Automatic 3-D Imaging and Measurement of Human Spines With a Robotic Ultrasound System

Cui Yang, Mingyao Jiang, Mianjie Chen, Maoqing Fu, Jianyi Li, and Qinghua Huang

Abstract—Comparing with X-ray, ultrasound imaging has the advantages of no radiation and easy operation, etc., which is expected to be an effective alternative in clinical assessment for the human spine. At present, freehand scanning is mainly used in ultrasound imaging field, and recently robot-assisted scanning has emerged as a promising way to gradually reduce human intervention. However, the automatic recognition and scanning, as well as the control strategy of robots that guarantee the tight coupling between a probe and human skin, are still challenges. Moreover, the current measurement method of the spine is usually conducted based on plain radiographs, which lacks spatial structure information of spine. In this article, a new radiation-free robotic ultrasound system for 3-D imaging and a measurement method of the spine based on 3-D imaging are presented. A fully convolutional network (FCN) named fuse-Unet based on RGB and depth images captured by an RGB-depth (RGB-D) sensor Kinect, is proposed to realize the automatic recognition of human spine area and scanning path pre-planning. Then a six-degree-of-freedom robotic arm plays the role of a doctor to complete the automatic scanning along the pre-planning path, during which a normal-vector-based method and two force sensors are used to ensure that the probe well fits the spine area. Finally, 3-D ultrasound reconstruction and visualization of the spine are realized, based on which cobb angles are calculated for assessing the morphological structure of the spine. Phantom and *in vivo* experiments are conducted for evaluating the performance of the proposed system. Experimental results validate the practicability and accuracy of the system, indicating its potential for clinical application.

Index Terms—3-D ultrasound imaging and measurement, automatic recognition and scanning, deep learning, RGB-depth (RGB-D) feature fusion, spine.

I. INTRODUCTION

SPINE is an important structure composed of different tissues with a wide range of clinical pathologies. Owing to the complexity of spine-related disease assessment, spine imaging is an essential tool for assessing spinal pathologies. With the help of spine imaging, doctors can observe and measure the structure of the spine directly, thereby achieving the diagnosis of spine deformities and injuries. Nowadays multiple ways of medical imaging, including magnetic resonance imaging (MRI), X-ray, computed tomography (CT), etc., have been widely used in the spine imaging field.

At present, X-ray is the most common spinal imaging technique, which can help doctors evaluate spinal deformities, e.g., scoliosis. However, the side effects of X-ray exposure can bring certain harm to the patients, particularly for the frequent and long-term imaging examinations of adolescents [1]. As an alternative to X-ray, MRI uses strong magnetic fields and radio frequency pulses to produce diagnostic images without ionizing radiation involved. MRI provides excellent soft-tissue contrast resolution, making it easy to distinguish the anatomical structures like nerve, intervertebral disk, and tendons [2]. The main drawback of MRI is that it is time-consuming. In addition, MRI is expensive and not compatible with certain pacemakers or aneurysm clips. Unlike MRI, X-ray, or CT, ultrasound imaging has the advantages of radiation-free, easy operation, low cost, and real-time visualization. In recent years, 3-D ultrasound imaging draws wide attention in spine-related disease assessment [3]–[6] due to the good spatial structure display of the entire spine offered by it, which suggests that ultrasound imaging could be a precise and safe alternative to X-ray for spine diagnose. So far 3-D ultrasound imaging is mainly obtained using two scanning approaches, freehand scanning [7], [8] and robot-assisted scanning [9]–[11].

Freehand scanning is a free-moving scan by the doctor's hand-held probe, the system records the position and orientation of the ultrasound probe through the tracking device, then constructs a 3-D ultrasound volume. Currently most tracking devices rely on optical tracking cameras [12]–[14] or electromagnetic (EM) sensors [5], [6], [15], alternative tracking methods use mechanical positioning such as linear

tracking systems [16], [17]. Cenni *et al.* [13] designed two optical tracking systems to verify the reliability and validity of a clinical 3-D freehand ultrasound. Different optical cameras were adopted in the two systems to track ultrasound images and achieve 3-D reconstruction. The authors conducted phantom experiments and carried out the reliability and validity analyses of extracting volumes and lengths. The results demonstrated reliable measurements and suitable performance for applications that do not require sub-mm and -ml accuracy. Zhou *et al.* [4] used freehand 3-D ultrasound imaging with EM sensors to image the scoliotic spine. The spinous process was detected according to the bony regions extracted from 3-D ultrasound volume projection image (VPI) and it is further used to extract the curve representing the spine. Based on extracted curve, the severity of scoliosis can be measured.

The freehand scanning can achieve a large scan range, and the scan position and orientation are arbitrarily controlled by the doctor, making it suitable for various application scenes. In addition, optical and EM sensors, which are commonly used in freehand scanning, have been proven to be precise and reliable; their main drawbacks are a line of sight and metal sensitivity, respectively. To obtain the technical and clinical advantages offered by ultrasound imaging in freehand scanning, highly qualified personnel trained to properly operate probes are required. Freehand scanning could also be a physical challenge for the doctors, as they need to hold probes in often awkward positions for prolonged periods of time, resulting in a high prevalence of musculoskeletal injuries and disorders [18]. With obese patients, doctors need to apply even higher forces on the probe to compensate for the signal loss, consequently increasing their risk of muscle injury. Despite all this, freehand scanning still attracts lots of attention from scholars.

To gradually reduce human intervention, robot-assisted scanning has emerged as a promising scan technique. Previous work has proven the superiority of robot-assisted scanning in 3-D ultrasound imaging [19]. The key advantages of robot-assisted scanning are stability, dexterity, repeatability, and high accuracy. Furthermore, robot-assisted scanning can avoid direct contact between doctors and patients, and hence prevent the spread of infectious diseases. Benefiting from the major advances mentioned above, robot-assisted scanning has drawn wide attention in recent years. Conti *et al.* [20] developed a collaborative robot-assisted scanning system. An operator was required to control the slave robot in the haptic mode or semiautomatic mode. A force controller was adopted to maintain the constant contact between the ultrasound probe and tissue surface. The system recorded the 2-D B-scans for real-time 3-D construction. Huang *et al.* [9] used a 3-D translating device to hold the ultrasound probe and scan the tissues in a fully automatic manner. To achieve automatic scanning, a pre-planning scan path was obtained based on the 3-D contour of the tissue captured by the RGB-depth (RGB-D) sensor Kinect. When scanning, the system could record the 2-D B-scans and their positional information for subsequent volume reconstruction. The same group thereafter developed an automatic scanning system equipped with a six-degree-of-freedom robotic arm [10]. In this system, a normal-vector-based method was adopted to make the probe

fit better to the tissue and thus obtain the high-quality 3-D volume.

In summary, a majority of existing robot ultrasound systems is devoid of intelligence. They either require user participation or lack a good method for path pre-planning. For example, the automatic scanning systems established in [9] and [10] can pre-plan the scan path correctly based on the tissue contour of regular shapes. However, the performance of these systems declines when working with certain human body parts, especially the spine area with various and irregular shapes that brings a challenge in automatic recognition. Furthermore, depth change of the spine area, together with the breathing of the patient, can directly cause inadequate contact between the probe and tissue surface, which is also a critical concern in a robot ultrasound system. In addition, scoliosis is one of the most common and harmful spine-related diseases, whose standard clinical measurement is calculating the Cobb angle based on plain radiographs [21]. Nevertheless, scoliosis is a 3-D pathology, which cannot be completely assessed by 2-D radiography [22]. Since the robot-assisted system has been proven to demonstrate high accuracy in spine level detection [23], [24], the measurement of scoliosis based on 3-D ultrasound volume should be taken into account to assist medical diagnosis.

In this article, we propose a robotic ultrasound system for 3-D imaging and measurement of the spine. An RGB-D sensor Kinect was used to capture the RGB image, depth image, and 3-D point cloud of the back of the patient. Based on RGB and depth image, a fully convolutional network (FCN) called fuse-Unet was proposed to achieve the automatic recognition and segmentation of the spine area. The fuse-Unet is an extended version of U-net that allows for a more efficient fusion of RGB features and depth features and hence more precise segmentation results. According to the segmentation result, a pre-planning scan path was extracted from the point cloud. Then the robotic arm carried the ultrasound probe to scan along the pre-planning path in a fully automatic manner. During the scanning, we adopted a normal-vector-based method and two force sensors to keep the probe in close contact with the spine area. At each scan point, the system recorded 2-D B-scan and corresponding positional information. After finishing the scan, the system realized 3-D reconstruction and visualization of the spine. And a measurement method for calculating the Cobb angle based on 3-D ultrasound volume is also demonstrated. We aim to bring the solution for enhancing robotic scanning for a case of irregular shapes, e.g., spine, and this is the first attempt to combine automatic human spine scanning with 3-D ultrasound imaging and measurement.

The main contributions of this work are: 1) intelligent recognition of the spine area and pre-planning of the scan path for the robot system without user-participation; 2) considering the complexity of scanning the spine area, two robot control strategies are adopted to guarantee the effective coupling of ultrasonic signal; and 3) comparing with traditional 2-D radiography, 3-D ultrasound visualization of the spine and a measurement method based on 3-D volume for scoliosis are offered to realize the spatial structure display and non-radiation examination.
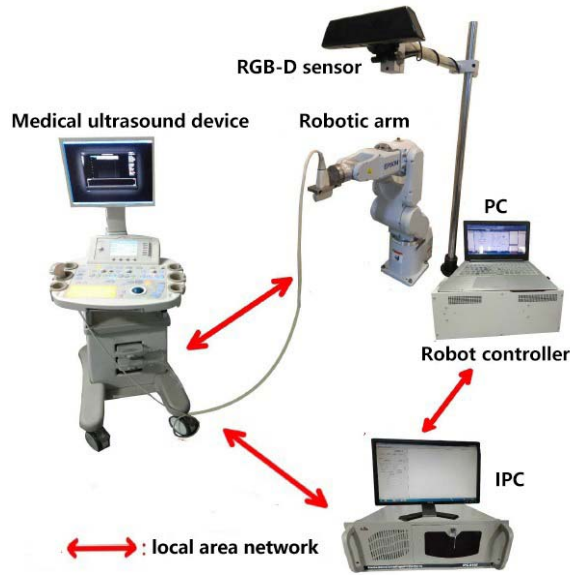
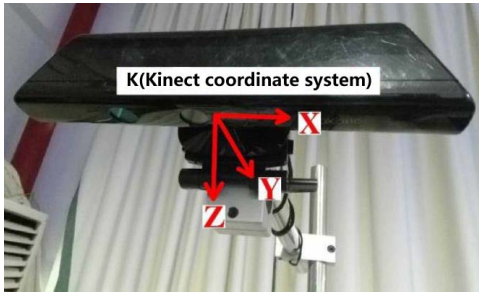Fig. 1.   Illustration of the proposed system.

Fig. 2.   Kinect coordinate system.

The article is organized as follows: Section II is a detailed description of about our proposed system and experimental method. Section III gives the phantom and *in vivo* experiment results. Discussions and conclusions are presented in Section IV.

II. METHODOLOGY

A. System Design and Calibrations

The proposed system, illustrated in Fig. 1, consists of a medical ultrasound device [Ultrasonix RP device (Ultrasonix Medical Corporation, Canada) equipped with a linear probe (L9-4/38) and a convex probe (C7-3/50)] for acquiring 2-D B-scans, a six-degree-of-freedom robotic arm with repetitive positioning accuracy of 0.02-mm (EPSON C4-A601S, Seiko Epson Corporation, Japan) for actuating the probe, an RGB-D sensor (Kinect v1, Microsoft Corporation, USA) for capturing the RGB image, depth image and 3-D point cloud of the back of the patient, an industrial computer (IPC, EVOC Intelligent Technology Company Ltd., China) for collecting all the B-scans and their positional information for the subsequent 3-D reconstruction, and a personal computer (PC) for completing the segmentation of spine area and controlling the robotic arm. The PC, IPC, and medical ultrasound device are connected through a local area network. The data transmission is based on TCP/IP protocol and the transmission rate is 100 Mb/s.

As seen in Fig. 1, the Kinect is fixed to the end of a steel bracket, which is 30-cm higher than the robotic arm. Two coordinate systems associating with the Kinect are image
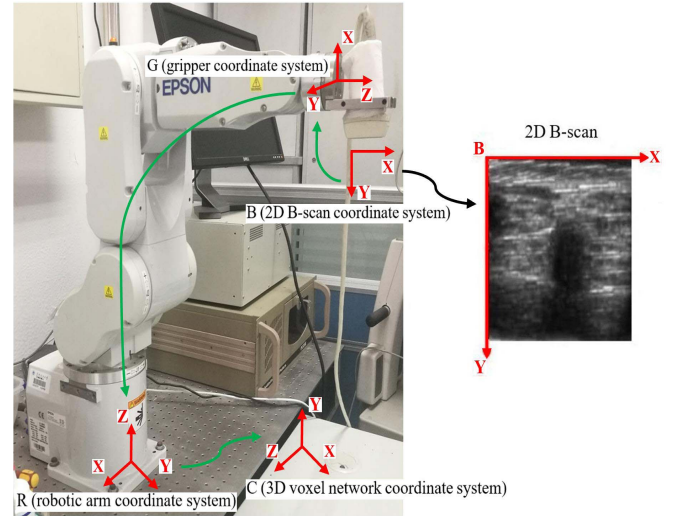

Fig. 3.   Conversion between the coordinate systems.

coordinate system I and Kinect coordinate system K (shown in Fig. 2). Both RGB image and depth image captured by Kinect have a resolution of 640×480. Taking the center point (320, 240) of the images as the origin, the transverse axis as the x-axis, and the vertical axis as the y-axis, the 2-D coordinate system I is established. In order to obtain the 3-D point cloud in the K, pixels in the I are transformed in the K using the following equations:

$$
\begin{aligned}
x_k &= (x_d - 320) \times z_d \div 525 \\
y_k &= (y_d - 240) \times z_d \div 525 \\
z_k &= z_d
\end{aligned}
\tag{1}
$$

where 525 mm is the focal length of the infrared camera that collects the depth image, (x_d, y_d) is the coordinate of the dth pixel in I, z_d is the corresponding depth value in depth image at (x_d, y_d). (x_k, y_k, z_k) is the 3-D coordinate of the pixel in K.

Fig. 3 gives four coordinate systems associated with the robotic arm, including a 2-D B-scans coordinate system B and three 3-D coordinate systems, i.e., gripper coordinate system G, robotic arm coordinate system R and 3-D voxel network coordinate system C. The ultrasound probe is fixed to a gripper that is attached to the end of the robotic arm. Under the probe, there is a B-scan coordinate system B. To achieve the 3-D ultrasound imaging, all the pixels in the 2-D B-scans should be projected onto a predefined 3-D voxel network, i.e., each pixel in $B \rightarrow G \rightarrow R \rightarrow C$, shown in Fig. 3. In our system, the real-time pose of the robotic arm could be read in IPC, and it is written as

$$
{}^{G}P_R = (X, Y, Z, U, V, W)
\tag{2}
$$

where ${}^{G}P_R$ represents the position of the gripper in R, which is composed of three translations (X, Y, Z) and three rotations (U, V, W). The order of rotations is as follows: rotate first about x-axis in R by angle W, then about y-axis in R by angle V, and finally about z-axis in R by angle U.

System calibration is a necessary step to ensure that all the independent devices work together. The robotic arm would temporarily pause at each scan point and collect a frame of B-scan steadily, and move to the next point only after the successful acquisition of the data. Hence temporal calibration

is no longer needed and two kinds of spatial calibration for the designed system are still required. The first spatial calibration is to find out the coordinate transformation from $\boldsymbol{B}$ to $\boldsymbol{C}$. Let $^{N}T_{M}$ denote the transformation matrix from $\boldsymbol{M}$ to $\boldsymbol{N}$. Thus, the whole conversion from $\boldsymbol{B}$ to $\boldsymbol{C}$ can be written as

$$P_C = {}^{C}T_R\ {}^{R}T_G\ {}^{G}T_B\ P_B \qquad (3)$$

where P_B is the pixel in $\boldsymbol{B}$, and P_C is the corresponding pixel in $\boldsymbol{C}$. As mentioned before, the $^{R}T_{G}$ can be obtained by directly reading the pose of the robotic arm. For simplicity, we defined the $^{C}T_{R}$ as $\boldsymbol{C}$ is the translation transformation of $\boldsymbol{R}$. Therefore, the first spatial calibration is aimed to estimate the $^{G}T_{B}$. A cross-wired phantom is used to conduct this spatial calibration according to [25].

The second spatial calibration is to solve the synergy problem of the Kinect and robotic arm. To be more specific, for each useful point in the point cloud collected by Kinect, the robotic arm has to know its exact position for scanning purpose, i.e., the coordinate transformation from $\boldsymbol{K}$ to $\boldsymbol{R}$ is needed in our system. According to [26], we estimate the $^{R}T_{K}$ by obtaining the coordinates of the same points in $\boldsymbol{K}$ and $\boldsymbol{R}$, respectively. A chessboard with 6×10 squares, each of which is of 1.75 cm $\times$ 1.75 cm, is used in this spatial calibration. Using Kinect, the 3-D coordinate values of 45 intersections on the chessboard in $\boldsymbol{K}$ are manually recorded. To obtain the coordinate values of the same points in $\boldsymbol{R}$, a steel needle is fixed to the gripper. Then we control the robotic arm to make the tip of the needle coincide with the 45 intersections one by one, and record the corresponding pose of the robotic arm at each intersection. After that, the coordinates of the tip at all intersections in $\boldsymbol{G}$ are converted to the 45 coordinate values in $\boldsymbol{R}$ according to

$$^{R}P_{T_i} = {}^{R}T_{G_i}\ {}^{G}P_T \qquad (4)$$

where $^{G}P_{T}$ represents the position of needle tip in $\boldsymbol{G}$, $^{R}P_{T_i}$ represents the position of the tip at ith intersection in $\boldsymbol{R}$ (i.e., the coordinate value of ith intersection in $\boldsymbol{R}$), and $^{R}T_{G_i}$ is read from the pose of robotic arm at ith intersection. With the coordinate values of 45 points in $\boldsymbol{K}$ and $\boldsymbol{R}$ obtained, we finally calculate the $^{R}T_{K}$ using Levenberg-Marquardt algorithm [27].

B. Automatic Segmentation of Human Spine Area Based on Deep Learning

For our system, the automatic recognition and segmentation of the spine area is the first step to realize automatic scanning. With the help of Kinect, the RGB and depth images of the back of the patient can be captured in real-time. As we have noticed, the spine area shows shadow texture feature in RGB image and differs from other areas of the back of the patient in-depth. Therefore, it is reasonable to believe a better segmentation result can be obtained if we fuse the RGB and depth feature.

However, using depth information for segmentation is considered a challenging task, because the unpredictable variation of scene illumination, as well as incomplete representation of objects due to complex occlusions, can bring noise into the depth image [28]. In order to successfully make use of depth information to increase the accuracy of segmentation, researchers have carried on a lot of explorations in RGB-D

fusion [29]–[32]. Recently, the state-of-the-art fusion method, which can be divided into two categories: low-level fusion and high-level fusion, usually employ two-stream FCNs by taking the RGB image and depth image as two input streams and fusing the two streams on a specific layer. Low-level features activate strongly on spatial information and high-level features contain more semantic information. Consequently, fusion on shallow layers, i.e., low-level fusion, suffers from the modality gap, and fusion on deep layers, i.e., high-level fusion, weakens the complementary spatial cues. In this article, we attempt to fuse the two features in multi-level to yield high-quality segmentation results.

Furthermore, a common consensus has been formed that for some specific applications, especially in the medical field, it is quite hard to collect a large amount of data. For those scenarios of small-scale data, simple networks should be chosen to avoid over-fitting. To cope with this challenge, a simple but effective network called Unet [33] is chosen as our basic architecture. Unet is a well-known FCNs that works fast and shows excellent performance on limited data, which is suitable for our scenarios.

Inspired by Hazirbas *et al.* [34], we modify the U-net into a two-stream multi-level fusion network named fuse-Unet, as shown in Fig. 4. The fuse-Unet retains the encoder-decoder structure of Unet. In the encoder part of fuse-Unet, instead of concatenating RGB and depth features as single input streams with four channel, depth feature and RGB feature are extracted separately as two input streams in the proposed network. The depth feature is fused into the RGB branch constantly, which introduces the multi-level fusion of the RGB-D feature. The fusion operation is simply implemented as an element-wise summation. In the decoder part, the high-level feature map is restored to original image resolution by up-sampling, meanwhile, the features from the encoder part are fused into the features in the decoder part through concatenation in multi-scale.

After image segmentation, a binary image with each pixel classified is obtained, shown in Fig. 5(a). Such a result cannot be directly regarded as the scan path, since the white part represents the entire contour of the spine area and possible segmentation errors can be observed outside the spine area. Therefore, we adopt a maximum connected domain algorithm to eliminate error segmentation results. Then, same as before, taking the center point (320, 240) of segmentation map as the origin, the transverse axis as the x-axis, and the vertical axis as the y-axis, the image coordinate system $\boldsymbol{I}$ is established. The scan point $(x_d, \bar{y}_d)$ in $\boldsymbol{I}$ is obtained by averaging all the y_d in spine area corresponding to x_d. By doing so, we have a pre-planning scan path in the segmentation map (i.e., in $\boldsymbol{I}$), as shown in Fig. 5(c). According to the formula (1), each scan point could be converted into Kinect coordinate system $\boldsymbol{K}$ with its depth coordinate z_k initiated by the corresponding depth value z_d in-depth image. Finally, the scan path in $\boldsymbol{K}$ could be converted into the robotic arm coordinate system $\boldsymbol{R}$, allowing the robotic arm to conduct the automatic scanning.

C. Robotic Arm Control Strategy

During automatic scanning, the robotic arm is driven to ensure the center point of the probe coincides with each point
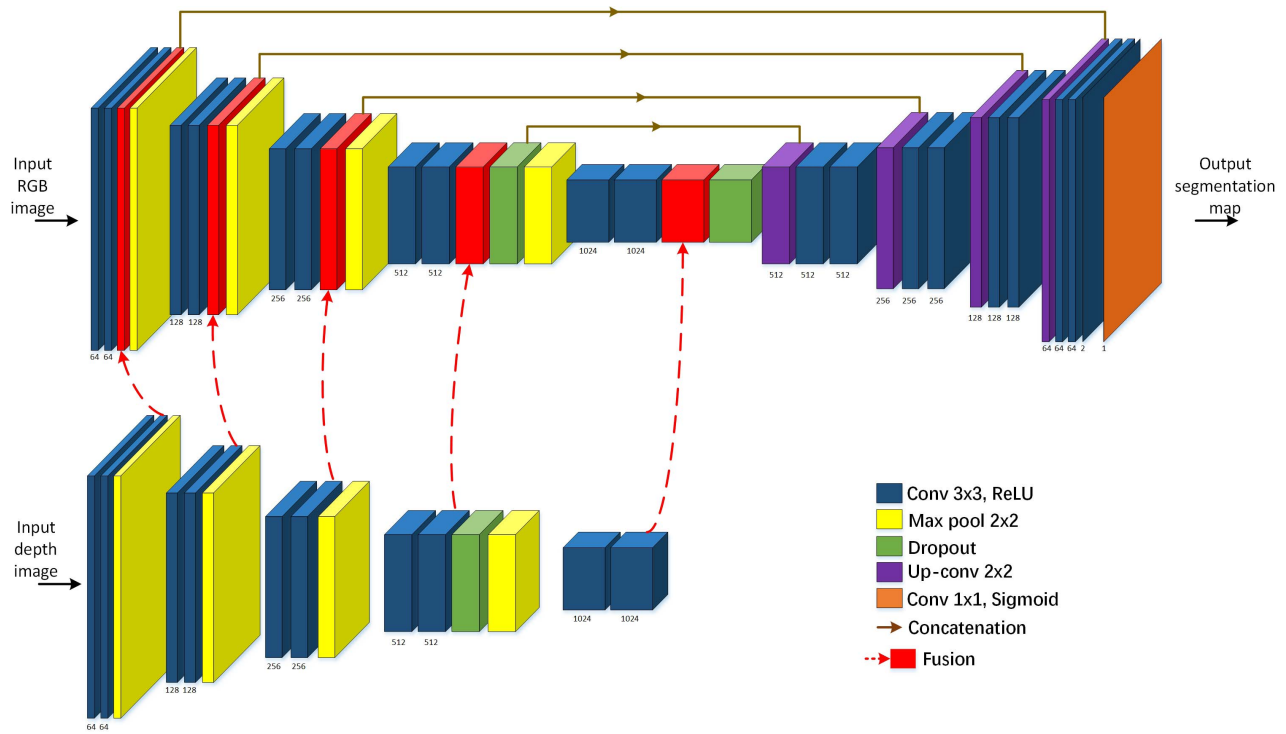
Fig. 4. Architecture of fuse-Unet. The numbers under the blocks indicate the numbers of the convolution kernels. In the encoder part of the fuse-Unet, the numbers of kernels are 64, 128, 256, 512, 1024, and in the decoder part, the numbers of kernels are 512, 256, 128, 64.
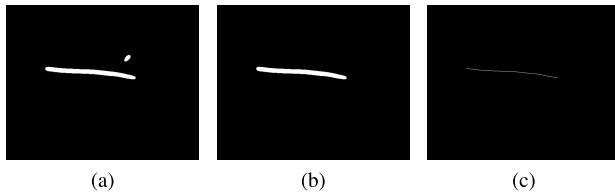

Fig. 5. Post processing on segmentation result. (a) Segmentation result. (b) After adopting maximum connected domain algorithm on segmentation result. (c) Scan path in I.
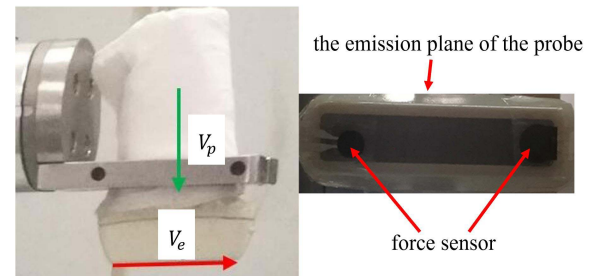

Fig. 6. Illustration of the probe with two force sensors. V_p and V_e are the vectors to define the pose of probe.

in the pre-planning scan path. However, high-quality imaging requires a good match between the center point of the probe and the scanning point, as well as the proper pose of the probe. In traditional freehand scanning, the doctor usually holds and adjusts the pose of the probe on the tissue surface to obtain the best ultrasound slices. Since the robotic arm plays the role of a doctor, the additional focus should be paid to the robotic arm control strategy.

A major consideration is to determine the pose of the probe properly, that is, the probe should be approximately perpendicular to the tissue surface and have close contact with the skin. We adopted the normal-vector-based method [10] to calculate the pose of the probe at each scan point. More specifically, the pose of the probe is defined by two vectors, i.e., V_p and V_e, in Fig. 6. For each scan point, the system determines the pose of the probe by making V_p parallel, meanwhile V_e perpendicular to the normal vector of the plane where the scan point lies. In order to find the normal vector, three points around the scan point are randomly selected to

form a triangle that can be approximately regarded as its local plane. Then the normal vector of the plane described by the triangle can be calculated for each scan point. In this way, the system is able to automatically adjust the center point of the probe to match well with each scan point in an appropriate pose.

Having determined the pose of the probe, we notice that the depth change in the spine area and the breath of subjects could lead to inadequate contact between probe and tissue, consequently, fine adjustment is necessary. Two single-freedom piezoresistive force sensors connected to a data acquisition card are attached to the left and right sides of the front face of the probe to get real-time force feedback, as shown in Fig. 6. There is a small gap between the force sensor and the probe to ensure that the pressure is nearly to zero in the suspended state. The thickness of each force sensor is 0.1–0.2 mm and

(a)

(b)

Fig. 9. Experimental result in phantom experiment. (a) Semantic segmentation result. (b) Top view and side view of the 3-D reconstructed spine phantom volume.

| RGB | Depth | Unet(RGB) | Unet(RGB-D) | FuseNet | fuse-Unet(Ours) |
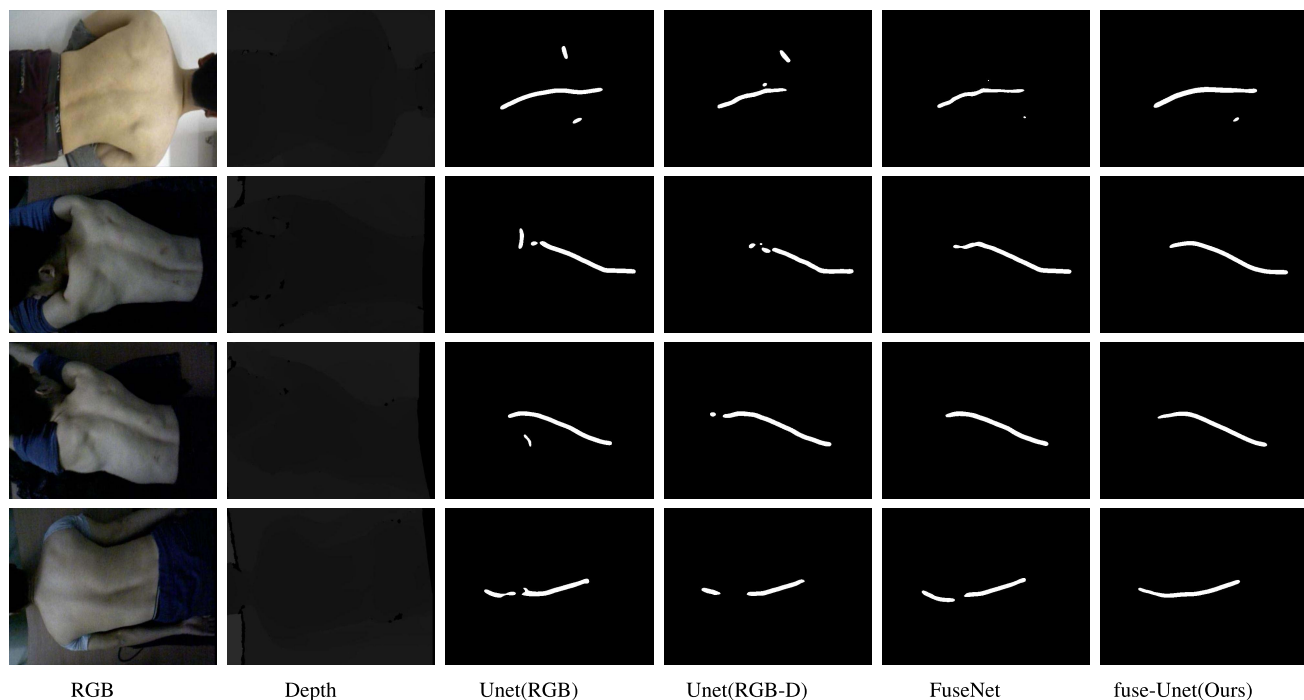
Fig. 10. Visualization examples of the semantic segmentation compared with Unet and FuseNet.

body poses and different scene illumination was adopted for the data collection. Using Kinect, we finally collected 50 sets of images, and manual labeling was conducted right after. In the scanning experiment, we used the linear probe and the convex probe to scan the back of the patient, respectively. The image field of the convex probe was 69°, its detecting depth range was 50–240 mm. In our scene, the detecting depth of the convex probe was set to 70 mm. The setup of the linear probe was the same as that of the phantom experiment.

The limited data collected in this article was far from the large-scale data requirements for deep learning. Hence, data augmentation based on geometric deformations and color deformations was applied to the available data. We primarily needed the robustness to the geometric deformations like shift, rotation, and scaling, as well as color deformations like brightness. Additionally, since the data for *in vivo* experiment was unable to simulate certain spine-related disease, especially scoliosis, we further adapted the image deformation based on
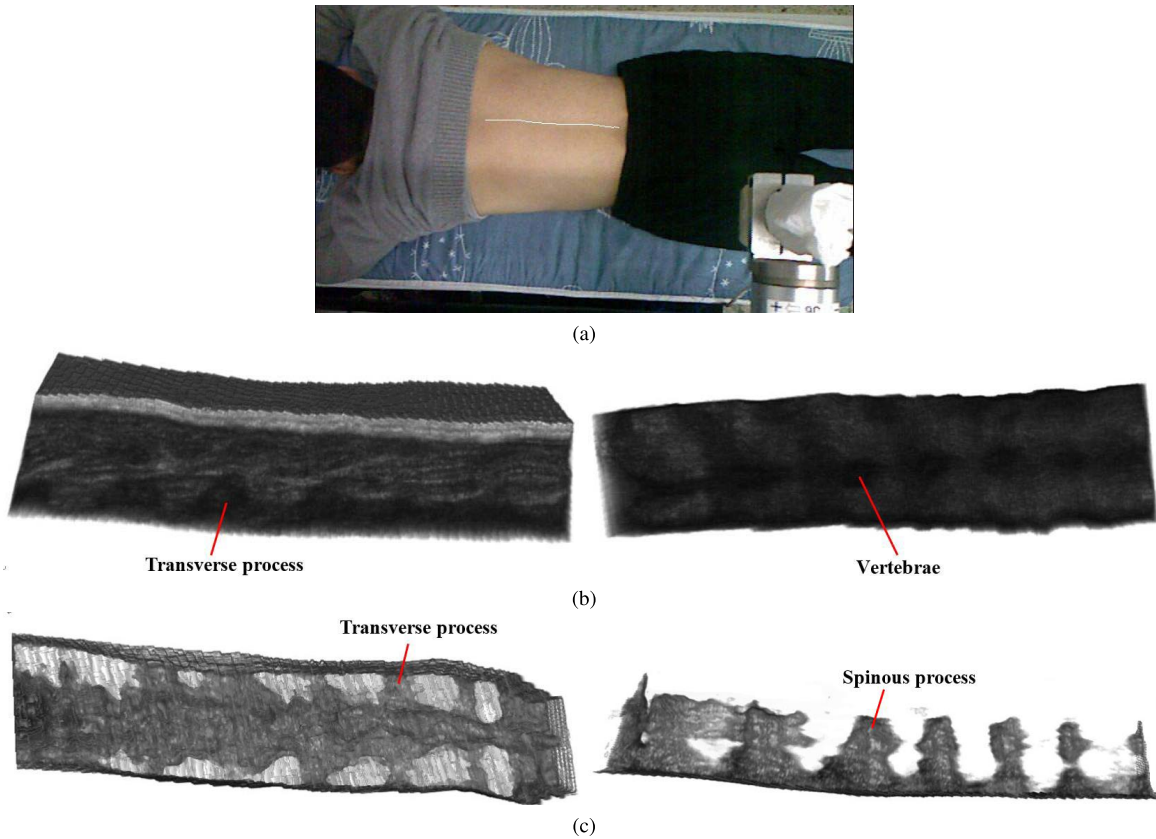
Fig. 11. Experimental results in *in vivo* experiment using the linear probe. (a) Semantic segmentation result. (b) Side view and bottom view of the original reconstructed 3-D spine volume. (c) Top view and side view of the spine volume with all the soft tissue removed.

moving least squares to teach the network the desired invariance. Moving least squares deformation [37] could distort the spine area while keeping the contour of the human body basically unchanged, shown in Fig. 8. After data augmentation, we generated 1500 datasets for the phantom experiment and 2500 datasets for *in vivo* experiment.

We trained the model via adaptive moment estimation (Adam). The input images were normalized and resized to 256×256 before feeding to the network. The training was carried out on an NVIDIA GeForce RTX 2080 GPU (8 GB). To minimize the overhead and make maximum use of the GPU memory, we preferred large input tiles over a large batch size and hence reduced the batch to a single set of images. For optimization, the initial learning rate was set to 10^{-4}. The weights in the convolution network were randomly initialized, and the binary cross entropy was taken as the loss function. To verify the performance of fuse-Unet in *in vivo* experiment, the comparison with Unet and FuseNet was conducted on the basis of ten-fold cross validation. We report the average Intersection over Union (IoU) of spine area in ten runs for the semantic segmentation evaluations. IoU is the standard metric for segmentation purposes, which can be formulated as the number of true positives (intersection) over the sum of true positives, false negatives, and false positives (union).

After realizing the automatic recognition and segmentation, a scan experiment was conducted for the phantom and *in vivo* experiment, respectively. In the phantom experiment, we chose to simulate various scoliotic conditions for accuracy verification of our measurement method. Scoliosis can be divided into three types according to the value of Cobb angle: mild ($<20°$), moderate ($20°\sim40°$), and severe ($>40°$). To verify the reliability of the proposed system in diagnosing scoliosis with different severities, the spine phantom was bent into 30 different Cobb angles, covering the three types of scoliosis and each type contained ten angles. Before scanning, the actual angle was measured and recorded. The spine phantom was scanned using a water-tank scanning approach. For each scan, the pre-planning path was obtained before water injection. The system completed the automatic scan and 3-D reconstruction of the spine phantom. Then Cobb angle measurement based on 3-D volume was conducted, and mean absolute error (MAE), mean relative error (MRE), and error range were calculated between measured angles and actual angles. Moreover, the measurement results obtained under two pre-planning paths could demonstrate the repeatability of the proposed methods using the intraclass correlation coefficient (ICC). ICC can be computed as follows:

$$\text{ICC} = \frac{\text{MS}_R - \text{MS}_E}{\text{MS}_R + (k-1)\text{MS}_E} \qquad (5)$$

where MS_R denotes the mean square for rows, MS_E denotes the mean square of error, and k denotes the number of pre-planning.

TABLE I
MEASUREMENT RESULTS IN PHANTOM EXPERIMENT

Measurement	Mild(<20°)	Moderate(20°~40°)	Severe(>40°)
Actual angle(°)	14.52±3.16	32.20±4.26	57.44±7.23
Measured angle(°)	14.29±2.49	33.12±4.33	59.21±7.95
MAE(°)	1.03±0.21	1.63±0.27	2.44±0.59
MRE(%)	7.35%	5.15%	4.29%
Error range(°)	0.48~1.80	1.02~2.67	1.15~3.72

(a)

(b)
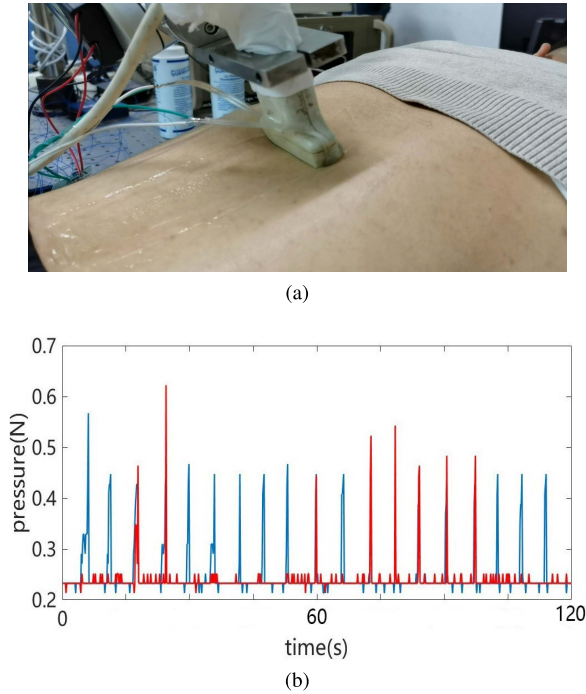
Fig. 12. Demonstration of the scan experiment with a force sensor in *in vivo*. (a) System is scanning the human spine automatically with the probe tightly attached to the spine area. (b) Variation of the force value during the scanning. Note that the red line is read from the left force sensor and the blue line is read from the right one.

In *in vivo* experiment, we enable the force sensors to fine-tune the pose when using the linear probe. The base force F and the floating size f are set to 1.9062 N and $f = 1.6737$ N, respectively. The force sensors are abandoned when using the convex probe for scanning. The subject's spine area was always coated with liquid ultrasound couplant during the scanning. By changing the subject's pose, different shapes of spine area were adopted to test the performance of segmentation, automatic scan, and 3-D reconstruction of the system. 3-D reconstruction and visualization were completed after scanning.

III. RESULTS

A. Results of Phantom Experiment

Phantom experiment validates the reliability of the proposed system. In this experiment, the system recognizes and scans the spine phantom automatically, and then reconstructs 3-D volume. Based on the 3-D volume, we calculate Cobb angles and then conduct statistical analysis between actual angles and measured angles. Table I gives the measurement results of three different scoliosis cases, respectively. The actual angles, measured angles, and MAE are represented in the form of Mean ± SD. As we can see in Table I, although the error increases with the Cobb angle, MAE, MRE, and error range are not much significant. In fact, the measured angles accord with the actual angles within the allowable error range of 5°, which is a clear manifestation of a high accuracy of our system. Fig. 9 gives segmentation result and 3-D reconstructed volume in the phantom experiment. Aided by the precise semantic segmentation, accurate 3-D spine phantom volume with a distinct spinous process and a transverse process can be obtained, thus it is more conducive to the calculation of Cobb angle. We further calculated the ICC value using the formula (5) with a two-way model to evaluate the inter-observer reproducibility. Through the statistics software (Statistical Product and Service Solutions), we obtained ICC = 99.6% at a 95% significant level. In general, our system is able to realize high-precision 3-D reconstruction and visualization of the whole spine with excellent repeatability, thereby assist doctors in the diagnosis of spine-related diseases.

B. Results of In Vivo Experiment

In this section, we evaluate the semantic segmentation performance of our fuse-Unet on the collected dataset. Results are reported in Table II. Note that Unet (RGB) takes only RGB image as its input, and Unet (RGB-D) takes the concatenation of RGB and depth image as its input. Our baseline Unet improves slightly when the depth feature is involved. Employing the two-stream multi-level feature fusion modules on our baseline attains an IoU of 69.2%, which outperforms the baseline by 4.8%. Compared with the top-performing RGB-D semantic segmentation model FuseNet [34], our fuse-Unet gains an improvement of 3%.

Fig. 10 illustrates the visualized comparisons of different semantic segmentation approaches on collected data. As we can see from Fig. 10, our fuse-Unet usually generates smoother and more reasonable results with less segmentation error. The results based on RGB information only reveal that without the aid of depth feature, the network tends to mistake some
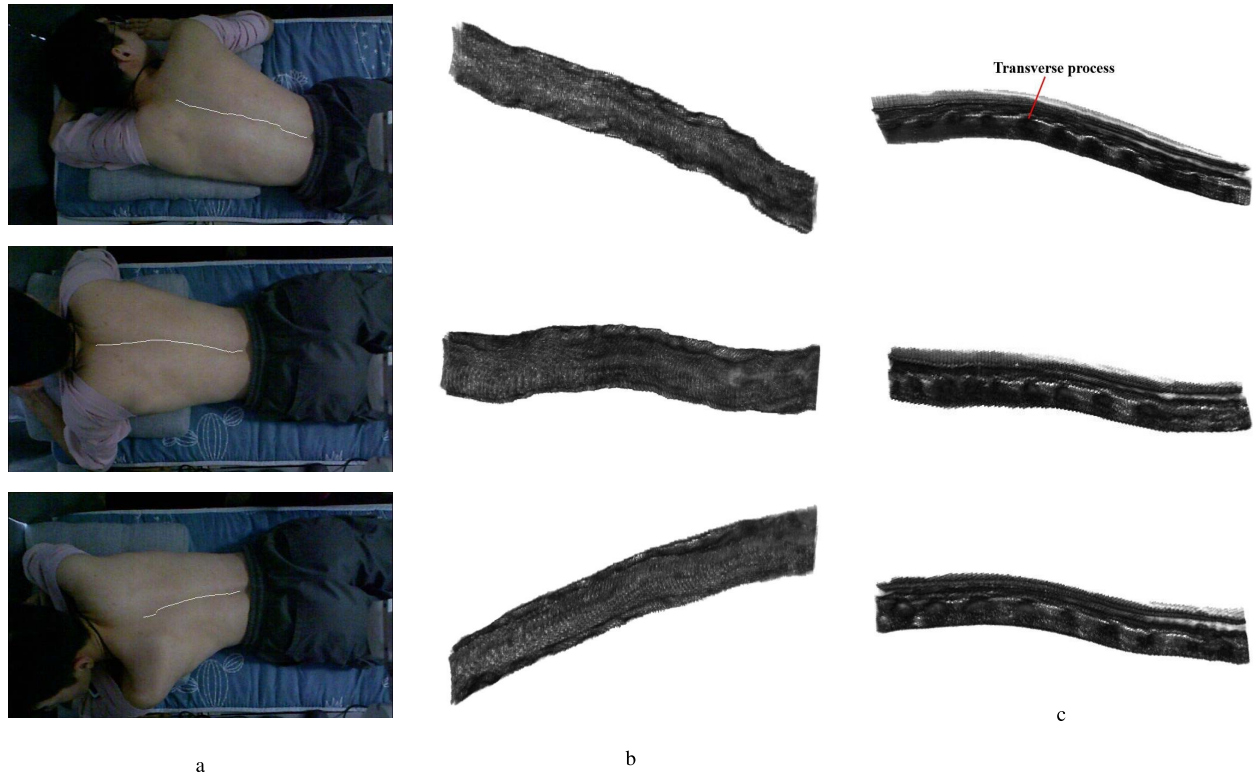
Fig. 13. 3-D reconstruction results under different pre-planning path in *in vivo* using the convex probe. (a) Semantic segmentation result. (b) Top view and (c) side view of the reconstructed 3-D spine volume.

TABLE II
PERFORMANCE COMPARISON OF FUSE-UNET WITH DIFFERENT
MODELS ON THE COLLECTED DATASET

Network	RGB	Depth	IoU
Unet(RGB)	$\sqrt{}$	$\times$	0.644
Unet(RGB-D)	$\sqrt{}$	$\sqrt{}$	0.651
FuseNet	$\sqrt{}$	$\sqrt{}$	0.662
fuse-Unet	$\sqrt{}$	$\sqrt{}$	**0.692**

non-spine areas for spine areas since they share similar shadow features. Moreover, the depth image is a kind of data that is quite different from RGB image, namely, each pixel value in-depth image represents the distance between object and camera, thus it is not an effective way of fusion by simply concatenating the RGB and depth images in channel dimension as a four channel input. In conclusion, experiment results demonstrate that our two-stream multi-level RGB-D feature fusion network fuse-Unet can integrate specific features in different scales and realize the information complementation between different modalities, thus achieving satisfactory performance in our application scene.

In *in vivo* experiment, two healthy male subjects are scanned to verify the performance of our system. We begin with the scan experiment using the linear probe, meanwhile, the pose of the probe is fine-tuned with the help of the two force sensors. The average time cost of automatic scanning is about 120 s and the 3-D reconstruction based on 100 to 150 frames

of B-scans can be completed within 10 s. Fig. 11 presents the experimental results in *in vivo* experiment. Fig. 11(a) is the semantic segmentation result of the human spine area, (b) illustrates the different views of the original 3-D spine volume scanned by the linear probe, from which we can observe the distinct bone structure of the spine. By setting the transparency of soft tissue to maximum, we have Fig. 11(c) for a clearer demonstration of the spine. In addition, Fig. 12 shows the scan experiment with a force sensor in *in vivo*. As seen from Fig. 12(b), the peaks above or under the minimum force 0.2325 N indicate the contact or separation of the probe and the tissue, which are mainly caused by the breathe of the subject. And because of the depth change of the spine area, the shapes of the two curves are not exactly the same. With the help of the fine adjustment strategy using force feedback, we can always ensure the pressure between the probe and tissue falls in the desired force range, hence obtain the satisfactory scanning shown in Fig. 12(a).

In the scan experiment using the convex probe, the average time cost of automatic scanning is about 60 s without enabling the force sensors, and 3-D reconstruction based on 100 to 150 frames takes within 10 s to complete. We conduct several scans on different postures of same subject. Fig. 13 gives the 3-D reconstruction results under different pre-planning paths. As seen in Fig. 13, the shape of the 3-D volume directly depends on the pre-planning of the scan paths. The large view of the convex probe can always cover the whole spine in 2-D B-scans, so regardless of the shape of the 3-D reconstruction results, the bone structure is always complete and clear. In conclusion, the shape of the 3-D volume changes
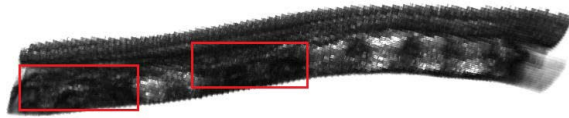
Fig. 14. Reconstructed volume scanned by the convex probe without enabling force sensor. Red rectangles indicate the empty voxels.

with the pre-planning path, and under the reasonable pre-planning, the bone structure is only referable to itself rather than the shape of the volume.

IV. DISCUSSION AND CONCLUSION

In this article, we focus on the clinical scenario of scanning, reconstruction, and measurement of the human spine, and develop an automatic scanning and 3-D ultrasound imaging system with a robotic arm for the human spine. To achieve fully automatic scanning of the spine, an RGB-D sensor Kinect is used to capture the RGB image and depth image of the back of the patient. Then, we propose a two-stream multi-level fusion network for RGB-D semantic segmentation to realize the automatic recognition and segmentation of the spine area. Based on the segmentation result, a pre-planning path is obtained and hence the robotic arm can carry the probe to scan along the path. With the collected raw B-scans, the system provides 3-D reconstruction, visualization, and measurement of the spine. Experimental results obtained from the phantom and *in vivo* experiment have validated the reliability of our system. We seek a solution for enhancing robotic scanning for a case of irregular shapes, e.g., spine, and firstly combine automatic human spine scanning with 3-D ultrasound imaging and measurement.

Although the preliminary work on the proposed system achieved promising results, further steps are to be taken for the clinical use of the system. Our *in vivo* experiment was conducted on the healthy male volunteers, nevertheless, the appearance of the spine in the ultrasound images might vary with the gender of the subjects and the pathological changes of the spine, thus tests on real patients in more challenging scenarios needs to be conducted in our future studies. Furthermore, the patient's breathe often leads to discontinuous reconstructed volume. In fact, when scan using the convex probe without enabling force sensor, occasional empty voxels could be found in the volume, shown in Fig. 14. Although the system can fine-tune the pose of the probe with the help of the force sensors, the scan time could double to nearly 120 s. To avoid the excessively long scan time, the step size of the fine-tune should be carefully chosen. We manually pick the points on the spinous process in 3-D slice to calculate the Cobb angle, in future works, we will consider detecting the bone surface structure automatically in B-scans [38]–[40] to further improve the accuracy of the measurement method and the automation of the system. And from a medical perspective, the scoliosis curve may not always be a perfect circular arc, hence fitting an arc of constant radius to the spine could introduce inaccuracies in the measurement. Hence different approaches to measure scoliosis could be tested on our

system [41]–[43]. Despite the fuse-Unet outperforms the state-of-the-art model FuseNet, more exploration should be carried on to improve the precision of the segmentation model. Last but not least, some inevitable errors, such as light interference, noise of the point cloud, and movement of patient's body, usually lead to the deviation of the pre-planning path from the spine area. In these conditions, the spine might fall out of the field of view of the probe. To overcome this problem, we will design a new fine adjustment strategy in future studies by tracking the spinous process in real-time B-scans and hence centering the spine in the view of the probe.

In conclusion, to the best of our knowledge, this is the first robot-assisted system that combines automatic human spine scanning with 3-D ultrasound imaging and measurement. Our system achieves the intelligent recognition of the spine area and pre-planning of the scan path for the robot system without user participation. And considering the complexity of scanning the spine area, two robot control strategies are adopted to guarantee the effective coupling of the ultrasonic signal. Comparing with traditional 2-D radiography, 3-D ultrasound visualization of the spine and a measurement method based on 3-D volume for scoliosis are offered to realize the spatial structure display and non-radiation examination. With more efforts made, we believe the proposed system is able to partially play the role of ultrasound radiologists and help doctors in the diagnosis of spine-related diseases. Except for spine, our system can be trained to realize the automatic scanning and 3-D reconstruction for other parts of the human body such as the knee joint [44], and thyroid.

REFERENCES

[1] P. Knott *et al.*, "SOSORT 2012 consensus paper: Reducing X-ray exposure in pediatric patients with scoliosis," *Scoliosis*, vol. 9, no. 1, p. 4, Dec. 2014.

[2] C. S. Ling, W. M. Diyana, W. Zaki, A. Hussain, and H. A. Hamid, "Semi-automated vertebral segmentation of human spine in MRI images," in *Proc. Int. Conf. Adv. Electr., Electron. Syst. Eng. (ICAEES)*, Nov. 2016, pp. 120–124.

[3] Q. N. Vo, E. H. M. Lou, and L. H. Le, "3D ultrasound imaging method to assess the true spinal deformity," in *Proc. 37th Annu. Int. Conf. IEEE Eng. Med. Biol. Soc. (EMBC)*, Aug. 2015, pp. 1540–1543.

[4] G.-Q. Zhou, W.-W. Jiang, K.-L. Lai, and Y.-P. Zheng, "Automatic measurement of spine curvature on 3-D ultrasound volume projection image with phase features," *IEEE Trans. Med. Imag.*, vol. 36, no. 6, pp. 1250–1262, Jun. 2017.

[5] Q. N. Vo, L. H. Le, and E. Lou, "A semi-automatic 3D ultrasound reconstruction method to assess the true severity of adolescent idiopathic scoliosis," *Med. Biol. Eng. Comput.*, vol. 57, no. 10, pp. 2115–2128, Oct. 2019.

[6] E. Garcia-Cano *et al.*, "A freehand ultrasound framework for spine assessment in 3D: A preliminary study," in *Proc. 42nd Annu. Int. Conf. IEEE Eng. Med. Biol. Soc. (EMBC)*, Jul. 2020, pp. 2096–2100.

[7] C. Herickhoff, J. Lin, and J. Dahl, "Low-cost sensor-enabled freehand 3D ultrasound," in *Proc. IEEE Int. Ultrason. Symp. (IUS)*, Oct. 2019, pp. 498–501.

[8] H. Chen, R. Zheng, E. Lou, and D. Ta, "Imaging spinal curvatures of AIS patients using 3D US free-hand fast reconstruction method," in *Proc. IEEE Int. Ultrason. Symp. (IUS)*, Oct. 2019, pp. 1440–1443.

[9] Q. Huang, B. Wu, J. Lan, and X. Li, "Fully automatic three-dimensional ultrasound imaging based on conventional B-scan," *IEEE Trans. Biomed. Circuits Syst.*, vol. 12, no. 2, pp. 426–436, Apr. 2018.

[10] Q. Huang, J. Lan, and X. Li, "Robotic arm based automatic ultrasound scanning for three-dimensional imaging," *IEEE Trans. Ind. Informat.*, vol. 15, no. 2, pp. 1173–1182, Feb. 2019.

[11] M. Tirindelli *et al.*, "Force-ultrasound fusion: Bringing spine robotic-US to the next 'level,'" *IEEE Robot. Autom. Lett.*, vol. 5, no. 4, pp. 5661–5668, Oct. 2020.

[12] P.-W. Hsu, R. W. Prager, A. H. Gee, and G. M. Treece, "Rapid, easy and reliable calibration for freehand 3D ultrasound," *Ultrasound Med. Biol.*, vol. 32, no. 6, pp. 823–835, Jun. 2006. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S030156290601489X

[13] F. Cenni, D. Monari, K. Desloovere, E. Aertbeliën, S.-H. Schless, and H. Bruyninckx, "The reliability and validity of a clinical 3D freehand ultrasound system," *Comput. Methods Programs Biomed.*, vol. 136, pp. 179–187, Nov. 2016. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S0169260716304874

[14] S.-W. Chung, C.-C. Shih, and C.-C. Huang, "Freehand three-dimensional ultrasound imaging of carotid artery using motion tracking technology," *Ultrasonics*, vol. 74, pp. 11–20, Feb. 2017.

[15] M. I. Daoud, A.-L. Alshalalfah, F. Awwad, and M. Al-Najar, "Freehand 3D ultrasound imaging system using electromagnetic tracking," in *Proc. Int. Conf. Open Source Softw. Comput. (OSSCOM)*, Sep. 2015, pp. 1–5.

[16] P. Toonkum, N. C. Suwanwela, and C. Chinrungrueng, "Reconstruction of 3D ultrasound images based on cyclic regularized Savitzky–Golay filters," *Ultrasonics*, vol. 51, no. 2, pp. 136–147, Feb. 2011. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S0041624X10001095

[17] Q.-H. Huang, Z. Yang, W. Hu, L.-W. Jin, G. Wei, and X. Li, "Linear tracking for 3-D medical ultrasound imaging," *IEEE Trans. Cybern.*, vol. 43, no. 6, pp. 1747–1754, Dec. 2013.

[18] D. Zhang and H. Huang, "Prevalence of work-related musculoskeletal disorders among sonographers in China: Results from a national Web-based survey," *J. Occupational Health*, vol. 59, no. 6, pp. 529–541, Nov. 2017.

[19] A. Priester, S. Natarajan, and M. Culjat, "Robotic ultrasound systems in medicine," *IEEE Trans. Ultrason., Ferroelectr., Freq. Control*, vol. 60, no. 3, pp. 507–523, Mar. 2013.

[20] F. Conti, J. Park, and O. Khatib, *Interface Design and Control Strategies for a Robot Assisted Ultrasonic Examination System*. Berlin, Germany: Springer, 2014, pp. 97–113.

[21] C.-W.-J. Cheung, G.-Q. Zhou, S.-Y. Law, K.-L. Lai, W.-W. Jiang, and Y.-P. Zheng, "Freehand three-dimensional ultrasound system for assessment of scoliosis," *J. Orthopaedic Transl.*, vol. 3, no. 3, pp. 123–133, Jul. 2015. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S2214031X1500039X

[22] R. Lechner, D. Putzer, D. Dammerer, M. Liebensteiner, C. Bach, and M. Thaler, "Comparison of two- and three-dimensional measurement of the cobb angle in scoliosis," *Int. Orthopaedics*, vol. 41, no. 5, pp. 957–962, May 2017.

[23] D. Ottacher, A. Chan, E. Parent, and E. Lou, "Positional and orientational accuracy of 3-D ultrasound navigation system on vertebral phantom study," *IEEE Trans. Instrum. Meas.*, vol. 69, no. 9, pp. 6412–6419, Sep. 2020.

[24] T. Lehmann, C. Rossa, N. Usmani, R. S. Sloboda, and M. Tavakoli, "Intraoperative tissue young's modulus identification during needle insertion using a laterally actuated needle," *IEEE Trans. Instrum. Meas.*, vol. 67, no. 2, pp. 371–381, Feb. 2018.

[25] R. W. Prager, R. N. Rohling, A. H. Gee, and L. Berman, "Rapid calibration for 3-D freehand ultrasound," *Ultrasound Med. Biol.*, vol. 24, no. 6, pp. 69–855, 1998.

[26] P. Bozek and E. Pivarciová, "Registration of holographic images based on the integral transformation," *Comput. Informat.*, vol. 31, no. 6, pp. 1369–1383, 2012.

[27] J. J. Moré, *The Levenberg-Marquardt Algorithm: Implementation and Theory* (Lecture Notes in Mathematics), vol. 630. Berlin, Germany: Springer, 1978, pp. 105–116.

[28] A. Garcia-Garcia, S. Orts-Escolano, S. Oprea, V. Villena-Martinez, and J. Garcia-Rodríguez, "A review on deep learning techniques applied to semantic segmentation," *CoRR*, vol. abs/1704.06857, pp. 1–23, Apr. 2017. [Online]. Available: http://arxiv.org/abs/1704.06857

[29] S. Gupta, R. B. Girshick, P. Arbelaez, and J. Malik, "Learning rich features from RGB-D images for object detection and segmentation," *CoRR*, vol. abs/1407.5736, pp. 345–360, Sep. 2014. [Online]. Available: http://arxiv.org/abs/1407.5736

[30] C. Couprie, C. Farabet, L. Najman, and Y. Lecun, "Toward real-time indoor semantic segmentation using depth information," *J. Mach. Learn.*, vol. 1, pp. 1–22, Oct. 2000.

[31] Y. Li, J. Zhang, Y. Cheng, K. Huang, and T. Tan, "Semantics-guided multi-level RGB-D feature fusion for indoor semantic segmentation," in *Proc. IEEE Int. Conf. Image Process. (ICIP)*, Sep. 2017, pp. 1262–1266.

[32] W. Shi *et al.*, "Multilevel cross-aware RGBD indoor semantic segmentation for bionic binocular robot," *IEEE Trans. Med. Robot. Bionics*, vol. 2, no. 3, pp. 382–390, Aug. 2020.

[33] O. Ronneberger, P. Fischer, and T. Brox, "U-Net: Convolutional networks for biomedical image segmentation," in *Medical Image Computing and Computer-Assisted Intervention—MICCAI 2015*, N. Navab, J. Hornegger, W. M. Wells, and A. F. Frangi, Eds. Cham, Switzerland: Springer, 2015, pp. 234–241.

[34] C. Hazirbas, L. Ma, C. Domokos, and D. Cremers, "FuseNet: Incorporating depth into semantic segmentation via fusion-based CNN architecture," in *Computer Vision—ACCV 2016*, S.-H. Lai, V. Lepetit, K. Nishino, and Y. Sato, Eds. Cham, Switzerland: Springer, 2017, pp. 213–228.

[35] Q. Huang, Y. Huang, W. Hu, and X. Li, "Bezier interpolation for 3-D freehand ultrasound," *IEEE Trans. Human-Mach. Syst.*, vol. 45, no. 3, pp. 385–392, Jun. 2015.

[36] Q. Huang, Q. Deng, L. Li, J. Yang, and X. Li, "Scoliotic imaging with a novel double-sweep 2.5-dimensional extended field-of-view ultrasound," *IEEE Trans. Ultrason., Ferroelectr., Freq. Control*, vol. 66, no. 8, pp. 1304–1315, Aug. 2019.

[37] S. Schaefer, T. McPhail, and J. Warren, "Image deformation using moving least squares," in *Proc. ACM SIGGRAPH Papers (SIGGRAPH)*, 2006, pp. 533–540.

[38] X. Feng, Q. Huang, and X. Li, "Ultrasound image de-speckling by a hybrid deep network with transferred filtering and structural prior," *Neurocomputing*, vol. 414, pp. 346–355, Nov. 2020. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S0925231220313990

[39] Q. Huang, Z. Miao, J. Li, L. Liu, and X. Li, "Classification of breast ultrasound with human-rating BI-RADS scores using mined diagnostic patterns and optimized neuro-network," *Neurocomputing*, vol. 417, pp. 536–542, Dec. 2020. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S0925231220313011

[40] Q. Huang, Y. Huang, Y. Luo, F. Yuan, and X. Li, "Segmentation of breast ultrasound image with semantic classification of superpixels," *Med. Image Anal.*, vol. 61, Apr. 2020, Art. no. 101657.

[41] T. Ungi *et al.*, "Spinal curvature measurement by tracked ultrasound snapshots," *Ultrasound Med. Biol.*, vol. 40, no. 2, pp. 447–454, Feb. 2014. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S0301562913010557

[42] R. Zheng, D. Hill, D. Hedden, M. Moreau, S. Southon, and E. Lou, "Assessment of curve progression on children with idiopathic scoliosis using ultrasound imaging method," *Eur. Spine J.*, vol. 27, no. 9, pp. 2114–2119, Sep. 2018.

[43] D.-S. Li, G.-Q. Zhou, Y.-K. He, P. Zhou, S.-Y. He, and Y.-P. Zheng, "Ultrasound sagittal projection imaging for the assessment of scoliosis," in *Proc. IEEE Int. Ultrason. Symp. (IUS)*, Sep. 2020, pp. 1–4.

[44] Q. Huang, J. Yao, J. Li, M. Li, M. R. Pickering, and X. Li, "Measurement of quasi-static 3-D knee joint movement based on the registration from CT to US," *IEEE Trans. Ultrason., Ferroelectr., Freq. Control*, vol. 67, no. 6, pp. 1141–1150, Jun. 2020.

Cui Yang received the B.S. and Ph.D. degrees in communication and information system from the South China University of Technology, Guangzhou, China, in 2005 and 2010, respectively.

She is currently an Associate Professor with the School of Electronic and Information, South China University of Technology. Her current research interests include signal processing, ultrasound imaging, and robotic ultrasound.

Mingyao Jiang received the B.E. degree in electronic communications and engineering from the Dalian University of Technology, Dalian, China, in 2018. He is currently pursuing the master's degree with the School of Electronic and Information Engineering, South China University of Technology, Guangzhou, China.

His research interest includes medical robotics.

Mianjie Chen received the B.E. degree in electronic information engineering from the College of Electronic Engineering (College of Artificial Intelligence), South China Agricultural University, Guangzhou, China, in 2020. He is currently pursuing the master's degree with the School of Electronic and Information Engineering, South China University of Technology, Guangzhou.

His research interest includes medical robotics.

Jianyi Li received the Ph.D. degree in human anatomy from Southern Medical University (formerly First Military Medical University), Guangzhou, China, in 2007.

He is currently a Professor with the Guangdong Provincial Key Laboratory of Medical Biomechanics, Guangdong Engineering Research Center for Translation of Medical 3-D Printing Application, Department of Anatomy, School of Basic Medical Sciences, Southern Medical University, and the Director of the Department of Research and Education, The Seventh Affiliated Hospital, Southern Medical University. His research interests include digital medicine, 3-D printing, and orthopedic biomechanics.

Maoqing Fu received the Ph.D. degree in human anatomy from Southern Medical University, Guangzhou, China, in 2017.

He is currently a full-time Spine Surgeon with The Seventh Affiliated Hospital, Southern Medical University. His research interests include minimally invasive spinal technology, treatment of chronic spinal pain, digital medicine, and orthopedic biomechanics.

Qinghua Huang received the B.E. and M.E. degrees in automatic control and pattern recognition from the University of Science and Technology of China, Hefei, China, in 1999 and 2002, respectively, and the Ph.D. degree in biomedical engineering from The Hong Kong Polytechnic University, Hong Kong, in 2007.

In 2008, he joined the School of Electronic and Information Engineering, South China University of Technology, Guangzhou, China, as an Associate Professor, where he was promoted to a Full Professor in 2013. He is currently a Distinguished Professor with Northwestern Polytechnical University, Xi'an, China. His research interests include ultrasonic imaging, medical image analysis, bioinformatics, and intelligent computation and its applications.