in detail in this paper.

2 Purpose

The purpose of this article is to investigate the effects that US image beamforming technique, phantom geometry, and material acoustic properties play in US calibration and what effect these factors have on overall US calibration accuracy. By quantifying these effects, US calibration for image-guided interventions can be optimized appropriately according to the beamforming technique and calibration phantom design.

3 Background

3.1 Error in Ultrasound Calibration

A common approach to US calibration is to image a phantom or set of fiducials with a known geometry and formulate a

marker-based registration problem relating the known geometry to the image through a coordinate transformation. This registration accuracy and, ultimately, the calibration accuracy are affected by a number of factors including the fiducial localization accuracy, fiducial configuration, and the number of fiducials used.^{13,14} As US calibration is inherently a registration problem,¹⁵ the metrics used for evaluating US calibration accuracy are the same as those used for registration processes, namely, the fiducial localization error (FLE), fiducial registration error (FRE), and target registration error (TRE).^{13,16}

The FLE is defined as the distance of the localized fiducial in the US image from the fiducial's true position in the image space. However, the true position of the fiducial in the image is "forever unknown."¹³ As a result, in practice, the true position of a fiducial is estimated as the mean location of fiducials identified in all the segmented images. FRE is the root-mean square distance between homologous fiducials after registration. The expected value of FRE depends on the number of fiducials used and the expected square value of FLE.¹³ Fitzpatrick et al.¹³ have shown that FRE is not a reliable predictor of registration accuracy. An error that is more of a concern in US calibration is the TRE, which is a more direct measure of registration error.¹⁷ TRE is defined as the distance between homologous points, other than the fiducials used for the registration process, after registration has been performed. TRE is strongly dependent on the fiducial configuration and the position of the measurement points with respect to registration fiducials.¹³ Therefore, to reduce TRE in US calibration, the fiducials should be placed throughout the entire field of view.

Although targets can be separate objects in practice, a surrogate metric, the leave-one-out TRE, is often used. Given a set of N imaged fiducials, X , with known locations, Y , the leave-one-out TRE can be calculated as follows:

$$\text{TRE}_{\text{Leave-one-out}} = \frac{1}{N} \sum_{i=1}^N \|Y_i - T_{-i}(X_i)\|, \quad (1)$$

where $T_{-i}(\cdot)$ is a calibration algorithm minimizing the FRE for the fiducial set $X \setminus \{X_i\}$, that is, all fiducials except the i th.

In this work, the leave-one-out TRE is used to evaluate and compare the calibration error in the proposed calibration techniques.

3.2 Rigid and Anisotropic-Scaled Calibration Transformations

Often, the transformation between the known geometry and calibration images is rigid, consisting only of rotation and translation. To specify a transformation, corresponding points between the geometry and its representation in the image must be known. In the case of Z-bar calibration,¹⁸ these correspondences can be approximated quickly allowing for the calibration transformation to be formulated as an “Orthogonal Procrustes” problem and solved analytically.^{19,20}

If point-to-point correspondences are not known *a priori*, the iterative closest point (ICP) algorithm²¹ can be used to simultaneously find these correspondences and the rigid transformation. ICP is a generic algorithm, independent of shape representation, that iteratively performs a correspondence and a transformation operation until a stable solution is obtained. The correspondence operation establishes a one-to-one correspondence between the measurement point-set and the model point-set using the shortest Euclidean distance, and the transformation operation finds the optimal rigid-body transform using a closed-form, least-squares method.^{19,20} In general, ICP converges monotonically but requires a fairly reasonable initialization to reach a global minimum.

We used anisotropically scaled ICP proposed by Chen et al.²² to solve for the US calibration transform. This method improves the accuracy of US calibration by solving for anisotropic pixel scaling in addition to translation and rotation. Deviations of the speed of sound from 1540 m/s (the speed assumed by most scanners) can result in anisotropic pixel scaling when using a linear array probe. These effects are particularly significant when performing US calibration in a room temperature water bath.¹⁵

Similar to traditional ICP, this algorithm converges monotonically.²² However, this increased flexibility has its disadvantages as a larger number of fiducials or images must be acquired to accurately determine the additional scaling parameters.²³ That is, although anisotropic-scaled calibration is guaranteed to have a lower FRE than rigid calibration, with fewer than seven calibration images, anisotropic-scaled calibration can result in a higher TRE.²³

3.3 Synthetic Aperture Imaging

SAI techniques, originally conceived for radar systems, have been investigated in medical US imaging since the early

1990s.²⁴ SAI provides dynamic focusing and improves resolution throughout the image via postprocessed beamforming, which in turn improves US image quality. Unlike traditional B-mode imaging, where transmit focal regions are predefined and fixed, SAI achieves focusing in both transmit and receive on every image point. Therefore, SAI allows for higher resolution and interpretability throughout the entire image, rather than only at prespecified foci. SAI has shown superior image quality in breast,²⁵ liver,²⁶ abdominal US imaging²⁷ due to increased resolution.

There are multiple methods for performing synthetic aperture beamforming. In the modified synthetic transmit aperture (STA) algorithm,²⁸ employed in this work, each element in the probe array transmits sequentially in an isolated manner. After each transmit, every crystal is used to receive echoes. The acquired complete dataset allows both dynamic transmit and receive focusing to be achieved. Then, image reconstruction is performed in two stages. In the first stage, an image is reconstructed after each transmit. As a result, the number of images reconstructed is equal to the number of transmissions, i.e., the number of elements in the array. Image reconstruction is performed using the delay-and-sum technique.²⁴ Coherent summation is performed by finding the geometric distance between the transmitting element to the image point and back to the receiving element, resulting in dynamic receive focusing. In STA, the angular directivity function, $f(\theta)$, of the transducer element is incorporated as weights in the delay-and-sum stage.²⁸ $f(\theta)$ describes the radiation pattern of the transducer element and can be calculated using the far-field approximation for the narrow-strip transducer element as follows:²⁹

$$f(\theta) = \frac{\sin(\pi d/\lambda \sin \theta)}{\pi d/\lambda \sin \theta} \cos \theta, \quad (2)$$

where d is the element width and λ is the wavelength. $f(\theta)$ is calculated for both the transmit and receive elements. During transmission, θ is the angle between the transmit element’s surface normal and a virtual line connecting the element’s center to the image point [Fig. 2(a)]. In receive, θ is defined as the angle between the receive element’s surface normal and a virtual line connecting the receive element’s center to the image point [Fig. 2(b)].

Due to the finite size of a single transducer element, its acoustic radiation pattern is not spherical and instead is angular dependent.²⁹ Therefore, images reconstructed in the first stage, referred to as low resolution images, have appreciable signal only over a limited field of view. In the second stage, the low resolution images are added together and a high resolution

Fig. 2 (a) The first transmit element’s directivity angle, θ_{r1} , for the image point P. (b) The i ’th receive element’s directivity angle, θ_{ri} , for the image point P.

image is obtained.²⁴ As a result, focusing in both transmit and receive is achieved (Fig. 3).

Segmentation of these STA images is greatly simplified due to the increased resolution, particularly for small features, such as echoes of wires in a calibration phantom. In the past, calibration has been restricted to conventional B-mode images. By performing calibration on STA images rather than B-mode

images, higher accuracy segmentation of the wire fiducials is possible.

4 Effect of Imaging Type on Ultrasound Calibration

Here, the performance of STA on Z-bar calibration accuracy is compared with that of conventional B-mode imaging.

Fig. 3 Modified synthetic transmit aperture imaging. A high resolution image is obtained in the second stage by adding images reconstructed in the first stage.

High-resolution STA images are expected to facilitate fiducial localization in the image, allowing for more accurate calibration of US probes.

4.1 Z-Bar Ultrasound Calibration

The basis of Z-bar calibration¹⁸ is the calibration phantom, which consists of two parallel plates a known distance apart with a grid of holes milled perpendicular to the faces of the plates. Wires, acting as line fiducials, are then threaded through the two plates to form a series of Z-shaped configurations. A magnetic tracker is rigidly affixed to the calibration phantom. These collinear line fiducials produce identifiable echoes in the US images. In the experiment presented in this section, the Z-bar phantom designed by the PERK lab, known as the fCal phantom,³⁰ was used. The tracked fCal phantom and an US image of it are shown in Fig. 4.

All of the line fiducials are localized in the US images. Because of the Z-shaped configuration, the intersections of the US image plane with the diagonal fiducials can be calculated based on similar triangles.⁶ Using the tracking information, these points are transformed into probe coordinates and registered with their corresponding pixel locations in the US image, yielding a calibration transform relating the US image to the tracking sensor rigidly fixed to the probe.

4.2 Experimental Setup

The calibration experiment consists of a SonixRP US Scanner and L14-5 linear probe, with an element pitch of 0.30 mm (BK Ultrasound, Analogic Corp., Massachusetts), a Z-bar calibration phantom, and an Aurora Magnetic Tracking System (Northern Digital Inc., Waterloo, Canada) arranged, as shown in Fig. 5. Both the US probe and calibration phantom are fixed in place using plastic supports to remove relative motion and are magnetically tracked, using six degrees-of-freedom (DoF) magnetic sensors, allowing for their tracking coordinates to be captured simultaneously with image acquisition. The root-mean-square error of the position and orientation measurements of the tracking system using a six DoF sensor is 0.48 mm and 0.30 deg, respectively.³¹ The Z-bar calibration phantom included two opposite 40-mm wide Z-shaped line fiducial configurations 5 mm apart in the depth direction.

Fig. 5 Experimental calibration setup comprising a magnetic tracking system, tracked ultrasound probe, tracked Z-bar calibration phantom, and plastic supports.

The US imaging settings for the calibration process are chosen to provide the highest resolution possible and cover the imaging depth that is intended for the image-guided procedure. In this study, the highest US frequency provided by the scanner for the linear array, 10 MHz, was used to acquire the highest resolution B-mode US images of the phantom. The image depth was set to 9 cm, the largest possible value at the chosen frequency. As US image quality is often compromised at larger depths, largely due to attenuation, choosing a large depth would ensure a challenging US calibration task. Two focal zones were chosen for B-mode imaging, at ~7 and 7.5 cm, as shown in Fig. 4, to improve image resolution especially at larger depths.

Calibration experiments were performed at an imaging depth of 9 cm with the calibration phantom placed 6 to 9 cm from the probe. This experiment involved acquiring nine images using both B-mode and STA imaging with the probe and calibration phantom physically secured to ensure correspondence between the images in each mode. The beamforming for STA imaging was performed using the Texo SDK (BK Ultrasound, Analogic Corp.), and the raw RF data were acquired using the SonixDAQ (BK Ultrasound, Analogic Corp.). The same center frequency as the one used for B-mode imaging was used for STA imaging, i.e., 10 MHz. The sampling frequency was set to 40 MHz. STA images were then reconstructed from the raw RF data, using an approximation of the speed of sound in water at room

Fig. 4 (a) Z-bar calibration phantom and (b) corresponding B-mode image.

temperature, 1480 m/s.³⁰ The line fiducials intersect the imaging plane at six locations, producing hyperintensities. Both STA and B-mode images were segmented by six users. The image segmentation was performed twice in each imaging mode by each user with the images displayed in a randomized order to minimize learning and fatigue effects. An anisotropic Orthogonal Procrustes solution based on singular value decomposition¹⁹ was then performed to find the relationship between the tracking coordinates and segmented line fiducials.

4.3 Results

Qualitatively, compared with the B-mode images, STA produced much more readily interpretable US images of the Z-bar calibration phantom, the line fiducials being clearer and easier to visually localize, as demonstrated in Fig. 6. STA imaging also reduced the number and severity of artifacts in the images, such as the reflections of the bottom of the water bath.

For evaluation purposes, the mean fiducial location identified by all participants was used as the estimate of the true physical location, assuming the localization follows a normal distribution without bias. The FLE was then calculated for each user. The mean absolute error was 0.15 and 0.21 mm for SAI and B-mode, respectively. SAI resulted in a 29% improvement in FLE compared with conventional B-mode imaging. For the sake of comparison, the axial and lateral dimensions of a pixel in the B-mode image were 0.175 and 0.176 mm, respectively. Calibration accuracy was determined by measuring both the FRE and the leave-one-out TRE. When these segmented fiducials were used for calibration, the resulting FRE also improved from 1.56 to 1.42 mm, i.e., an improvement of 9%, with SAI. The resulting mean leave-one-out TRE improved from 2.00 to 1.80 mm, corresponding to a 10% improvement with SAI. These values are recorded in Table 1.

The improvements in FLE, FRE, and leave-one-out TRE are all statistically significant ($p < 0.01$). It is important to note that the error due to magnetic tracking is a major contributor to the calibration error. As a result, the large improvement in line fiducial localization translated into a more modest but still significant improvement in the final calibration accuracy.

Fig. 6 The (a) STA and (b) corresponding B-mode image of the Z-bar phantom.

Table 1 Calibration error.

	SAI			B-mode		
	FLE (mm)	FRE (mm)	TRE _{Leave-one-out} (mm)	FLE (mm)	FRE (mm)	TRE _{Leave-one-out} (mm)
	0.14	1.36	1.74	0.20	1.67	2.15
	0.14	1.34	1.67	0.23	1.60	2.02
	0.14	1.37	1.74	0.16	1.78	2.31
	0.25	1.64	2.07	0.29	1.46	1.88
	0.14	1.44	1.82	0.20	1.40	1.78
	0.16	1.45	1.84	0.24	1.46	1.90
	0.16	1.43	1.83	0.21	1.50	1.89
	0.11	1.34	1.68	0.15	1.58	1.97
	0.12	1.36	1.76	0.21	1.67	2.16
	0.20	1.52	1.89	0.24	1.71	2.24
	0.13	1.42	1.80	0.15	1.41	1.84
	0.13	1.35	1.72	0.21	1.47	1.87
Mean	0.15	1.42	1.80	0.21	1.56	2.00
STD	0.04	0.09	0.11	0.04	0.13	0.18

5 Phantom Material and Geometric Properties on Ultrasound Calibration

There is considerable uncertainty regarding the localization of the centroid of wire fiducials.⁶ As shown in Fig. 1, wire fiducials appear larger than their true size and/or blurred in B-mode US images, especially when placed at an oblique angle with the image plane or outside the focal region of the transducer. Such widespread echoes in the image are not representative of the true cross-section of the line fiducials with the image plane. As a result, the true intersection of the US beam's mid-plane with these line fiducials is not robustly identifiable, which decreases segmentation accuracy.

While changing the gain and dynamic range of the B-mode image could improve the image quality, not all the artifacts can be eliminated and fiducial localization remains a challenging task. For the same gain and dynamic range settings, fiducials with lower echogenicity lessen the impact of the aforementioned artifacts in US, such as the comet tail artifact and spreading at oblique angles. In this section, the effect of such materials on US calibration accuracy is investigated.

5.1 Phantomless Ultrasound Calibration

The phantomless calibration⁹ technique has the advantage of using tracked surgical tools, such as a needle or a stylus, in lieu of a dedicated phantom to perform calibration. US phantomless calibration can be formulated as a point-to-line registration problem,¹⁶ which was solved using the ASICP algorithm mentioned in Sec. 3.2. An added benefit of this approach is that, in any image, only a single calibration object needs to

be identified, which simplifies the process of automatic segmentation by eliminating any correspondence problems. Also, this framework allows for the use of a calibration phantom that can be modeled as a line or cylinder, similar to a surgical tool. This facilitates integration of different materials in the phantom construction.

As with Z-bar calibration, these phantoms and the US probe are tracked using a magnetic tracking system (Aurora, Northern Digital, Canada). The US probe is calibrated through the acquisition of at least three images showing a cross-section of the calibration object. However, a larger number of images (7 to 12) were acquired to ensure a higher accuracy calibration representative of clinical use. The phantoms used in this study and automated segmentation procedures are explained in the following sections.

5.2 Phantom Material

In this study, calibration phantoms were modeled as hollow plastic straws of different diameters, 7 and 3 mm, or solid tubes. The material used for making hypoechoic solid tubes was polyvinyl alcohol cryogel (PVA-C), a well-known tissue mimicking material in US imaging.³² These phantoms were made by pouring PVA-C into molds, specifically, plastic straws with two different diameters, and then subjecting them to three freeze-thaw cycles. The phantoms were then fixed in a tracked frame. For comparison purposes, a braided wire phantom, with a diameter of 0.3 mm, was used as a line fiducial. The braided wire was placed in an fCal phantom.³³ Images of the phantoms are shown in Fig. 7.

In our experience, braided wires have better visibility in US images compared with smooth wires. Both plastic straws and PVA-C tubes are less reflective than wire fiducials traditionally used in US calibration. Conventional B-mode and STA images of these phantoms are shown in Fig. 8.

5.3 Object Segmentation from Ultrasound Images

In this work, automatic segmentation of user defined region of interest was used to localize the centroid of the phantoms. The cross-sectional views of the phantoms used are ellipsoidal, the

Fig. 7 (a) The phantom frame was used to hold straws and PVA-C tubes. (b) The fCal phantom is shown. The braided wire was threaded through the holes of this frame.

eccentricity of the ellipse varying with the angle of insonation relative to the orientation of the object. Therefore, the automatic segmentation process used involved automatic detection of best-fit ellipse. The first step of the segmentation is edge detection, where thresholding based on Otsu's method³⁴ was used to binarize the B-mode image. Morphological operations were performed to thin and smooth the binary images, so the edges could be extracted. For STA images, with high resolution and thin edges of the fiducials in the image, the edge detection was performed using a Canny edge detector.³⁵ The second step of the segmentation was ellipse fitting, where Simonovsky's MATLAB implementation was employed. First, two points on the edge mask were chosen and defined as the major axis. The minor axis was then determined through the Hough transform. The algorithm iterated over all pairs subject to reasonable exclusion criteria, such as bounds on the major axis length, and the final ellipses were constrained to meet criteria such as bounds on the eccentricity. Segmented fiducials in both conventional B-mode and STA images are displayed in Fig. 8.

5.4 Experimental Setup

The phantom and US probe were held stationary in a water bath. US images of the phantoms and the raw RF data were acquired using a SonixRP US scanner and the SonixDAQ. The US probe and phantom were tracked using an Aurora tracking system. For every phantom position, first, the tracking information of the US probe and the phantom were recorded. Next, the B-mode image with a fixed focus (at 5.47 cm) was acquired directly from the scanner. Finally, STA beamforming and image reconstruction were performed, as described in Sec. 4.2.

5.5 Results

Overall, B-mode displayed better visibility of large phantoms due to reduced interference from shadowing, blurring, and ringing effects along object boundaries. STA images displayed fewer of these artifacts with higher resolution and improved interpretability of small phantoms. However, most of the STA images suffered from low contrast and signal-to-noise ratio (SNR) because of reflections from the sides of the phantom frame and low transmit power as only one transducer element is used to transmit at a time.

Quantitatively, each calibration resulted in a visual alignment of the cross-section of the target and the US image, as shown in Fig. 9.

For the purposes of evaluation, calibration accuracy was determined by measuring the leave-one-out TRE. These values are recorded in Table 2.

For statistical analysis, first, multifactorial ANOVA was used to determine the significance of three factors in the experiment. The first factor (I) is the imaging type, either B-mode or STA. The second factor (M) is the phantom material, either a PVA-C tube or a hollow straw. The third factor (T) was the diameter of the phantom, either large (7 mm) or small (3 mm). The results are summarized in Table 3 with an asterisk (*) indicating significant findings. Then, an unpaired t -test was performed to compare the braided wire fiducial with the phantom and imaging type that gave the lowest calibration error with statistical significance based on the ANOVA test.

The resulting model from ANOVA only showing significant components ($p < 0.05$) is as follows:

Fig. 8 Automatically segmented fiducials in conventional B-mode (top row) and STA (bottom row) images. Phantoms from left to right: plastic straw (7 mm in diameter), small plastic straw (3 mm in diameter), large PVA-C tube, small PVA-C tube, and braided wire (0.3 mm in diameter).

Fig. 9 Visual representation of the straw phantom registered to the US image after calibration.

Table 2 US calibration error using different phantoms in both B-mode and STA images.

Phantom	Imaging technique	Mean TRE _{Leave-one-out} (mm)	Standard deviation (mm)
Large straw	B-mode	0.70	0.85
	STA	1.95	0.95
Small straw	B-mode	1.63	0.77
	STA	0.79	0.64
Large PVA-C tube	B-mode	0.70	0.48
	STA	0.96	0.63
Small PVA-C tube	B-mode	1.20	0.83
	STA	0.56	0.39
Braided wire	B-mode	1.22	0.77
	STA	0.98	0.60

Table 3 Multivariate ANOVA.

Factor	Sum Sq.	d.f.	Mean Sq.	<i>F</i>	<i>p</i>
<i>I</i>	0.81	1	0.81	1.56	0.22
<i>M</i>	2.67	1	2.66	5.15	0.03*
<i>T</i>	0.03	1	0.03	0.06	0.82
<i>I</i> * <i>M</i>	2.16	1	2.16	4.19	0.04*
<i>I</i> * <i>T</i>	2.35	1	2.35	4.56	0.04*
<i>M</i> * <i>T</i>	0.001	1	0.001	0	0.96
Error	36.65	71			
Total	48.04	77			

$$\text{TRE}_{\text{Leave-one-out}} = c + f(M) + g(I, M) + h(I, T) + \epsilon, \quad (3)$$

where

$$c = 1.11 \text{ mm}$$

$$g(I, M) = \begin{cases} 0.25 \text{ mm}, & \text{if } I = \text{Bmode and } M = \text{PVA-C} \\ 0.25 \text{ mm}, & \text{if } I = \text{STA and } M = \text{Straw} \\ -0.25 \text{ mm}, & \text{if } I = \text{Bmode and } M = \text{Straw} \\ -0.25 \text{ mm}, & \text{if } I = \text{STA and } M = \text{PVA-C} \end{cases}$$

$$f(M) = \begin{cases} 0.25 \text{ mm}, & \text{if } M = \text{Straw} \\ -0.25 \text{ mm}, & \text{if } M = \text{PVA-C} \end{cases}$$

$$h(I, T) = \begin{cases} 0.23 \text{ mm}, & \text{if } I = \text{Bmode and } T = \text{Small} \\ 0.23 \text{ mm}, & \text{if } I = \text{STA and } T = \text{Large} \\ -0.23 \text{ mm}, & \text{if } I = \text{Bmode and } T = \text{Large} \\ -0.23 \text{ mm}, & \text{if } I = \text{STA and } T = \text{Small} \end{cases}$$

ANOVA analysis indicates that all three factors had some influence on the resulting leave-one-out TRE either directly

or through interaction. The US image type showed significant interaction effects with both factors regarding phantom construction. This analysis shows that the lowest error with statistical significance is associated with SAI using a thin PVA-C tube as it corresponds with decreased error in each function, $f(M)$, $g(I, M)$, and $h(I, T)$. Thus, the calibration error obtained with thin PVA-C tube and SAI was compared with that obtained with the braided wire line fiducial in STA. The result of an unpaired t -test strongly suggests that the smaller PVA-C tube in STA imaging improved calibration accuracy compared with the braided wire in STA imaging, but it was not statistically significant due to the relatively low sample size ($p = 0.06$).

The results showed that calibration error based on conventional B-mode imaging was reduced for larger diameter phantoms. This is due to a more accurate segmentation of the centroid of the fiducials. For smaller phantoms, shadowing effects and blurring of the fiducials resulted in less accurate automatic segmentation of the centroids, which led to higher calibration error. With SAI, segmentation of the smaller phantoms was improved. However, segmentation of larger objects in SAI did not improve compared with B-mode imaging. In all cases, calibration error from a wire phantom, which is commonly used, was higher than larger phantoms with lower echogenicity.

To decrease US calibration error, we recommend using SAI and small PVA-C tubes. If only B-mode US is available, we recommend the use of a large straw over the conventional braided wire for its accessibility in addition to improved calibration accuracy.

6 Discussion

We have shown that the optimal phantom construction, both in terms of size and material acoustic properties, is dependent on the US reconstruction and beamforming algorithms. The purpose of US calibration is to find the transformation between the coordinate system of the tracking sensor attached to the probe and that of the image plane. It is important to note that the location of the image plane with respect to the probe, or a sensor attached to it, is independent of US beamforming. Therefore, the US beamforming technique used for US calibration is not required to be the same as the imaging mode used for a clinical procedure following US calibration. As a result, for the purpose of US calibration, an US beamforming that provides a high resolution image of the phantom should be acquired to reduce FLE, improving the overall calibration accuracy. As more advanced US beamforming techniques, such as SAI, become available, we can only expect that the considerations in calibration phantom design remain and that there will not be a single calibration phantom that performs optimally for all techniques.

In this work, we evaluated the performance of SAI in US calibration with different phantom materials and geometries. SAI allows dynamic focusing leading to higher resolution images and improved interpretation compared with conventional B-mode imaging. The improved localization capability improves the accuracy achievable in US calibration. However, one of the limitations of this work is the specific SAI technique used, i.e., STA, as it has a low transmit power due to transmission from a single element of the array transducer, which limits its applicability in calibration at large depths in which the SNR and contrast are poor. These issues could be overcome by employing other SAI techniques, such as multielement SAI³⁶

and synthetic aperture sequential beamforming,³⁷ where multiple elements are used for transmission, ensuring sufficient transmitted energy especially at larger depths. Coded excitations can also be used to improve the SNR in SAI.^{38–40} Other advanced US beamforming techniques, such as coherent plane-wave imaging, also provide a higher SNR compared with STA.⁴¹ Although SAI techniques are currently limited by clinical accessibility, open-architecture US scanners including SonixTouch (BK Ultrasound, Analogic Corp.) and ECUBE (Alpinion, Washington) have been developed, allowing for customized beamforming, and a number of commercial medical US manufacturers, such as Siemens,⁴² Philips,⁴³ and Zonare,⁴⁴ have recently incorporated similar advanced imaging modes in their US scanners.

Most US calibration procedures, such as those employed in this study, are performed in water using a highly reflective calibration phantom.¹⁵ Therefore, image speckle is not a major concern in the choice of US beamforming and processing. However, in the case of *in-situ* calibration, where calibration is performed inside the surgical scene,⁴⁵ additional aspects of image quality aside from resolution become important. For example, speckle reduction techniques such as SAI compounding⁴⁶ may improve fiducial localization. In addition, SA tissue harmonic imaging has also been shown to improve image quality.⁴⁷ The impact of said techniques in *in-situ* calibration and their impact on FLE and TRE need future investigation. Moreover, SAI in tissue is susceptible to motion and phase aberration artifacts, caused by spatial variations of the speed of sound in tissue, which may lead to errors in focusing calculations in SAI and deteriorate image quality.⁴⁸ The effects of different techniques for target displacement estimation⁴⁹ and the correction of motion^{24,50,51} and phase aberration⁵² in SAI for *in-situ* calibration need further investigation.

While 3-D US calibration has been performed for B-mode imaging,⁸ doing so with SAI is hindered by the limited access to raw RF data from 3-D probes and to customized beamforming. However, once such limitations are addressed, 3-D SA images can be reconstructed using 2-D probes.^{53,54} Current commercial 3-D US scanners are optimized for acceptable imaging frame rates as opposed to depth and resolution, which are of higher interest in calibration with a stationary phantom. Barring these difficulties, 3-D probe calibration may be constructed from the combination of biplane calibrations.⁴⁵

This work can be extended in a number of ways. First, the phantom materials evaluated in this paper have been constrained to braided wire, plastic straws, and PVA-C tubes. In the future, the range of phantom materials can be extended to include materials used specifically in certain calibration frameworks. For example, in the phantomless calibration technique,⁹ the calibration phantom can be replaced by the surgical tool, which is often metallic. It would be interesting to evaluate the performance of advanced US beamforming techniques, such as SAI, on US calibration using such materials. Second, this paper focused largely on the use of cylindrical or line phantoms, which form the majority of US calibration procedures.¹⁵ However, there has been an increasing interest in different phantom geometries for calibration including single-wall phantoms,⁷ which may show similar performance differences based on material type and image reconstruction strategy. Lastly, another area of future work will be incorporating TRE prediction models with the proposed phantoms. TRE prediction models are important in guiding how calibration has been performed. Early

point-to-point TRE prediction models, such as that developed by Fitzpatrick et al.,^{13,14} have seen increasing generality in terms of imaging characteristics, incorporating anisotropic⁵⁵ and heteroscedastic⁵⁶ FLE. Spatial stiffness-based models have been developed for point-to-point⁵⁷ and recently point-to-line¹⁶ TRE prediction, which can be directly applied to the phantom geometries examined in this paper.

7 Conclusions

Accurate US calibration is a fundamental step for inclusion of US as a real-time modality for use in many image-guided interventions. With increased availability of a diverse array of US beamforming techniques, it is important to consider phantom material and geometry in designing optimal calibration procedures. Using conventional B-mode imaging, we showed that calibration can be improved using large diameter phantoms, compared with smaller diameter plastic straws, PVA-C tubes, or wire fiducials, due to improved fiducial localization. With SAI on the other hand, smaller diameter phantoms with lower echogenicity, i.e., small PVA-C tubes, resulted in better US calibration overall. In the case of a Z-bar calibration, STA significantly reduced the FLE, FRE, and TRE compared with B-mode imaging. This study of the effects of calibration phantom material and thickness has the potential to alter how calibration phantoms are designed cognizant of the image reconstruction and beamforming technique.

Disclosures

The authors have no relevant financial interests in the manuscript and no other potential conflicts of interest to disclose.

Acknowledgments

The authors would like to thank Dr. Tamie Poepping for making her laboratory and equipment including SonixRP and SonixDAQ available to us. This project was funded by the Canadian Institutes of Health Research (MOP 119447), the Natural Sciences and Engineering Research Council of Canada (RPGIN 2014-04504), and the Canadian Foundation for Innovation (20094).

References

1. A. Bhatia and R. Brull, "Is ultrasound guidance advantageous for interventional pain management? A systematic review of chronic pain outcomes," *Anesth. Analg.* **117**(1), 236–251 (2013).
2. M. Kolios, "Biomedical ultrasound imaging from 1 to 1000 MHz," *Can. Acoust.* **37**(3), 35–43 (2009).
3. T. Peters and K. Cleary, *Image-Guided Interventions: Technology and Applications*, Springer Science & Business Media, New York (2008).
4. T. M. Peters and C. A. Linte, "Image-guided interventions and computer-integrated therapy: Quo vadis?," *Med. Image Anal.* **33**, 56–63 (2016).
5. R. Walsh et al., "Design of a tracked ultrasound calibration phantom made of lego bricks," *Proc. SPIE* **9036**, 90362C (2014).
6. T. K. Chen et al., "A real-time freehand ultrasound calibration system with automatic accuracy feedback and control," *Ultrasound Med. Biol.* **35**, 79–93 (2009).
7. R. W. Prager et al., "Rapid calibration for 3-D freehand ultrasound," *Ultrasound Med. Biol.* **24**(6), 855–869 (1998).
8. E. C. Chen et al., "Solving for free-hand and real-time 3D ultrasound calibration with anisotropic orthogonal procrustes analysis," *Proc. SPIE* **9036**, 90361Z (2014).
9. A. Khamene and F. Sauer, "A novel phantom-less spatial and temporal ultrasound calibration method," in *Medical Image Computing and Computer-Assisted Intervention–MICCAI 2005*, pp. 65–72, Springer (2005).
10. P.-W. Hsu et al., *Freehand 3D Ultrasound Calibration: A Review*, Springer Verlag, Berlin Heidelberg (2009).
11. G. Ameri et al., "Synthetic aperture imaging in ultrasound calibration," *Proc. SPIE* **9036**, 90361I (2014).
12. G. Ameri et al., "Line fiducial material and thickness considerations for ultrasound calibration," *Proc. SPIE* **9415**, 941529 (2015).
13. J. M. Fitzpatrick, J. B. West, and C. R. Maurer Jr., "Predicting error in rigid-body point-based registration," *IEEE Trans. Med. Imaging* **17**(5), 694–702 (1998).
14. J. M. Fitzpatrick and J. B. West, "The distribution of target registration error in rigid-body point-based registration," *IEEE Trans. Med. Imaging* **20**(9), 917–927 (2001).
15. L. Mercier et al., "A review of calibration techniques for freehand 3-D ultrasound systems," *Ultrasound Med. Biol.* **31**(2), 143–165 (2005).
16. E. C. Chen, T. M. Peters, and B. Ma, "Guided ultrasound calibration: where, how, and how many calibration fiducials," *Int. J. CARS* **11**(6), 889–898 (2016).
17. J. M. Fitzpatrick, "Fiducial registration error and target registration error are uncorrelated," *Proc. SPIE* **7261**, 726102 (2009).
18. R. M. Comeau et al., "Intraoperative ultrasound for guidance and tissue shift correction in image-guided neurosurgery," *Med. Phys.* **27**(4), 787–800 (2000).
19. P. H. Schönemann, "A generalized solution of the orthogonal procrustes problem," *Psychometrika* **31**(1), 1–10 (1966).
20. B. K. Horn, "Closed-form solution of absolute orientation using unit quaternions," *J. Opt. Soc. Am. A* **4**(4), 629–642 (1987).
21. P. J. Besl and N. D. McKay, "Method for registration of 3-D shapes," *Proc. SPIE* **1611**, 586–606 (1992).
22. E. C. Chen et al., "Registration of 3D shapes under anisotropic scaling: anisotropic-scaled iterative closest point algorithm," *Int. J. CARS* **10**(6), 867–878 (2015).
23. E. C. Chen et al., "An iterative closest point framework for ultrasound calibration," in *Augmented Environments for Computer-Assisted Interventions*, pp. 69–79, Springer International Publishing, Switzerland (2015).
24. J. A. Jensen et al., "Synthetic aperture ultrasound imaging," *Ultrasonics* **44**, e5–e15 (2006).
25. W. H. Kim et al., "Synthetic aperture imaging in breast ultrasound: a preliminary clinical study," *Acad. Radiol.* **19**(8), 923–929 (2012).
26. A. H. Brandt et al., "Clinical evaluation of synthetic aperture harmonic imaging for scanning focal malignant liver lesions," *Ultrasound Med. Biol.* **41**(9), 2368–2375 (2015).
27. M. C. Hemmsen et al., "In vivo evaluation of synthetic aperture sequential beamforming," *Ultrasound Med. Biol.* **38**(4), 708–716 (2012).
28. Y. Tasinkevych et al., "Modified synthetic transmit aperture algorithm for ultrasound imaging," *Ultrasonics* **52**, 333–342 (2012).
29. A. Selfridge, G. Kino, and B. Khuri-Yakub, "A theory for the radiation pattern of a narrow-strip acoustic transducer," *Appl. Phys. Lett.* **37**(1), 35–36 (1980).
30. N. Bilaniuk and G. S. Wong, "Speed of sound in pure water as a function of temperature," *J. Acoust. Soc. Am.* **93**(3), 1609–1612 (1993).
31. "Aurora specifications (planar field)," 2017, <http://www.ndigital.com/medical/products/aurora/>.
32. K. Surry et al., "Poly (vinyl alcohol) cryogel phantoms for use in ultrasound and MR imaging," *Phys. Med. Biol.* **49**(24), 5529–5546 (2004).
33. A. Lasso et al., "Plus: open-source toolkit for ultrasound-guided intervention systems," *IEEE Trans. Biomed. Eng.* **61**(10), 2527–2537 (2014).
34. N. Otsu, "A threshold selection method from gray-level histograms," *Automatica* **11**(285–296), 23–27 (1975).
35. J. Canny, "A computational approach to edge detection," *IEEE Trans. Pattern Anal. Mach. Intell.* **PAMI-8**(6), 679–698 (1986).
36. M. Karaman, P.-C. Li, and M. O'Donnell, "Synthetic aperture imaging for small scale systems," *IEEE Trans. Ultrason. Ferroelectr. Freq. Control* **42**(3), 429–442 (1995).
37. J. Kortbek, J. A. Jensen, and K. L. Gammelmark, "Sequential beamforming for synthetic aperture imaging," *Ultrasonics* **53**(1), 1–16 (2013).
38. M. O. Donnell and Y. Wang, "Coded excitation for synthetic aperture ultrasound imaging," *IEEE Trans. Ultrason. Ferroelectr. Freq. Control* **52**(2), 171–176 (2005).

39. R. Y. Chiao and L. J. Thomas, "Synthetic transmit aperture imaging using orthogonal golay coded excitation," in *2000 IEEE Ultrasonics Symp.*, Vol. 2, pp. 1677–1680, IEEE (2000).
40. K. L. Gammelmark and J. A. Jensen, "Multielement synthetic transmit aperture imaging using temporal encoding," *IEEE Trans. Med. Imaging* **22**(4), 552–563 (2003).
41. G. Montaldo et al., "Coherent plane-wave compounding for very high frame rate ultrasonography and transient elastography," *IEEE Trans. Ultrason. Ferroelectr. Freq. Control* **56**(3), 489–506 (2009).
42. C. Bradley, "Retrospective transmit beam formation. ACUSON SC2000 volume imaging ultrasound system," Mountain View, California, https://www.siemens.com.tr/it/Assets/saglik/Whitepaper_Bradley.pdf (18 January 2017).
43. K. Thiele et al., *Exploring nSIGHT Imaging-A Totally New Architecture for Premium Ultrasound*, White paper, Koninklijke Philips NV (2013).
44. D. Napolitano et al., "Zone-based b-mode imaging," in *2003 IEEE Symp. on Ultrasonics*, Vol. 1, pp. 25–28, IEEE (2003).
45. A. Faraz et al., "Automatic real-time intra-operative 2D bi-plane ultrasound calibration during in situ minimal invasive heart surgery," *Int. J. Comput. Assisted Radiol. Surg.* **10**(Suppl. 1), S112 (2015).
46. J. M. Hansen and J. A. Jensen, "Compounding in synthetic aperture imaging," *IEEE Trans. Ultrason. Ferroelectr. Freq. Control* **59**(9), 2054–2065 (2012).
47. M. C. Hemmsen, J. H. Rasmussen, and J. A. Jensen, "Tissue harmonic synthetic aperture ultrasound imaging," *J. Acoust. Soc. Am.* **136**(4), 2050–2056 (2014).
48. C. H. Frazier and W. D. O'Brien, "Synthetic aperture techniques with a virtual source element," *IEEE Trans. Ultrason. Ferroelectr. Freq. Control* **45**(1), 196–207 (1998).
49. S. Korukonda and M. M. Doyley, "Visualizing the radial and circumferential strain distribution within vessel phantoms using synthetic-aperture ultrasound elastography," *IEEE Trans. Ultrason. Ferroelectr. Freq. Control* **59**(8), 1639–1653 (2012).
50. K. L. Gammelmark and J. A. Jensen, "Duplex synthetic aperture imaging with tissue motion compensation," in *2003 IEEE Symp. on Ultrasonics*, Vol. 2, pp. 1569–1573, IEEE (2003).
51. G. E. Trahey and L. F. Nock, "Synthetic receive aperture imaging with phase correction for motion and for tissue inhomogeneities. II. Effects of and correction for motion," *IEEE Trans. Ultrason. Ferroelectr. Freq. Control* **39**(4), 496–501 (1992).
52. M. Karaman, H. S. Bilge, and M. O'Donnell, "Adaptive multi-element synthetic aperture imaging with motion and phase aberration correction," *IEEE Trans. Ultrason. Ferroelectr. Freq. Control* **45**(4), 1077–1087 (1998).
53. M. F. Rasmussen and J. A. Jensen, "Comparison of 3-D synthetic aperture phased-array ultrasound imaging and parallel beamforming," *IEEE Trans. Ultrason. Ferroelectr. Freq. Control* **61**(10), 1638–1650 (2014).
54. S. I. Nikolov and J. A. Jensen, "Investigation of the feasibility of 3D synthetic aperture imaging," in *2003 IEEE Symp. on Ultrasonics*, Vol. 2, pp. 1903–1906 (2003).
55. A. D. Wiles et al., "A statistical model for point-based target registration error with anisotropic fiducial localizer error," *IEEE Trans. Med. Imaging* **27**(3), 378–390 (2008).
56. A. Danilchenko and J. M. Fitzpatrick, "General approach to first-order error prediction in rigid point registration," *IEEE Trans. Med. Imaging* **30**(3), 679–693 (2011).
57. B. Ma et al., "Estimation of optimal fiducial target registration error in the presence of heteroscedastic noise," *IEEE Trans. Med. Imaging* **29**(3), 708–723 (2010).

Golafsoun Ameri is a PhD candidate at the Imaging Research Laboratories at Robarts Research Institute. She received her bachelor of engineering degree in electrical engineering and a master of science degree in biomedical physics both from Ryerson University, Toronto, Canada. Her current research focus is on ultrasound-guided needle interventions.

John S. H. Baxter is a PhD candidate at the Imaging Research Laboratories at Robarts Research Institute in London, Canada. He received his bachelor's in software engineering from the University of Waterloo, Waterloo, Canada, in 2012. His research focuses on the use of variational optimization in the definition of and solution to medical image processing tasks, such as segmentation and image enhancement.

A. Jonathan McLeod is a PhD candidate at the Imaging Research Laboratories at Robarts Research Institute in London, Canada. He received his bachelor's in electrical engineering at the University of Western Ontario, London, Canada, in 2011. His research is focused on image-guided surgery.

Terry M. Peters is a scientist in the Imaging Research Laboratories at the Robarts Research Institute (RRI), London, Ontario, Canada, and professor in the Department of Medical Imaging and Medical Biophysics at Western University London, Canada, as well as a member of the graduate programs in neurosciences and biomedical engineering. He directs a research laboratory with a focus on research and development in the field of image-guided surgery and therapy.

Elvis C. S. Chen obtained his PhD from School of Computing, Queen's University (Kingston, Canada) in 2007. He works in the field of image-guided interventions, by applying techniques from robotics, computer vision, and computer graphics to the field of surgery. His research interests include joint kinematics, ultrasound-guided interventions, tool calibration and tracking, and vision-guided laparoscopy. Currently, he is a research associate at Robarts Research Institute with cross appointment (assistant professor) at Western University (London, Canada).