Honorable Mention Poster Award

Fusion of electromagnetic tracking with speckle-tracked 3D freehand ultrasound using an unscented Kalman filter

Andrew Lang[a], Parvin Mousavi[b], Gabor Fichtinger[a,b,c,d], Purang Abolmaesumi[a,b,d]

[a]Department of Electrical and Computer Engineering, Queen's University, Kingston, ON, Canada;
[b]School of Computing, Queen's University, Kingston, ON, Canada;
[c]Department of Mechanical and Materials Engineering, Queen's University, Kingston, ON, Canada;
[d]Department of Surgery, Queen's University, Kingston, ON, Canada

ABSTRACT

Freehand 3D ultrasound (US) using a 2D US probe has the advantage over conventional 3D probes of being able to collect arbitrary 3D volumes at a lower cost. Conventionally, optical and electromagnetic (EM) sensors are used to keep track of the US probe position. Optical tracking provides more accuracy but requires line-of-sight which can be a problem for many applications. Conversely, EM tracking does not have any line-of-sight restrictions, but it has lower accuracy and measurement jitter, and is susceptible to metallic distortions. Ultrasound imaging has the advantage that the speckle inherent in all images contains relative position information due to the decorrelation of speckle over distance. However, tracking the position of US images using speckle information alone suffers from drifts caused by tissue inconsistencies, and overall lack of accuracy. In our work, we examine the possibility for overcoming the limitations of both EM US tracking and freehand, speckle-based US image tracking, through the fusion of these techniques. Even though positions found through speckle-based tracking have very little jitter, the overall error is large, due to drifts in position estimation. By combining the EM and speckle-based tracking information using an Unscented Kalman Filter, we are able to reduce the drift errors as well as to eliminate high-frequency jitter noise from the EM tracker positions. Such fusion produces a smooth and accurate 3D reconstruction superior to those using the EM tracker alone. In addition, we look at the effect of metallic distortions on our fusion and demonstrate improvements over the EM tracker reconstruction.

Keywords: Freehand ultrasound, speckle tracking, electromagnetic tracking, sensor fusion, unscented Kalman filter

1. INTRODUCTION

The generation of three-dimensional (3D) ultrasound (US) volumes requires the use of either a stand-alone 3D US probe or a tracked 2D US probe in what is known as freehand 3D US.[1] Since many existing US systems already have a 2D US probe, forgoing the added cost of purchasing a 3D probe is desired. Freehand 2D US reconstructions are created by sweeping a 2D US probe over a volume of interest and tagging every image in the sequence with a position and orientation in space. This tracking can be done by attaching a position sensor directly to the US probe and finding the rigid transformation between the US image coordinate system and the position sensor coordinate system. This transformation can be found through a calibration technique.[2] Optical and electromagnetic (EM) tracking are among the two common position tracking systems with individual benefits and shortcomings. Optical tracking systems have high tracking accuracy; however, they pose a line-of-sight restriction. Conversely, EM tracking systems have an order of magnitude lower accuracy than optical tracking but without a line-of-sight restriction. Additionally, the EM tracking accuracy significantly deteriorates in the presence of metallic distortions.[3]

Send correspondences to Parvin Mousavi, School of Computing, Queen's University, Kingston, Ontario, Canada, K7L 3N6. E-mail: pmousavi@cs.queensu.ca.

Medical Imaging 2009: Ultrasonic Imaging and Signal Processing, edited by Stephen A. McAleavey, Jan D'hooge
Proc. of SPIE Vol. 7265, 72651A · © 2009 SPIE · CCC code: 1605-7422/09/$18 · doi: 10.1117/12.813879

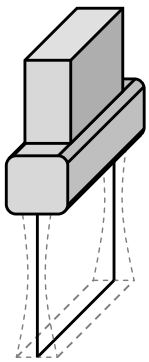


Figure 1. Each ultrasound image contains information from a three-dimensional area. Subsequent frames will contain overlapping information.

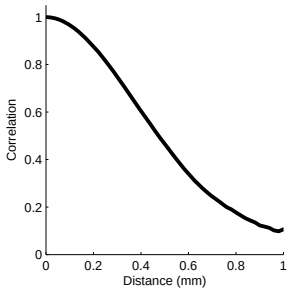


Figure 2. Decorrelation curve relating the distance to the correlation coefficient value.

An alternative to using an external physical positioning and tracking system for freehand 3D US reconstruction is to obtain the tracking information using the speckle contained in the images themselves. Reconstruction of 3D US volumes using speckle tracked freehand US requires finding the in-plane and out-of-plane transformation parameters between image frames independently. The in-plane transformation parameters can be found using image registration techniques[4,5] while the out-of-plane parameters can be determined using speckle decorrelation.[6,7]

The speckle decorrelation method takes advantage of the inherently 3D nature of ultrasound images (Figure 1). As the US probe is moved in an elevational direction*, speckle that is fully developed† decorrelates at a rate that is related to the distance between frames. Speckle decorrelation curves can then be created that relate correlation to distance (Figure 2). Since the speckle decorrelates completely within about 1 mm, the image acquisition frame-to-frame distance must not exceed this value. Recently, Housden et al.[8] developed a robust algorithm for untracked reconstruction of 3D US volumes that allows for direction changes and frame intersections during image acquisition, as well as optimal frame selection. Laporte and Arbel[9] extended the speckle decorrelation model by using a probabilistic model for the decorrelation curves.

The largest source of error in untracked freehand US is a drift in the out-of-plane parameters away from the true values. A reason that drift errors arise is that the tissue being scanned does not contain fully developed speckle. Speckle that is not fully developed will produce different decorrelation curves, thereby decreasing the tracking accuracy.[10] Gee et al.[11] have demonstrated a method of correcting for different speckle types by using additional information contained in the axial and lateral decorrelation curves. Probe rotation is also a source of error as the decorrelation model assumes only elevational motion.[12] Despite these significant problems, untracked freehand US has been shown to have good small scale accuracy, with very little jitter in the measurements.

*Given the standard x and y axes in the plane of an image, the elevational direction will be defined by the z axis. Additionally, the lateral and axial directions will be defined by the x and y axes, respectively.

†Fully developed speckle refers to speckle that exhibits a Rayleigh amplitude distribution

Studying the differences between electromagnetic tracking and speckle tracking, we can see that they perform very differently. Over a large scale reconstruction, the EM tracker performs very well without any drift. However, an EM tracked reconstructed US volume will show a noticeable amount of jitter when the volume is examined at a small scale. Conversely, a speckle tracked US reconstruction will show the opposite characteristics as the EM reconstruction, with a considerable amount of drift and almost no jitter. These differences lend well to the possibility of fusing EM tracked positions with those of the speckle tracking. This fusion is the focus of our work.

Previously, Housden et al.[13] proposed a method for freehand 3D US reconstruction through combining speckle tracking measurements with those of a six degree-of-freedom (dof) tracker as well as a three dof tracker. The fusion framework that was used consisted of rigidly fitting the speckle tracked transformations to the secondary tracker in addition to performing a constrained non-rigid deformation correction that minimizes the error between the two transformations. A disadvantage of this fusion method is that is does not take into account the relative accuracies of the two trackers. Inaccurate or noisy tracker positions can significantly affect the results. Additionally, the fusion is not designed with real-time acquisition in mind, as the fusion takes place after an entire scan is taken.

For an alternative method of fusion, we look at the problem of fusing speckle tracking with an EM tracker in a similar manner to that seen in the problem of fusing inertial measurement units (IMUs) with global positioning systems (GPS).[14] GPS, much like EM tracking, provides very reliable absolute measurement of position, but with low precision on a small-scale. On the other hand, IMUs, similar to speckle tracking, offer very good short term, small-scale accuracy but suffer from a severe drift. Together, the large-scale accuracy of the GPS and the small-scale accuracy of the IMU can provide better positioning than either sensor individually. This fusion framework often includes the use of nonlinear Kalman filtering, whether it be an extended Kalman filter or an unscented Kalman filter.[15]

In our work, we have developed a fusion approach similar to GPS/IMU fusion, where we use an unscented Kalman filter. This approach is well adapted for the real-time fusion of the speckle tracking and EM tracking. We additionally study the effect of metallic distortions on the accuracy of 3D volume reconstruction using our proposed methodology. The rest of this paper is organized as follows: in Section 2 we provide our methodology, in Section 3 we give our experimental results, and in Section 4 we present our conclusions.

2. METHODOLOGY

An accurate freehand 3D US reconstruction first begins with calibrating the US images to the tracking device. We first calibrate the US images to an optical tracker, which will be used as a ground truth tracking. Next, the EM tracker is calibrated to the US images. A freehand sweep of the US probe is then done over a volume of interest, after which a 3D reconstruction can be obtained either by using the calibrated tracking information or through speckle tracking. After each reconstruction is acquired, the EM tracked reconstruction and the speckle tracked reconstruction are combined using a UKF fusion method.

2.1 EM Tracker Calibration

The calibration of the optical tracker coordinate system to the US probe coordinate system was performed using an N-wire phantom in a water bath.[16] Instead of calibrating the EM tracker directly using the N-wire phantom, we chose to take advantage of the fixed transformation between the optical digital reference body (DRB) and EM sensor that are both rigidly mounted on the US probe. The transformation between the DRB and the EM sensor can be solved for, using techniques commonly applied in robotics, namely a hand-eye calibration method proposed by Tsai and Lenz.[17] Following this method, we use the relation

$$^{EM_2}\mathbf{T}_{EM_1}\,{}^{EM}\mathbf{T}_D = {}^{EM}\mathbf{T}_D\,{}^{D_2}\mathbf{T}_{D_1} \tag{1}$$

where $^{EM_2}\mathbf{T}_{EM_1}$ represents the transform between two EM sensor positions, $^{D_2}\mathbf{T}_{D_1}$ represents the transform between two DRB positions and $^{EM}\mathbf{T}_D$ represents the rigid transform between the DRB and the EM sensor which can be solved for with enough measurements (see Figure 3). Note that $^{EM}\mathbf{T}_D$ is equivalent to $^{EM_1}\mathbf{T}_{D_1}$ or $^{EM_2}\mathbf{T}_{D_2}$ as this relation is constant between measurements. We have chosen to solve for $^{EM}\mathbf{T}_D$ using a closed-form solution that is outlined in Tsai and Lenz.[17] As this proposed solution does not perform optimally under

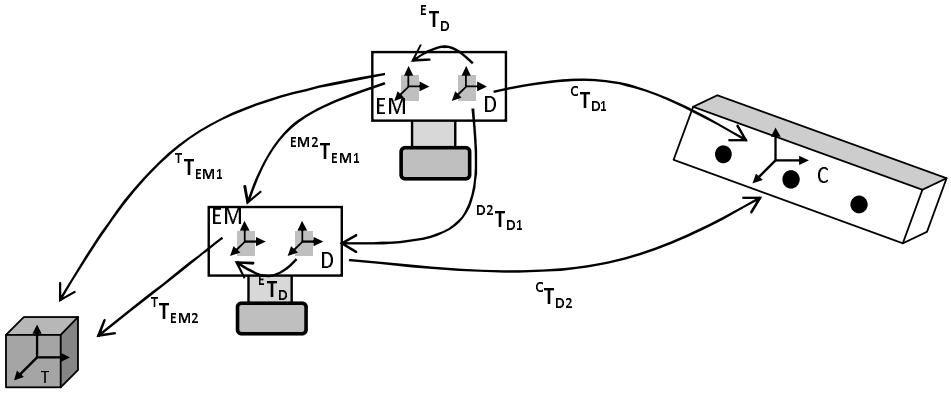


Figure 3. Between any two probe locations in space, the relation ${}^{EM_2}\mathbf{T}_{EM_1}\,{}^{EM}\mathbf{T}_D = {}^{EM}\mathbf{T}_D\,{}^{D_2}\mathbf{T}_{D_1}$ holds. T represents the EM transmitter, EM represents the EM sensor, D represents the optical DRB and C represents the optical tracker.

noisy measurements, it was used only as an initial guess. A non-linear least-squares optimization, CMA-ES,[18] was then applied to minimize the error between the two sides of Equation (1):

$$\sum_{i=1}^{N}(\mathbf{t}_i^l - \mathbf{t}_i^r)^T(\mathbf{t}_i^l - \mathbf{t}_i^r) + K(\mathbf{r}_i^l - \mathbf{r}_i^r)^T(\mathbf{r}_i^l - \mathbf{r}_i^r) \tag{2}$$

where $\mathbf{t}_i^l$ and $\mathbf{t}_i^r$ represent the three translational parameters from the translational part of the matrices formed by the left and right hand side of Equation (1), $\mathbf{r}_i^l$ and $\mathbf{r}_i^r$ represent the vector of the three Rodrigues parameters formed from the rotational part of the matrices formed by the left and right hand side of Equation (1), and K represents a scaling parameter that weights the error of the translational and rotational components together. A value of $K = 10$ was used since the EM tracker was found to have a static rotational measurement accuracy (in degrees) that was about 10 times smaller than that of the translational measurements (in mm).

2.2 Speckle Tracking

We have applied a very simple speckle tracking algorithm for freehand US reconstruction. Each RF frame is divided into a 11x7 grid of patches, not including a small border around the edges, as the edge information is much less reliable. Figure 4 shows an example B-scan US image that has been divided into a grid of patches.

Before any reconstructions, an initial speckle calibration is performed where the speckle decorrelation curves are created. During the reconstruction, the in-plane transformations between frames are found, followed by finding the out-of-plane transformations. The final step consists of frame selection for optimal reconstruction.

2.2.1 Speckle calibration

To calibrate our speckle decorrelation curves, initially a speckle phantom placed on a linear stage setup is scanned. The stage is then moved in the elevational direction in 0.02 mm increments for a total of 4 mm. Elevational decorrelation curves are created for each patch in the grid by averaging the each patch's curve over all of the images.

2.2.2 Determining in-plane transforms

To find the in-plane transformation parameters (axial and lateral translations and roll rotation about the elevational axis), a block matching technique is used. Between two frames, the axial and lateral translations are found for each block in the images. These translations are calculated for each block by setting up a search region in the second image, and identifying the maximum correlation value between the first image block and a block on the second image that is within the search region. After the correlation values are found for each block within the search region, a 2D matrix of correlation values is created where the peak value corresponds to

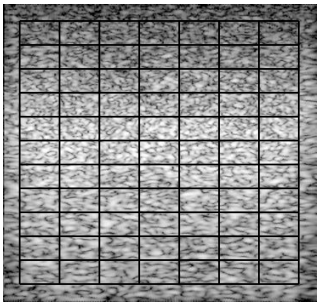


Figure 4. An example of a B-scan image from an RF frame divided into a grid of 11×7 patches. The first and last 2 rows were not used in the speckle tracking.

the final pixel level translation. To estimate the sub-pixel level translation values, a method similar to that of Housden et al.[4] is used. In this method, only the lateral translations between two images are found by fitting an offset Gaussian function to the five correlation values around the peak value (two values on each side of the peak value). The fitting is done using a Levenberg-Marquardt optimization, after which the sub-pixel translation values can be taken as the mean offset of the Gaussian function. We extend this theory to finding both the axial and lateral sub-pixel translation values by fitting a two dimensional offset Gaussian function to the 3×5 matrix of correlation values surrounding the peak value. We use five samples in the lateral direction and three samples in the axial direction as a result of the axial resolution being far greater than the lateral resolution. Again, Levenberg-Marquardt optimization is used to determine the sub-pixel translation in both the axial and lateral directions.

After finding the axial and lateral translation for each block, a least squares solution to find the overall in-plane transformation parameters is used. A method similar to that of Arun et al.[19] is used to find the in-plane 2D transformation matrix.

2.2.3 Determining out-of-plane transforms

The out-of-plane transformation parameters (elevational translation and pitch and yaw rotations about the axial and lateral axes, respectively) are also found through a block based method. Finding the axial and lateral translations for each block through an offset Gaussian function fitting also allows the sub-pixel peak correlation value to be calculated. Knowing the correlation value between the two blocks, the pre-calibrated speckle decorrelation curve is used to determine the distance between the two blocks. Given a distance value between the two images for all of the blocks, a plane is fit using a least squares method[20]

$$\mathbf{w} = (\mathbf{M}^T\mathbf{M})^{-1}\mathbf{M}^T\mathbf{z} \tag{3}$$

where $\mathbf{M}$ is an n×3 matrix containing the axial and lateral position of all of the blocks in the first two columns and a column of 1's for the last column. Also, $\mathbf{z}$ is an n×1 vector of elevational block distances. The resulting vector, $\mathbf{w}$, contains the parameters of the plane equation

$$z = w_1x + w_2y + w_3 \tag{4}$$

To find the corresponding elevational translation and pitch and yaw rotations, we must first form a normal vector to the plane

$$\mathbf{n} = \frac{\begin{bmatrix} w_1 & w_2 & 1 \end{bmatrix}}{\left\| \begin{bmatrix} w_1 & w_2 & 1 \end{bmatrix} \right\|} \tag{5}$$

and the three out-of-plane parameters between the frames become

$$t_z = w_3 \tag{6}$$

$$r_{pitch} = \arctan\frac{n_1}{n_3} \tag{7}$$

$$r_{yaw} = \arcsin n_2 \tag{8}$$

2.2.4 Speckle tracking reconstruction

The speckle tracked reconstruction is created by starting with the first frame of the sequence. The speed of the scan determines which frame is used next by the speckle tracking algorithm to determine the relative frame location. For slower scans, every three or four frames are skipped, whereas for faster scans, no frames are skipped. Frames are skipped to keep the frames from being too close to each other, which would cause additional errors in the elevational distance between the frames as the decorrelation curves are much flatter at closer distances due to the Gaussian shape of the curve. This frame selection continues to place all subsequent frames that are the number of frames to skip apart. For example, if three frames are skipped, the frame sequence would go 1, 4, 7, 10, etc. until the end of the sequence is reached.

2.3 Sensor Fusion

The fusion framework that was employed to combine the EM tracking with speckle tracking takes a very similar form to the fusion of GPS and INS systems. An unscented Kalman filter[21] is used for the fusion in a manner very similar to techniques used by van der Merwe and Wan.[15] Also, considerations for using an UKF with quaternion orientations were taken as done by Kraft.[22] Details of our process and measurement model follow. Further details on implementation of the UKF can be found in Ref. 21.

2.3.1 Process model

For our process model, a first order constant velocity model is used. The process state vector is

$$\mathbf{x} = \begin{bmatrix} \mathbf{p} & \mathbf{v} & \mathbf{q} & \boldsymbol{\omega} & \mathbf{b_p} & \mathbf{b_q} \end{bmatrix}^T \tag{9}$$

where $\mathbf{p} = \begin{bmatrix} p_x & p_y & p_z \end{bmatrix}^T$ and $\mathbf{v} = \begin{bmatrix} v_x & v_y & v_z \end{bmatrix}^T$ are the position and velocity vectors for the lateral, axial and elevational translations, $\mathbf{q} = \begin{bmatrix} q_0 & q_1 & q_2 & q_3 \end{bmatrix}^T$ and $\boldsymbol{\omega} = \begin{bmatrix} \omega_x & \omega_y & \omega_z \end{bmatrix}^T$ are the quaternion orientation and rotational velocity for the roll, pitch and yaw rotations, and $\mathbf{b_p} = b_{v_z}$ and $\mathbf{b_q} = \begin{bmatrix} b_{\omega_y} & b_{\omega_z} \end{bmatrix}^T$ are the bias terms for the out-of-plane transformation parameters. Note that the bias terms are not incorporated into the state model for the in-plane parameters as the speckle-tracking does not exhibit drift in those parameters. Additionally, since a first order velocity model is used, any acceleration will be considered as noise. Thus, we have as our process noise vector

$$\mathbf{n} = \begin{bmatrix} \mathbf{a_p} & \mathbf{a_q} & \mathbf{n_{b_p}} & \mathbf{n_{b_q}} \end{bmatrix}^T \tag{10}$$

Looking at the discrete-time update models for each variable in our state vector we have

$$\mathbf{p}_{k+1} = \mathbf{p}_k + T\mathbf{v}_k + \frac{T^2}{2}\mathbf{a_p} \tag{11}$$

$$\mathbf{v}_{k+1} = \mathbf{v}_k + T\mathbf{a_p} \tag{12}$$

$$\mathbf{q}_{k+1} = \mathbf{q}_k \otimes \mathbf{q}_\Delta \tag{13}$$

$$\boldsymbol{\omega}_{k+1} = \boldsymbol{\omega}_k + T\mathbf{a_q} \tag{14}$$

$$\mathbf{b_p}_{k+1} = \mathbf{b_p}_k + \mathbf{n_{b_p}} \tag{15}$$

$$\mathbf{b_q}_{k+1} = \mathbf{b_q}_k + \mathbf{n_{b_q}} \tag{16}$$

where T represents the time between samples and $\otimes$ represents a quaternion multiplication. The delta quaternion, $\mathbf{q}_\Delta$, is a quaternion representing the differential rotation using the rotational velocity and acceleration noise.[22] The quaternion can be created by first determining the axis-angle representation of the rotation

$$\alpha_\Delta = \left| \boldsymbol{\omega}_k + \frac{T^2}{2}\mathbf{a_q} \right| \tag{17}$$

$$\mathbf{e}_\Delta = \frac{\boldsymbol{\omega}_k + \frac{T^2}{2}\mathbf{a_q}}{\left| \boldsymbol{\omega}_k + \frac{T^2}{2}\mathbf{a_q} \right|} \tag{18}$$

where (17) is the angle and (18) is the axis, from which we can create the quaternion

$$\mathbf{q}_\Delta = \begin{bmatrix} \cos\frac{\alpha_\Delta}{2}, & \mathbf{e}_\Delta \sin\frac{\alpha_\Delta}{2} \end{bmatrix}^T \tag{19}$$

The bias terms are represented as a random-walk process affected by zero-mean Gaussian noise.

2.3.2 Measurement model

The measurement vector of our system consists of measurements made by the EM tracker and the speckle tracking:

$$\mathbf{y}_k = \begin{bmatrix} \mathbf{p_{EM}} & \mathbf{q_{EM}} & \mathbf{v_s} & \boldsymbol{\omega}_s \end{bmatrix}^T \tag{20}$$

where $\mathbf{p_{EM}}$ and $\mathbf{q_{EM}}$ are the position of the US frame made by the EM tracker and $\mathbf{v_s}$ and $\boldsymbol{\omega}_s$ are the difference in position and rotation between the current frame and the previous frame as measured by the speckle tracking. The bias term in the state vector is subtracted from the speckle tracking measurements for the measurement update. Each of the different measurements also has an associated measurement noise that is chosen empirically based on the accuracy of their reconstructions.

3. EXPERIMENTS AND RESULTS

A Sonix RP research ultrasound machine (Ultrasonix Medical Corp., Richmond, Canada) was used to collect radio frequency (RF) ultrasound frames. An RF frame contained 256 lines with 2064 samples each, that were subsequently passed through an envelope detector. US frames were captured at a depth of 4 cm with a probe frequency of 10 MHz using a single focal point at 2 cm. Tracking was done using an EM tracker (pciBird, Ascension Technology Corp., Burlington, VT, USA) and an optical tracker (Optotrak Certus, Northern Digital Inc., Waterloo, Canada). EM sensor measurements were filtered using the pciBird's internal AC notch and DC filters to reduce the amount of noise in the measurements. Additionally, a DRB (Traxtal Technologies Inc., Toronto, Canada) was rigidly mounted to the probe in addition to the EM sensor.

In total, 26 different scans of a speckle phantom were done. The first 10 scans were performed using a micron-precision linear stage. The phantom was placed on top of the stage, with the US probe placed rigidly above the phantom. A layer of water covered the phantom to prevent any contact between the US probe and the phantom during scanning. The last 16 scans were done as completely freehand scans, where the US probe was slowly moved across the phantom.

For each set scans, the first half were in a distortion-free environment and the second half were taken with the electromagnetic field undergoing a distortion caused by placing the phantom and the EM transmitter on a table largely made up of metal. The transmitter was approximately 2 cm from the table while the EM sensor attached to the probe was approximately 15 cm from the table above the phantom.

For evaluating the accuracy of the linear stage driven scans, the speckle tracked, EM tracked and fused reconstruction transformations are compared to known stage positions. For freehand scans, the reconstructions are compared to the optical tracking transformations. For all cases, the inter-frame RMS error was used to define the error of the individual reconstructions. The orientation error is taken as the error in the roll, pitch and yaw angles.

Table 1 demonstrates the errors for the scans using the stage with and without distortion. As seen from the error values for the undistorted case, it is clear that that sensor fusion has improved upon the EM tracking in both the translational and rotational components of the reconstruction. In order to determine whether or not the improvement of the fusion is significant, we ran a t-test to compare the EM tracking errors against the Fusion tracking errors. For the translational components, p-values of 0.0052 and 0.0096 were reported for the undistorted and distorted cases, respectively, making the difference significant. However, for the rotational components, p-values of 0.19 and 0.01 were found for each case, meaning that under no distortion, the difference fails to show to be significant. Figures 5 and 6 show example plots of each of the six different transformation parameters from one data set in the non-distortion and distortion cases. The fusion results can be seen to have improved the in-plane translations significantly in both cases over the EM measurements by following the speckle tracking closely. The rotational component improvements can also be seen as a result of the smoothing of the EM measurements by the UKF.

The errors of freehand scans that use an optical tracker as the ground truth measurement are given in Table 2. Similar results to the stage tracked errors are observed. Sensor fusion again results in significant improvements over the EM tracking for the translational components with p-values of 0.03 and 0.0006 for the no distortion and distortion cases, respectively. The small difference between the EM tracking and fusion tracking translational

Table 1. RMS error of the speckle tracking, EM tracking and fusion tracking using a stage for translations.

No distortion

	Translational Error (mm)	Rotational Error (degrees)
Speckle Tracking	0.50 ± 0.17	0.64 ± 0.17
EM Tracking	0.45 ± 0.17	0.069 ± 0.028
UKF Fusion Tracking	0.23 ± 0.14	0.061 ± 0.022

Distortion

	Translational Error (mm)	Rotational Error (degrees)
Speckle Tracking	0.22 ± 0.11	0.92 ± 0.10
EM Tracking	0.77 ± 0.14	0.17 ± 0.054
UKF Fusion Tracking	0.49 ± 0.15	0.15 ± 0.047

Table 2. RMS error of the speckle tracking, EM tracking and fusion tracking using a freehand scans.

No distortion

	Translational Error (mm)	Rotational Error (degrees)
Speckle Tracking	1.49 ± 0.71	1.12 ± 0.49
EM Tracking	0.31 ± 0.11	0.057 ± 0.016
UKF Fusion Tracking	0.30 ± 0.11	0.063 ± 0.016

Distortion

	Translational Error (mm)	Rotational Error (degrees)
Speckle Tracking	1.10 ± 0.82	1.09 ± 1.19
EM Tracking	1.90 ± 0.43	0.32 ± 0.08
UKF Fusion Tracking	1.53 ± 0.31	0.30 ± 0.089

error under no distortion came out as significant due to the fact that the fusion consistently had a smaller error value as compared to the EM tracking, however small that difference was. Looking at the rotational components, we see that the EM tracking was better than the sensor fusion under no distortion and the fusion was better under distortion with p-values of 0.0001 and 0.025, respectively. One possibility that the EM tracking had better accuracy in rotations for the freehand non-distorted case is that the speckle tracking did not follow the curves of the optical tracking very well on a small scale, thus pulling the fusion away from the actual values. Figure 6 shows the plots for the out-of-plane parameters of the reconstructions for a single scan with and without distortion. Again, we see the fusion parameters following the EM parameters closely due to the bias model used.

From Figures 6 and 7, it is important to note that the distortion does not increase the amount of noise present in the EM tracking by much; rather it causes a minor drift in the reconstructions. This minor drift results in the superior performance of sensor fusion compared to the EM tracking alone, especially for the translational components. Since in our Kalman filter, the in-plane translations are weighted more towards the speckle tracking translations, the fusion involving these parameters will be significantly improved over the EM tracking as long as the speckle tracking accuracy is better than the EM tracking.

We also see in Tables 1 and 2 that, in the distortion case, the translational error of the sensor fusion is higher than that of the speckle tracking. The reason for this difference is that it is assumed by the sensor fusion that there is no drift in the EM tracking translational values. Hence, the fused translations will be very close to the EM values, with only small differences due to the added speckle information. Since, under distortion the EM tracker translations exhibit a drift, the resulting fused translations will exhibit a drift as well. If the drift of the EM tracking is more than that of the speckle tracking, then the fused results will be worse that the speckle tracking. For the rotational case, the distortion caused only a minor amount of drift, allowing the fusion to maintain its accuracy. Much of the improvement in accuracy in the rotational parameters through fusion can be

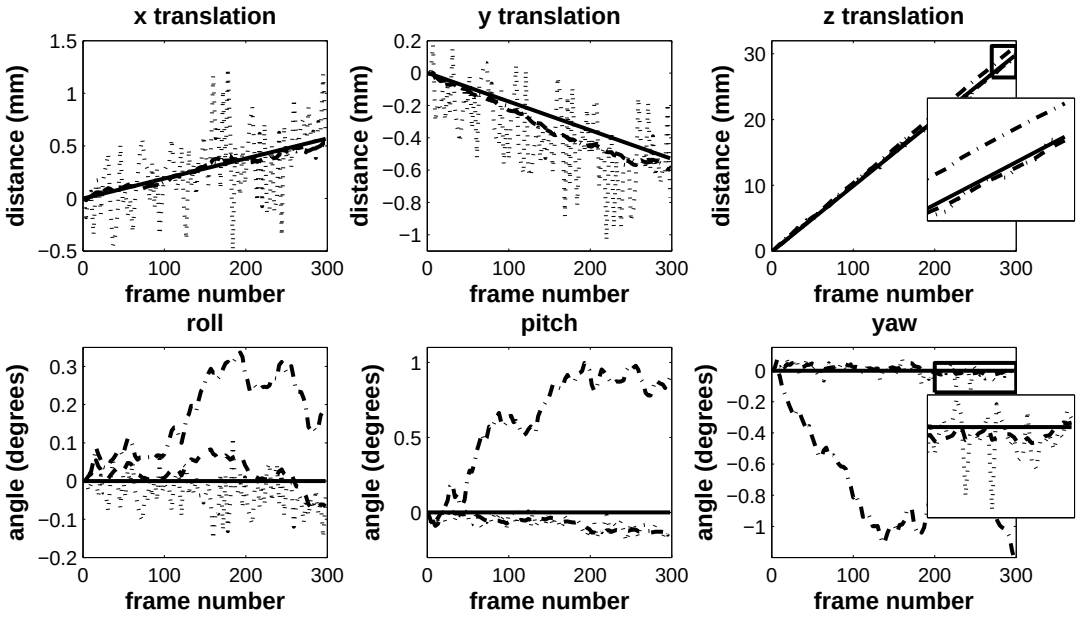


Figure 5. Plots of the six transformation parameters from one scan using the linear-stage under no distortion for the stage tracking (solid line), EM tracking (dotted line), speckle tracking (dash-dotted line) and fusion (dashed line).

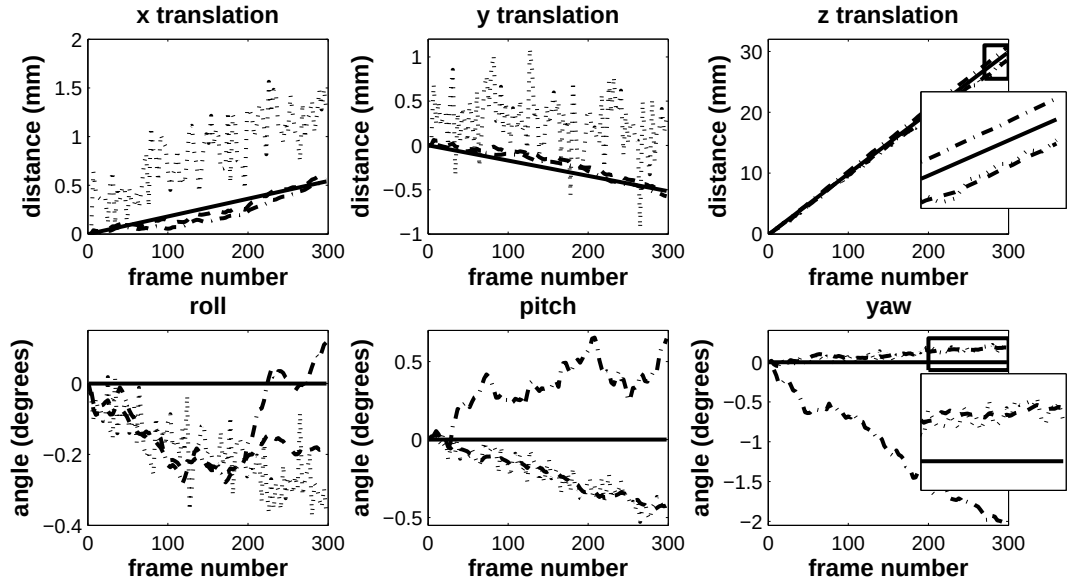
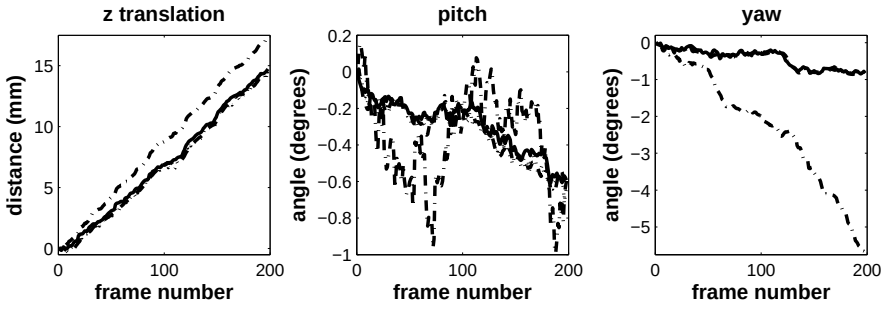
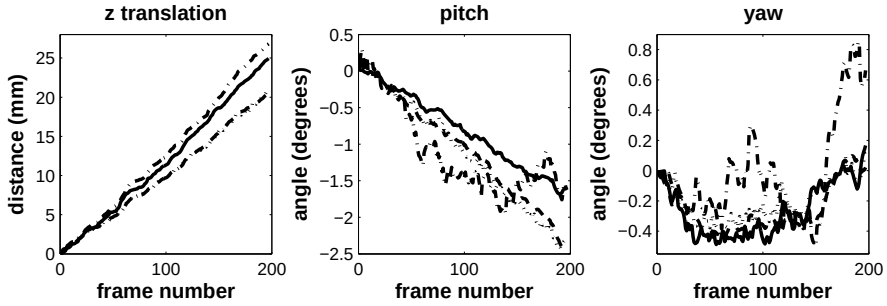


Figure 6. Plots of the six transformation parameters from one scan using the linear-stage under distortion for the stage tracking (solid line), EM tracking (dotted line), speckle tracking (dash-dotted line) and fusion (dashed line).



(a) No distortion



(b) Distortion

Figure 7. Plots of the three out-of-plane transformation parameters for one freehand scan data without (a) and with (b) added distortion for the optical tracking (solid line), EM tracking (dotted line), speckle tracking (dash-dotted line) and fusion (dashed line).

attributed to the low-pass filtering effect of the EM tracker generating a smoother curve, as well as the addition of the speckle tracking information, which often exhibits peaks and valleys in the plots that line up with the optical tracker reconstruction, only with a significant drift error.

The volumetric reconstructions of a single ultrasound scan is plotted in Figure 8. A close inspection of these figures reveals that the fusion reconstruction is much smoother than the EM tracker reconstruction, and very closely resembles that of the optical tracking with the only exception of the drift in the elevational error caused by the combined drift of the EM tracker and the speckle tracking. This elevational error leads us to perhaps the largest downfall of this fusion technique, in that given a drift in both the speckle tracking and the EM tracking, the fused result will often exhibit a drift that is close to that of the EM tracking.

4. DISCUSSION AND CONCLUSIONS

It is clear that using speckle tracking in and of itself is not accurate enough for 3D ultrasound reconstructions. It exhibits far too much drift, especially when longer scans are required. However, it does provide good small scale accuracy with little jitter. By combining the speckle tracking with electromagnetic tracking, we are able to use the strengths of each to improve the reconstruction accuracy by combining them in a sensor fusion framework.

The results presented here offer insight to an optimal way of fusing EM and speckle-tracked position information. In evaluating the results of the fusion, it is better to study the translational and rotational components of the reconstruction separately. Looking simply at the translational component of the reconstructions, we see the fusion exhibit improvements over the EM tracking under both the non-distortion and distortion cases. The accuracy of the axial and lateral translations of the speckle tracking account for much of this improvement, as the elevational translations sometimes have a drift that cannot be corrected for if both the speckle tracking and EM tracking show drift. If there is no drift in the EM tracking elevational positions, a much more significant improvement over both the speckle tracking and the EM tracking is seen through the fusion, as shown by the

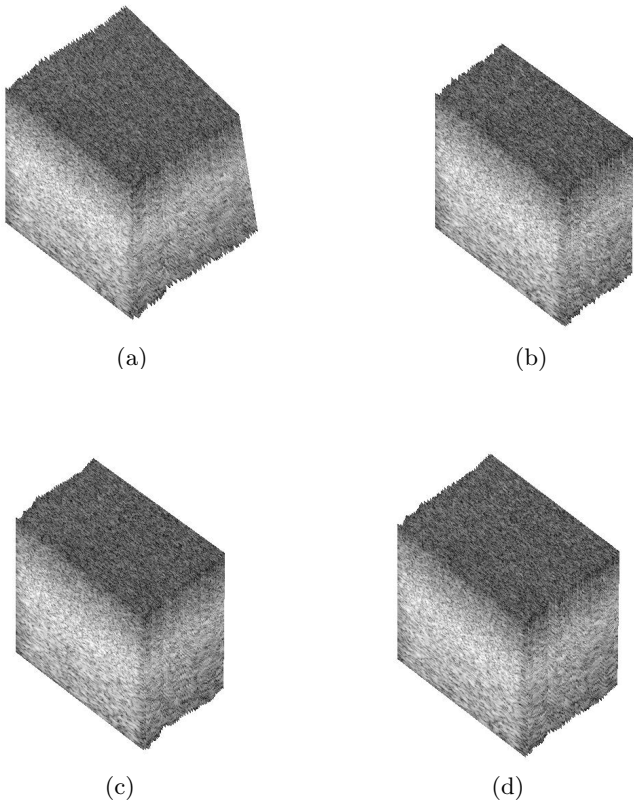


Figure 8. Views of 3D reconstructions from one freehand scan data under distortion for (a) speckle tracking, (b) EM tracking, (c) sensor fusion and (d) optical tracking.

stage-tracked results. Also, studying the rotational components of the fusion, we observe results that are very similar to that of the EM tracking alone. The fusion shows small improvements, especially when distortion is added, partially due to the filtering effect of the UKF and partially due to the addition of the speckle tracking information.

One of the more important aspects of the sensor fusion technique presented here is that much of the jitter exhibited in the EM tracker reconstruction is removed. Also, since the rotational components of the EM tracking are not as affected by distortion, we will see at a minimum an improvement in five of the six transformation parameters, with improvement in the elevational parameter dependant on the amount of drift in the EM tracker. It is also possible that we can further improve the fusion accuracy, by improving on the speckle tracking accuracy. A very simple real-time speckle tracking algorithm was used for this work; however, more advanced speckle tracking algorithms could be used to improve upon the accuracy, such as that of Housden et al.[8] and Laporte and Arbel.[9] Further, using a rotational decorrelation model in addition to the elevational decorrelation model could improve the speckle tracking accuracy, as Housden et al. have done.[12] Another source of inaccuracy, especially in the freehand scans, is the calibration error of the transformation between the US image and the optical tracker. Calibration is not perfect, and thus small errors in calibration will lead to small errors in the differences between the speckle tracking and optical tracking. This becomes more apparent with a speed of sound mismatch between the calibration in water and that of the phantom. Care must be taken to minimize this source of error; however, small differences will always exist because of the imperfect focusing of the US beam.

ACKNOWLEDGMENTS

This research was supported in part by the Natural Sciences and Engineering Research Council (NSERC) and the Canadian Institutes of Health Research (CIHR).

REFERENCES

[1] Fenster, A. and Downey, D. B., "Three-dimensional ultrasound imaging," *Annual Review of Biomedical Engineering* **2**, 457–475 (2000).

[2] 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).

[3] Nixon, M. A., McCallum, B. C., Fright, W. R., and Price, N. B., "The effects of metals and interfering fields on electromagnetic trackers," *Presence: Teleoperators & Virtual Environments* **7**(2), 204–218 (1998).

[4] Housden, R. J., Gee, A. H., Treece, G. M., and Prager, R. W., "Subsample interpolation strategies for sensorless freehand 3D ultrasound," *Ultrasound in Medicine & Biology,* **32**(12), 1897–1904 (2006).

[5] Geiman, B. J., Bohs, L. N., Anderson, M. E., Breit, S. M., and Trahey, G. E., "A novel interpolation strategy for estimating subsample speckle motion," *Physics in Medicine and Biology* **45**(6), 1541–1552 (2000).

[6] Tuthill, T., Krucker, J., Fowlkes, J., and Carson, P., "Automated three-dimensional US frame positioning computed from elevational speckle decorrelation," *Radiology* **209**(2), 575–582 (1998).

[7] Weng, L., Tirumalai, A., Lowery, C., Nock, L., Gustafson, D., Behren, P. V., and Kim, J., "US extended-field-of-view imaging technology," *Radiology* **203**(3), 877–880 (1997).

[8] Housden, R. J., Gee, A. H., Treece, G. M., and Prager, R. W., "Sensorless reconstruction of unconstrained freehand 3D ultrasound data," *Ultrasound in Medicine & Biology,* **33**(3), 408–419 (2007).

[9] Laporte, C. and Arbel, T., "Combinatorial and probabilistic fusion of noisy correlation measurements for untracked freehand 3-D ultrasound," *IEEE Transactions on Medical Imaging* **27**(7), 984–994 (2008).

[10] Hassenpflug, P., Prager, R. W., Treece, G. M., and Gee, A. H., "Speckle classification for sensorless freehand 3-D ultrasound," *Ultrasound in Medicine & Biology,* **31**(11), 1499–1508 (2005).

[11] Gee, A. H., Housden, R. J., Hassenpflug, P., Treece, G. M., and Prager, R. W., "Sensorless freehand 3D ultrasound in real tissue: Speckle decorrelation without fully developed speckle," *Medical Image Analysis,* **10**(2), 137–149 (2006).

[12] Housden, R. J., Gee, A. H., Prager, R. W., and Treece, G. M., "Rotational motion in sensorless freehand three-dimensional ultrasound," *Ultrasonics* **48**(5), 412–422 (2008).

[13] Housden, R., Treece, G., Gee, A., and Prager, R., "Calibration of an orientation sensor for freehand 3D ultrasound and its use in a hybrid acquisition system," *BioMedical Engineering OnLine* **7**(1), 5 (2008).

[14] Knight, D., "Achieving modularity with tightly coupled GPS/INS," in *Proceedings of the IEEE PLANS '92*, 426–432 (1992).

[15] van der Merwe, R., Wan, E., and Julier, S., "Sigma-point Kalman filters for nonlinear estimation and sensor-fusion: Applications to integrated navigation," in *Proceedings of the AIAA Guidance, Navigation & Control Conference*, 16–19 (2004).

[16] Chen, T. K., Abolmaesumi, P., Thurston, A. D., and Ellis, R. E., "Automated 3D freehand ultrasound calibration with real-time accuracy control," in *Proc. Medical Image Computing and Computer Assisted Intervention (MICCAI'06), Part I*, *Lecture Notes in Computer Science* **4190**, 899–906, Springer, Copenhagen, Denmark (2006).

[17] Tsai, R. Y. and Lenz, R. K., "A new technique for fully autonomous and efficient 3D robotics hand-eye calibration," in *IEEE Transactions on Robotics and Automation*, **5**, 345–358 (1989).

[18] Hansen, N., "The CMA evolution strategy: a comparing review," in *Towards a new evolutionary computation. Advances on estimation of distribution algorithms*, Lozano, J. A., Larranaga, P., Inza, I., and Bengoetxea, E., eds., 75–102, Springer (2006).

[19] Arun, K. S., Huang, T. S., and Blostein, S. D., "Least-squares fitting of two 3-D point sets," *IEEE Trans.Pattern Anal.Mach.Intell.* **9**(5), 698–700 (1987).

[20] Krupa, A., Fichtinger, G., and Hager, G., "Full motion tracking in ultrasound using image speckle information and visual servoing," in *IEEE Int. Conf. on Robotics and Automation, ICRA'07*, 2458–2464 (2007).

[21] Julier, S. J., Uhlmann, J. K., and Durrant-Whyte, H. F., "A new approach for filtering nonlinear systems," *Proceedings of the American Control Conference, 1995* **3**, 1628–1632 (1995).

[22] Kraft, E., "A quaternion-based unscented Kalman filter for orientation tracking," in *Proceedings of the Sixth International Conference on Information Fusion*, **1**, 47–54 (2003).