Measurement-Free Calibration of Freehand Ultrasound Systems

Siddharth Trehan

Abstract

I describe a system for the spatial calibration of freehand ultrasound systems, which is used to correlate the position of features within ultrasound images to their actual location in 3D. This becomes important in 3D reconstruction, where a localization device is placed on the ultrasound probe, and a transformation must be determined from the 2D coordinate system of an individual ultrasound image to the 3D world coordinate system before a 3D volume can be reconstructed. The approach I develop determines this transformation from a free-hand ultrasound scan of a water tank without requiring any measurements other than readings from the localization device.

1. Introduction

When an ultrasound probe is fitted with a localization device that reports 6-DOF position and orientation, the question still remains of how exactly each pixel within an ultrasound frame maps to a 3D location in the real world. The localization device can be mounted with different positions and orientations relative to captured ultrasound frame and while the transformation from the localization device coordinate system to a world coordinate system is known, unless it is mounted in a known and controlled way, the transformation from a 2D ultrasound image coordinate system to the localization device coordinate system is unknown. The purpose of calibration is to figure out this unknown transformation.

Typically calibration is done by running an ultrasound scan on a water tank. Here, ultrasound picks out a single planar interface, which is seen as a line in each ultrasound image. For each ultrasound image, the location of the line is determined and a pose reading is captured from the localization device. Then, a correlation between the pose estimate and the line position and orientation is used to determine the transformation between an ultrasound frame and the localization device.

In previous work, calibration has required either the fixing of a global coordinate system or a known configuration of the interface in real-world coordinates– some sort of external measurement. With the system I propose, however, no measurement has to be taken. An algorithm posed in an optimization framework can figure out the optimal transformation that explains the location of each line and the pose reading from the localization device for each frame.

2. Problem Statement

A pixel at location (i, j) within an image maps to real-world coordinates (x, y, z) through the following transformation:

$$\begin{pmatrix} x \\ y \\ z \end{pmatrix} = T_{w\leftarrow s} T_{s\leftarrow u} \begin{pmatrix} i \\ j \\ 1 \end{pmatrix} \tag{1}$$

Here, $T_{w\leftarrow s}$ is the 3-by-4 transformation matrix from the sensor coordinate system to the world coordinate system (including a 3-by-3 basis matrix for the three orthogonal axes of this coordinate system, and a 3-by-1 translation offset vector for the zero of the coordinate system), and $T_{s\leftarrow u}$ is the 4-by-3 transformation matrix from the ultrasound image coordinate system to the sensor coordinate system (including a 3-by-2 basis matrix for the two axes of the ultrasound image, a 3-by-1 translation offset vector for the position of the upper left pixel of the image in the sensor coordinate system, and a 1-by-3 augmented row that preserves the 1 in the last row of the position vector). As we capture data, $T_{w\leftarrow s}$ is different for each ultrasound image (it is the pose reading from the sensor), while $T_{s\leftarrow u}$ remains constant because the sensor is mounted with a fixed position and orientation on the probe. The matrix $T_{s\leftarrow u}$ is unknown.

Meanwhile, the metal interface we image through ultrasound is entirely planar. That means for each point on a bright line in the image, the real world point corresponding to it lies on the plane of this metal interface. We can parametrize this plane with the inverse-normal 3-by-1 vector $\vec{z}$ so that for all points on this plane:

$$\vec{z}^T \begin{pmatrix} x \\ y \\ z \end{pmatrix} = 1 \tag{2}$$

This vector $\vec{z}$ is unknown as well. Our objective is, through sufficient images and pose readings from the sensor, determine the transformation matrix $T_{s\leftarrow u}$ without any knowledge of $\vec{z}$.

3. Analysis

We describe the two transformation matrices in Equation 1 internally as follows:

$$T_{w\leftarrow s} = \begin{pmatrix} R_{w\leftarrow s} & \vec{t}_{w\leftarrow s} \end{pmatrix} \tag{3}$$

And:

$$T_{s\leftarrow u} = \begin{pmatrix} R_{s\leftarrow u} & \vec{t}_{s\leftarrow u} \\ 0 \quad 0 & 1 \end{pmatrix} \tag{4}$$

We will also for convenience break down $R_{s\leftarrow u}$ internally into its two 3-by-1 columns:

$$R_{s\leftarrow u} = \begin{pmatrix} \vec{r}_1 & \vec{r}_2 \end{pmatrix} \tag{5}$$

We can then rewrite Equation 1 as follows:

$$\begin{pmatrix} x \\ y \\ z \end{pmatrix} = R_{w\leftarrow s} R_{s\leftarrow u} \begin{pmatrix} i \\ j \end{pmatrix} + R_{w\leftarrow s} \vec{t}_{s\leftarrow u} + \vec{t}_{w\leftarrow s} \tag{6}$$

We can then combine Equation 6 and Equation 2 and rearrange into the following parametrization of the equation for the line on the image:

$$
\begin{aligned}
\vec{z}^T \begin{pmatrix} x \\ y \\ z \end{pmatrix} &= \vec{z}^T R_{w\leftarrow s} R_{s\leftarrow u} \begin{pmatrix} i \\ j \end{pmatrix} + \vec{z}^T R_{w\leftarrow s} \vec{t}_{s\leftarrow u} + \vec{z}^T \vec{t}_{w\leftarrow s} = 1 \\
&\underbrace{\left(\frac{\left(R_{w\leftarrow s} R_{s\leftarrow u}\right)^T \vec{z}}{1 - \vec{z}^T R_{w\leftarrow s} \vec{t}_{s\leftarrow u} - \vec{z}^T \vec{t}_{w\leftarrow s}} \right)^T}_{\vec{y}^T} \begin{pmatrix} i \\ j \end{pmatrix} = 1
\end{aligned} \tag{7}
$$

Note that the 2-by-1 vector $\vec{y}$ is the inverse-normal that parametrizes the position and orientation of the line in the ultrasound image. The numerator of this expression for $\vec{y}$ is a vector, while the denominator is a scaling factor. Therefore, the numerator, which depends on the unknowns $R_{s\leftarrow u}$ and $\vec{z}$, determines the orientation of the line, while the denominator, which depends on all three unknowns $R_{s\leftarrow u}$, $\vec{z}$, and $\vec{t}_{s\leftarrow u}$, determines the position of the line. To reduce the dimensionality of the optimization problem, therefore, we will consider the orientation and the position of the lines in the images separately– we will use the information from the observed orientation of lines to compute half of the unknowns, and the information from the observed position of lines and the values of the computed unknowns from orientation to compute the other half of the unknowns.

3.1. Estimation from Orientation of Lines

We can remove the magnitude of $\vec{y}$ and observe only orientation by taking the ratio of its two components, $\rho(\vec{y}) = y_1/y_2$. By comparing the observed ratio in the image data ($\rho(\vec{y}_{obs})$) to the projected ratio given "guessed" values of $\vec{z}$, $\vec{r_1}$, and $\vec{r_2}$ ($\rho(\vec{y}_{proj})$), we can iteratively converge onto the true values of these unknowns.

First, we determine the projected value of this ratio by manipulating Equation 7. The denominator only scales the vector, so we can ignore it. The value of this ratio becomes:

$$\rho(\vec{y}_{proj}) = \frac{\vec{z}^T R_{w\leftarrow s} \vec{r}_1}{\vec{z}^T R_{w\leftarrow s} \vec{r}_2} \tag{8}$$

Note, however, that there are many redundancies in unknowns in this equation. First, the magnitude of $\vec{z}$ cancels out, so it doesn't affect ρ. Next, the relative magnitudes of $\vec{r}_1$ and $\vec{r}_2$ affect ρ, but the individual magnitudes themselves do not. Therefore, we parametrize $\vec{z} = \|z\|\hat{z}$, $\vec{r}_1 = \|r_1\|\hat{r}_1$, and $\vec{r}_2 = s\|r_1\|\hat{r}_2$. Then:

$$\rho(\vec{y}_{proj}) = \frac{1}{s} \frac{\hat{z}^T R_{w\leftarrow s} \hat{r}_1}{\hat{z}^T R_{w\leftarrow s} \hat{r}_2} \tag{9}$$

Then, because $\hat{z}$, $\hat{r}_1$, and $\hat{r}_2$ are normalized and parametrized by two parameters each, this equation in total varies with 7 unknown parameters. Furthermore, if there is no skew in the ultrasound image coordinate system (as is the case with most ultrasound images), $\hat{r}_1$ and $\hat{r}_2$ will be orthogonal, which removes 1 parameter and makes the equation vary with 6 unknown parameters total.

Then, we can find these 6 unknown parameters through the following optimization program over all observed lines and all sensor poses over all ultrasound frames:

$$\begin{aligned} \underset{s,\hat{z},\hat{r}_1,\hat{r}_2}{\operatorname{argmin}} \sum \left\| \rho(\vec{y}_{proj}) - \rho(\vec{y}_{obs}) \right\|^2 &= \underset{s,\hat{z},\hat{r}_1,\hat{r}_2}{\operatorname{argmin}} \sum \left\| \frac{1}{s} \frac{\hat{z}^T R_{w\leftarrow s} \hat{r}_1}{\hat{z}^T R_{w\leftarrow s} \hat{r}_2} - \frac{y_1}{y_2} \right\|^2 \\ &= \boxed{\underset{s,\hat{z},\hat{r}_1,\hat{r}_2}{\operatorname{argmin}} \sum \left\| y_2 \hat{z}^T R_{w\leftarrow s} \hat{r}_1 - y_1 s \hat{z}^T R_{w\leftarrow s} \hat{r}_2 \right\|^2} \end{aligned} \tag{10}$$

3.2. Estimation from Position of Lines

Once we compute estimates of the six numbers parameterizing s, $\hat{z}$, $\hat{r}_1$ and $\hat{r}_2$, we combine these estimates with observations of the positions of lines to compute five remaining parameters: $\|z\|$, $\|r_1\|$, and the 3-by-1 vector $\vec{t}_{s\leftarrow u}$. We will determine these values by directly comparing $\vec{y}_{proj}$ with $\vec{y}_{obs}$.

For convenience, define the vector 2-by-1 vector $\vec{u}$, the 3-by-1 vector $\vec{v}$, and the scalar w in the following ways, noting that these three variables can all be computed from the six unknown parameters defined earlier:

$$\begin{aligned} \vec{u}^T &= \hat{z}^T R_{w\leftarrow s} \begin{pmatrix} \hat{r}_1 & s\hat{r}_2 \end{pmatrix} \\ \vec{v}^T &= \hat{z}^T R_{w\leftarrow s} \\ w &= \hat{z}^T \vec{t}_{w\leftarrow s} \end{aligned} \tag{11}$$

Then, Equation 7 yields the following optimization program for the values of the five unknown parameters:

$$\begin{aligned} \underset{\|z\|,\|r_1\|,\vec{t}_{s\leftarrow u}}{\operatorname{argmin}} \sum \|\vec{y}_{proj} - \vec{y}_{obs}\|^2 &= \underset{\|z\|,\|r_1\|,\vec{t}_{s\leftarrow u}}{\operatorname{argmin}} \sum \| \frac{\|z\|\|r_1\|\vec{u}}{1 - \|z\|\vec{v}^T\vec{t}_{s\leftarrow u} - \|z\|w} - \vec{y}_{obs}\|^2 \\ &= \boxed{\underset{\|z\|,\|r_1\|,\vec{t}_{s\leftarrow u}}{\operatorname{argmin}} \sum \|\|r_1\|\vec{u} + \vec{v}^T\vec{t}_{s\leftarrow u}\vec{y}_{obs} + w\vec{y}_{obs} - \frac{1}{\|z\|}\vec{y}_{obs}\|^2} \end{aligned} \tag{12}$$

Note that this equation is a positive semidefinite quadratic in $\frac{1}{\|z\|}$, $\|r_1\|$, and $\vec{t}_{s\leftarrow u}$, which we can leverage during the optimization process. As a result, once these five parameters are computed, all eleven unknown parameters characterizing the unknown matrix $T_{s\leftarrow u}$ and the unknown plane configuration $\vec{z}$ have been determined via a measurement-free algorithm for freehand ultrasound calibration using a water tank.

4. Computation

Solving the two optimization problems computationally is straightforward because they are not very complex, but the process is described here for ease of replication.

An eleven-dimensional vector parametrizes all the unknowns to be determined in the system:

$$
\begin{aligned}
\vec{x} &= \begin{pmatrix} \vec{\theta}_1^T & \vec{\theta}_k^T & \vec{\theta}_z^T & s_T & s_z^{-1} & \vec{t}_{s\leftarrow u}^T \end{pmatrix}^T \\
\vec{\theta}_1 &= \begin{pmatrix} \theta_1 & \phi_1 \end{pmatrix}^T \\
\vec{\theta}_k &= \begin{pmatrix} \theta_k & \phi_k \end{pmatrix}^T \\
\vec{\theta}_z &= \begin{pmatrix} \theta_z & \phi_z \end{pmatrix}^T \\
\vec{t}_{s\leftarrow u} &= \begin{pmatrix} t_1 & t_2 & t_3 \end{pmatrix}^T
\end{aligned}
\tag{13}
$$

Using two spherical angles each, $\vec{\theta}_1$, $\vec{\theta}_k$, and $\vec{\theta}_z$ parametrize three different normal vectors: $\hat{r}_1$, $\hat{r}_k$, and $\hat{z}$, respectively. From $\hat{r}_1$ and $\hat{r}_k$, $\vec{r}_2$ (with a magnitude relative to $\hat{r}_1$) can be constructed so that it is always orthogonal to $\hat{r}_1$ and its magnitude relative to $\hat{r}_1$ varies between negative and positive infinity:

$$
\begin{aligned}
\hat{r}_2 &= \frac{\hat{r}_1 \times \hat{r}_k}{\|\hat{r}_1 \times \hat{r}_k\|} \\
\frac{\|\vec{r}_2\|}{\|\vec{r}_1\|} &= \tan\left(\frac{\pi}{2} \cdot \|\hat{r}_1 \times \hat{r}_k\|\right)
\end{aligned}
\tag{14}
$$

Finally, s_T and s_z^{-1} are scaling factors for the entries of the $T_{s\leftarrow u}$ matrix and the $\vec{z}$ vector:

$$
\begin{aligned}
\vec{r}_1 &= s_T \cdot \hat{r}_1 \\
\vec{r}_2 &= s_T \cdot \frac{\|\vec{r}_2\|}{\|\vec{r}_1\|} \cdot \hat{r}_2 \\
\vec{z} &= \frac{1}{s_z^{-1}} \cdot \hat{z}
\end{aligned}
\tag{15}
$$

Then, the basin-hopping algorithm is used to compute estimates of the corresponding parts of the $\vec{x}$ vector for the two optimization problems posed in Sections 3.1 and 3.2. It is found to perform better than other algorithms because when error is introduced into batches of images, it produces many small local minima scattered throughout the search space– basin-hopping ensures that a wide and deep minimum is found. The initial guess for $\vec{x}$ is generated randomly, and it has been found that this approach to the optimization problem produces the most reliable results. Furthermore, the initial guess is generated and the basin-hopping algorithm run multiple times to pick the most optimal result.