A low-cost multi-modal medical imaging system with fringe projection profilometry and 3D freehand ultrasound

Jhacson Meza[a], Pedro Simarra[a], Sara Contreras-Ojeda[a], Lenny A. Romero[b], Sonia H. Contreras-Ortiz[a], Fernando Arámbula Cosío[c], and Andrés G. Marrugo[a]

[a]Facultad de Ingeniería, Universidad Tecnológica de Bolívar, Cartagena, Colombia
[b]Facultad de Ciencias Básicas, Universidad Tecnológica de Bolívar, Cartagena, Colombia
[c]Institute of Applied Mathematics and System, Universidad Nacional Autónoma de México, Ciudad de México, 04510.

ABSTRACT

The growing need to perform surgical procedures, monitoring, and intervention of greater precision have led to the development of multimodal medical imaging systems. Multimodal images are a strategy to overcome the limitations of medical imaging technologies by combining the strengths of individual modalities or technologies. In this work, we propose a low-cost multimodal system that combines 3D freehand ultrasound with fringe projection profilometry to obtain information from the external and the internal structure of an object of interest. Both modalities are referred to a single coordinate system defined in the calibration to avoid post-processing and registration of the acquired images. The freehand ultrasound calibration results are similar to those previously reported in the literature using more expensive infrared tracking systems. The calibration reproducibility at the center point of the ultrasound image was 0.6202 mm for 8 independent calibrations. We tested our system on a breast phantom with tumors. Encouraging results show the potential of the system for applications in intraoperative settings.

Keywords: Ultrasound imaging, fringe projection profilometry, 3D imaging, 3D calibration, multi-modal imaging

1. INTRODUCTION

In recent years, multimodal imaging techniques have been used to overcome the limitations of conventional medical imaging technologies by combining the strengths of individual modalities. This is how systems have been developed that allow the integration, for example, of X-ray computed tomography with positron emission tomography (CT/PET) or monophotonic computed tomography with X-ray computed tomography (SPECT/CT), among others.[1] X-ray-CT is one of the most widely used medical imaging technologies because of its high spatial resolution and a good contrast of the underlying structures. However, CT does not yield information on soft tissue, function or biological activity, and its usefulness in intraoperative environments is limited.

There are other techniques free of ionizing radiation that allow obtaining structural and functional information, even in real time, which facilitates its use in intraoperative environments for the guidance of surgical procedures and monitoring. Ultrasound (US) imaging is one of the most frequently used modalities for diagnosis and therapy. Its applications include cardiology, urology, obstetrics, and gynecology.[2] US 3D allows the visualization and measurement of the volume of 3D structures of interest, however the displayed volume is limited to the size of the probe.[3] For this reason, most of the procedures for guiding, for example, needles for taking biopsies are carried out using US 2D. There are even techniques that allow obtaining 3D volumes from US 2D images by tracking the probe and the corresponding record of the images.[4–6] US has also recently been used in other procedures, which conventionally use CT, for example for the evaluation of scoliosis,[7] for the guidance of molecular fluorescence tomography,[8] among others.

Further author information: (Send correspondence to J.M., and A.G.M.)
J.M.: E-mail: jhacsonmeza@outlook.com
A.G.M.: E-mail: agmarrugo@utb.edu.co

Despite the success of 2D US, it has certain disadvantages. It does not allow the visualization of planes parallel to the skin, the assessment of the evolution of a disease is underestimated by the difficulty in positioning the probe in the same position during different tests, and in the case of 3D US, the displayed volume is very small. Ultrasound may be combined with other imaging techniques to overcome some of these disadvantages. There are different optical techniques that using visible or infrared light obtain images for diagnosis and guidance of medical procedures. For example, the use of structured light systems through which the topography of a surface is obtained, have allowed the non-contact evaluation of different medical conditions like scoliosis,[9] to perform dermatological tests,[10] and even the development of multimodal systems for intraoperative guidance combined with molecular fluorescence.[11]

In general, multimodal image systems require image registration, especially when they are acquired independently and not synchronized. The registration of multimodal images is difficult since the correspondence problem is not guaranteed, and it is required to have characteristics that are present in all the modalities as points of reference to estimate the appropriate geometric transformation. On the other hand, the techniques that are acquired simultaneously require special calibration to be able to obtain the different modalities or images in the same frame of reference. This circumstance is a major challenge because the different image systems can have different resolutions, field of observation, signal-to-noise ratio, among other parameters. In this work, we propose a low-cost multimodal system that combines freehand-US with fringe projection profilometry to obtain information from the external and internal structure of an object of interest with a robust performance. In the following sections, we briefly describe the background of the different imaging modalities and the theory behind the different calibration strategies. Finally, we present preliminary results of the implemented multimodal system.

2. BACKGROUND

2.1 3D Freehand Ultrasound

3D freehand ultrasound is a non-invasive, inexpensive, safe, and portable medical imaging technique[12] used in many clinical applications such as quantification of plaque volume and morphology,[13] brain shift evaluation during neurosurgery,[14] radiotherapy planning,[15] prostate biopsy,[16] surveillance of arteriovenous fistulae,[17] etc. Freehand ultrasound imaging technique consists of acquiring 2D ultrasound images (B-scans) and tracking the position and orientation of a probe in space with a position sensor or an optical or electromagnetic tracker system. Thus, knowing the position of the probe and its relation with the ultrasound image plane, these B-scans can be located in a volume or 3D space in a fixed global coordinate system. We can establish this relation (rigid-body transformation) through a calibration procedure using a phantom, that is, an object with known geometry which helps us to calculate the transformation matrix by identifying its characteristics in the US images, and in the physical phantom space[18] using an iterative optimization method.[19] There are many phantoms proposed in the literature: cross-wire, three-wire, single-wall, Cambridge phantom, among others.[19] But the simplest is a point target[20] scanned from different positions and orientations, which we can build with cross-wires, using a spherical object, or the tip of a stylus.

2.2 Fringe Projection Profilometry

Optical 3D shape measurement based on digital fringe projection profilometry (FPP) is a well-known optical metrology technique with many biomedical, industrial, and entertainment applications due to its high-speed data processing, accuracy, and flexibility.[21] In FPP, we project a periodic structured light pattern (typically a sinusoidal intensity distribution) on the surface of an object which due to the topographic variations is distorted. We register the deformed pattern with a camera and use it to calculate its phase distribution since the depth information is encoded in it. There are two main methods to retrieve the phase from the fringe images: phase shifting[22] and Fourier transform.[23] On the one hand, phase-shifting allows us to pixel-by-pixel estimate a phase-map using at least three patterns, achieving high-accuracy, high-resolution and surface reflectivity robustness results, especially for irregular object surfaces.[24] On the other hand, Fourier transform profilometry (FTP) is used to calculate phase-maps with a single structured light pattern through Fourier analysis, which makes it faster than phase-shifting and suitable for dynamic scene measurements, but this method is quite sensitive to complex surfaces, i.e., we are only able to measure smooth topographies of objects with FTP.

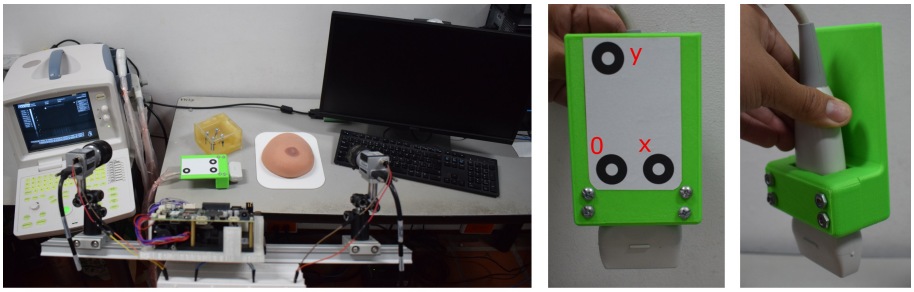


Figure 1. (a) Experimental setup. (b)-(c) Target attached to the probe to track it with a piece created in a 3D printer.

Phase obtained with both phase-shifting and Fourier transform is a wrapped phase map since its values range from $-\pi$ to π with 2π discontinuities because of the arctangent function used in the phase calculation. Due to the above, an unwrapping procedure is necessary to estimate a continuous or unwrapped phase-map identifying and removing the discontinuities by adding or subtracting 2π integer numbers in the phase. Algorithms for phase unwrapping are classified into two groups: spatial[25] and temporal.[26] The spatial algorithms detect the 2π jumps using a single wrapped phase-map based on the neighboring pixels. This condition implies that the surface cannot have 2π discontinuities in the wrapped phase. The temporal algorithms, in contrast, allow the unwrapping of the phase-map of isolated objects or with abrupt topography changes using more than one discontinuous phase-map. Finally, with an unwrapped phase distribution, we need to map phase values to physical metric values, and for that, we need our FPP system to be calibrated.

3. MATERIALS AND METHOD

Our experimental setup shown in Fig. 1(a), consists of two monochromatic CMOS cameras (Basler acA1300-200um; 1280 x 1024; 203 fps), a DLP projector (LightCrafter 4500), a B-mode ultrasound machine (Biocare, model iS 20) with a linear and curvilinear probe, a VGA to USB capture card for US image acquisition (Epiphan AV.io HD), and a breast phantom with tumors (3B SONOtrain P125). Moreover, we use a cross-wire phantom for freehand US calibration, a black and white (B/W) checkerboard with 10×10 mm squares for cameras and projector calibration, a black and white target of three coplanar circles attached to the linear probe for the tracking, and a PC workstation.

3.1 Freehand Ultrasound System and Calibration

We use an ultrasound machine Biocare iS 20 to acquire B-mode US images, and the linear transducer configured at 7.5 MHz frequency and a 7 cm depth. For tracking the transducer position and orientation, we use a stereo-vision system and an in-house developed acquisition software where we simultaneously acquire US images and images from the two cameras.

3.1.1 Transducer tracking

To track the probe pose, we can use a mechanical, optical, or electromagnetic system. In many recent works commercial optical tracker systems have been extensively used, but often these devices are expensive.[16,17,27,28] Here, we propose a low-cost optical strategy for tracking the transducer pose using a stereo-vision system composed of two conventional monochromatic cameras, and a target formed by three coplanar sets of two concentric circles (one black and one white) attached to the probe with a piece created in a 3D printer as shown in Fig. 1(b)-(c). The proposed tracking procedure was carried out through triangulation and using epipolar geometry constraints.[29]

With the estimated centroids of the three coplanar circles in both images, we need to match them to calculate their 3D coordinates. We exploit the epipolar geometry constraints to find the correct correspondences between the points. The epipolar geometry is the intrinsic projective geometry between two views. It does not depend on the structure of the scene but in the relative pose of the cameras and their intrinsic parameters. Given a point $\mathbf{x}_1$ in camera 1, we know how $\mathbf{x}_2$ is constrained, i.e., point $\mathbf{x}_2$ in the image plane 2 lies in the epipolar line $\mathbf{l}_2$ (Fig. 2(a)).

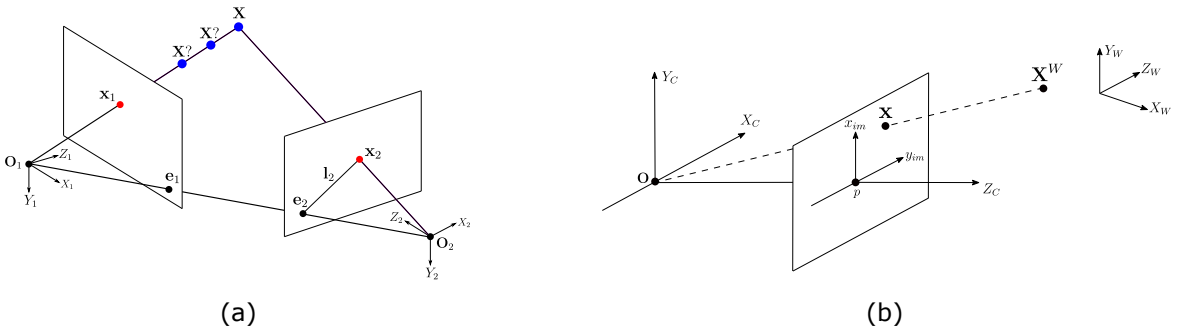


Figure 2. (a) Epipolar geometry: the intrinsic geometry of two views. (b) Pinhole camera model.

The epipolar geometry is described by the fundamental matrix $\mathbf{F}$, a 3×3 matrix that relates the correspondences of points

$$\mathbf{x}_2^{\mathsf{T}} \mathbf{F} \mathbf{x}_1 = 0 \ . \tag{1}$$

Then, given a center $\mathbf{x}_1$ in pixels of our coplanar circles in homogeneous coordinates in image 1, the respective epipolar line in the image 2 is

$$\mathbf{l}_2 = \mathbf{F} \mathbf{x}_1 \ . \tag{2}$$

With $\mathbf{l}_2$ we search for the center in image 2 closer to the epipolar line calculated, that is, we measure the distance between all the centers in camera 2 and the epipolar line, and we select the closest point to $\mathbf{l}_2$ as the corresponding point of $\mathbf{x}_1$. We perform the above for all the centers in the first view.

Once the correspondences are established, we estimate the 3D coordinates of the three points through triangulation. Based on the pinhole camera model in Fig. 2(b), the projection in homogeneous coordinates in the image plane of a 3D point $\mathbf{X}^W = [X,\, Y,\, Z,\, 1]^{\mathsf{T}}$ given in a world coordinate system can be described as

$$s\mathbf{x} = \mathbf{K}\, {}^{C}\mathbf{M}_W\, \mathbf{X}^W \ , \tag{3}$$

where s is a scale factor, $\mathbf{x} = [x,\, y,\, 1]^{\mathsf{T}}$ is the homogeneous 2D coordinates of the projected 3D point,

$$\mathbf{K} = \begin{bmatrix} f_x & \gamma & x_0 \\ 0 & f_y & y_0 \\ 0 & 0 & 1 \end{bmatrix} \ , \tag{4}$$

is the intrinsic matrix which contains the intrinsic parameters of the camera, and ${}^{C}\mathbf{M}_W = [\mathbf{R} \mid \mathbf{t}]$ is the extrinsic matrix which describes the orientation with a 3×3 rotation matrix $\mathbf{R}$ and position with a 3×1 translation vector $\mathbf{t}$ of the world frame $(X_W,\, Y_W,\, Z_W)$ relative to the camera frame $(X_C,\, Y_C,\, Z_C)$. From now on, we refer to $\mathbf{P} = \mathbf{K}\, {}^{C}\mathbf{M}_W$ as the projection matrix. Note that vectors $\mathbf{x}$ and $\mathbf{P}\mathbf{X}^W$ are not equal, they have the same direction (are parallel) but differ by a scale factor s.

In the case of two views (camera 1 and camera 2) that see the same point $\mathbf{X}$, based on Eq. 3 we have

$$s_1 \mathbf{x}_1 = \mathbf{P}_1 \mathbf{X} \ , \qquad s_2 \mathbf{x}_2 = \mathbf{P}_2 \mathbf{X} \ , \tag{5}$$

and we need to solve for $\mathbf{X}$. The homogeneous (DLT) solution to this problem is given by:

$$\begin{bmatrix} x_1 \mathbf{p}_1^{3\mathsf{T}} - \mathbf{p}_1^{1\mathsf{T}} \\ y_1 \mathbf{p}_1^{3\mathsf{T}} - \mathbf{p}_1^{2\mathsf{T}} \\ x_2 \mathbf{p}_2^{3\mathsf{T}} - \mathbf{p}_2^{1\mathsf{T}} \\ y_2 \mathbf{p}_2^{3\mathsf{T}} - \mathbf{p}_2^{2\mathsf{T}} \end{bmatrix} \mathbf{X} = \mathbf{0} \ , \tag{6}$$

where $\mathbf{p}_1^i$ and $\mathbf{p}_2^i$ represent the i-th row of $\mathbf{P}_1$ and $\mathbf{P}_2$ respectively, as column vectors.

As seen in Fig. 1(b)-(c) the target of the three circles is in a plane attached to the probe with a piece created in a 3D printer. As these three points lie in a plane, we can define a coordinate system with them, which we

refer to as the target or probe frame. These three circles are labeled as shown in Fig. 1(b), where we define the center of the circle labeled with 0 as the origin of the target frame. Therefore, the 3D position of this point represents the translation vector $\mathbf{t}$ of the probe frame respect to the cameras. Furthermore, the circles labeled with x and y represent the direction of the x-axis and y-axis of the target coordinate system, respectively. With the 3D points 0-label and x-label we can estimate the unit vector $\hat{\mathbf{x}}$ of the transducer frame, and likewise using the 3D points 0-label and y-label we estimate the unit vector $\hat{\mathbf{y}}$. The unit vector in direction of the z-axis is calculated using the cross product: $\hat{\mathbf{z}} = \hat{\mathbf{x}} \times \hat{\mathbf{y}}$. Finally, the rotation matrix of the target frame relative to the stereo-vision system is $\mathbf{R} = [\hat{\mathbf{x}} \;\; \hat{\mathbf{y}} \;\; \hat{\mathbf{z}}]$. In this way we track the position $\mathbf{t}$ and orientation $\mathbf{R}$ of the target or transducer coordinate system base on the stereo images, but for that we need to calibrate the two cameras in order to know their intrinsic parameters and the extrinsic of the second camera (pose relative to camera 1) using the camera 1 as the world frame. For this stereo calibration, we use the B/W checkerboard.[30]

3.1.2 Transducer calibration

For the calibration procedure, we used a point phantom because of its ease of construction and precise and accurate calibration results.[12,19,31] The phantom is formed by a pair of cross-wires built with two cotton threads submerged in a $10 \times 10 \times 7.1$ cm water bath. The point phantom has several limitations. Many images had to be acquired from different perspectives of the point target to estimate the unknown transformation matrices and scale factors involved in the spatial calibration. Additionally, the alignment of the ultrasound image plane with the target location is difficult due to the finite thickness of the ultrasound beam.[20] Furthermore, automatic segmentation of the cross-wire point in the B-scans it is also tough, and instead, manual segmentation is often adopted.[12,32]

In Figure 3, we illustrate the spatial relationships of the five coordinate systems involved in the probe calibration. Calibration in a 3D freehand ultrasound system requires tracking the position and orientation of the probe in space and finding its geometric relationship with the US image plane. In other words, using the transformation matrix ${}^{W}\mathbf{T}_{T}$ from the transducer coordinate system $\{T\}$ to a stationary world coordinate frame $\{W\}$, we need to estimate the transformation ${}^{T}\mathbf{T}_{I}$ from the image frame $\{I\}$ to the transducer frame $\{T\}$. Furthermore, we also need to calculate the x and y scales of the US image s_x and s_y in millimeters per pixel, in order to convert a pixel of the B-scan to metric units.

We use the camera 1 of a stereo-vision system as the world reference frame $\{W\}$ and with the second camera $\{Cam_2\}$ we estimate the transformation ${}^{W}\mathbf{T}_{T}$ through triangulation. The cross-wire phantom coordinate system $\{F\}$ is placed in such a way that its origin coincides with the point of crossing wires as seen in Fig. 3. If we acquire a B-scan of the cross-wire phantom, we can map the segmented image coordinate of the point target (x, y) in pixels to the physical phantom frame $\{F\}$ as

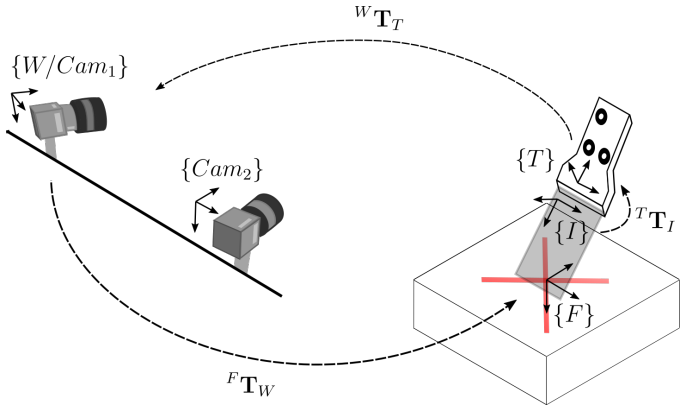


Figure 3. Transducer calibration: geometric relations and reference frames.

$$\begin{bmatrix} 0 \\ 0 \\ 0 \\ 1 \end{bmatrix} = {}^{F}\mathbf{T}_{W}\,{}^{W}\mathbf{T}_{T}\,{}^{T}\mathbf{T}_{I} \begin{bmatrix} s_x x \\ s_y y \\ 0 \\ 1 \end{bmatrix} \quad , \tag{7}$$

where, using the X-Y-Z Euler fixed-angles scheme, each rigid transformation matrix can be written as

$${}^{j}\mathbf{T}_{i}(t_x, t_y, t_z, \alpha, \beta, \gamma) = \begin{bmatrix} \cos\alpha\cos\beta & \cos\alpha\sin\beta\sin\gamma - \sin\alpha\cos\gamma & \cos\alpha\sin\beta\cos\gamma + \sin\alpha\sin\gamma & t_x \\ \sin\alpha\cos\beta & \sin\alpha\sin\beta\sin\gamma + \cos\alpha\cos\gamma & \sin\alpha\sin\beta\cos\gamma - \cos\alpha\sin\gamma & t_y \\ -\sin\beta & \cos\beta\sin\gamma & \cos\beta\cos\gamma & t_z \\ 0 & 0 & 0 & 1 \end{bmatrix} \quad . \tag{8}$$

The expression in Eq. 7 give us three equations with 11 unknowns: two scale factors s_x and s_y, three translations and three rotations angles from the ${}^{T}\mathbf{T}_{I}$ matrix, and only three translations in ${}^{F}\mathbf{T}_{W}$ since we can set the rotation angles to zero because we do not need to estimate the orientation of the coordinate system $\{F\}$. Note that the transformation ${}^{W}\mathbf{T}_{T}$ is known and is given by the stereo system using the transducer tracking method described in Section 3.1.1. If we use a total of N B-scans of the point target in different positions and orientations, we calculate these 11 unknowns by minimizing

$$\sum_{i=1}^{N} \left| {}^{F}\mathbf{T}_{W}\,{}^{W}\mathbf{T}_{Ti}\,{}^{T}\mathbf{T}_{I}\,\mathbf{X}_{i}^{I} \right| \quad , \tag{9}$$

where $\mathbf{X}_{i}^{I} = [s_x x_i,\ s_y y_i,\ 0,\ 1]^{\mathsf{T}}$ is the i-th coordinate of the point phantom given in the frame $\{I\}$. This is a non-linear minimization problem which can be solved using an iterative optimization method as the Levenberg-Marquardt algorithm.[33,34]

After estimating the unknowns involved in the calibration, we can discard the transformation matrix ${}^{F}\mathbf{T}_{W}$, which is only used to calculate the other parameters. Moreover, with the scale factors and the matrix ${}^{T}\mathbf{T}_{I}$ we can map a point from the US image to the world coordinate frame

$$\begin{bmatrix} X \\ Y \\ Z \\ 1 \end{bmatrix} = {}^{W}\mathbf{T}_{T}\,{}^{T}\mathbf{T}_{I} \begin{bmatrix} s_x x \\ s_y y \\ 0 \\ 1 \end{bmatrix} \quad . \tag{10}$$

3.2 Fringe Projection System and Calibration

The digital fringe projection system used consists of a camera and a projector. As seen in Fig. 1(a) of our experimental setup, we have two cameras and a projector. We use these two cameras for the transducer pose tracking, but for the FPP system, we use only camera 1 and the DLP projector for the 3D reconstruction system, that is, the same camera defined as the world frame in the 3D freehand ultrasound system (Fig. 3). Also, we use this camera as the world coordinate frame in the fringe projection system to obtain 3D data from both freehand ultrasound and fringe projection profilometry in the same absolute coordinate system.

Here we introduce a new frame, the digital projector coordinate system $\{P\}$, which with the camera 1 constitutes a stereo-vision system since we can consider the projector as an inverse camera.[35] This consideration is viable because of the projector projects images instead of capturing them, and moreover, optically both camera and projector are the same.

3.2.1 Image formation and fringe analysis

A typical sinusoidal fringe pattern generated with a computer to project onto an object can be described as

$$I^{P}(x, y) = a^{P} + b^{P}\cos[2\pi x/p_0] \quad , \tag{11}$$

where a^P is the mean intensity value, b^P is the amplitude or projector modulation, and p_0 is the fringe (period) pitch or the number of pixels for the fringe period. Commonly a^P and b^P are both set to 255/2.

Once the structured light $I^P(x,y)$ is projected onto the surface, the deformed sinusoidal pattern captured by the camera can be mathematically represented as

$$I^C(x,y) = A^C(x,y) + B^C(x,y)\cos[\phi(x,y)] \ , \tag{12}$$

where $A^C(x,y)$ is the average intensity or DC component of the fringe image, $B^C(x,y)$ is the intensity modulation or half of the peak-to-valley intensity modulation, and $\phi(x,y)$ is the phase distribution of the distorted fringe, which is related to the temporal phase difference of $I^C(x,y)$ relative to the projected image $I^P(x,y)$.

The phase $\phi(x,y)$ encodes the depth information and can be regarded as quantification of fringe distortion, thereby in FPP the phase retrieval procedure is an important step. For calibration and 3D reconstruction, we use the standard N-step phase-shifting method to estimate the phase map $\phi(x,y)$.[24] In Eq.12 we have only three unknowns: $A^C(x,y)$, $B^C(x,y)$ and $\phi(x,y)$, and we can use a set of at least three fringe images with phase-shifted patterns to solve for $\phi(x,y)$. Then, we introduce a phase shift of δ_k in the projected images and the k-th distorted fringe pattern acquired by the camera is

$$I^C_k(x,y) = A^C(x,y) + B^C(x,y)\cos[\phi(x,y) + \delta_k] \ , \tag{13}$$

for $k = 1, 2, 3, ..., N$. These N images give us N equations with three unknowns. If $\delta_k = 2\pi k/N$ a least-squares solution for the phase is

$$\phi(x,y) = -\arctan\left(\frac{\sum_{k=1}^{N} I^C_k \sin\delta_k}{\sum_{k=1}^{N} I^C_k \cos\delta_k}\right) \ . \tag{14}$$

As a result of the arctangent function, the phase calculated with Eq. 14 is a wrapped phase map, due to its values range from $-\pi$ to π with 2π discontinuities. A phase unwrapping algorithm where we remove these 2π jumps needs to be applied to obtain a continuous phase map for accurate measurements. These discontinuities are removed by adding or subtracting integer multiples of 2π in the 2π jumps locations. Thus, the unwrapped phase map $\Phi(x,y)$ is given by

$$\Phi(x,y) = \phi(x,y) + 2\pi k(x,y) \ , \tag{15}$$

where $k(x,y)$ are the integer numbers which are regarded as the fringe order. If $k(x,y)$ is uniquely estimated for each pixel, then $\Phi(x,y)$ represents the absolute phase map. We use a centerline image and spatial phase unwrapping to retrieve absolute phase at each pixel.[35–37]

3.2.2 Stereo-vision system calibration

The calibration strategy carried out is mainly based on the proposed by Zhang and Huang,[35] where we regard the projector as an inverse of a camera, which allows us to calibrate the projector in the same manner as a camera. With this approach, the camera and projector can be modeled using stereo-vision principles, and the calibration process is performed like that of a traditional stereo system using a B/W checkerboard placed in arbitrary positions and orientations.

To use the traditional stereo-vision calibration approach, we need to establish point correspondences between the camera and the projector using the absolute phase. We project eight vertical phase-shifted images (Fig. 4(a)) to estimate the vertical wrapped phase, and a centerline (Fig. 4(b)) for absolute phase-retrieval $\Phi^C_v(x,y)$ where we assume that points in this line have zero absolute phase. Also, eight horizontal phase-shifted patterns (Fig. 4(c)) and a horizontal centerline (Fig. 4(d)) were projected for the absolute phase estimation $\Phi^C_h(x,y)$. Finally, we project white light (Fig. 4(e)) in order to use this image for checkerboard corners estimation.

With the horizontal and vertical absolute phase maps, we can find the corresponding point in the projector $(x^P,\, y^P)$ of a given point $(x^C,\, y^C)$ in the camera

$$x^P = \frac{p_0}{2\pi}\Phi^C_v(x^C, y^C) + \frac{w}{2} \ , \tag{16}$$

$$y^P = \frac{p_0}{2\pi}\Phi^C_h(x^C, y^C) + \frac{h}{2} \ , \tag{17}$$

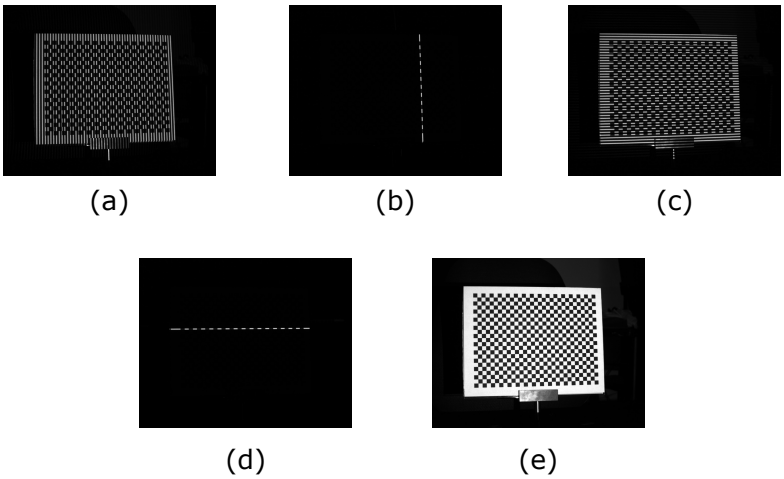


Figure 4. Images acquired in each checkerboard pose: (a) vertical fringes, (b) vertical centerline, (c) horizontal fringes, (d) horizontal centerline, and (e) checkerboard with pure white image projection.

where p_0 is the pitch of the vertical and horizontal fringes. w is the width and h is the height of the DMD projector sensor, or the number of columns and rows, respectively.

With Eq. 16 and Eq. 17 we establish the correspondences of points between the detected corners from the B/W checkerboard in the camera image and the projector corners in the DMD image. With the matching points we calculate the intrinsic matrix $\mathbf{K}$ of the camera and the projector, and moreover, the extrinsic matrix ${}^{P}\mathbf{M}_{W}$ relative to the projector by aligning the world coordinate system with the camera frame as discussed in Section. 3.1.1.

We also compensate for lens distortion which is not considered in the pinhole camera model (Fig. 2(b)), due to the lens distortions can introduce errors in the stereo-vision model.[38,39] The mathematical model for the lens distortion is given by

$$\begin{bmatrix} x_d \\ y_d \end{bmatrix} = (1 + k_1 r_n^2 + k_2 r_n^4 + k_3 r_n^6) \begin{bmatrix} x_d \\ y_d \end{bmatrix} + \begin{bmatrix} 2p_1 x_n y_n + p_2(r_n^2 + 2x_n^2) \\ 2p_2 x_n y_n + p_1(r_n^2 + 2y_n^2) \end{bmatrix} \quad , \tag{18}$$

where $r_n^2 = x_n^2 + y_n^2$. k_1, k_2 and k_3 are the radial distortion coefficients and p_1 and p_2 are the tangential coefficients. $[x_d,\, y_d]^{\mathsf{T}}$ are the distorted image coordinates and $[x_n,\, y_n]^{\mathsf{T}}$ the normalized image coordinates.

3.2.3 Phase-to-coordinates mapping

With the calibrated intrinsic and extrinsic parameters of the camera and the projector, we can estimate the metric 3D coordinates of an object by projecting phase-shifted fringes and centerline images onto the surface to be reconstructed.

If we align the world coordinate system with the camera coordinate system we have: ${}^{C}\mathbf{M}_{W} = [\mathbf{I} \mid \mathbf{0}]$, where $\mathbf{I}$ is the 3×3 identity matrix and $\mathbf{0}$ is a 3×1 zero vector. And ${}^{P}\mathbf{M}_{W} = [\mathbf{R} \mid \mathbf{t}]$, where $\mathbf{R}$ and $\mathbf{t}$ represents the position and orientation of the world relative to the projector frame.

Based on the absolute phase map Φ^C of the surface of an object, estimated with fringe patterns with a pitch p_0 and a centerline, we can relate each camera point $(x^C,\, y^C)$ with a line with the same absolute phase value on the projector with the expression

$$x^P = \frac{p_0}{2\pi} \Phi^C(x^C, y^C) + \frac{w}{2} \quad . \tag{19}$$

Using Eq. 3 of the projection of a 3D point in the 2D sensor, $s^C\mathbf{x}^C = \mathbf{P}^C\mathbf{X}$, $s^P\mathbf{x}^P = \mathbf{P}^P\mathbf{X}$, we have

$$s^C \begin{bmatrix} x^C \\ y^C \\ 1 \end{bmatrix} = \begin{bmatrix} p_{11}^C & p_{12}^C & p_{13}^C & p_{14}^C \\ p_{21}^C & p_{22}^C & p_{23}^C & p_{24}^C \\ p_{31}^C & p_{32}^C & p_{33}^C & p_{34}^C \end{bmatrix} \begin{bmatrix} X \\ Y \\ Z \\ 1 \end{bmatrix} \quad , \quad s^P \begin{bmatrix} x^P \\ y^P \\ 1 \end{bmatrix} = \begin{bmatrix} p_{11}^P & p_{12}^P & p_{13}^P & p_{14}^P \\ p_{21}^P & p_{22}^P & p_{23}^P & p_{24}^P \\ p_{31}^P & p_{32}^P & p_{33}^P & p_{34}^P \end{bmatrix} \begin{bmatrix} X \\ Y \\ Z \\ 1 \end{bmatrix} \quad . \tag{20}$$

Finally, with Eq. 19 and Eq. 20 we have a system of seven unknowns (X, Y, Z, s^C, s^P, x^P, y^P) and seven equations, where we can determine uniquely $\mathbf{X}$ with the expression

$$\begin{bmatrix} X \\ Y \\ Z \end{bmatrix} = \begin{bmatrix} p_{11}^C - p_{31}^C x^C & p_{12}^C - p_{32}^C x^C & p_{13}^C - p_{33}^C x^C \\ p_{21}^C - p_{31}^C x^C & p_{22}^C - p_{32}^C x^C & p_{23}^C - p_{33}^C x^C \\ p_{11}^P - p_{31}^P x^P & p_{12}^P - p_{32}^P x^P & p_{13}^P - p_{33}^P x^P \end{bmatrix}^{-1} \begin{bmatrix} p_{34}^C x^C - p_{14}^C \\ p_{34}^C y^C - p_{24}^C \\ p_{34}^P x^P - p_{14}^P \end{bmatrix} \quad . \tag{21}$$

4. EXPERIMENTS AND RESULTS

4.1 Freehand Calibration Assessment

To evaluate the transducer quality calibration, we measure the precision with the calibration reproducibility (CR) using the approach proposed by Prager et al.[19] and by Hsu et al.[32] Calibration reproducibility measures the dispersion in the reconstructed 3D position of a point from the ultrasound image using N calibration parameters estimated through N different calibrations. As a result of these different calibrations, we have N scale factors s_x and s_y and N transformations ${}^T\mathbf{T}_I$. Thus, a specific point (x, y) in the B-scan is reconstructed from the image plane to the transducer coordinate system, and in this frame, we measure the variability of this point.

Using the approach proposed by Prager et al.,[19] they use the bottom right pixel (x_{max}, y_{max}) of the B-scan to measure CR under two different calibrations with the expression

$$\mu_{CR1} = \left| {}^T\mathbf{T}_{I1}\, \mathbf{X}_1^I - {}^T\mathbf{T}_{I2}\, \mathbf{X}_2^I \right| \quad , \tag{22}$$

where $\mathbf{X}_i^I = [s_{xi} x_{max}, s_{yi} y_{max}, 1, 0]^\intercal$. We calculate CR with this expression using all possible pairs of calibration parameters and finally the mean of this values is reported as the final precision.

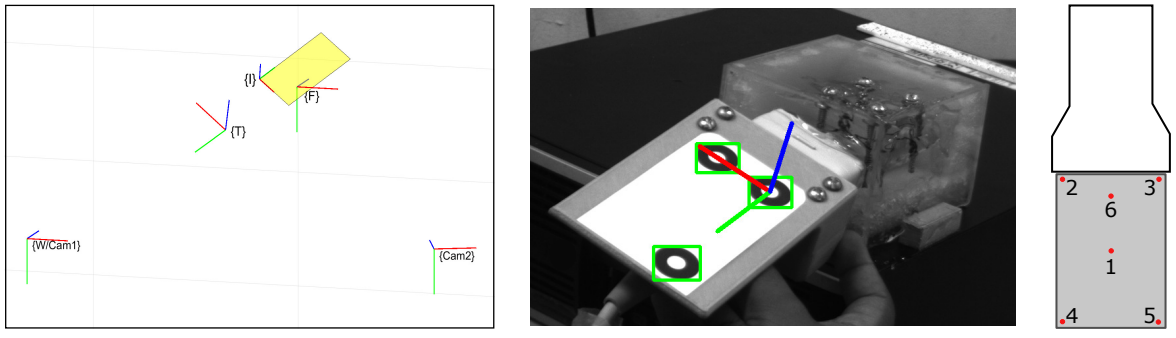
With the proposed strategy by Hsu et al.,[32] calibration reproducibility is assessed as

$$\mu_{CR2} = \frac{1}{N} \sum_{i=1}^{N} \left| {}^T\mathbf{T}_{Ii}\, \mathbf{X}_i^I - \bar{\mathbf{X}}^T \right| \quad , \tag{23}$$

where $\bar{\mathbf{X}}^T$ is the mean in each dimension of all the reconstructed points $\mathbf{X}_i^I$ in the transducer $\{T\}$ coordinate system. As trial points, authors use the center and the four corners of the image.

In this work we carried out a total of eight calibrations with 50 images each one using a linear transducer set to 7 cm depth. In Fig. 5(a) we show an example of the 3D reconstructed coordinate systems involved in the probe calibration using the transducer pose shown in Fig. 5(b).

Table 1 shows the results of the calibration precision using Eq. 22 and Eq. 23. In Fig. 5(c) we see the trial points used in the precision estimation. We report the dispersion at the center of the B-scan (P_1), the



(a) (b) (c)

Figure 5. (a) Reconstruction of the coordinate systems involved in the transducer calibration using the probe pose in (b). (b) Probe pose in the acquisition of a B-scan of the cross-wire phantom drawn in the image. (c) Trial points used in the precision estimation.

Trial point	μ_{CR1} (mm)	μ_{CR2} (mm)
Center	0.9928	0.6202
Bottom right (x_{max}, y_{max})	3.0350	1.9195
Mean (center and four corners)	1.9040	1.2163
Mid-point at 1.5 cm depth	0.2616	0.4019

Table 1. Precision assessment: calibration reproducibility results with 8 calibrations.

dispersion at the four corners ($P_2 - P_5$) and center (P_1), and at a mid-point (P_6) at approximately 1.5 cm depth. Pixel coordinates in the B-scan at this depth was estimated using the final scale factors calculated with the transducer calibration. The aim of use this point is to be compared with the precision reported in Ref. 32 of 0.27 mm measured in the center of the image, where authors use a point phantom and the probe at 3 cm depth. Furthermore, Lindseth et al.[40] report a CR at the center of the B-scan of 0.62 mm with a point phantom and a linear probe at 8 cm depth.

Finally, we also evaluate the RMS error of all the equations obtained in each calibration at the optimal solution found with the Levenberg–Marquardt algorithm. We report a mean RMS error through each calibration of 0.5031 mm.

4.2 Freehand US and 3D Reconstruction

We carried out an experiment where we use both 3D imaging techniques (FPP and Freehand US) with the two-camera stereo system and the projector. The combined system was fixed through both acquisitions. Camera 1 acted as the world coordinate system. First, we projected a total of eight phase-shifted fringe images onto the surface of the breast phantom for the phase-retrieval with phase-shifting. We used an additional centerline image for absolute phase estimation. The 3D reconstruction in the reference frame $\{W\}$ is shown in Fig. 6(a). Afterward, we acquire a B-scan from the phantom of a tumor located near the nipple of the breast, and we map this image in the world frame $\{W\}$ using Eq. 10 and the pose of the target estimated with the stereo-vision cameras. The result of both reconstructions relative to $\{W\}$ is shown in Fig. 6(b). Note that both reconstructions are in the same coordinate system, and the reconstructed B-scan is totally within the phantom surface. The result is in agreement with the expected position of the tumor.

5. CONCLUSIONS

In this work, we have proposed a low-cost multi-modal medical imaging system capable of acquiring the 3D surface of an object via fringe projection profilometry and detecting the internal structure via freehand ultrasound. Both modalities are referred to a single coordinate system defined in the calibration to avoid post-processing and

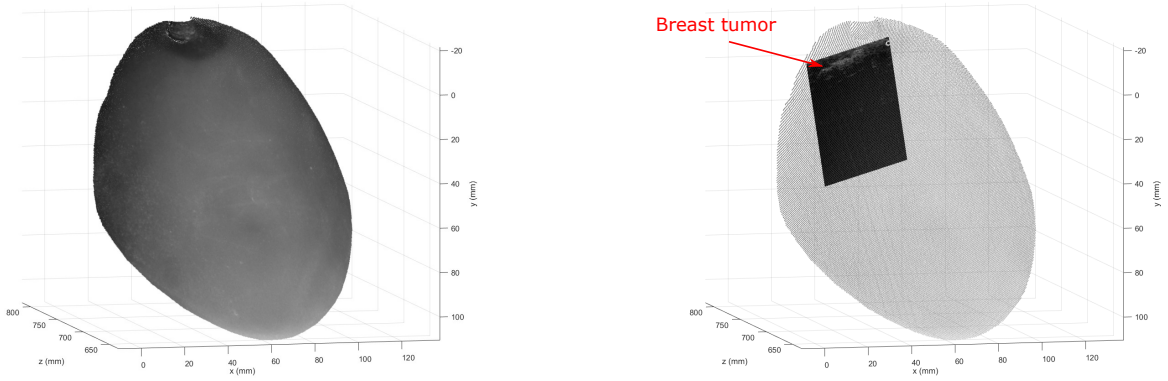


Figure 6. Reconstruction results using FPP and Freehand US: (a) 3D reconstructed surface of the breast phantom. (b) B-scan of a tumor from the breast mapped to the world frame with the 3D phantom surface in the same coordinate system.

registration of the acquired images. The freehand ultrasound calibration results are similar to those previously reported in the literature using more expensive infrared tracking systems. We carried out an experiment in a breast phantom with tumors. Encouraging results show the potential of the system, and future work involves quantitative evaluation of the multi-modal precision. The complementary nature of both methods allows for potential applications in intraoperative settings.

ACKNOWLEDGMENTS

This work has been partly funded by Universidad Tecnológica de Bolívar projects C2018P005 and C2018P018. J. Meza thanks Universidad Tecnológica de Bolívar for a post-graduate scholarship.

REFERENCES

[1] James, M. L. and Gambhir, S. S., "A Molecular Imaging Primer: Modalities, Imaging Agents, and Applications," *Physiological Reviews* **92**(2), 897–965 (2012).

[2] Ortiz, S. H. C., Chiu, T., and Fox, M. D., "Ultrasound image enhancement: A review," *Biomedical Signal Processing and Control* **7**(5), 419–428 (2012).

[3] Krücker, J. F., Meyer, C. R., LeCarpentier, G. L., Fowlkes, J. B., and Carson, P. L., "3D spatial compounding of ultrasound images using image-based nonrigid registration," *Ultrasound in Medicine and Biology* **26**(9), 1475–1488 (2000).

[4] Torres, F., Fanti, Z., Lira, E., García-Segundo, C., Reyes-Ramirez, B., Hazan, E. J., Gerson, R., and Arambula-Cosio, F., "Rastreo de Imágenes y Reconstrucción de Volúmenes de Ultrasonido Médico," *Revista mexicana de ingeniería biomédica* **33**(2), 101–115 (2008).

[5] Fanti, Z., Torres, F., and Arámbula Cosío, F., "Preliminary results in large bone segmentation from 3D freehand ultrasound," in [*IX International Seminar on Medical Information Processing and Analysis*], Brieva, J. and Escalante-Ramírez, B., eds., 89220F–9, SPIE (2013).

[6] Cenni, F., Monari, D., Schless, S.-H., Aertbeliën, E., Desloovere, K., and Bruyninckx, H., "Efficient image based method using water-filled balloons for improving probe spatial calibration in 3d freehand ultrasonography," *Ultrasonics* **94**, 124–130 (2019).

[7] Cheung, C.-W. J., Zhou, G.-Q., Law, S.-Y., Lai, K.-L., Jiang, W.-W., and Zheng, Y.-P., "Freehand three-dimensional ultrasound system for assessment of scoliosis," *Journal of Orthopaedic Translation* **3**(3), 123–133 (2015).

[8] Li, B., Berti, R., Abran, M., and Lesage, F., "Ultrasound guided fluorescence molecular tomography with improved quantification by an attenuation compensated born-normalization and in vivopreclinical study of cancer," *Review of Scientific Instruments* **85**(5), 053703 (2014).

[9] Pino-Almero, L., Mínguez-Rey, M. F., Sentamans-Segarra, S., Salvador-Palmer, M. R., Anda, R. M. C.-O. d., and La O, J. L.-d., "Quantification of topographic changes in the surface of back of young patients monitored for idiopathic scoliosis: correlation with radiographic variables," *Journal of Biomedical Optics* **21**(11), 116001 (2016).

[10] Marrugo, A. G., Romero, L. A., Pineda, J., Vargas, R., Altamar-Mercado, H., Marrugo, J., and Meneses, J., "Toward an automatic 3D measurement of skin wheals from skin prick tests," in [*Dimensional Optical Metrology and Inspection for Practical Applications VIII*], Zhang, S. and Harding, K. G., eds., 3–11, SPIE (2019).

[11] Quang, T. T., Kim, H.-Y., Bao, F. S., Papay, F. A., Edwards, W. B., and Liu, Y., "Fluorescence Imaging Topography Scanning System for intraoperative multimodal imaging," *PLoS ONE* **12**(4), e0174928 (2017).

[12] Mercier, L., Langø, T., Lindseth, F., and Collins, D. L., "A review of calibration techniques for freehand 3-d ultrasound systems," *Ultrasound in medicine & biology* **31**(4), 449–471 (2005).

[13] Fenster, A., Landry, A., Downey, D. B., Hegele, R. A., and Spence, J. D., "3d ultrasound imaging of the carotid arteries," *Current Drug Targets-Cardiovascular & Hematological Disorders* **4**(2), 161–175 (2004).

[14] Comeau, R. M., Sadikot, A. F., Fenster, A., and Peters, T. M., "Intraoperative ultrasound for guidance and tissue shift correction in image-guided neurosurgery," *Medical physics* **27**(4), 787–800 (2000).

[15] Gee, A. H., Treece, G. M., Prager, R. W., Cash, C. J., and Berman, L., "Rapid registration for wide field of view freehand three-dimensional ultrasound," *IEEE transactions on medical imaging* **22**(11), 1344–1357 (2003).

[16] Hu, Y., Kasivisvanathan, V., Simmons, L. A., Clarkson, M. J., Thompson, S. A., Shah, T. T., Ahmed, H. U., Punwani, S., Hawkes, D. J., Emberton, M., et al., "Development and phantom validation of a 3-d-ultrasound-guided system for targeting mri-visible lesions during transrectal prostate biopsy," *IEEE Transactions on Biomedical Engineering* **64**(4), 946–958 (2016).

[17] Colley, E., Carroll, J., Thomas, S., Varcoe, R. L., Simmons, A., and Barber, T., "A methodology for non-invasive 3-d surveillance of arteriovenous fistulae using freehand ultrasound," *IEEE Transactions on Biomedical Engineering* **65**(8), 1885–1891 (2017).

[18] De Lorenzo, D., Vaccarella, A., Khreis, G., Moennich, H., Ferrigno, G., and De Momi, E., "Accurate calibration method for 3d freehand ultrasound probe using virtual plane," *Medical physics* **38**(12), 6710–6720 (2011).

[19] Prager, R. W., Rohling, R. N., Gee, A., and Berman, L., "Rapid calibration for 3-d freehand ultrasound," *Ultrasound in medicine & biology* **24**(6), 855–869 (1998).

[20] Hsu, P.-W., Treece, G. M., Prager, R. W., Houghton, N. E., and Gee, A. H., "Comparison of freehand 3-d ultrasound calibration techniques using a stylus," *Ultrasound in medicine & biology* **34**(10), 1610–1621 (2008).

[21] Gorthi, S. S. and Rastogi, P., "Fringe projection techniques: whither we are?," *Optics and lasers in engineering* **48**(2), 133–140 (2010).

[22] Zuo, C., Feng, S., Huang, L., Tao, T., Yin, W., and Chen, Q., "Phase shifting algorithms for fringe projection profilometry: A review," *Optics and Lasers in Engineering* **109**, 23–59 (2018).

[23] Takeda, M. and Mutoh, K., "Fourier transform profilometry for the automatic measurement of 3-d object shapes," *Applied optics* **22**(24), 3977–3982 (1983).

[24] Zhang, S., [*High-Speed 3D imaging with digital fringe projection techniques*], CRC Press (2016).

[25] Ghiglia, D. C. and Pritt, M. D., [*Two-dimensional phase unwrapping: theory, algorithms, and software*], vol. 4, Wiley New York (1998).

[26] Zuo, C., Huang, L., Zhang, M., Chen, Q., and Asundi, A., "Temporal phase unwrapping algorithms for fringe projection profilometry: A comparative review," *Optics and Lasers in Engineering* **85**, 84–103 (2016).

[27] Cenni, F., Monari, D., Desloovere, K., Aertbeliën, E., Schless, S.-H., and Bruyninckx, H., "The reliability and validity of a clinical 3d freehand ultrasound system," *Computer methods and programs in biomedicine* **136**, 179–187 (2016).

[28] Hu, Y., Gibson, E., Lee, L.-L., Xie, W., Barratt, D. C., Vercauteren, T., and Noble, J. A., "Freehand ultrasound image simulation with spatially-conditioned generative adversarial networks," in [*Molecular imaging, reconstruction and analysis of moving body organs, and stroke imaging and treatment*], 105–115, Springer (2017).

[29] Hartley, R. and Zisserman, A., [*Multiple view geometry in computer vision*], Cambridge university press (2003).

[30] Zhang, Z., "A flexible new technique for camera calibration," *IEEE Transactions on pattern analysis and machine intelligence* **22** (2000).

[31] Blackall, J. M., Rueckert, D., Maurer, C., Penney, G. P., Hill, D. L., and Hawkes, D. J., "An image registration approach to automated calibration for freehand 3d ultrasound," in [*International Conference on Medical Image Computing and Computer-Assisted Intervention*], 462–471, Springer (2000).

[32] Hsu, P.-W., Prager, R. W., Gee, A. H., and Treece, G. M., "Freehand 3d ultrasound calibration: a review," in [*Advanced imaging in biology and medicine*], 47–84, Springer (2009).

[33] Marquardt, D. W., "An algorithm for least-squares estimation of nonlinear parameters," *Journal of the society for Industrial and Applied Mathematics* **11**(2), 431–441 (1963).

[34] Moré, J. J., "The levenberg-marquardt algorithm: implementation and theory," in [*Numerical analysis*], 105–116, Springer (1978).

[35] Zhang, S. and Huang, P. S., "Novel method for structured light system calibration," *Optical Engineering* **45**(8), 083601 (2006).

[36] Hu, Q., Huang, P. S., Fu, Q., and Chiang, F.-P., "Calibration of a three-dimensional shape measurement system," *Optical Engineering* **42**(2), 487–494 (2003).

[37] Vargas, R., Marrugo, A. G., Pineda, J., Meneses, J., and Romero, L. A., "Camera-projector calibration methods with compensation of geometric distortions in fringe projection profilometry: A comparative study," *Opt. Pura Apl* **51**(3), 50305–1 (2018).

[38] Vargas, R., Marrugo, A. G., Pineda, J., and Romero, L. A., "A flexible and simplified calibration procedure for fringe projection profilometry," in [*Modeling Aspects in Optical Metrology VII*], **11057**, 110571R, International Society for Optics and Photonics (2019).

[39] Pineda, J., Vargas, R., Romero, L. A., Meneses, J., and Hernández, A. G. M., "Fringe quality map for fringe projection profilometry in labview," *Optica pura y aplicada* **51**(4), 7 (2018).

[40] Lindseth, F., Tangen, G. A., Langø, T., and Bang, J., "Probe calibration for freehand 3-d ultrasound," *Ultrasound in medicine & biology* **29**(11), 1607–1623 (2003).