Boost Calibration for Dual-arm Co-robotic Ultrasound System

Shengtai Yao, Yixuan Wu, Russell Taylor, Emad M. Boctor
Johns Hopkins University, Baltimore MD 21218, USA
eboctor@jhu.edu

***Abstract*—Ultrasound tomography (UST) is an emerging imaging modality for prostate cancer detection, often requiring a dual-arm robotic system. Accurate calibration of the robotic system is essential for successful UST. In this paper, we propose a novel method, termed Boost Calibration, which performs the calibration using only 2D data. Additionally, we introduce a robust and efficient error estimation method in simulation. Simulation experiments demonstrate the effectiveness and robustness of the proposed method.**

***Index Terms*—Calibration, Dual-arm, Ultrasound tomography, Robotic ultrasound**

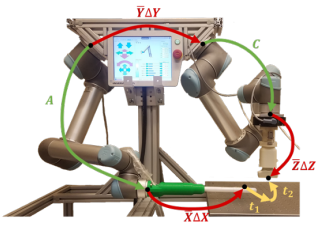


Fig. 1: Dual-arm co-robotic ultrasound system

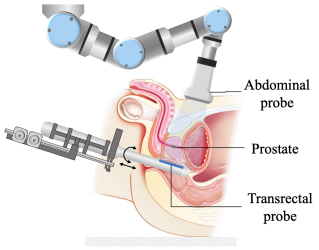


Fig. 2: Clinical setup

I. Introduction

In 2022, prostate cancer was the second most commonly diagnosed cancer and the fifth leading cause of cancer death among men worldwide, with an estimated 1.5 million new cases and 397,000 deaths [1]. Early detection of aggressive prostate cancer is crucial since 5-year survival rate drops from nearly 100% in localized or regional stages to about 34% once the disease metastasizes [2]. Although ultrasound imaging is widely accessible, cost-effective, and capable of real-time visualization, conventional ultrasound techniques have limited sensitivity and specificity in detecting prostate cancer [3]. Ultrasound tomography (UST) reconstructs 3D tissue structure and provides quantitative parameters (e.g., speed of sound) [4], which enables quantification of mechanical properties of biological tissues. Pilot studies have shown that UST has comparable prostate cancer detection accuracy to multiparametric magnetic resonance imaging [10]–[12].

Previously we introduced a dual robotic arm UST imaging setup to accommodate prostate anatomy [8], [9], where one arm holds a transrectal probe and the other holds an abdominal probe to achieve aligned transmit-receive at different view angles, as shown in Fig. 1-2.

To accurately track ultrasound image, it is important to obtain the fixed transformation between ultrasound beam and the robot end effector, which is known as "ultrasound calibration"; and the transformation between the base frames of the dual robotic arms. This dual-arm calibration process involves three unknown rigid transformations $\bar{X}, \bar{Y}, \bar{Z} \in SE(3)$, as illustrated in Fig. 1. A standard approach is sequentially solving the BXp problem [5] for each arm and then combining them together.

Emad Boctor was supported by NSF CAREER AWARD [1653322]. Yixuan Wu was supported by the Graduate Partnerships Program and Fellows Award for Research Excellence, National Institutes of Health.

However, in practice, calibrating both arms independently and directly combining the transformations often results in misalignment errors larger than the image resolution cell (around 0.75 mm assuming 1 MHz central frequency), leading to artifacts in UST images.

To address this issue, we propose a novel method called "Boost Calibration" to combine these transformations by solving the nonlinear equation $AXt_1 = YCZt_2$. Furthermore, a fast simulation method is developed to estimate the upper bound of error under high-dimensional configurations in dual arms system. And note that unlike the $AXB = YCZ$ problem [6] requiring two stereo cameras for 3D data, the proposed method works with only 2D inputs, like standard RGB cameras or ultrasound transducers.

II. Methodology

A. Calibration Algorithm

Boost Calibration, as the name suggests, performs two stages of calibration. The first stage follows the standard approach described above, in which each robotic arm is individually calibrated using BXp solving framework to obtain the ultrasound calibration matrices $\bar{X}$ and $\bar{Z}$. In practice, this step often does not require additional effort, as the ultrasound calibrations for each arm may already be available when the dual-arm system is assembled. And then, both arms simultaneously observe several common target points, from which the transformation between the two robot base frames, denoted as $\bar{Y}$, is estimated. But this method yields large misalignment errors.

Our approach is to simultaneously estimate the three unknown transformations, while leveraging the prior calibration results as initial estimates. This strategy aims to enable more

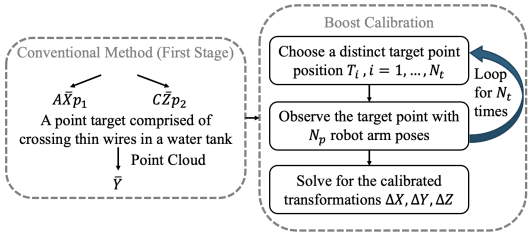


Fig. 3: The process of boost calibration method

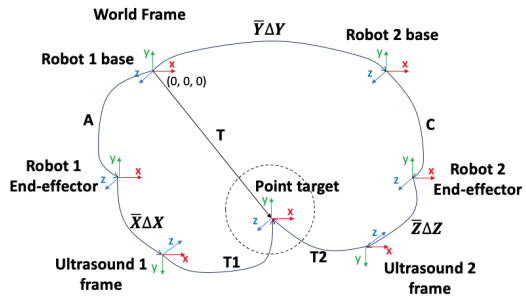


Fig. 4: Kinematic model of the dual-arm ultrasound calibration

accurate and efficient calibration when assembling the dual-arm system. Based on this idea, we proceed to the second stage. The whole process of the method can be seen in Fig. 3.

Assume the three initial estimates $\bar{X}, \bar{Y}$ and $\bar{Z}$, and assume that there exists a small transformation between these estimates and the true calibration matrices, which ensures consistency within the entire dual-arm system:

$$A\bar{X}(\Delta X)\bar{t}_1 = \bar{Y}(\Delta Y)C\bar{Z}(\Delta Z)\bar{t}_2, \tag{1}$$

where $\bar{t}_1 = [t_1, 1], \bar{t}_2 = [t_2, 1] \in \mathbb{R}^4$ are the homogeneous representation of the point target's coordinate in the abdominal and transrectal ultrasound frames, respectively. Due to the use of two-dimensional ultrasound imaging, the translational component in the out-of-plane (elevational) direction is always zero. All other terms in the equation are rigid transformations in $SE(3)$ and can be homogeneously represented by a 4×4 matrix. Using the exponential coordinates for rotation $R_{\bar{X}}$ [7], the transformation $\bar{X}$ can be expressed as:

$$\begin{aligned}\bar{X} &= \begin{bmatrix} R_{\bar{X}} & p_{\bar{X}} \\ 0 & 1 \end{bmatrix} = \begin{bmatrix} \exp\left((a_{\bar{X}})^\wedge \theta_{\bar{X}}\right) & p_{\bar{X}} \\ 0 & 1 \end{bmatrix} \\ &= \begin{bmatrix} I + (a_{\bar{X}})^\wedge \sin\theta_{\bar{X}} + \left((a_{\bar{X}})^\wedge\right)^2 (1 - \cos\theta_{\bar{X}}) & p_{\bar{X}} \\ 0 & 1 \end{bmatrix},\end{aligned} \tag{2}$$

where $a_{\bar{X}} \in \mathbb{R}^3$ with $\|a_{\bar{X}}\| = 1$ is the unit rotation axis, $\theta_{\bar{X}} \in \mathbb{R}$ is the rotation angle, and $p_{\bar{X}} \in \mathbb{R}^3$ is the translation vector. The operator $(\cdot)^\wedge$ denotes the skew-symmetric matrix associated with cross product.

Assuming small errors between $\bar{X}$ and the potential true transformation X in the rotation axis, angle, and translation (denoted by $\Delta a_X \in \mathbb{R}^3, \Delta\theta_X \in \mathbb{R}, \Delta p_X \in \mathbb{R}^3$, respectively), we can derive the small transformation ΔX, such that $\bar{X}(\Delta X)$ is a first order approximation to X,

$$\Delta X = \begin{bmatrix} I + (a_{\bar{X}})^\wedge \Delta\theta_X + (\Delta a_X)^\wedge \theta_{\bar{X}} & R_{\bar{X}}^{-1} \Delta p_X \\ 0 & 1 \end{bmatrix}, \tag{3}$$

where we assume $\Delta a_X \perp a_{\bar{X}}$ to ensure the rotation axis of X, $\Delta a_X + a_{\bar{X}}$ remains unit up to the first order approximation.

Using (2) and (3), the small transformations $\Delta Y, \Delta Z$ can be expressed in a similar form. Substituting these into the dual-arm system in (1), and defining transformation

$$U := \bar{Y}^{-1} A \bar{X}, \quad W := C\bar{Z},$$

we obtain the following linearized system with 21 unknown parameters and 3 constraints:

$$\begin{aligned} &A_1\Delta a_X + A_2\Delta a_Y + A_3\Delta a_Z + \\ &B_1\Delta\theta_X + B_2\Delta\theta_Y + B_3\Delta\theta_Z + \\ &C_1\Delta p_X + C_2\Delta p_Y + C_3\Delta p_Z = D, \\ \text{subject to} \quad &\Delta a_X \perp a_{\bar{X}}, \quad \Delta a_Y \perp a_{\bar{Y}}, \quad \Delta a_Z \perp a_{\bar{Z}}, \end{aligned} \tag{4}$$

where,

$$\begin{aligned} &A_1 = -R_U(t_1)^\wedge \theta_{\bar{X}}, \quad A_2 = (R_W t_2 + p_W)^\wedge \theta_{\bar{Y}}, \\ &A_3 = R_W(t_2)^\wedge \theta_{\bar{Z}}, \quad B_1 = R_U(a_{\bar{X}})^\wedge t_1, \\ &B_2 = -(a_{\bar{Y}})^\wedge (R_W t_2 + p_W), \quad B_3 = -R_W(a_{\bar{Z}})^\wedge t_2, \\ &C_1 = R_U R_{\bar{X}}^{-1}, \quad C_2 = -R_{\bar{Y}}^{-1}, \quad C_3 = -R_W R_{\bar{Z}}^{-1}, \\ &D = -R_U t_1 - p_U + R_W t_2 + p_W. \end{aligned}$$

Given a fixed target point, we change the configurations of the robot arm: specifically, poses A, C and ultrasound observations t_1, t_2. Each configuration yields a set of parameters in (4). By stacking these parameter matrices across multiple configurations, we formulate a linear constrained least squares problem, which can be solved by Lagrange multiplier, and obtain the gap between the initial estimates $\bar{X}, \bar{Y}, \bar{Z}$ and the potential true transformations X, Y, Z.

It should be noted the vector $R_W t_2 + p_W$ is just the target point coordinates in world frame under the initial estimated transformations. As a result, varying the robot configurations leads to similar values for this vector. Moreover, operator $(\cdot)^\wedge$ produces a singular skew-symmetric matrix. Consequently, stacking the parameter matrices across all configurations yields a large matrix with a high condition number, particularly due to the columns stacked by A_2. This makes the solution of Δa_Y to be highly sensitive to noise. To mitigate this issue, it is necessary to calibrate the point target at multiple positions.

B. Error Estimation

It is important to evaluate the effectiveness of the boost calibration algorithm. A natural approach is to measure the deviation between the desired target point and the actual reached point. To ensure robust calibration across all configurations, we focus on the worst-case error.

One configuration of a robotic arm can be interpreted as a combination of the pose of the robot arm (A or C) and the ultrasound observation (t_1 or t_2). In practice, the target point

lies within a constrained region, allowing us to characterize a configuration by two components: the coordinates of the target point in the world frame, and the pose of the ultrasound frame relative to the target point frame.

As illustrated in Fig. 4, we take robot 1 as an example. Let T denotes the transformation (actually a translation) from world frame (i.e., Robot 1 base) to target point frame, and denote $\bar{p}_T = [p_T, 1]$ as the homogeneous coordinates in world frame. Let T_1 denote the transformation from ultrasound 1 frame towards target point frame. By definition, the translation component of T_1 is $p_{T_1} = t_1$. Given T_1 and T, we can calculate the robot pose A can be calculated by $TT_1^{-1}(\bar{X}\Delta X)^{-1}$ and the actual reached point is given by $AX\bar{t}_1$ in world frame. A similar argument applies to Robot 2 and we get the calibration error on both sides:

$$E_1 = \left\| TT_1^{-1}(\bar{X}\Delta X)^{-1} X\bar{t}_1 - \bar{p}_T \right\|, \tag{5}$$

$$E_2 = \left\| Y(\bar{Y}\Delta Y)^{-1} TT_2^{-1}(\bar{Z}\Delta Z)^{-1} Z\bar{t}_2 - \bar{p}_T \right\|. \tag{6}$$

The above describes the calibration error under a single configuration, but our goal is to evaluate the worst-case calibration error. The configuration is parameterized by 8 variables: transformation T, parameterized by 3 variables; transformation T_1, which requires 5 variables since we use 2D ultrasound probe. Directly sampling over all 8 variables results in high computational cost. But fortunately, since the rotation matrices preserve the Euclidean norm of vectors, we can derive the following upper bound on calibration error:

$$E_1 \le \max_{t_1} \|(Q_X - I)\bar{t}_1\| := EU_1, \tag{7}$$

$$E_2 \le \max_{t_2} \|(Q_Z - I)\bar{t}_2\| + \max_{p_T} \|(Q_Y - I)\bar{p}_T\| := EU_2, \tag{8}$$

where $Q_X = (\bar{X}\Delta X)^{-1}X$, $Q_Y = Y(\bar{Y}\Delta Y)^{-1}$, $Q_Z = (\bar{Z}\Delta Z)^{-1}Z$. By applying (7) and (8), we observe that it suffices to sample only t_1 (2 variables) for Robot 1; and independently sample t_2 (2 variables) and p_T (3 variables) for Robot 2. This significantly reduces the computational cost compared to sampling the full 8-dimensional configuration space. Moreover, it is notable that the upper bound on the calibration error of the dual-arm robotic system only depends on the ultrasound observation t_1, t_2 and the position of the target point in the world frame p_T.

C. Validation

A set of ground truth transformations (i.e., X, Y, Z) is first provided. The validation process is outlined in the following algorithm.

Algorithm 1 Validation process for boost calibration

for $i = 1$ **to** N_e **do**
- Generate perturbed transformations $\bar{X}, \bar{Y}, \bar{Z}$
- **for** $j = 1$ **to** N_t **do**
  - Sample a target point
  - Select N_p observing ultrasound poses
  - (Add noise to t_1, t_2 for each pose)
- Calibrate via (4)
- Calculate error bound via (7)–(8)

In the algorithm, N_e denotes the number of experiments conducted. For each experiment, we generate perturbed transformations served as initial estimates, apply the boost calibration method, and estimate the resulting calibration error. Several points are worth noting:

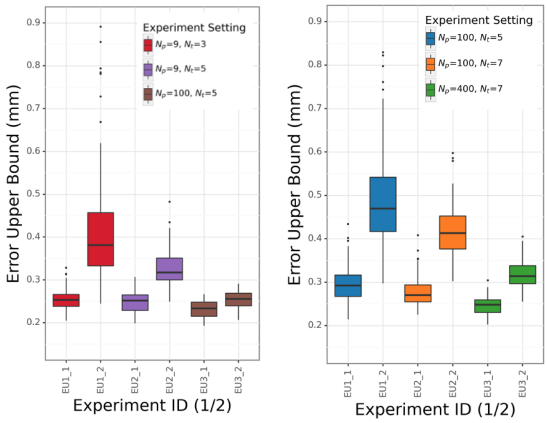
- **Configuration Selection:** Each experiment selects several configurations, defined by the target point in the world frame and ultrasound poses relative to it, which is equivalent to specifying the robot arm poses and the ultrasound observations.
- **Ultrasound inaccuracy modeling:** Due to the 2D nature of the probe and signal distortion, ultrasound observations are less precise. To simulate this and test the robustness of boost calibration, noise can be added to t_1 and t_2.

III. Results

We randomly add $1\%, 0.5°$ and 1% perturbations to the rotation axis, angle, and translation, respectively. For each case, we conduct $N_e = 100$ experiments. The simulation results are presented in Fig. 5.

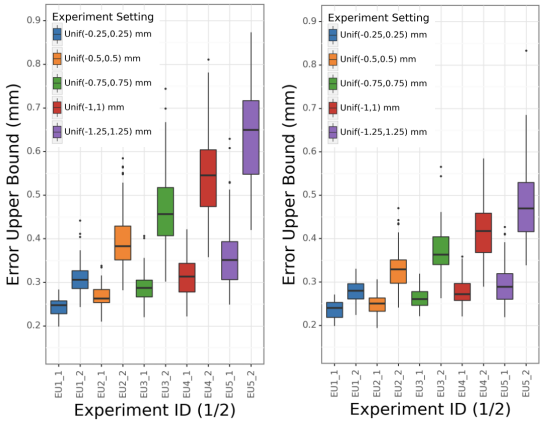
From Fig. 5, we observe that as the number of target points N_t increases, the calibration results become more stable, which is consistent with the discussion in Section II-A. With 5 distinct target points and 100 poses (essentially 10 poses per arm) for each point, the calibration achieves high precision, with errors for both arms around 0.25 mm, well below the ultrasound image resolution of 0.75 mm.

After considering ultrasound inaccuracy by adding noise $\text{Unif}(-1, 1)$ mm to t_1 and t_2, the method still achieves accurate result as N_t, N_p increase, as shown in Fig. 5b. With 7 distinct target points and 400 poses (essentially 20 poses per arm) for each point, the upper bounds on calibration error reach approximately 0.3 mm,which remains below the ultrasound image resolution of 0.75 mm. In Fig. 6, we present experiments under different noise levels. The results show that



(a) Without Ultrasound noise (b) With Ultrasound noise

Fig. 5: Simulation-based validation of the boost calibration method. (a) Without ultrasound noise; (b) Unif(-1,1) mm ultrasound noise added to t_1, t_2.



(a) $N_t = 6, N_p = 50$ (b) $N_t = 7, N_p = 100$

Fig. 6: Results with different noise level added

TABLE I: Absolute errors between two ultrasound frames before and after boost calibration, for A ($N_p = 9, N_t = 3$) and B ($N_p = 100, N_t = 7$) under ultrasound noise Unif($-1, 1$) mm, averaged over 100 experiments with the transrectal frame y-axis rotated from $-45°$ to $45°$.

Pose component	First stage	Boost calibrated	
		A	B
Roll (degree)	1.0	2.8e-2	2.5e-3
Pitch (degree)	0.4	3.3e-2	2.1e-2
Yaw (degree)	1.1	1.2e-2	4.9e-3
Translation distance (mm)	3.5	1.3e-2	5.3e-3

with $N_t = 6$ and $N_p = 50$, the error upper bound remains below 0.75 mm when the noise follows Unif$(-0.75, 0.75)$ mm. Furthermore, with $N_t = 7$ and $N_p = 100$, the method can tolerate noise up to Unif$(-1.25, 1.25)$ mm. This illustrates both the effectiveness and robustness of the boost calibration method.

We then conduct a simulation similar to that in real clinical use. The transrectal probe is fixed, while the abdominal probe is positioned just ahead of the transrectal probe, with the z-axis oriented in the opposite direction and the yz-plane aligned (since we are using 2D probes). The distance between the origins of the two ultrasound frames is set to 10 cm. We analyze the actual translation between these two ultrasound frames against the ground truth described above, before and after boost calibration. As shown in Tab. I, the "boost step" significantly increase the precision. Note that the variance across 100 rotation experiments is extremely small and is therefore omitted.

IV. Discussion and Conclusion

In summary, we propose a novel boost calibration method for dual-arm robotic systems using ultrasound, along with an efficient approach to estimate calibration error. Simulations show high accuracy despite ultrasound inaccuracy: without ultrasound noise, the upper bound of the calibration error is around 0.25 mm; with Unif(-1,1) mm noise in ultrasound observation, it reaches 0.3 mm. Assuming the ultrasound transducer operates at 1 MHz, its image resolution would be 0.75 mm, hence, the maximum error is well below half of the image resolution. Simulations further show that with 6 target points and 50 robot arm poses, the method tolerates Unif$(-0.75, 0.75)$ mm noise while keeping the error bound below the resolution; with 7 target points and 100 poses, it tolerates Unif$(-1.25, 1.25)$ mm noise.

A world frame must first be defined in the calibration process. In our setup, it was initially attached to the left robot arm holding the transrectal probe, which increased errors for the other arm due to the base-to-base transformation. Since the transrectal probe moves only within a limited range, it is more appropriate to attach the world frame to its robot base. The abdominal probe allows for more varied poses, enabling a more accurate base-to-base transformation.

This approach uses only 2D data, making it suitable for standard RGB cameras and ultrasound transducers. It can also be applied when assembling a dual-arm system from two pre-calibrated arms, or for recalibration after long-term use, particularly in clinical settings with collision risks. However, a limitation is the complexity of the calibration process, as it requires multiple target point positions and numerous robot poses for each position, even though this is almost inevitable when high precision in ultrasound transducer localization is desired. Another limitation is that we do not account for the poses that the robot arm cannot reach due to spatial constraints and clinical conditions. In practice, this restriction may actually reduce the calibration precision.

Our future work includes testing the algorithm on real robots and performing ultrasound tomography for prostate cancer detection.

References

[1] Bray F, Laversanne M, Sung H, et al. Global cancer statistics 2022: GLOBOCAN estimates of incidence and mortality worldwide. CA Cancer J Clin, 2024.

[2] National Cancer Institute. Cancer stat facts: Prostate cancer, 2024.

[3] Dias A B, O'Brien C, Correas J M, et al. Multiparametric ultrasound and micro-ultrasound in prostate cancer: a comprehensive review. Br J Radiol, 2022.

[4] Li C, Duric N, Littrup P, et al. In vivo breast sound-speed imaging with ultrasound tomography. Ultrasound Med Biol, 2009.

[5] Zhang H K, Cheng A, Kim Y, et al. Phantom with multiple active points for ultrasound calibration. J Med Imaging, 2018.

[6] Wang G, Li W, Jiang C, et al. Simultaneous calibration of multicoordinates for a dual-robot system by solving the AXB=YCZ problem. IEEE Trans Robot, 2021.

[7] Murray R M, Li Z, Sastry S S. A mathematical introduction to robotic manipulation. CRC Press, 2017.

[8] Gilboy K M, Wu Y, Wood B J, et al. Dual-robotic ultrasound system for in vivo prostate tomography. In: Int Workshop on Advances in Simplifying Medical Ultrasound, 2020.

[9] Aalamifar F, Khurana R, Cheng A, et al. Enabling technologies for robot assisted ultrasound tomography. Int J Med Robot Comput Assist Surg, 2017.

[10] Wiskin J W, Enders J, Williams C, et al. Imaging of prostate cancer with 3D ultrasound tomography. In: Medical Imaging: Ultrasonic Imaging and Tomography, 2022.

[11] Williams C, Daneshvar M, Wu Y, et al. Prostate ultrasound tomography: correlation with MRI and whole mount histopathology. J Urol, 2021.

[12] Parikh S H, Hesswani C, Azar W S, et al. 3D quantitative transmission ultrasound tomography for prostate cancer detection: an ex vivo study. J Urol, 2024.