Robotic Needle Insertion with 2D Ultrasound – 3D CT Fusion Guidance

Long Lei, Baoliang Zhao*, Xiaozhi Qi, Rui Mi, Hai Ye, Peng Zhang, Qiong Wang, Pheng-Ann Heng, Ying Hu*

***Abstract*—Puncture robots pave a new way for stable, accurate and safe percutaneous liver tumor puncture operation. However, affected by respiratory motion, intraoperative accurate location of the tumor and its surrounding anatomical structures remains a difficult problem in existing robot-assisted puncture operations. In this paper, a dual-arm robotic needle insertion system with guidance of intraoperative 2D ultrasound (US) and preoperative 3D computed tomography (CT) fusion is proposed, addressing the shortcomings of existing puncture robots. To deal with the challenge of cross-modal and cross-dimensional registration between 2D US and 3D CT, a decoupled two-stage registration approach combining initial vessel structure-based 3D US – 3D CT registration with intraoperative intensity-based 2D US - 3D US registration is proposed. To achieve fast and robust ultrasound probe calibration, a method based on an improved N-wire phantom is proposed. Twenty puncture experiments are performed in different breath-holding positions on a respiratory motion simulation platform, and experimental results show that the mean puncture error is 2.48 mm, which can meet the requirements in a wide of clinical scenarios.**

This work was supported by the National Key Research and Development Program of China (No. 2022YFB4703500), National Natural Science Foundation of China (Nos. 62273328, U21A20489), Guangdong Basic and Applied Basic Research Foundation (Nos. 2021B1515130003, 2021B1515120011, 2021A1515012264), Shenzhen Science and Technology Program (Nos. JCYJ20200109115201707, JCYJ20220818101408019), Research Grants Council of the Hong Kong Special Administrative Region (No. T45-401/22-N), and Hong Kong ITF Fund (No. GHP/080/20SZ). This work is also supported by CAS Key Laboratory of Human-Machine Intelligence-Synergy Systems, Shenzhen Institutes of Advanced Technology. (Corresponding author: Baoliang Zhao and Ying Hu.)

Long Lei is with Shenzhen Institute of Advanced Technology, Chinese Academy of Sciences, Shenzhen 518055, China, he is also with Department of Computer Science and Engineering, The Chinese University of Hong Kong, Hong Kong 999077, China (e-mail: long.lei@siat.ac.cn).

Baoliang Zhao, Xiaozhi Qi, Peng Zhang, Qiong Wang and Ying Hu are with Shenzhen Institute of Advanced Technology, Chinese Academy of Sciences, Shenzhen 518055, China (bl.zhao@siat.ac.cn; xz.qi@siat.ac.cn; zhangpeng@siat.ac.cn; wangqiong@siat.ac.cn; ying.hu@siat.ac.cn).

Rui Mi and Hai Ye are with Department of Radiology, Shenzhen University General Hospital, Shenzhen University Clinical Medical Academy, Shenzhen 518055, China (e-mail: mirror@szu.edu.cn; yehai@szu.edu.cn).

Pheng-Ann Heng is with Department of Computer Science and Engineering, The Chinese University of Hong Kong, Hong Kong 999077, China, he is also with Guangdong-Hong Kong-Macao Joint Laboratory of Human Machine Intelligence-Synergy Systems, CUHK, Hong Kong 999077, China (e-mail: pheng@cse.cuhk.edu.hk).

***Note to Practitioners*—In clinical percutaneous liver tumor puncture operation, due to the lack of real-time and clear image guidance, it is difficult to locate the tumor and its surrounding vital anatomical structures. In addition, the stability and accuracy of manual operation are poor. The development of a puncture robot is an effective solution for these problems. However, existing CT and magnetic resonance imaging (MRI) guided robots do not consider the tumor localization errors caused by inconsistent breath-holding positions between preoperative scan period and intraoperative puncture period, and US guided robots are limited by the poor image quality and the narrow field of vision. In this paper, a dual-arm robotic needle insertion system with guidance of intraoperative 2D US and preoperative 3D CT fusion is proposed. This system can take advantage of the real-time ultrasound and clear CT images at the same time, and can provide real-time, clear and all-round guidance for percutaneous liver tumor puncture operation, which has obvious advantages over the existing puncture robots. Phantom experiments have been completed and animal experiments will be carried out in the future.**

***Index Terms*—Percutaneous puncture, puncture robot, respiratory motion, ultrasound - CT registration, ultrasound probe calibration.**

I. Introduction

PERCUTANEOUS image-guided needle insertion is a basic operation during the diagnosis and treatment of liver cancer. Whether in biopsy or ablation, needles are required to be accurately inserted into the tumor along the preoperative planned path, thereby ensuring the operation accuracy and safety [1]. In practice, the liver is located in the abdominal cavity and moves under the influence of respiration, and the displacement in the cranial-caudal (CC) direction can reach 23 mm [2]. However, clinical puncture operation lacks clear and real-time image guidance. Although CT and MRI can clearly image the anatomical area of the target, they cannot provide real-time image guidance during needle insertion. Ultrasound imaging can capture the motion of the target anatomical area in real time, while the images are with low signal-to-noise ratio, poor soft tissue contrast and 2D surgical field of view. These problems of image guidance make it difficult to determine the position of the liver tumor and its surrounding anatomical structures during the clinical needle insertion operation. In addition, the accuracy and stability of manual operation are highly dependent on the experience and

attention of the doctor, especially for the ultrasound guided puncture, since bimanual operation has extremely high requirements on the two hands coordination and the hand-eye coordination of the doctor.

With the development of robotics, puncture robots pave a new way for stable, accurate and safe puncture operation [3]. To reduce the impact of respiratory motion on puncture accuracy, most existing CT/MRI guided robots [4, 5] use a breath-holding puncture strategy. Preoperative scan of the patient in an end-expiratory breath-holding state is required to perform the image-to-robot calibration and plan the needle insertion path, then the patient holds his or her breath again, and the robot autonomously aligns the needle with the planned path and completes needle insertion. Since the breath-holding is subjectively controlled by the patient, it is difficult to ensure that the preoperative scanning phase and the intraoperative puncture phase are in the same breath-holding position. As a result, the preoperative calibration results are not applicable to the intraoperative puncture, which further leads to puncture errors. The commercial robot system XCAT ACE [6] overcomes the problem of target drift caused by inaccurate breath-holding position by step-by-step check and active puncture angle correction. However, step-by-step check increases the risk of radiation exposure for patients, and intraoperative puncture angle adjustment can cause tissue pulling and increase surgical trauma. There are also some CT/MRI guided puncture robots [7, 8] that adopt a respiratory gating approach, and the needle is advanced only during the static portion of the respiratory cycle. When the organ motion is detected, the needle will be released and move freely with the organ [7], or the robot actively follows the organ movements to avoid the needle pulling the tissue [8]. However, the gating strategy prolongs the operation time and increases the risk of surgery. In our previous work [9], the tumor is intraoperatively located by developing a model that correlates the tumor motion with the position of a surface marker. However, the modeling requires multiple preoperative CT scans, which increases the radiation exposure for patients.

For the puncture robots under the guidance of real-time 2D US, needle insertion is generally limited in the imaging plane to see the target tumor and the needle at the same time from images [10, 11]. There is a trade-off between the optimal needle insertion plane and the optimal imaging plane, and the vital anatomical structures around the target tumor are easy to be mistakenly penetrated. Although flexible steerable needle puncture robots [12] can insert the needle according to complex trajectories to avoid obstacles such as blood vessels, rotating the needle shaft inevitably causes additional trauma to the patient when the needle is bent. In addition, real-time tracking of anatomical structures [13] and needle tips [14, 15] remains challenging due to the inherent defects of low signal-to-noise ratio and limited field of view of ultrasound images.

Fusion of preoperative 3D CT/MRI images with intraoperative 2D US images based on image registration technology can provide real-time and all-round clear guidance for needle-based interventions [16]. 2D US – 3D CT/MRI registration belongs to the cross-modal and cross-dimensional image registration problem [17], and the images to be aligned are very different in terms of not only appearance but also fields of view. In commercial navigation systems, such as GE V-Nav [18], Siemens eSieFusion [19] and Samsung S-fusion [20], a rigid transformation that optimally aligns an US slice with a preoperative CT/MRI volume is firstly determined based on some manually identified anatomic fiducials. Then the transformation is updated in real time according to the pose change of the probe, which is usually measured by an optical or electromagnetic tracking system. Although this method has good real-time performance, the intraoperative real-time registration relies on the initial transformation. Once the target organ position changes due to breathing, the initial transformation no longer applies, which will result in alignment errors. Another solution is seeking a reliable cross-modal similarity measure criterion that quantifies the level of alignment between the images to be registered. Several complex cross-modal similarity criterions have been proposed, such as the modality independent neighborhood descriptor (MIND) [21], contextual conditioned mutual information (CoCoMI) [22] and their extensions. However, their robustness in the 2D US – 3D CT/MRI registration needs to be further verified due to the fact that the useful information in an ultrasound image is sparse and noise-corrupted. In addition, Wein et al. [23] converted the 2D US – 3D CT registration to the monomodal registration between an US image and a simulated US volume generated from a given CT volume. However, seeking a reliable simulation method of US images from CT has become a new challenge [24]. Recently, the explosive development of deep learning has paved a promising path for the 2D US – 3D CT/MRI registration [25]. For example, Markova et al. [26] used two separate feature extraction networks to learn dense keypoint descriptors, and RANSAC is deployed to align them to obtain the pose of the input ultrasound image relative to the input MRI volume. Guo et al. [27] developed an ultrasound frame-to-volume registration pipeline based on end-to-end deep convolutional neural networks. In addition, deep convolutional neural networks are also used to learn the cross-modal similarity criterion [28] and the modal translation process [29]. Even though superior efficiency and competitive accuracy are reported, deep learning models often require a large amount of data and time for training, and the generalization ability of the models are hard to guarantee. In a word, the 2D US – 3D CT/MRI registration remains a popular research direction in the field of medical image analysis and requires further study.

US probe calibration is also a key step to determine the accuracy of the robotic needle insertion system with guidance of 2D US – 3D CT/MRI fusion [30]. US probe calibration is to obtain the coordinate transformation between the US image and the probe locator. Existing calibration methods are based on an artificial phantom with known geometry and with

a set of fiducials that can be localized both in the US images and by the probe locator [31]. Among these calibration models, the N-wire phantom [32] and its variants [33-35] have become the most common phantoms due to their advantages of simple structure, easiness to manufacture and operator-independence. Chen et al. proposed a closed-form solution of calibration parameters [33], while only the middle line in N-wire is used to compute the calibration transformation matrix, which is with low efficiency, and the detection accuracy of the fiducials in ultrasound images may lead to a skewed transformation matrix. Carbajal et al. proposed a cost function to take into account all the lines in N-wire phantom, and the cost function was iteratively optimized to obtain a rigid transformation matrix [34]. However, this method is sensitive to initial value and easy to get stuck at locally optimal value.

This study aims to achieve stable, accurate and safe percutaneous liver tumor puncture, and focuses on solving the intraoperative positioning problem of tumors and surrounding anatomical structures in robot-assisted puncture operations. The main contributions of this article are listed below.

i) To solve the problems of existing CT/MRI/US guided puncture robots, a dual-arm robotic needle insertion system with the guidance of intraoperative 2D US and preoperative 3D CT fusion is proposed.

ii) To address the challenge of the cross-modal and cross-dimensional 2D US – 3D CT registration, a decoupled two-stage approach combining vessel structure-based 3D US – 3D CT registration with intensity-based 2D US - 3D US registration is proposed, and a three-level pyramid registration framework based on Markov Random Field (MRF) is adopted to optimally align the 3D US with the 2D US images.

iii) To improve the efficiency and robustness of ultrasound probe calibration, a novel N-wire phantom with multiple parallel-distributed interphase lines is developed, and the closed-form solution of optimal transformation is obtained.

iv) To test the performance of the proposed system, puncture experiments are performed at different breath-holding positions on a respiratory motion simulation platform.

The rest of this paper is organized as follows. Section II presents the overview of the proposed dual-arm robotic needle insertion system with guidance of intraoperative 2D US and preoperative 3D CT fusion. The proposed US probe calibration and 2D US – 3D CT registration methods are introduced in Section III and Section IV, respectively. Experiments and results are presented in Section V. Finally, the conclusion is presented in Section VI.

II. METHOD OVERVIEW AND SYSTEM SETUP

A. System Setup

Fig. 1 shows the proposed dual-arm robotic needle insertion system with guidance of intraoperative 2D US and preoperative 3D CT fusion. In this system, the accurate positioning of the target tumor and its surrounding anatomical structures under the influence of respiratory motion is achieved by means of the 2D US – 3D CT registration, and the puncture stability is achieved by replacing manual operation with a dual-arm robot.

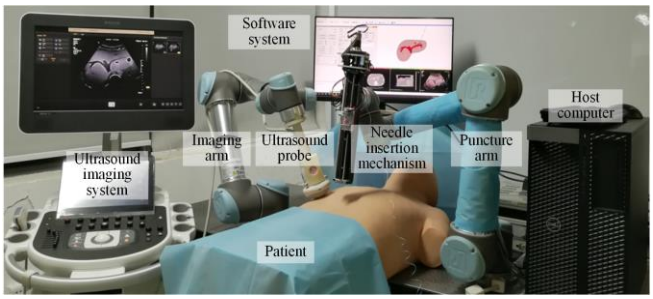


Fig. 1. The developed dual-arm robotic needle insertion system with guidance of intraoperative 2D US and preoperative 3D CT fusion.

The hardware of the system mainly includes two lightweight robot arms (UR5, Universal Robots, Denmark), a customized needle insertion mechanism, an ultrasound imaging system (Affiniti 30, Philips Inc, Holland) with a volume probe (V6-2, Philips Inc, Holland), a frame grabber (OK_VGA41A-4E+, Beijing JoinHope Image Technology Co., Ltd, China) and a host computer. The UR5 robot arm has 6 degrees of freedom (DOFs) and is with repeated positioning accuracy of ± 0.1 mm, it can flexibly and accurately adjust the position and orientation of the probe and needle as needed. The volume probe can perform high-resolution 2D imaging of the human abdomen at a frequency of up to 66 Hz, it also can perform high-resolution 3D imaging at a frequency of 2 Hz. The frame grabber is connected on the video output of the ultrasound imaging system to collect and transfer live image data to the host computer. The needle insertion mechanism [9] can achieve a needle insertion length of up to 300 mm with a repeated positioning accuracy of ± 0.005 mm.

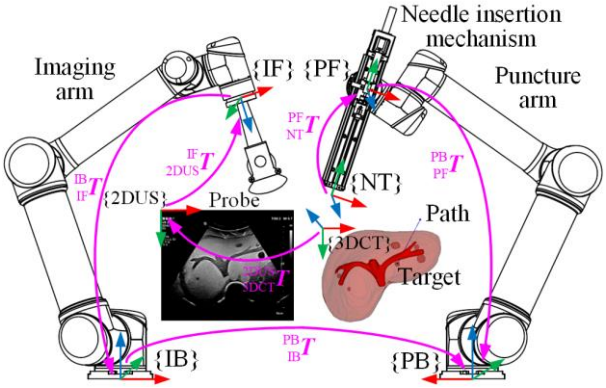


Fig. 2. The transformations of the coordinate systems in the developed system.

B. System Working Principle

The transformations of coordinate systems in the developed system are illustrated in Fig. 2. {3DCT} represents the CT image coordinate system, {2DUS} represents the ultrasound image coordinate system, {IF} represents the end flange coordinate system of the imaging arm, {IB} represents the base coordinate system of the imaging arm, {PB}

represents the base coordinate system of the puncture arm, {PF} represents the end flange coordinate system of the puncture arm, {NT} represents the needle tip coordinate system. ${}^{2DUS}_{3DCT}\boldsymbol{T}$, ${}^{IF}_{2DUS}\boldsymbol{T}$, ${}^{IB}_{IF}\boldsymbol{T}$, ${}^{PB}_{IB}\boldsymbol{T}$, ${}^{PB}_{PF}\boldsymbol{T}$ and ${}^{PF}_{NT}\boldsymbol{T}$ represent the rigid transformations between related coordinate systems. Among these transformations, ${}^{IF}_{2DUS}\boldsymbol{T}$, ${}^{PB}_{IB}\boldsymbol{T}$ and initial ${}^{PF}_{NT}\boldsymbol{T}$ are preoperatively calibrated, ${}^{2DUS}_{3DCT}\boldsymbol{T}$ is intraoperatively obtained based on the 2D US – 3D CT registration, ${}^{IB}_{IF}\boldsymbol{T}$ and ${}^{PB}_{PF}\boldsymbol{T}$ are provided by the real-time robot feedback. As a result, a point in {3DCT} can be mapped into the puncture robot space as in Equation (1).

$$ {}^{PB}\boldsymbol{P} = {}^{PB}_{IB}\boldsymbol{T}\,{}^{IB}_{IF}\boldsymbol{T}\,{}^{IF}_{2DUS}\boldsymbol{T}\,{}^{2DUS}_{3DCT}\boldsymbol{T}\,{}^{3DCT}\boldsymbol{P} \tag{1} $$

where ${}^{3DCT}\boldsymbol{P}$ represents a point in {3DCT}, and ${}^{PB}\boldsymbol{P}$ represents the corresponding point in {PB}.

C. Software Framework

The developed software of the robotic system consists of four submodules: an image archiving and communication module, a system calibration module, an image analysis and visualization module and a needle insertion control module. The software is developed with the open-source Qt toolkit, and the data transmission among the four submodules is achieved with the open-source OpenIGTLink protocol [36]. Details about each of the four modules are given below.

1) Image archiving and communication module. The US images in DICOM format can only be manually sent from the DICOM server of the ultrasound imaging system to the DICOM client of the host computer through TCP/IP protocol, while the frame grabber can only collect the images in JPG format in real time through the display port (DP) of the ultrasound system, thus the image archiving and communication module is developed to convert live JPG images into DICOM images according to the tag information in the single DICOM image received by the DICOM client, and then the synthesized DICOM images are archived and sent to other modules. This module is developed with the open-source DICOM Toolkit (DCMTK) [37]. To improve the speed of image processing, multithreading and parallel processing are adopted. Finally, the stable storage and transmission frequency of 10 Hz can be realized.

2) System calibration module. This module is used to conduct US probe calibration, dual-arm calibration and needle tip calibration, and to obtain the transformations ${}^{IF}_{2DUS}\boldsymbol{T}$, ${}^{PB}_{IB}\boldsymbol{T}$ and initial ${}^{PF}_{NT}\boldsymbol{T}$. This module is developed with the Public Software Library for Ultrasound (PLUS) [38].

3) Image analysis and visualization module. This module is used to perform needle insertion path planning and 2D US – 3D CT registration, and to monitor the needle insertion process by visualizing the needle, target tumor and its surrounding anatomical structures. This module is developed with the open-source 3D Slicer [39] and its extension SlicerIGT [40]. The refresh rate of the navigation interface is 10Hz.

4) Needle insertion control module. This module is used to communicate with the two robot arms and needle insertion mechanism, sending the control commands and obtaining their motion status. The refresh frequency of the main thread of this module is 20 Hz.

III. Ultrasound Probe Calibration

A. The Overall Framework of Calibration Method

To obtain the transformation ${}^{IF}_{2DUS}\boldsymbol{T}$, a novel US probe calibration method based on an improved N-wire phantom is proposed in this paper. The calibration setup and transformations of coordinate systems involved are shown in Fig. 3(a). The calibtaion phantom is fixed, the imaging arm holds the US probe to image the fiducials in the calibration phantom for several times with different positions, and the corresponding tansformation ${}^{IF}_{IB}\boldsymbol{T}$ (the inverse of ${}^{IB}_{IF}\boldsymbol{T}$) is recorded for each US image, denoted as ${}^{IF}_{IB}\boldsymbol{T}_i$. Denote ${}^{2DUS}\boldsymbol{P}_{ij}$ as the the j-th fiducial in the i-th US image, its position vector in the calibration phantom coordinate system {CP}is ${}^{CP}\boldsymbol{P}_{ij}$, and its position vector in the coordiante system {IF} ${}^{IF}\mathbf{P}_{ij}$ can be calculated by two transformation routes, as shown in Equation (2), where, ${}^{IB}_{CP}\boldsymbol{T}$ is the constant transformation from the calibration phantom coordinate system to the base coordinate system of the imaging arm.

$$ {}^{IF}_{2DUS}\boldsymbol{T}\begin{bmatrix} {}^{2DUS}\boldsymbol{P}_{ij} \\ 1 \end{bmatrix} = \begin{bmatrix} {}^{IF}\boldsymbol{P}_{ij} \\ 1 \end{bmatrix} = {}^{IF}_{IB}\boldsymbol{T}_i\,{}^{IB}_{CP}\boldsymbol{T}\begin{bmatrix} {}^{CP}\boldsymbol{P}_{ij} \\ 1 \end{bmatrix} \tag{2} $$

Thus, once the pair-wise point set $\{({}^{2DUS}\boldsymbol{P}_{ij}, {}^{CP}\boldsymbol{P}_{ij})|i=1,...,M, j=1,...,N\}$ and ${}^{IB}_{CP}\boldsymbol{T}$ are obtained, the paired point set $\{({}^{2DUS}\boldsymbol{P}_{ij}, {}^{IF}\boldsymbol{P}_{ij})|i=1,...,M, j=1,...,N\}$ can be obtained based on the right side route in Equation (2). Then, ${}^{IF}_{2DUS}\boldsymbol{T}$ can be optimized to align the two point sets $\{{}^{2DUS}\boldsymbol{P}_{ij}|i=1,...,M, j=1,...,N\}$ and $\{{}^{IF}\boldsymbol{P}_{ij}|i=1,...,M, j=1,...,N\}$ in the least squares sense, as in Equation (3). In our work, a singular value decomposition (SVD)-based method [41] is adopted to obatin the closed-form solution of the optimal rigid ${}^{IF}_{2DUS}\boldsymbol{T}$, improving the robustness of the calibration.

$$ {}^{IF}_{2DUS}\boldsymbol{T} = \underset{{}^{IF}_{2DUS}\boldsymbol{R}\in SO(3),\ {}^{IF}_{2DUS}\boldsymbol{t}\in\mathbb{R}^3}{\arg\min} \sum_{i}^{M}\sum_{j}^{N}\left\| \left({}^{IF}_{2DUS}\boldsymbol{R}\,{}^{2DUS}\boldsymbol{P}_{ij} + {}^{IF}_{2DUS}\boldsymbol{t} \right) - {}^{IF}\boldsymbol{P}_{ij} \right\|^2 \tag{3} $$

where ${}^{IF}_{2DUS}\boldsymbol{R}$ and ${}^{IF}_{2DUS}\boldsymbol{t}$ are the rotation matrix and translation vector corresponding to the rigid transformation ${}^{IF}_{2DUS}\boldsymbol{T}$, separately.

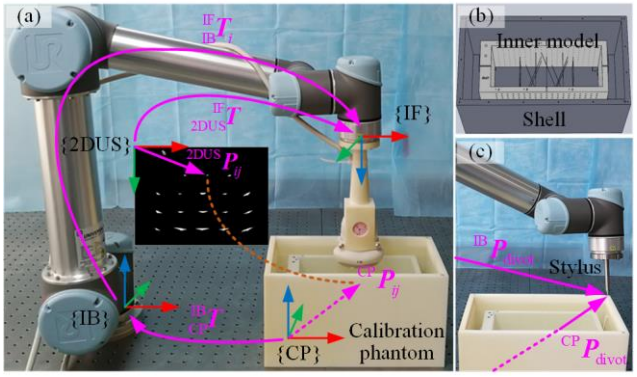


Fig. 3. (a) The US probe calibration setup and transformations of coordinate systems, (b) the designed calibration phantom, and (c) the calibration of ${}^{IB}_{CP}\boldsymbol{T}$, ${}^{IB}\boldsymbol{P}_{\text{divot}}$ is the divot position in {IB}, its corresponding position in {CP} is ${}^{CP}\boldsymbol{P}_{\text{divot}}$.

B. N-Wire Based Calibration Phantom

In order to obtain the pair-wise point set $\{({}^{2DUS}\boldsymbol{P}_{ij}, {}^{CP}\boldsymbol{P}_{ij}) | i=1,...,M, j=1,...,N\}$ and the transformation ${}^{IB}_{CP}\boldsymbol{T}$, a novel calibration phantom is designed (Fig. 3(b)). The calibration phantom consists of a shell and an inner model, several small divots are unevenly distributed on the shell, and these divots are used to calculate the transformation ${}^{IB}_{CP}\boldsymbol{T}$. The inner model has some small holes with a diameter of 0.6 mm evenly distributed on its front and back walls, and these holes are used to arrange nylon wires with a diameter of 0.5 mm to construct N-wire structures. The shell and inner model are printed by a 3D printer (Lite600HD, Shanghai Union Technology Corporation, China) with an accuracy of ±0.1 mm, and there is a tight fit between them.

As shown in Fig. 3(c), the ${}^{IB}_{CP}\boldsymbol{T}$ is calculated as follows. Firstly, a stylus is installed on the imaging arm flange, and its tip is set as the tool center point of the robot arm. Then the positions of at least four divots in {IB} are measured by the stylus. The positions of these divots in {CP} are also known in the designed calibration phantom. With the corresponding positions of these divots in the frame {CP} and the frame {IB}, optimal rigid ${}^{IB}_{CP}\boldsymbol{T}$ can be finally obtained by the SVM-based method [41].

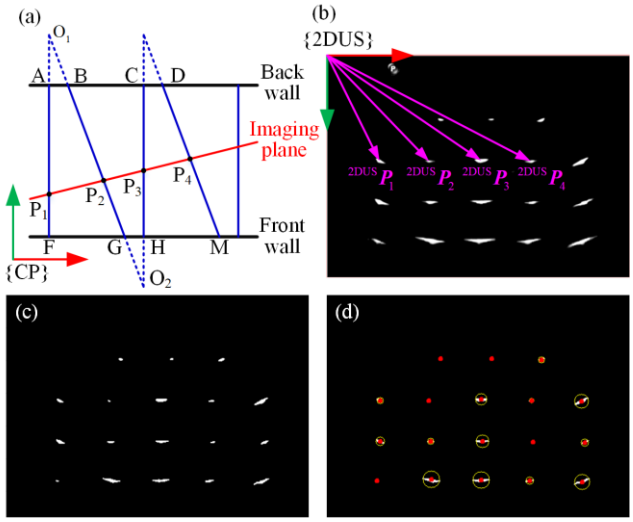


Fig. 4. (a) Sketch of the N-wires in the calibration phantom, (b) original intersection points between the N-wires and the US imaging plane, (c) bright spots after threshold segmentation and (d) the centers (red dots) of their minimum envelope circles (yellow circles).

The N-wires in the inner model are assigned as in Fig. 4(a), and the interphase lines are arranged in parallel so that any three adjacent lines can form an N-wire structure. During calibration, due to the presence of ultrasound imaging artifacts, the intersection points between the N-wires and the imaging plane appear as elongated bright spots in the ultrasound image, as shown in Fig. 4(b). Generally, the greater the angle between the nylon wire and the normal of the imaging plane, or the farther the intersection point is from the probe, the more obvious the artifacts. For the obtained ultrasound images, the central area of each bright spot is firstly extracted by performing personalized threshold segmentation to reduce the influence of artifacts on the detection accuracy of the intersection points, as shown in Fig. 4(c). Then, after detecting the contours of these central areas, the minimum envelope circles are calculated, and their centers are finally regarded as the intersection points, as shown in Fig. 4(d).

For the N-wire that consists of the wires AF, BG and CH, which intersects with the imaging plane at points P_1, P_2 and P_3, these intersection locations in the ultrasound image are detected as ${}^{2DUS}\boldsymbol{P}_1$, ${}^{2DUS}\boldsymbol{P}_2$ and ${}^{2DUS}\boldsymbol{P}_3$, respectively. According to the similar triangle principle [34], as shown in Equation (4), ${}^{CP}\boldsymbol{P}_2$ can be obtained. For the N-wire that consists of the wires BG, CH and DM, given ${}^{2DUS}\boldsymbol{P}_2$, ${}^{2DUS}\boldsymbol{P}_3$ and ${}^{2DUS}\boldsymbol{P}_4$, ${}^{CP}\boldsymbol{P}_3$ can be obtained in the same way. In this way, the intersection points between all wires and the imaging plane can be maximally utilized to improve the calibration efficiency.

$$\frac{|O_1P_2|}{|O_1O_2|-|O_1P_2|}=\frac{\left\|{}^{2DUS}\boldsymbol{P}_1-{}^{2DUS}\boldsymbol{P}_2\right\|}{\left\|{}^{2DUS}\boldsymbol{P}_2-{}^{2DUS}\boldsymbol{P}_3\right\|} \tag{4}$$

where $|\cdot|$ represents the length of the corresponding line segment.

IV. Intraoperative 2D US - Preoperative 3D CT Registration

A. General Framework

Given a preoperatively acquired CT volume $I_{3DCT}: \Omega_{3DCT} \in \mathbb{R}^3 \rightarrow \mathbb{R}$, and an intraoperative ultrasound image $I_{2DUS}: \Omega_{2DUS} \in \mathbb{R}^3 \rightarrow \mathbb{R}$, in order to obtain the transformation ${}^{2DUS}_{3DCT}\boldsymbol{T}$ that optimally aligns I_{3DCT} with I_{2DUS}, a novel registration method based on decoupling strategy is proposed.

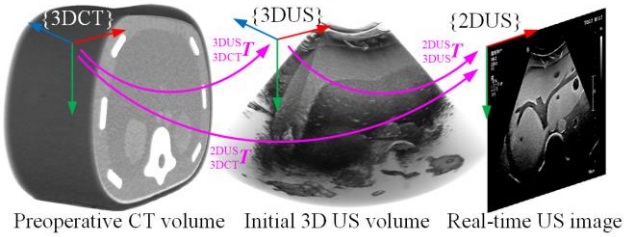


Fig. 5. Intraoperative 2D US - preoperative 3D CT registration.

As shown in Fig. 5, by initially performing the 3D US imaging of target organ, denoting the obtained ultrasound volume as $I_{3DUS}: \Omega_{3DUS} \in \mathbb{R}^3 \rightarrow \mathbb{R}$, the 2D US – 3D CT registration is decoupled as cross-modal 3D US - 3D CT registration and cross-dimensional 2D US - 3D US registration. The 3D US - 3D CT registration is conducted based on segmented intrahepatic vessel structures to obtain the transformation ${}^{3DUS}_{3DCT}\boldsymbol{T}$, the 2D US - 3D US registration is conducted based on gray-level feature to obtain ${}^{2DUS}_{3DUS}\boldsymbol{T}$. The initial 3D US volume can be obtained by robotic 3D ultrasound scan and reconstruction [42-44] or a 3D volume probe.

B. Initial 3D US - 3D CT Registration

Initial 3D US - 3D CT registration is performed based on intrahepatic vessel structures to bypass the need for cross-modal similarity measure criterion, and to ensure the registration robustness, and the workflow is shown in Fig. 6.

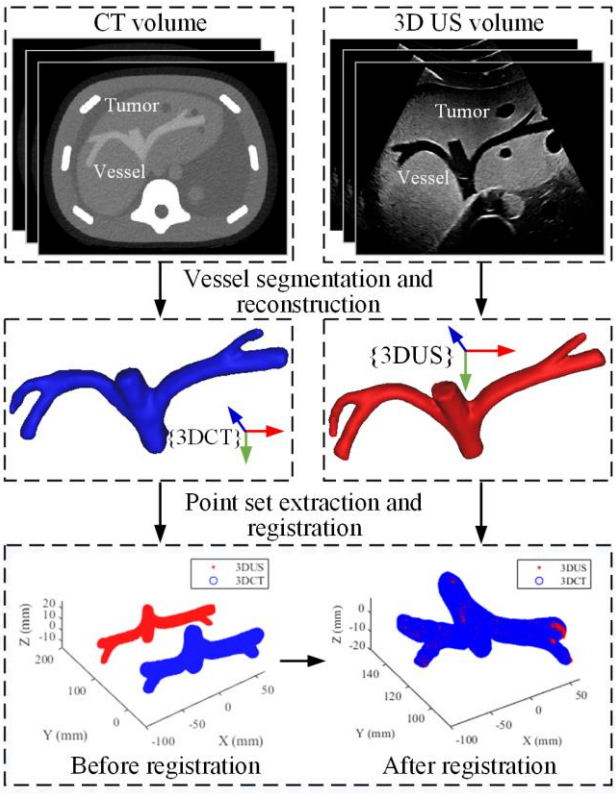


Fig. 6. Schematic diagram of initial 3D US – 3D CT registration based on the point sets of vessel surface in CT and 3D US volumes.

The intrahepatic vessels in CT images are segmented by the U-Net [45] mainly trained on the public MSD dataset [46]. For the vessel segmentation in ultrasound images, a cross-modal transfer learning method proposed in our previous work [29] is adopted. That is, ultrasound-style images are firstly simulated from a mass of labeled CT images to construct a transition dataset, which is used to bridge the huge domain discrepancy between CT and ultrasound images, then the U-Net is pre-trained with the transition set, and finally the trained network is transferred to the real ultrasound dataset with small data size for fine tuning. In this way, the problem of insufficient labeled ultrasound data in the task of intrahepatic vessel segmentation is solved. On the basis of the intrahepatic vessel segmentation, the intrahepatic vessels are reconstructed [1], and the points on the vessel surface in CT and 3D US images are evenly extracted to construct the point sets to be registered. Then the coherent point drift algorithm [47] is adopted to optimally align the two point sets, and the rigid transformation ${}^{3\mathrm{DUS}}_{3\mathrm{DCT}}\boldsymbol{T}$ is obtained.

C. Intraoperative 2D US - 3D US Registration

In the field of image registration, the methods combing graph model with discrete optimization have the nature of global search [17]. In order to obtain the rigid transformation ${}^{2\mathrm{DUS}}_{3\mathrm{DUS}}\boldsymbol{T}$ that can optimally align the initial 3D US with the intraoperative 2D US, inspired by Zikic et al. [48] and Porchetto et al. [49], a three-level pyramid registration framework based on the Markov Random Field (MRF) is adopted and multiple iterations are performed at each level.

We specify the transformation ${}^{2\mathrm{DUS}}_{3\mathrm{DUS}}\boldsymbol{T}$ by $para = (r_\mathrm{x}, r_\mathrm{y}, r_\mathrm{z}, t_\mathrm{x}, t_\mathrm{y}, t_\mathrm{z})$, in which $(t_\mathrm{x}, t_\mathrm{y}, t_\mathrm{z})$ is a translation vector, and the Euler angles $(r_\mathrm{x}, r_\mathrm{y}, r_\mathrm{z})$ represents a composition of sequential rotations about the axes Z, X, Y. In each iteration, we formulate the registration as a discrete labeling problem on a pairwise MRF with six nodes $G = \langle V, E\rangle$, as shown in Fig. 7, where each node represents a discrete random variable $v_i \in V, i = 1, \ldots, 6$, which is associated with a parameter of $para$. Given an initial transformation $para^0 = (r_\mathrm{x}^0, r_\mathrm{y}^0, r_\mathrm{z}^0, t_\mathrm{x}^0, t_\mathrm{y}^0, t_\mathrm{z}^0)$, for each variable v_i, set it a k_i-divided symmetric space L_i of radius ρ_i centered on its initial value, that is, using the label space $L = \{L_i \mid i = 1, \ldots, 6\}$ to discretize the continuous parameter space of $para$. Each edge $(v_i, v_j) \in E,\ i, j = 1, \ldots, 6, i \neq j$ represents the interdependency between the corresponding parameters. For the mono-modal image registration, a pairwise potential $p(v_i, v_j)$ in Equation (5) is associated to each edge $(v_i, v_j) \in E$ to quantify the sensitivity of registration to the variations of the related parameters. Based on the constructed MRF, the optimal rigid transformation $\widehat{para}$ is approximately obtained by searching for the minimum value of Equation (6) in the label space L, which is solved by the Fast-PD optimization algorithm proposed by Komodakis et al. [50]. After each iteration, ρ_i is reduced by a factor λ_i, that is, $\rho_i = \rho_i(1 - \lambda_i)$. When the sum of squared differences (SSD) [49] between the $I_{2\mathrm{DUS}}$ and the extracted slice from $I_{3\mathrm{DUS}}$ specified by the current parameters is no longer reduced, or the maximum number of iterations is reached, the iteration is terminated and the iteration in the next pyramid level is started. The pseudo-code of the intraoperative 2D US – 3D US registration method is shown in **Algorithm 1**.

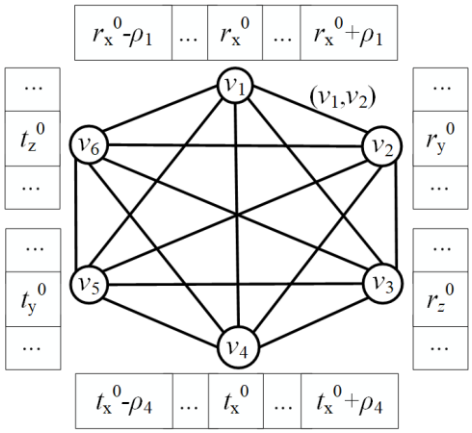


Fig. 7. Pairwise Markov Random Field with six nodes $G = \langle V, E\rangle$.

$$p(v_i, v_j) = \mathrm{SSD}\left(I_{2\mathrm{DUS}}, para_{i,j}[I_{3\mathrm{DUS}}]\right) \tag{5}$$

where $para_{i,j}$ denotes the parameters after only v_i and v_j are updated from $para^0$, $para_{i,j}[I_{3\mathrm{DUS}}]$ denotes the extracted slice from the $I_{3\mathrm{DUS}}$ specified by the $para_{i,j}$, $\mathrm{SSD}(\cdot,\cdot)$ is to quantify the dissimilarity between the image $I_{2\mathrm{DUS}}$ and the extracted slice $para_{i,j}[I_{3\mathrm{DUS}}]$.

$$\hat{para} = \underset{L}{\operatorname{argmin}} \sum_{(v_i, v_j)\in E} p(v_i, v_j) \tag{6}$$

Algorithm 1: Intraoperative 2D US – 3D US registration.

Input: intraoperative 2D US image $I_{2\text{DUS}}$; 3D US volume $I_{3\text{DUS}}$; sampling rates of three pyramid levels $\{r_l | l = 1,2,3\}$; initial label space radii of six transformation parameters for the three levels $\{\rho_{i,l}^0 | i = 1,\dots,6;\ l = 1,2,3\}$; label space reduction factors of the six parameters $\{\lambda_i | i = 1,\dots,6\}$; maximum number of iterations for the three levels $\{N_l | l = 1,2,3\}$; maximum number of continuous iterations during which the registration performance does not improve C.

Output: optimal transformation parameters $para = (r_x, r_y, r_z, t_x, t_y, t_z)$

$para \leftarrow para^0 = (r_x^0, r_y^0, r_z^0, t_x^0, t_y^0, t_z^0)$
for $l = 1,2,3$ **do**
  $I_{2\text{DUS}}, I_{3\text{DUS}} \leftarrow \text{Sampling}(I_{2\text{DUS}}, I_{3\text{DUS}}, r_l)$
  $c \leftarrow 0; n \leftarrow 0$
  $\rho_i \leftarrow \rho_{i,l}^0, i = 1,\dots,6$
  while $c \neq C$ or $n < N_l$ **do**
    $S \leftarrow \text{SSD}(I_{2\text{DUS}}, para[I_{3\text{DUS}}])$
    $L \leftarrow \text{Construct\_Label\_Space}(para, \rho_i)$
    $\widehat{para} \leftarrow \text{Fast-PD\_Optimization}(G, L)$
    $S^* \leftarrow \text{SSD}(I_{2\text{DUS}}, \widehat{para}[I_{3\text{DUS}}])$
    if $S^* < S$ **do**
      $para \leftarrow \widehat{para}; S \leftarrow S^*; c \leftarrow 0$
    else do
      $c \leftarrow c + 1$
    end
    $\rho_i \leftarrow \rho_i(1 - \lambda_i)$
  end
end

When the patients hold their breath in a position, the 3D US imaging and 2D US imaging are performed successively, and there is an inherent coordinate transformation relationship between the single image obtained by the 2D US imaging and the volume obtained by the 3D US imaging, which is denoted as ${}^{2\text{DUS}}_{3\text{DUS}}\boldsymbol{T}_0$. Then, if the patients hold their breath in a different position, the relationship between the real-time obtained 2D US slice and the previously obtained 3D US volume will change, and the deviation equals to the displacement of organ. We use the maximum amplitude of liver motion as the prior information to limit the range of the discrete label space, which ensures the accuracy and efficiency of the registration.

V. PHANTOM EXPERIMENTS AND RESULT ANALYSIS

A. Ultrasound Probe Calibration Evaluation

In the experiments, the imaging depth is set as 16 cm, the ultrasound image consists of 1024 × 768 pixels, and the pixel spacing is 0.25 mm × 0.25 mm. The inner model of the calibration phantom is arranged with four layers of N-wires, and the layer interval is 2 cm. The first layer has only one N-wire structure, other layers contain three N-wire structures composed of five lines (Fig. 4a), thus 10 calibration points can be obtained in each image. A total of 10 independent datasets are collected, and each set consists of 10 images with corresponding transformation ${}^{\text{IF}}_{\text{IB}}\boldsymbol{T}$. Based on these data, 10 calibration experiments are independently performed, and 10 calibration transformations are obtained.

The calibration reproducibility (CR) [34] is adopted to quantify the precision of the method. Specifically, the center and four corners of the ultrasound image are selected as the feature points to reflect the generality. The point reconstruction accuracy (PRA) [34, 35] is adopted to evaluate the accuracy. The calibration points in the latter dataset are used to verify the calibration matrix obtained based on the current dataset, and so on, the first dataset is used to test the last calibration matrix.

TABLE I
PRECISION EVALUATION RESULTS OF THE PROPOSED METHOD

Points (mm)	(128, 96)	(0, 0)	(256, 0)	(0, 192)	(256, 192)	Mean
CR (mm)	0.78	2.52	2.16	2.79	2.36	2.12

TABLE II
ACCURACY EVALUATION RESULTS OF THE PROPOSED METHOD

Experiments	1	2	3	4	5	6	7	8	9	10	Mean
PRA (mm)	1.21	1.61	1.53	1.11	1.02	1.42	1.05	1.55	1.17	1.09	1.28

The precision and accuracy evaluation results of the proposed method are summarized in Tables I and II, respectively. Experimental results show that the mean CR of 2.12 mm and the mean PRA of 1.28 mm are achieved with 10 images per calibration. Compared with the previous study, Carbajal et al. obtained a mean CR about 2.00 mm and a mean PRA around 1.40 mm with 400 images per calibration [34], while our method can realize similar precision and accuracy with fewer images. Obviously, the calibration efficiency of our method is higher, which is benefited from the optimized layout of multiple N-wires in our calibration phantom. In addition, our method obtains the closed-form solution of the rigid transformation by adopting the SVD-based method, which ensures the robustness of the calibration matrix calculation with the presence of calibration point artifacts in ultrasound images.

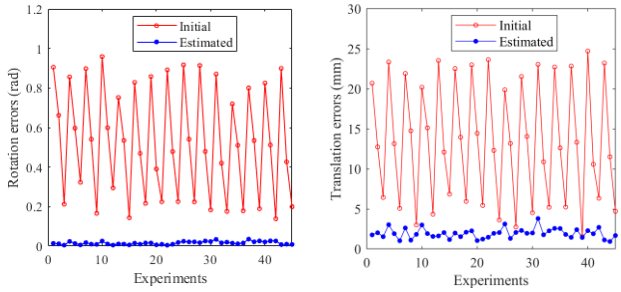
B. 2D US - 3D US Registration Evaluation

To evaluate the performance of the 2D US - 3D US registration method, independent registration experiments on a human abdomen model (057A, CIRS, USA) have been conducted. By scanning the model at different poses, 15 pairs of 2D US - 3D US data with true coordinate transformations are obtained. The region of interest (ROI) is extracted based on center crop method in physical space to speed up the registration and reduce the interference from irrelevant pixels. The final 3D US volume includes 460 × 375 × 128 voxels, and the pixel spacing is 0.34 mm × 0.34 mm × 0.66 mm. The final 2D US image includes 600 × 500 pixels, and the pixel spacing is 0.25 mm × 0.25 mm. For each pair of 2D US - 3D US data, three registration experiments are conducted with different initialization. In each experiment, transformation parameters are initialized by adding random deviations to the true values. For t_x, t_y and t_z, the deviations are in the range of 15 mm, and for r_x, r_y and r_z, the deviations are in the range of

0.6 rad. Finally, in total 45 independent registration experiments are conducted.

In the experiments, a three-level pyramid is adopted, and the sampling rates from top to bottom are 1/4, 1/2, and 1, respectively. For all the parameters t_x, t_y and t_z, the initial label space radii from top to bottom are 5 mm, 3 mm and 1 mm, for all r_x, r_y and r_z, the initial radii are 0.2 rad, 0.15 rad, 0.1 rad. $\lambda_i = 0.1, i = 1, \ldots, 6$, and $k_i = 5, i = 1, \ldots, 6$. The maximum iteration number of Fast-PD optimization algorithm is set as 100 times. In each level, the iteration is stopped when the registration performance is no longer improved within 30 consecutive iterations or the total iteration number reaches 200 times.

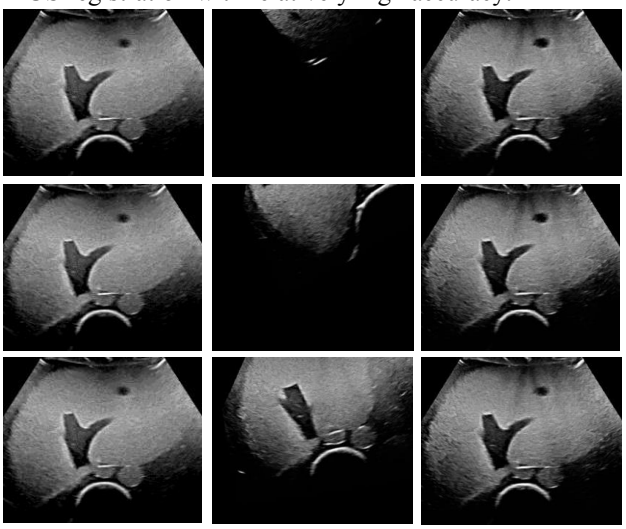
All experiments are performed on a computer with 8 CPU cores (Intel Core i7-9700K CPU @ 3.60 GHz) and 64GB RAM. To evaluate the registration results quantitatively, the Euclidean distance between the estimated and true rotation vectors (r_x, r_y, r_z) is denoted as rotation error, the Euclidean distance between the estimated and true translation vectors (t_x, t_y, t_z) is denoted as translation error. The rotation and translation errors of 45 independent registration experiments are shown in Fig. 8. With the registration method, mean rotation error is reduced from the initial 0.526 rad to 0.016 rad, and the mean translation error is reduced from the initial 13.39 mm to 1.94 mm.



(a) Rotation errors (b) Translation errors

Fig. 8. Quantitative results of 45 2D US - 3D US registration experiments.

Fig. 9 shows the registration results of the first three experiments, corresponding to the first pair of 2D US - 3D US data with the true coordinate transformation parameters (0, 0, 0, -126.85 mm, 34.39 mm, 0). In the figure, the three rows are the registration results under three different initialization conditions. The initial slice represents the slice extracted from the inputted 3D US according to the initial transformation parameters, and the estimated slice is extracted based on the obtained parameters by registration. In the first row, the initial parameters are (0.496 rad, 0.513 rad, 0.557 rad, -114.49 mm, 44.90 mm, 12.85 mm), while the obtained parameters by registration are (-0.012 rad, 0.005 rad, 0.004 rad, -127.19 mm, 33.37 mm, 1.39 mm). In this experiment, the initial rotation error is 0.905 rad, and the estimated rotation error is 0.014 rad. The initial translation error is 20.70 mm, and the estimated translation error is 1.75 mm. In the second row, after registration, the transformation parameters change from (0.395 rad, 0.369 rad, 0.382 rad, -120.67 mm, 41.81 mm, 8.31 mm) to (0.001 rad, -0.012 rad, 0.001 rad, -127.45 mm, 33.82 mm, -1.85mm). That is, the rotation error is reduced from 0.662 rad to 0.012 rad, and the translation error is reduced from 12.74 mm to 2.02 mm. In the third experiment, the initial parameters are (0.089 rad, 0.129 rad, 0.143 rad, -125.38 mm, 38.29 mm, 4.91 mm), corresponding to a rotation error of 0.212 rad and a translation error of 6.44 mm, and the estimated parameters are (0.002 rad, -0.004 rad, 0.002 rad, -127.35 mm, 33.61 mm, -1.21 mm), corresponding to a rotation error of 0.005 rad and a translation error of 1.53 mm. These results show that our method can complete the 2D US - 3D US registration with relatively high accuracy.



2D US image Initial slice Estimated slice

Fig. 9. Visualization of the first 2D US - 3D US registration experiment.

In addition, the registration time in all the experiments is about 2 minutes, this registration speed cannot meet the clinical requirements. In fact, the registration efficiency can be improved by using GPU acceleration methods and minimizing the size of the input images while preserving key anatomical information. In future, we will strive to decrease the registration time, such that the image registration and needle insertion operation are completed within one breath-holding period.

C. Robotic Needle Insertion Experiments

In order to evaluate the effectiveness of the proposed method and the performance of the developed system, needle insertion experiments are conducted on a respiratory motion simulation platform.

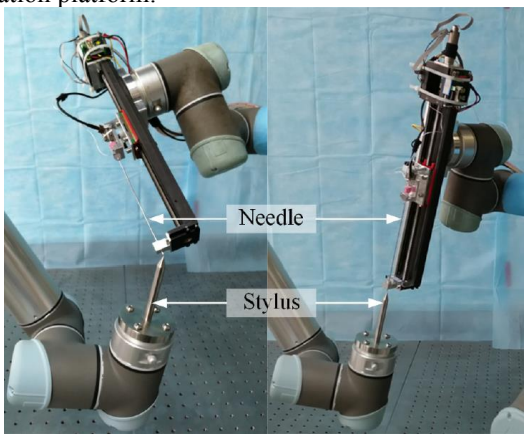


Fig. 10. The scenarios of setting the tool center point for puncture arm.

Before experiments, the system calibration mainly including dual-arm calibration and US probe calibration is

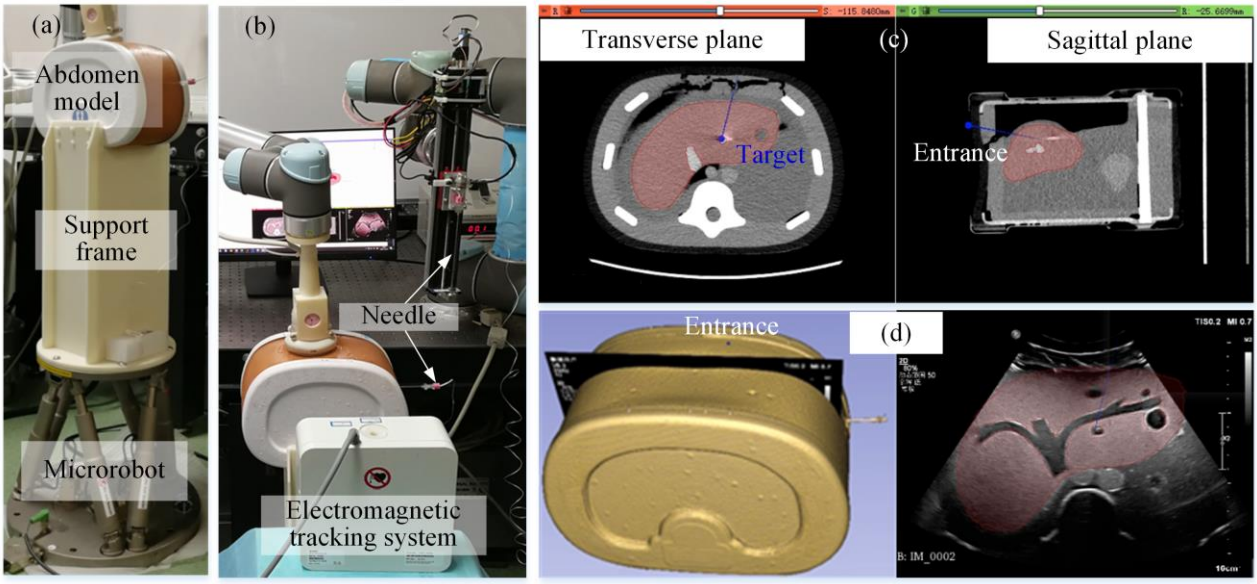


Fig. 11. (a) Respiratory motion simulation platform, (b) experimental setup for robotic puncture with 2D US – 3D CT fusion guidance, (c) needle insertion path planning, and (d) intraoperative 2D US - preoperative 3D CT fusion result.

firstly performed, the calibration flow is as follows. Firstly, a specially designed steel stylus is installed on the flange of the imaging arm, and the needle insertion mechanism is returned to zero position. Secondly, as shown in Fig. 10, the imaging arm is dragged so that its end stylus is in position, the puncture arm is dragged to make the puncture needle touch the stylus tip in four different postures, and the needle tip is set as the tool center point of the puncture arm based on an ordinary four-point tool center point calibration algorithm adopted by the UR5 robots, that is, the initial ${}^{\mathrm{PF}}_{\mathrm{NT}}\boldsymbol{T}$ is determined. In the same way, the imaging arm is dragged to make the stylus tip touch the fixed needle tip, and the stylus tip is set as the tool center point of the imaging arm. Next, the two arms are dragged to make their tool tips contact at four asymmetric points, and their positions in two robot base coordinate systems are recorded in pairs. Based on these pairwise coordinates, the transformation ${}^{\mathrm{PB}}_{\mathrm{IB}}\boldsymbol{T}$ is obtained by the SVD-based closed solution method [39], and the dual-arm calibration is completed. Finally, as described in Section III, after the ${}^{\mathrm{IB}}_{\mathrm{CP}}\boldsymbol{T}$ is obtained, the stylus is removed and ultrasound probe is installed, and the US probe calibration is conducted to compute ${}^{\mathrm{IF}}_{\mathrm{2DUS}}\boldsymbol{T}$.

The respiratory motion simulation platform consists of a hexapod microrobot (H-840.D2, Physik Instrumente, German), a human abdomen model (057A, CIRS, USA) and a nylon support frame, as shown in Fig. 11(a). A 3D translational motion in Equation (7) [51] is applied to the microrobot to mimics the displacement of the liver during breathing.

$$\boldsymbol{P}(t) = \boldsymbol{a}\cos^2\left(\frac{\pi}{T}t - \frac{\pi}{2}\right) \tag{7}$$

where t represents the time, $\boldsymbol{a} \in \mathbb{R}^3$ represents the extent of the motion. As reported in [52], the average motion amplitudes of liver in the cranial-caudal, left-right and anterior-posterior directions are 15 mm, 7.5 mm and 10 mm, respectively, thus $\boldsymbol{a} = [15,\ 7.5,\ 10]^{\mathrm{T}}$. T represents the period of the motion, which is set to 4 s.

Experimental setup is shown in Fig. 11(b). In order to measure the puncture error, two electromagnetic positioning needles are used. One is inserted into the abdomen model in advance, and its tip is regard as the target tumor, and the other is used as the puncture needle. Both needle tips are tracked by an electromagnetic tracking system (Aurora, Northern Digital, Canada). After puncture, the distance between the two needle tips is taken as the puncture error.

Before experiments, CT scan of the abdomen model is firstly conducted. Then needle insertion path is planned on the CT model, as shown in Fig. 11(c). At the reference breath-holding state t=0, the 3D US imaging of the abdomen model is conducted, and the 3D US – 3D CT registration is performed to obtain ${}^{\mathrm{3DUS}}_{\mathrm{3DCT}}\boldsymbol{T}$. The centroids of six tumors visible in both CT and 3D US volume are regarded as landmarks, and the mean Euclidean distance between the landmarks in 3D US volume and transformed corresponding landmarks in CT volume is defined as 3D US - 3D CT registration error, which is 1.32 mm. To test the performance of the proposed method, puncture experiments are performed in different breath-holding states.

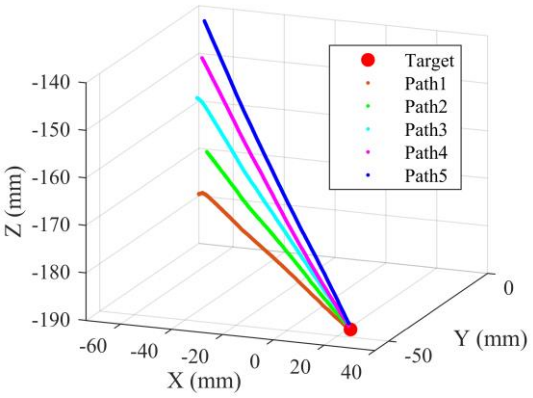


Fig. 12. The target and five different puncture paths at t = 0.5 s.

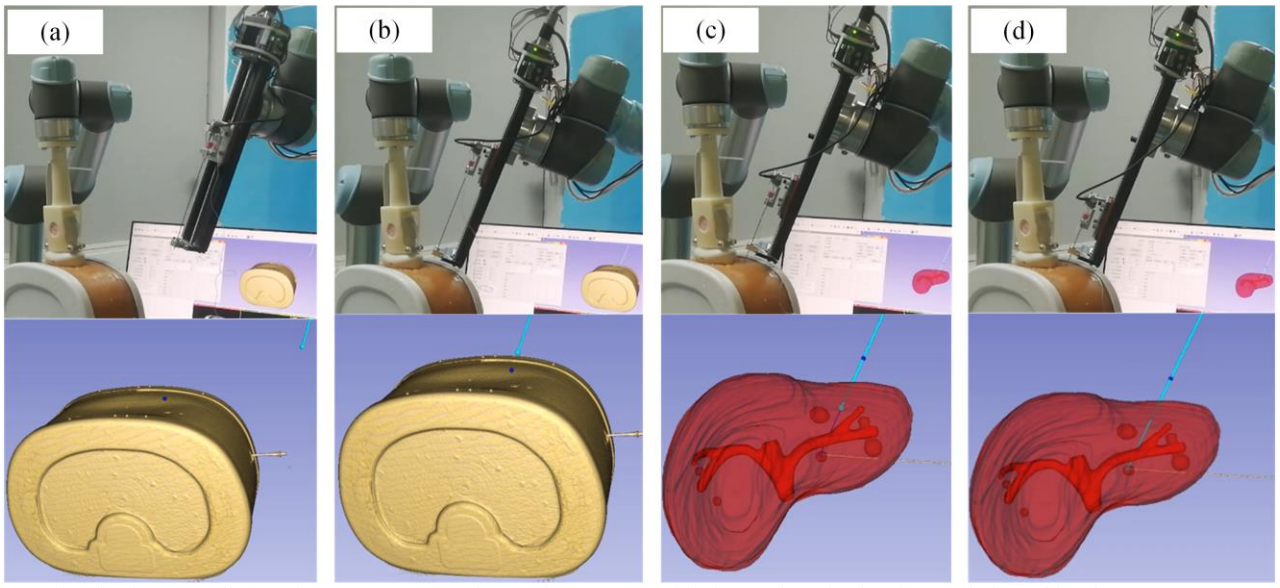


Fig. 13. Visualization of needle insertion process, (a) initial state, (b) entrance positioning, (c) inserting into liver and (d) reaching the target.

At the breath-holding state t=0.5 s, which is with deviation 0.15$\boldsymbol{a}$, the 2D US imaging is firstly conducted, and ${}^{2\mathrm{DUS}}_{3\mathrm{DUS}}\boldsymbol{T}$ is updated from ${}^{2\mathrm{DUS}}_{3\mathrm{DUS}}\boldsymbol{T}_0$ based on the 2D US - 3D US registration. The visualization of the intraoperative 2D US – preoperative 3D CT registration is shown in Fig. 11(d). Then, the planned puncture entrance point and target point in {3DCT} are mapped into the puncture robot space as in Equation (8), and the insertion depth and needle posture are calculated as in Equation (9)-(11). Finally, the needle insertion is conducted with the guidance of 2D US – 3D CT fusion. With the same procedure, needle insertion experiments are conducted in another three different breathing-holding states t = 1 s, 1.5 s and 2 s (with deviations 0.5$\boldsymbol{a}$, 0.85 $\boldsymbol{a}$ and $\boldsymbol{a}$). At each breath-holding state, five punctures are repeated along different paths, and totally 20 experiments are performed.

$$ {}^{\mathrm{PB}}\boldsymbol{P} = {}^{\mathrm{PB}}_{\mathrm{IB}}\boldsymbol{T}\,{}^{\mathrm{IB}}_{\mathrm{IF}}\boldsymbol{T}\,{}^{\mathrm{IF}}_{\mathrm{2DUS}}\boldsymbol{T}\,{}^{\mathrm{2DUS}}_{\mathrm{3DUS}}\boldsymbol{T}\,{}^{\mathrm{3DUS}}_{\mathrm{3DCT}}\boldsymbol{T}\,{}^{\mathrm{3DCT}}\boldsymbol{P} \tag{8} $$

where ${}^{\mathrm{3DCT}}\boldsymbol{P}$ represents a point in {3DCT}, and ${}^{\mathrm{PB}}\boldsymbol{P}$ represents the corresponding point in {PB}.

$$ dep = \left\| {}^{\mathrm{PB}}\boldsymbol{P}_{\mathrm{entrance}} - {}^{\mathrm{PB}}\boldsymbol{P}_{\mathrm{target}} \right\|_2 \tag{9} $$

$$ {}^{\mathrm{PB}}\hat{\boldsymbol{Y}}_{\mathrm{NT}} = ({}^{\mathrm{PB}}\boldsymbol{P}_{\mathrm{entrance}} - {}^{\mathrm{PB}}\boldsymbol{P}_{\mathrm{target}}) / dep \tag{10} $$

$$ {}^{\mathrm{PB}}\hat{\boldsymbol{X}}_{\mathrm{NT}} = ({}^{\mathrm{PB}}\boldsymbol{P}_{\mathrm{entrance}} \times {}^{\mathrm{PB}}\boldsymbol{P}_{\mathrm{target}}) / \left\| {}^{\mathrm{PB}}\boldsymbol{P}_{\mathrm{entrance}} \times {}^{\mathrm{PB}}\boldsymbol{P}_{\mathrm{target}} \right\|_2 \tag{11} $$

where ${}^{\mathrm{PB}}\boldsymbol{P}_{\mathrm{entrance}}$ and ${}^{\mathrm{PB}}\boldsymbol{P}_{\mathrm{target}}$ are the position vectors of the entrance point and target point in the coordinate system {PB}, dep is the insertion depth, ${}^{\mathrm{PB}}\hat{Y}_{\mathrm{NT}}$ is the Y-axis direction vector of the coordinate system {NT} in {PB}, and ${}^{\mathrm{PB}}\hat{X}_{\mathrm{NT}}$ is the X-axis direction vector of the coordinate system {NT} in {PB}, which means that the X-axis of {NT} is perpendicular to the plane determined by the entrance point, target point and the origin of the coordinate system {PB}.

The target and five different puncture paths at t = 0.5 s are shown in Fig. 12. Visualization of the needle insertion process in the first experiment at t = 0.5 s is shown in Fig. 13, the spatial relative positions between the target, the needle, and the obstacles such as blood vessels are visualized in real time, which enables doctors to monitor the surgical process, and ensures the safety of the robotic needle insertion. The boxplot of the puncture errors in all experiments is shown in Fig. 14, the mean puncture errors for each breath-holding state are summarized in Table III. For all experiments, the mean puncture error is 2.48 mm, and the maximum puncture error is 3.19 mm. As described in [53], the puncture accuracy can meet the requirements in a wide of clinical scenarios.

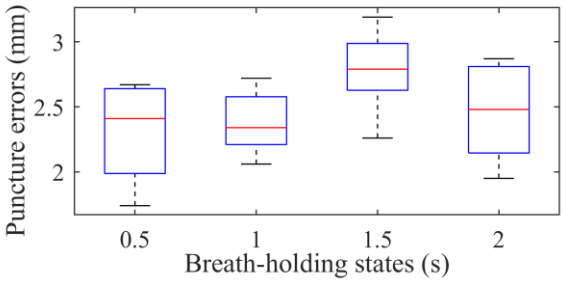


Fig. 14. Boxplot of puncture errors at four breath-holding states.

TABLE III
Mean puncture errors at four breath-holding states

Breath-holding state (s)	0.5	1	1.5	2	Mean
Mean puncture errors (mm)	2.30	2.38	2.78	2.46	2.48

Overall, the key innovation of the developed robotic system lies in the fusion guidance of intraoperative 2D ultrasound and preoperative CT. The significance of the fusion guidance can be explained from two perspectives: respiratory motion compensation and ultrasound image enhancement. From the perspective of respiratory motion compensation, based on the proposed 2D US – 3D CT registration method, the spatial correspondence and transformation relationship between preoperative CT images and intraoperative anatomical structures is established by using the intraoperative ultrasound images as a bridge, so the developed robot can avoid puncture errors caused by inconsistent breath-holding position before and during surgery existing in the literature [4, 5] and [53], in which the 3D US – 3D CT registration is used to locate the intraoperative tumor and its surrounding critical structures. Compared with the robot system XCAT ACT [6], no

step-by-step check is required, and compared with the robot in our previous work [9], multiple preoperative CT scans are not required, so the radiation exposure of patients should be less when using the robot developed in this paper. In addition, after the preparatory work is completed, our robot inserts the needle without interruption. Therefore, compared with the robots adopting respiratory gating strategy [7, 8]，the actual needle insertion time is shorter and the corresponding surgical risk is lower. From the perspective of ultrasound image enhancement, fusing clear 3D CT images with intraoperative ultrasound images can provide real-time, clear, and all-round guidance for puncture surgery, so there is no need to limit the puncture path in the ultrasound imaging plane, thereby separating the puncture path planning from the selection of the ultrasound imaging plan, this is where our robot is superior to the robots in [10, 11]. In addition, our method does not require intraoperative adjustment of the puncture trajectory by rotating the needle shaft [12], avoiding additional trauma to the tissue.

VI. CONCLUSION

In this work, in order to solve the problems of position drift of tumor and poor accuracy/stability of doctors' manual operation in traditional percutaneous liver puncture, a dual-arm robotic needle insertion system with guidance of intraoperative US and preoperative CT fusion is proposed, achieving stable, accurate and safe puncture. Specifically, a 2D US - 3D CT registration method that decouples cross-modality and cross-dimension is proposed, which can compensate for the liver motion in the case of inconsistent breath-holding positions between the preoperative CT scan phase and intraoperative puncture phase. Besides, a fast and robust ultrasound probe calibration method based on an improved N-wire phantom is proposed. To test the performance of the proposed system, puncture experiments are performed in different breath-holding states on a respiratory motion simulation platform, and experimental results show that the mean puncture error is 2.48 mm, which can meet the requirements in a wide of clinical scenarios.

The main limitation of the proposed system is that ultrasound imaging is completed by the operator dragging the robotic arm. The efficiency of finding the imaging plane depends on the operator's experience. Due to the damping of the robotic arm, it is difficult to control the appropriate contact force between the probe and the skin. Therefore, the next step is to study the robot's automatic ultrasound imaging method [54] to improve efficiency and ensure imaging quality and patient comfort.

REFERENCES

[1] L. Lei, L. Huang, B. Zhao, *et al*., "Diffeomorphic respiratory motion estimation of thoracoabdominal organs for image‐guided interventions," *Med. Phys.*, vol. 48, no. 8, pp. 4160-4176, Jun. 2021.

[2] JC. Park, SH. Park, JH. Kim, *et al*., "Liver motion during cone beam computed tomography guided stereotactic body radiation therapy," *Med. Phys.*, vol. 39, no. 10, pp. 6431-6442, Oct. 2012.

[3] F. J. Siepel, B. Maris, M. K. Welleweerd, *et al*., "Needle and biopsy robots: a review," *Curr. Robot. Rep.*, vol. 2, no. 1, pp. 73-84, Jan. 2021.

[4] Y. Koethe, S. Xu, G. Velusamy, *et al*., "Accuracy and efficacy of percutaneous biopsy and ablation using robotic assistance under computed tomography guidance: a phantom study," *Eur. Radiol.*, vol. 24, no. 3, pp. 723-730, Nov. 2014.

[5] E. Franco, D. Brujic, M. Rea, *et al*., "Needle-guiding robot for laser ablation of liver tumors under MRI guidance," *IEEE-ASME T. Mech.*, vol. 21, no. 2, pp. 931-944, Sep. 2015.

[6] E. Ben-David, M. Shochat, I. Roth, *et al*., "Evaluation of a CT-guided robotic system for precise percutaneous needle insertion," *Journal of Vascular and Interventional Radiology*, vol. 29, no. 10, pp. 1440-1446, 2018.

[7] M. J. Musa, K. Sharma, K. Cleary, and Y. Chen, "Respiratory compensated robot for liver cancer treatment: Design, fabrication, and benchtop characterization," *IEEE/ASME Trans. Mechatronics*, vol. 27, no. 1, pp. 268–279, Feb. 2022.

[8] L. Zheng, H. Wu, L. Yang, Y. Lao, Q. Lin, and R. Yang, "A Novel Respiratory Follow-Up Robotic System for Thoracic-Abdominal Puncture," *IEEE Transactions on Industrial Electronics*, vol. 68, no. 3, pp. 2368-2378, Mar. 2021.

[9] L. Lei, H. Tang, J. Zhang, Y Wu, B Zhao, and B. Li, "Automatic registration and precise tumour localization method for robot‐assisted puncture procedure under inconsistent breath‐holding conditions," *Int. J. Med. Robot.*, vol. 17, no. 6, pp. e2319, Aug. 2021.

[10] Y. Gao, X. Liu, X. Zhang, Z. Zhou, W. Jiang, L. Chen, Z. Liu, D. Wu, and W. Dong, "A dual-armed robotic puncture system: Design, implementation and preliminary tests," *Electronics*, vol. 11, no. 5, pp. 740, 2022.

[11] S. Chen, F. Wang, Y. Lin, Q. Shi, and Y. Wang, "Ultrasound-guided needle insertion robotic system for percutaneous puncture," *Int. J. Comput. Assisted Radiol. Surg.*, vol. 16, no. 3, pp. 475-484, Jan. 2021.

[12] Y. Zhao, C. Wen, Y. Zhang, and H. Zhang, "Needle tip pose estimation for ultrasound-guided steerable flexible needle with a complicated trajectory in soft tissue," *IEEE Robotics and Automation Letters*, vol. 7, no. 4, pp. 11705-11712, Oct. 2022.

[13] J. Tan, J. Li, Y. Li, B. Li, Y. Leng, Y. Rong, and C. Fu, "Autonomous Trajectory Planning for Ultrasound-Guided Real-Time Tracking of Suspicious Breast Tumor Targets," *IEEE Transactions on Automation Science and Engineering*, 2023.

[14] Z. Yang, L. Yang, M. Zhang, C. Zhang, S. C. H. Yu, and L. Zhang, "Ultrasound-guided catheterization using a driller-tipped guidewire with combined magnetic navigation and drilling motion," *IEEE/ASME Transactions on Mechatronics*, vol. 27, no. 5, pp. 2829-2840, Oct. 2022.

[15] W. Yan, Q. Ding, J. Chen, K. Yan, R. S. Y. Tang, and S. S. Cheng, "Learning-based needle tip tracking in 2D ultrasound by fusing visual tracking and motion prediction," *Medical Image Analysis*, vol. 88, pp. 102847, 2023.

[16] R. Liao, L. Zhang, Y. Sun, *et al*., "A review of recent advances in registration techniques applied to minimally invasive therapy," *IEEE T. Multimedia*, vol. 15, no. 5, pp. 983-1000, Feb. 2013.

[17] E. Ferrante, N. Paragios, "Slice-to-volume medical image registration: A survey," *Med. Image Anal.*, vol. 39, pp. 101-123, Jul. 2017.

[18] E. M. Jung, C. Friedrich, P. Hoffstetter, *et al*., "Volume navigation with contrast enhanced ultrasound and image fusion for percutaneous interventions: first results," *PloS One*, vol. 7, no. 3, pp. e33956, Mar. 2012.

[19] C. J. Burke, J. Bencardino, R. Adler, "The potential use of ultrasound‐magnetic resonance imaging fusion applications in musculoskeletal intervention," *J. Ultras. Med.*, vol. 36, no. 1, pp. 217-224, Dec. 2017.

[20] A. Y. Kim, M. W. Lee, D. I. Cha, *et al*., "Automatic registration between real-time ultrasonography and pre-procedural magnetic resonance images: a prospective comparison between two registration methods by liver surface and vessel and by liver surface only," *Ultrasound Med. Biol.*, vol. 42, no. 7, pp. 1627-1636, Jul. 2016.

[21] M. P. Heinrich, M. Jenkinson, M. Bhushan, *et al*., "MIND: Modality independent neighbourhood descriptor for multi-modal deformable registration," *Med. Image Anal.*, vol. 16, no. 7, pp. 1423-1435, Oct. 2012.

[22] H. Rivaz, Z. Karimaghaloo, V. S. Fonov, *et al*., "Nonrigid registration of ultrasound and MRI using contextual conditioned mutual information," *IEEE T. Med. Imaging*, vol. 33, no. 3, pp. 708-725, Dec. 2013.

[23] W. Wein, S. Brunke, A. Khamene, *et al*., "Automatic CT-ultrasound registration for diagnostic imaging and image-guided intervention," *Med. Image Anal.*, vol. 12, no. 5, pp. 577-585, Oct. 2008.

[24] B. Fuerst, W. Wein, M. Müller, *et al*., "Automatic ultrasound–MRI registration for neurosurgery using the 2D and 3D LC2 Metric," *Med. Image Anal.*, vol. 18, no. 8, pp. 1312-1319, Dec. 2014.

[25] A. Hering, L. Hansen, T. C. Mok, A. C. Chung, H. Siebert, S. Häger, A. Lange, *et al*., "Learn2Reg: comprehensive multi-task medical image registration challenge, dataset and evaluation in the era of deep learning," *IEEE Transactions on Medical Imaging*, vol. 42, no. 3, pp. 697-712, 2022.

[26] V. Markova, M. Ronchetti, W. Wein, O. Zettinig, and R. Prevost, "Global multi-modal 2D/3D registration via local descriptors learning," in *International Conference on Medical Image Computing and Computer-Assisted Intervention*, Springer, pp. 269-279, Sep. 2022.

[27] H. Guo, X. Xu, X. Song, S. Xu, H. Chao, *et al.*, "Ultrasound Frame-to-Volume Registration via Deep Learning for Interventional Guidance," *IEEE Transactions on Ultrasonics, Ferroelectrics, and Frequency Control*, 2022.

[28] G. Haskins, J. Kruecker, U. Kruger, *et al*., "Learning deep similarity metric for 3D MR–TRUS image registration," *Int. J. Comput. Assisted Radiol. Surg.*, vol. 14, no. 3, pp. 417-425, Oct. 2019.

[29] Y. Song, J. Zheng, L. Lei, Z. Ni, B. Zhao, and Y. Hu, "CT2US: Cross-modal transfer learning for kidney segmentation in ultrasound images with synthesized data," *Ultrasonics*, vol. 122, pp. 106706, May 2022.

[30] J. Xiong, C. Xu, K. Ibrahim, *et al*., "A mechanism-image fusion approach to calibration of an ultrasound-guided dual-arm robotic brachytherapy system," *IEEE-ASME T. Mech.*, vol. 26, no. 6, pp. 3211-3220, Feb. 2021.

[31] L. Mercier, T. Langø, F. Lindseth, *et al*. "A review of calibration techniques for freehand 3-D ultrasound systems," *Ultrasound Med. Biol.*, vol. 31, no. 2, pp. 449-471, Feb. 2005.

[32] R. M. Comeau, A. Fenster, T. M. Peters, "Integrated MR and ultrasound imaging for improved image guidance in neurosurgery," in *Proc. Medical Imaging 1998: Image Processing*, SPIE, 1998, vol. 3338, pp. 747-754.

[33] T. K. Chen, A. D. Thurston, R. E. Ellis, *et al*., "A real-time freehand ultrasound calibration system with automatic accuracy feedback and control," *Ultrasound Med. Biol.*, vol. 35, no. 1, pp. 79-93, Jan. 2009.

[34] G. Carbajal, A. Lasso, Á. Gómez, *et al*., "Improving N-wire phantom-based freehand ultrasound calibration,". *Int. J. Comput. Assisted Radiol. Surg.*, vol. 8, no. 6, pp. 1063-1072, Jul. 2013.

[35] C. Shen, L. Lyu, G. Wang, *et al*., "A method for ultrasound probe calibration based on arbitrary wire phantom," *Cogent Eng.*, vol. 6, no. 1, pp. 1592739, Mar. 2019.

[36] J. Tokuda, G. S. Fischer, X. Papademetris, *et al*., "OpenIGTLink: an open network protocol for image‐guided therapy environment," *Int. J. Med. Robot. Comp.*, vol. 5, no. 4, pp. 423-434, Jul. 2009.

[37] M. Eichelberg, J. Riesmeier, T. Wilkens, *et al*., "Ten years of medical imaging standardization and prototypical implementation: the DICOM standard and the OFFIS DICOM toolkit (DCMTK)," in *Proc. Medical Imaging 2004: PACS and Imaging Informatics,* SPIE, 2004, vol. 5371, pp. 57-68.

[38] A. Lasso, T. Heffter, A. Rankin, *et al*., "PLUS: open-source toolkit for ultrasound-guided intervention systems," *IEEE T. Biomed. Eng.*, vol. 61, no. 10, pp. 2527-2537, May 2014.

[39] A. Fedorov, R. Beichel, J. Kalpathy-Cramer, *et al*., "3D Slicer as an image computing platform for the quantitative imaging network," *Magn. Reson. Imaging*, vol. 30, no. 9, pp. 1323-1341, Nov. 2012.

[40] T. Ungi, A. Lasso, G. Fichtinger, "Open-source platforms for navigated image-guided interventions," *Med. Image Anal.*, vol. 33, pp. 181-186, Oct. 2016.

[41] O. Sorkine-Hornung, M. Rabinovich, "Least-squares rigid motion using svd," *Computing*, vol. 1, no. 1, pp. 1-5, Jan. 2017.

[42] Z. Jiang, Y. Zhou, Y. Bi, M. Zhou, T. Wendler, and N. Navab, "Deformation-Aware Robotic 3D Ultrasound," *IEEE Robotics and Automation Letters*, vol. 6, no. 4, pp. 7675-7682, Oct. 2021.

[43] Z. Jiang, Y. Zhou, D. Cao, and N. Navab, "DefCor-Net: Physics-aware ultrasound deformation correction," *Medical Image Analysis*, vol. 90, pp. 102923, 2023.

[44] Z. Jiang, F. Duelmer, and N. Navab, "DopUS-Net: Quality-Aware Robotic Ultrasound Imaging Based on Doppler Signal," *IEEE Transactions on Automation Science and Engineering*, 2023.

[45] O. Ronneberger, P. Fischer, T. Brox, "U-net: Convolutional networks for biomedical image segmentation," in *International Conference on Medical Image Computing and Computer-Assisted Intervention,* Springer, 2015, pp. 234-241.

[46] M. Antonelli, A. Reinke, S. Bakas, *et al*., "The medical segmentation decathlon," *Nature communications*, vol. 13, no. 1, pp. 4128, 2022.

[47] A. Myronenko, X. Song, "Point set registration: Coherent point drift," *IEEE T. Pattern Anal. Mach. Intell.*, vol. 32, no. 12, pp. 2262-2275, Mar. 2010.

[48] D. Zikic, B. Glocker, O. Kutter, *et al*., "Linear intensity-based image registration by Markov random fields and discrete optimization," *Med. Image Anal.*, vol. 14, no. 4, pp. 550-562, Aug. 2010.

[49] R. Porchetto, F. Stramana, N. Paragios, *et al*., "Rigid slice-to-volume medical image registration through Markov random fields," in *Medical Computer Vision and Bayesian and Graphical Models for Biomedical Imaging,* Cham, Springer, 2016, pp. 172-185.

[50] N. Komodakis, G. Tziritas, N. Paragios, "Fast, approximately optimal solutions for single and dynamic MRFs," in *Proc. IEEE Conference on Computer Vision and Pattern Recognition*, 2007, pp. 1-8.

[51] J. Chevrie, N. Shahriari, M. Babel, *et al*., "Flexible needle steering in moving biological tissue with motion compensation using ultrasound and force feedback," *IEEE Robot. Autom. Let.*, vol.3, no.3, pp. 2338-2345, Feb. 2018.

[52] A. Kirilova, G. Lockwood, P. Choi, *et al*., "Three-dimensional motion of liver tumors using cine-magnetic resonance imaging," *Int. J. Radiat. Oncol. Biol. Phys.*, vol. 71, no. 4, pp. 1189-1195, Jul. 2008.

[53] R. Kojcev, B. Fuerst, O. Zettinig, J. Fotouhi, S. C. Lee, B. Frisch, R. Taylor, E. Sinibaldi, and N. Navab, "Dual-robot ultrasound-guided needle placement: closing the planning-imaging-action loop," *Int. J. Comput. Assisted Radio. Surg*., vol. 11, no. 6, pp. 1173-1181, Apr. 2016.

[54] Z. Jiang, S.E. Salcudean, and N. Navab, "Robotic ultrasound imaging: State-of-the-art and future perspectives," *Medical Image Analysis*, pp. 102878, July 2023.

Long Lei received the Ph.D. degree from Harbin Institute of Technology, Shenzhen, China, in 2022. He is currently a postdoctoral fellow in Department of Computer Science and Engineering, The Chinese University of Hong Kong. He is also an assistant research fellow in Shenzhen Institute of Advanced Technology, Chinese Academy of Science, Shenzhen, China. His research interests include surgical robot, medical image analysis, and biomechanics.

Baoliang Zhao received the Ph.D. degree from University of Nebraska-Lincoln, Lincoln, NE, USA, in 2015. He is currently a senior engineer in Shenzhen Institute of Advanced Technology, Chinese Academy of Sciences, Shenzhen, China. His research interests include medical robots, image navigation and intelligent control technology.

Xiaozhi Qi received the Ph.D. degree from Harbin Institute of Technology, Shenzhen, China, in 2017. He is currently an associate professor in Shenzhen Institute of Advanced Technology, Chinese Academy of Science, Shenzhen, China. His current research interests include planning, navigation and control of the medical assistant robots.

Rui Mi received the M.S. degree from Shenzhen University Clinical Medical Academy, Shenzhen, China, in 2017. He is currently a radiologist in Department of Radiology of Shenzhen University General Hospital, Shenzhen, China. His research interests include abdominal imaging, functional MRI, and radiomics.

Hai Ye received the M.S. degree from Anhui Medical University, Hefei, China, in 2019. He is currently a radiologist in Department of Radiology of Shenzhen University General Hospital, Shenzhen, China. His research interests include functional MRI and the application of AI in image diagnosis.

Peng Zhang received the Ph.D. degree from Harbin Institute of Technology, Harbin, China, in 2010. He is currently a senior engineer in Shenzhen Institute of Advanced Technology, Chinese Academy of Sciences, China. His current research interests include medical/surgical robots, medical automation equipment and virtual simulation technology.

Qiong Wang received the Ph.D. degree from Chinese University of Hong Kong, Hong Kong, China, in 2012. She is currently a professor in Shenzhen Institute of Advanced Technology, Chinese Academy of Sciences, Shenzhen, China. Her research interests include surgical robots, visualization, medical imaging, human–computer interaction, and computer graphics.

Pheng-Ann Heng received the Ph.D. degree in computer science from Indiana University in 1992. He is currently a professor in Department of Computer Science and Engineering, The Chinese University of Hong Kong. His research interests include AI and VR for medical applications, surgical simulation, visualization, graphics and human-computer interaction.

Ying Hu received the Ph.D. degree in mechanical engineering from Harbin Institute of Technology, Shenzhen, China, in 2007. She is currently a Professor in Shenzhen Institute of Advanced Technology, Chinese Academy of Sciences, China. Her research interests include parallel robots, medical assistant robots, and mobile robots.