Mathematical Biosciences and Engineering

MBE, 21(5): 5947–5971.
DOI: 10.3934/mbe.2024262
Received: 18 December 2023
Revised: 13 January 2024
Accepted: 24 January 2024
Published: 15 May 2024

http://www.aimspress.com/journal/MBE

Research article

A prostate seed implantation robot system based on human-computer interactions: Augmented reality and voice control

Xinran Zhang[1], Yongde Zhang[1,2,*], Jianzhi Yang[1] and Haiyan Du[1]

[1] Key Laboratory of Advanced Manufacturing and Intelligent Technology, Harbin University of Science and Technology, Harbin 150080, China
[2] Foshan Baikang Robot Technology Co., Ltd., Foshan 528237, China

* **Correspondence:** Email: zhangyd@hrbust.edu.cn.

Abstract: The technology of robot-assisted prostate seed implantation has developed rapidly. However, during the process, there are some problems to be solved, such as non-intuitive visualization effects and complicated robot control. To improve the intelligence and visualization of the operation process, a voice control technology of prostate seed implantation robot in augmented reality environment was proposed. Initially, the MRI image of the prostate was denoised and segmented. The three-dimensional model of prostate and its surrounding tissues was reconstructed by surface rendering technology. Combined with holographic application program, the augmented reality system of prostate seed implantation was built. An improved singular value decomposition three-dimensional registration algorithm based on iterative closest point was proposed, and the results of three-dimensional registration experiments verified that the algorithm could effectively improve the three-dimensional registration accuracy. A fusion algorithm based on spectral subtraction and BP neural network was proposed. The experimental results showed that the average delay of the fusion algorithm was 1.314 s, and the overall response time of the integrated system was 1.5 s. The fusion algorithm could effectively improve the reliability of the voice control system, and the integrated system could meet the responsiveness requirements of prostate seed implantation.

Keywords: prostate seed implantation robot; voice control; augmented reality; medical robot; three-dimensional registration; brachytherapy; prostate cancer

1. Introduction

In the traditional prostate seed implantation surgery, doctors use transrectal ultrasound (TRUS) to obtain medical images to judge the spatial relationship between prostate tissue and the needle [1]. The medical images obtained by TRUS have limited vision space, insufficient image quality and lack of depth information, which makes the surgical procedure extremely dependent on the professional knowledge and experience of doctors [2,3]. With the development of science and technological innovation, robot-assisted seed implantation has been widely used. However, the visualization problem of prostate seed implantation exists, which makes the process of robot-assisted prostate seed implantation more unstable, and the curative effect is severely limited [4].

The development of augmented reality technology has become a reality to provide doctors with high-quality visualization functions in the surgical field. Augmented reality technology can fuse real scenes and virtual scenes in a more natural way, which ensures the intuitiveness and real-time of visual information, and plays a role in intraoperative navigation [5–7]. The three-dimensional virtual model reconstructed by two-dimensional tomography can make up for the shortcomings of traditional medical images. Pre-operative planning in the three-dimensional model can help doctors obtain a more accurate surgical path and a more intuitive spatial relationship.

Robot-assisted prostate seed implantation can meet the technical and application requirements of the surgery [8,9]. In the known medical robots, the control methods are usually traditional methods such as mouse and keyboard. These control methods are all contact control, which easily leads to two-way pollution between doctors and instruments. Voice control technology is a new non-contact control method. Combined with the set voice control instruction, the surgical visualization system can be organically combined with the robot motion control system, so as to ensure the corresponding relationship between the visualization information and the robot motion information [10]. Li et al. [11] designed a pneumatic robot system for prostate seed implantation which can work in MRI environment.

In terms of prostate seed implantation robot, Fichtinger et al. [12] developed a robot-assisted prostate brachytherapy system, which successfully completed clinical feasibility and safety trials. Djohossou et al. [13] studied a compact and lightweight delta robot to improve the accuracy of prostate brachytherapy. Halima et al. [14] developed a 6-DOF compact and lightweight cooperative robot for prostate brachytherapy. Wang et al. [15] improved the structure of the RRT cantilever articulated seed implantation robot and designed a multi-needle soft tissue stable insertion device. The device can fix the prostate so that the needle can be inserted accurately. Chen et al. [16] proposed a robot system for prostate brachytherapy guided by TRUS. The system adopts a new automatic needle release mechanism to realize multi-needle insertion into prostate according to real-time dose measurement. Jiang et al. [17] proposed a new transperineal prostate brachytherapy robot system. The puncture positioning error of the robot can be limited within 0.7 mm. Dai et al. [18] designed an 11-DOF transrectal prostate seed implantation robot. The attitude adjustment module based on double parallelogram mechanism is designed, so that the ultrasonic probe can do centering motion. Bakouri et al. [19] designed a robot wheelchair based on convolutional neural network voice control. Tran et al. [20] developed a soft robot exoskeleton controlled by voice to assist hand movement, and evaluated the system.

Augmented reality technology provides visual navigation function for seed implantation surgery, which can make up for the limitations of two-dimensional imaging. Watanabe et al. [21] proposed a perspective neuronavigation system. The system uses augmented reality technology and motion

capture system to ensure navigation functions with minimum blind spots. Cohen et al. [22] developed an AR guidance system for robot-assisted laparoscopic surgery. Yamamoto et al. [23] have developed an interactive interface for remote minimally invasive surgical robots, which provides augmented reality visual feedback using three-dimensional graphics overlay. Song et al. [24] introduced the HMD-based AR prototype into endodontic surgery for the first time, and the system can be used to dynamically guide dentists to operate surgical instruments for posterior teeth treatment. Gîrbacia et al. [25] proposed an AR system using RGB-D camera. Lee et al. [26] adopted a vision-based augmented reality (AR) tracking system. The system helps doctors locate recurrent laryngeal nerves (RLN) during robotic thyroid surgery. Xu et al. [27] developed a surgical navigation system based on space augmented reality to assist the placement of pedicle screws in minimally invasive spinal surgery. Samei et al. [28] proposed a robot image guidance system to assist laparoscopic radical prostatectomy. The system has been successfully applied in the operation of 12 patients. Schiavina et al. [29] performed radical prostatectomy on 26 patients using a robotic surgical system guided by real-time augmented reality, and evaluated the surgical results.

Voice control technology is a more direct and convenient way of human-computer interaction. According to the actual needs, doctors can control the corresponding equipment by voice. As early as the 1990s, voice control was used in medical robots. Reichenspurner et al. [30] controlled the robotic arm to locate the endoscope through voice control technology. Zinchenko et al. [31] developed a voice control method of endoscope stand for long-distance movement control. Gundogdu et al. [32] developed a voice control system to control the robotic arm. The control system has been applied to RRRR robotic arm. The developed system improves the efficiency of voice recognition. Ruzaij et al. [33] proposed a multi-mode voice controller for rehabilitation. Pramanik et al. [34] designed a fully voice-controlled robot that can be used as a 24-hour hospital assistant. The voice control system has a training system, which can guide robots in multiple languages. Matarneh et al. [35] proposed a set of voice instructions for intracranial surgery. Doctors control the medical robots through voice instructions.

In this paper, a prostate seed implantation robot system based on AR and voice control is proposed, which provides a more friendly visualization system and control system. The MRI medical image of prostate is denoised and segmented, and then the three-dimensional virtual model of prostate is reconstructed by surface rendering technology. Based on the obtained three-dimensional virtual model of prostate, the architecture and scene of augmented reality system are constructed. An improved singular value decomposition (SVD) three-dimensional registration algorithm based on iterative closest point (ICP) is proposed, and the accuracy experiment, real-time experiment, and repetitive experiment of the three-dimensional registration algorithm are used to verify the improved algorithm. The scheme and process of voice control are defined, and the non-real-time problem of voice control system is optimized by three-dimensional registration technology. Finally, a voice denoising fusion algorithm based on spectral subtraction and back propagation neural network (BPNN) is proposed to denoise and optimize the voice signal. The experimental platform of voice control is built, and the reliability of the fusion algorithm is verified by the voice signal-to-noise ratio experiment, denoising delay experiment, voice denoising waveform diagram experiment, and voice recognition experiment.

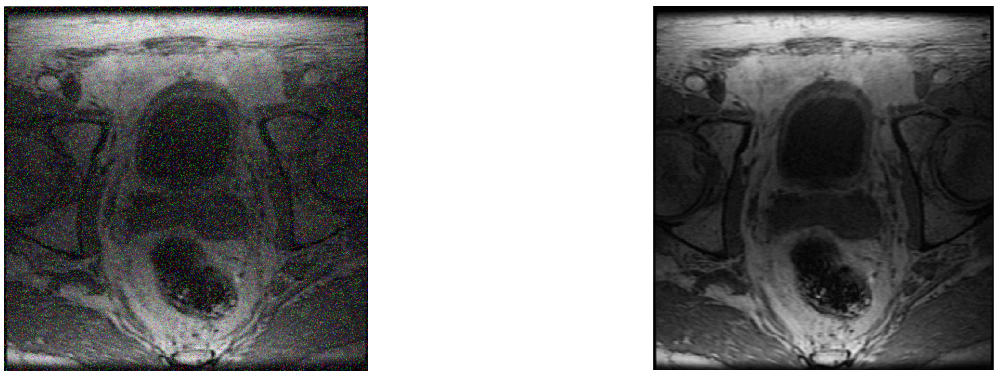
2. Materials and methods

2.1. Augmented reality system of prostate seed implantation

2.1.1. Three-dimensional model reconstruction

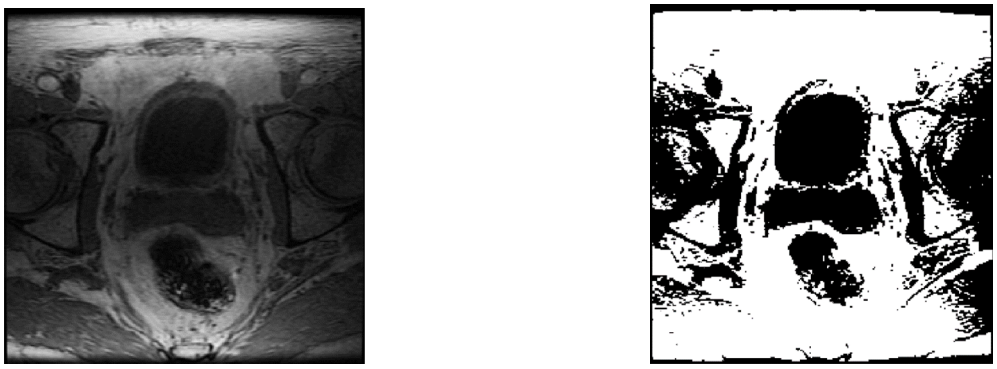
Median filtering has a good effect on eliminating salt and pepper noise and it can avoid the damage to the edge structure of image features. Therefore, median filtering method is used to denoise the medical image of the prostate. The comparison of prostate magnetic resonance imaging (MRI) before and after denoising are shown in Figure 1.

According to the characteristics of prostate MRI sequence images, a region-based threshold segmentation algorithm is used to segment prostate medical images. After selecting an appropriate threshold, the segmentation results of important physiological features in prostate medical images are shown in Figure 2. After image segmentation, the obvious characteristics of organs and tissues mainly include rectum, prostate, bladder, femur and pubic bone.



(a) (b)

Figure 1. Comparison of prostate MRI images before and after denoising: (a) original image, (b) denoised image.



(a) (b)

Figure 2. Comparison of prostate MRI images before and after segmentation: (a) pre-segmentation image, (b) segmented image.

The surface rendering method of Marching Cubes algorithm is adopted to complete the rendering of 3D virtual model [36]. In this paper, 60 MRI sequence images of prostate with 1.5 mm interval are used as the original data of 3D virtual model reconstruction. After the image denoising and image segmentation, it is necessary to render the medical image to reconstruct the 3D virtual model, as shown in Figure 3.

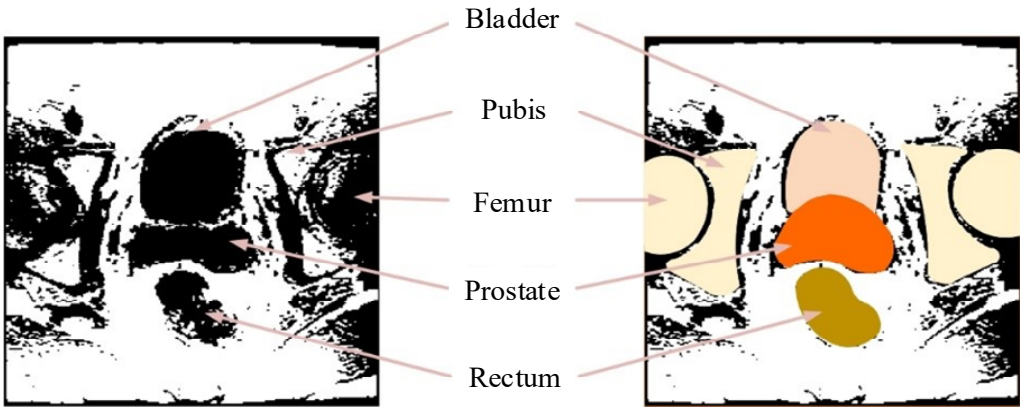
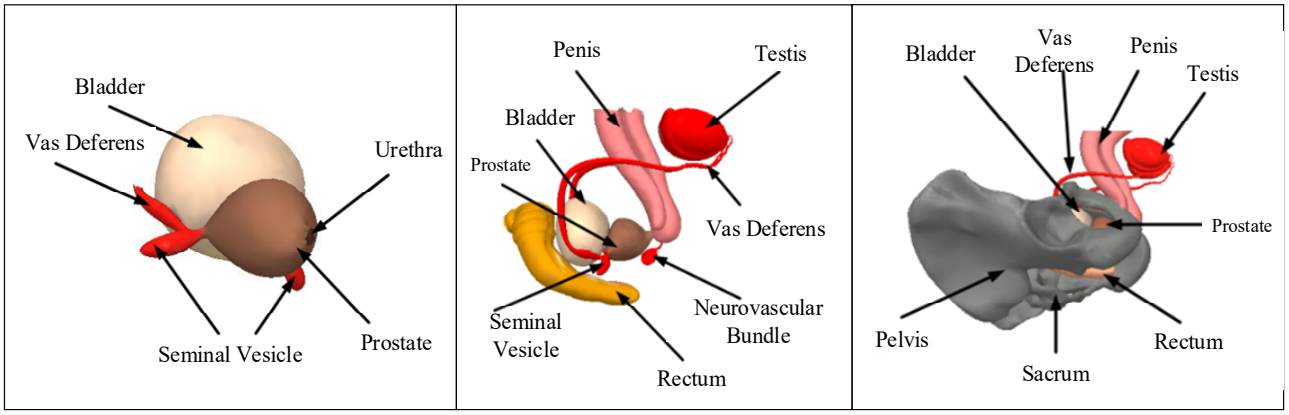


Figure 3. Image rendering: in the rendered medical image, the target features are more obvious, and the characteristic contours of physiological tissues are clearer.

The rendered medical image is imported into the 3DSlicer open-source platform, and the 3D virtual model reconstruction program is written by marching cubes algorithm, so as to reconstruct the 3D virtual model of prostate and its surrounding tissues and organs, as shown in Figure 4.



(a) (b) (c)

Figure 4. Three-dimensional model of prostate and its surrounding organs: (a) prostate tissue model, (b) prostate urinary organ model, (c) pelvic model.

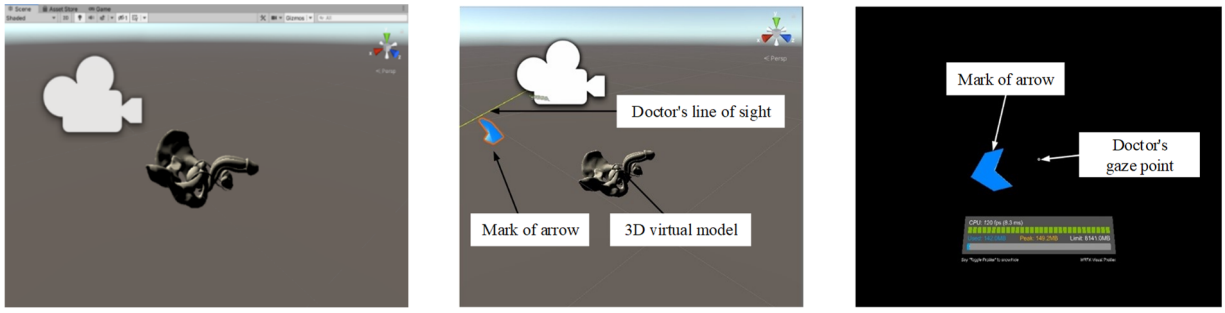
2.1.2. Augmented reality scene construction

Augmented reality system is a system that simultaneously meets real-time interaction, virtual-real integration and three-dimensional display technology [37]. The architecture of augmented reality

system for prostate seed implantation is completed from two aspects: Hardware and software. The main 3D display device of augmented reality system is Microsoft HoloLens 2 head-mounted display. The software system adopts unity, visual studio and Hololens 2 simulator.

The augmented reality scene is built in Unity. According to the actual needs, interactive setting, boundary setting, and spatial perception setting are made. Auxiliary orientation determination of augmented reality scene is shown in Figure 5. In order to provide better visualization and interactive control effect, the boundary and collision volume of the model are set, as shown in Figure 6.

In the augmented reality scene, doctors cannot only adjust the position of the three-dimensional virtual model in the visual field by changing the head and gaze direction, but also directly adjust the position of the model itself in the visual field through the preset interactive control mode, as shown in Figure 7. The interactive control methods such as gestures, gaze and voice can be used in simulators or hardware devices to manipulate the model.



(a) (b) (c)

Figure 5. (a) Augmented reality scene construction in Unity. Auxiliary orientation determination of augmented reality scene: (b) spatial relations in Unity and (c) doctor's visual field in augmented reality environment.

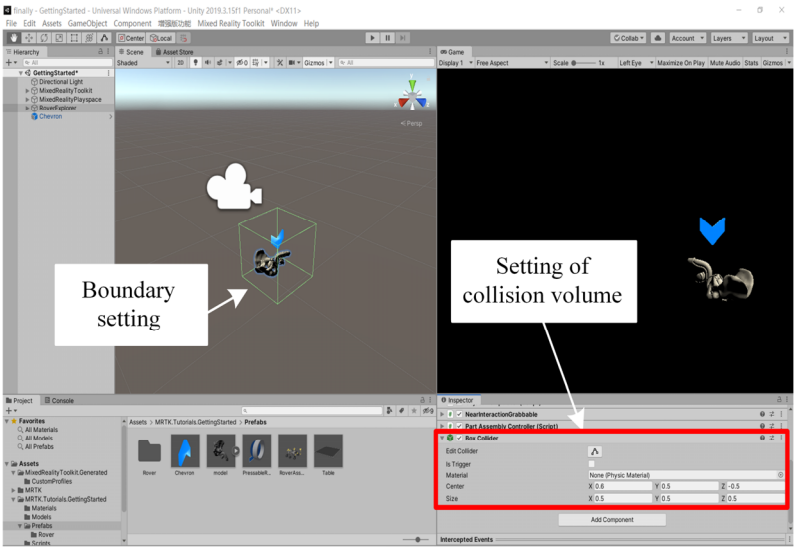
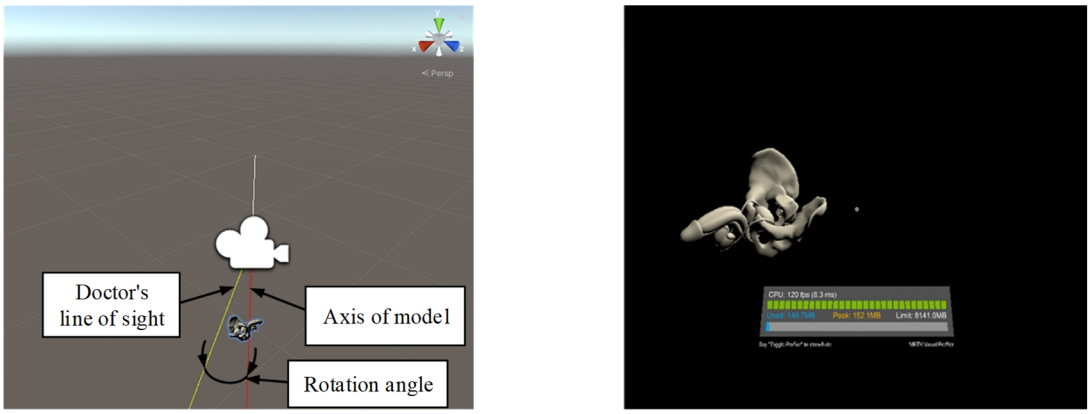


Figure 6. Boundary and collision body setting.

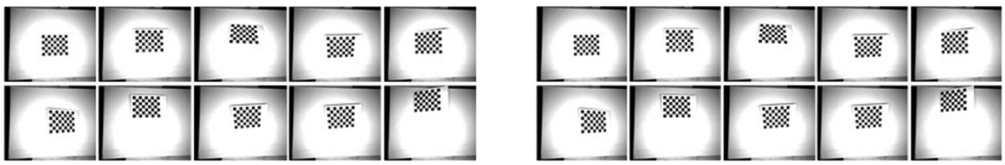


(a) (b)

Figure 7. Interactive control mode: (a) Unity scene for controlling model rotation and (b) doctor's visual field in augmented reality environment.

2.2. Three-dimensional registration technology

A total of 20 photos of checkerboard calibration board were taken by binocular camera at different angles, as shown in Figure 8. The photos obtained by left camera and right camera and the corresponding calibration error histograms were exported, as shown in Figure 9. Then, the camera calibration and distortion correction are realized using these 20 photos and the camera calibration toolbox of Matlab, and the internal parameters of the camera are solved.



(a) (b)

Figure 8. The photos obtained by left camera and right camera: (a) left camera photo, (b) right camera photo.

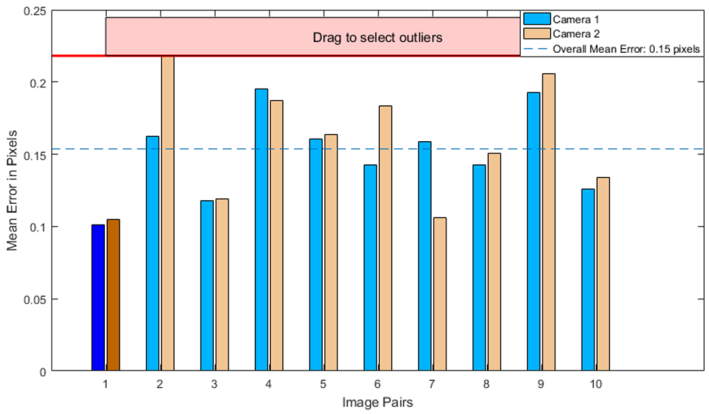


Figure 9. Calibration error histogram of binocular camera.

A three-dimensional registration conversion relationship based on point-to-point data is established between the virtual world coordinate system and the real-world coordinate system. $\{v_i\}$ and $\{r_i\}$ are the coordinate point sets of the patient's body surface markers in the virtual world coordinate system and the real-world coordinate system. *i* is the number of markers. Four-dimensional homogeneous matrix T_v^r is the transformation matrix between the two coordinate point sets. The conversion relationship between two coordinate point sets can be expressed by Eq (1).

$$r_i = R'v_i + T', \tag{1}$$

where R' is the rotation matrix and T' is the translation matrix. The solved R' and T' should satisfy the least square Eq (2).

$$E = min \sum_i^n (r_i - R'v_i - T')^2 \tag{2}$$

It is necessary to solve the optimal solution of R' and T' satisfying the least square Eq (2). An improved SVD algorithm based on ICP is proposed to optimize the 3D registration technology. The solution process is as follows.

1). Solve centroids $\bar{v}$ and $\bar{r}$ of point sets $\{v_i\}$ and $\{r_i\}$:

$$\bar{v} = \frac{1}{n}\sum_i^n v_i, \tag{3}$$

$$\bar{r} = \frac{1}{n}\sum_i^n r_i. \tag{4}$$

2). Establish new point sets $\{v_{ci}\}$ and $\{r_{ci}\}$ with centroid as coordinate origin:

$$v_{ci} = v_i - \bar{v}, \tag{5}$$

$$r_{ci} = r_i - \bar{r}. \tag{6}$$

3). Substitute new point sets $\{v_{ci}\}$ and $\{r_{ci}\}$ to optimize the least square Eq (2):

$$E = \sum_i^n (r_{ci} - R'v_{ci})^2 = \sum_i^n ({r_{ci}}^T r_{ci} + {v_{ci}}^T v_{ci} - 2{r_{ci}}^T R' v_{ci}). \tag{7}$$

4). Solve the rotation matrix R':

According to the Eq (7), when the value of the last item $2{r_{ci}}^T R' v_{ci}$ is larger, the value of *E* will be smaller, that is, the condition $max(R'H)$ needs to be met. *H* can be expressed by Eq (8).

$$H = \sum_i^n ({r_{ci}}^T v_{ci}) \tag{8}$$

By singular value decomposition of *H*, unitary matrix *U*, semi-positive definite diagonal matrix Λ and conjugate transposed unitary matrix V^T can be obtained.

$$H = U\Lambda V^T \tag{9}$$

The rotation matrix R' satisfying condition $max(R'H)$ can be expressed by Eq (10).

$$R' = VU^T \tag{10}$$

5). Solve the translation matrix T':

$$T' = \bar{r} - R'\bar{v}. \tag{11}$$

After calculating with SVD algorithm, the rotation matrix R' and translation matrix T' satisfying the least square Eq (2) can be obtained. Because there is no iterative optimization, the accuracy of the results cannot be guaranteed. Therefore, it is necessary to make accuracy judgment and iterative optimization on the results calculated by SVD algorithm, so as to obtain the optimal solution. The iteration conditions are set as iteration times N_{max} and iteration accuracy ε.

6). The rotation matrix R' and translation matrix T' obtained by SVD algorithm are substituted into Eq (1), and the calculation result is set as a new virtual coordinate point set $\{v_i\}$.

7). The maximum distance between the real coordinate point set $\{r_i\}$ and the new virtual coordinate point set $\{v_i\}$ is solved, as shown in Eq (12).

$$d = max\|v_i - r_i\|^2 \tag{12}$$

8). If the calculation results do not meet the iterative conditions, the SVD algorithm needs to be used for re-calculation, and the optimal solution will be obtained according to the iterative process. If the calculation results meet the iterative conditions, the calculated rotation matrix R' and translation matrix T' are the optimal solutions.

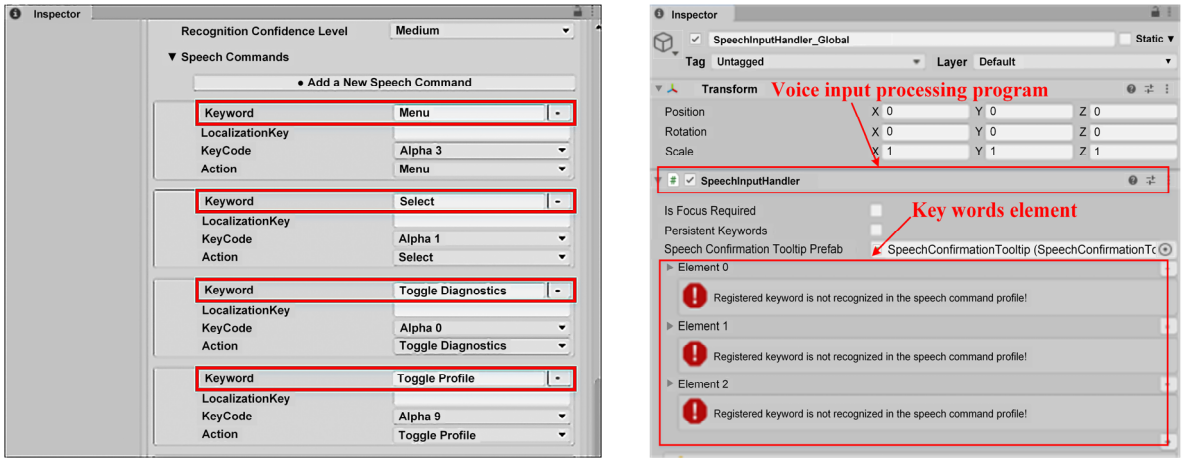
2.3. Voice control, optimization and denoising

2.3.1. Control

Prostate seed implantation surgery usually includes multiple implantation paths and multiple implantation targets. In order to observe more intuitively and reduce interference information, the displayed augmented reality information is set as a single implantation target in a single implantation path. Doctors can switch the display information of the implantation path and the implantation target in real time through voice control instructions.

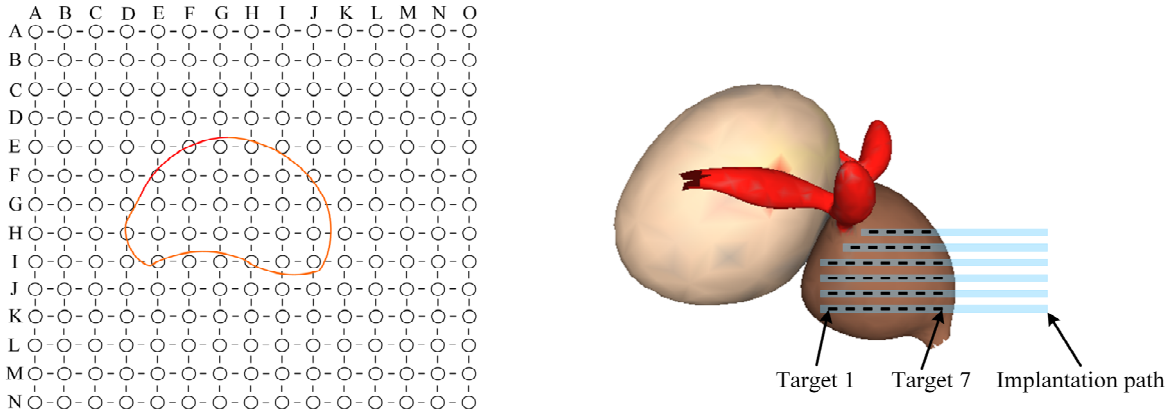
Microsoft HoloLens 2 is equipped with a microphone to receive voice control instructions, and the MRTK toolkit provides the software development kit (SDK) for voice control. The "MRTK Project Configurator" is set, the microphone function is enabled, and then voice control instructions are added to the keyword list in the inspector window, as shown in Figure 10(a).

Seed implantation path and seed implantation target components are added to the 3D virtual model. Blank items are created in the hierarchical structure window, and the voice input handler component and keyword elements are added to the blank items, as shown in Figure 10(b). By issuing the voice control instruction, the specific implantation path and target information can be displayed or hidden in the three-dimensional virtual model of the prostate. Voice control instructions contain specific meanings for doctors, not simple sequential voice control instructions.



(a) (b)

Figure 10. (a) Keyword list of voice control instruction and (b) input voice processing program according to keyword elements.



(a) (b)

Figure 11. (a) Seed implantation guide plate and (b) schematic diagram of seed implantation navigation.

The voice control instruction contains the information of seed implantation path and seed implantation target. In the operation of prostate seed implantation, the doctor will use the guide plate to assist the needle insertion. By numbering the pinholes distributed in the guide plate, the two-dimensional spatial positions of the prostate are described, as shown in Figure 11(a). The first instruction of voice control instruction represents the horizontal coordinate of prostate, and the second instruction represents the vertical coordinate of prostate. The depth information is added to describe the distribution position of seed targets. Targets distributed on the seed implantation path can be referred to by numbers 1, 2, and 3, as shown in Figure 11(b). The voice control instructions of the seed implantation can be set by instructions such as AA2, BC7, etc.

2.3.2. Optimization of voice control instruction based on 3D registration algorithm

The robot motion control instructions planned before operation can only guarantee the position

relationship between the implantation path and the implantation target, but cannot guarantee the position relationship between the patient and the robot. Therefore, the set positional relationship needs a series of conversion operations to make the actual motion of the robot under voice control meet the intraoperative requirements.

As shown in Figure 12, the doctor can preliminarily set the distribution information of the implantation path and the implantation target in the robot base coordinate system before operation, and the coordinate relationship between the path and the target can be represented by the virtual world coordinate $\{v_i\}$. During the operation, the doctor controls the robot in teaching mode so that the end of the implantation needle touches the markers attached to the abdominal surface of the patient. According to the information feedback result, the computer can obtain the coordinates of real markers in the robot base coordinate system, and use the real-world coordinates $\{r_i\}$ to represent them.

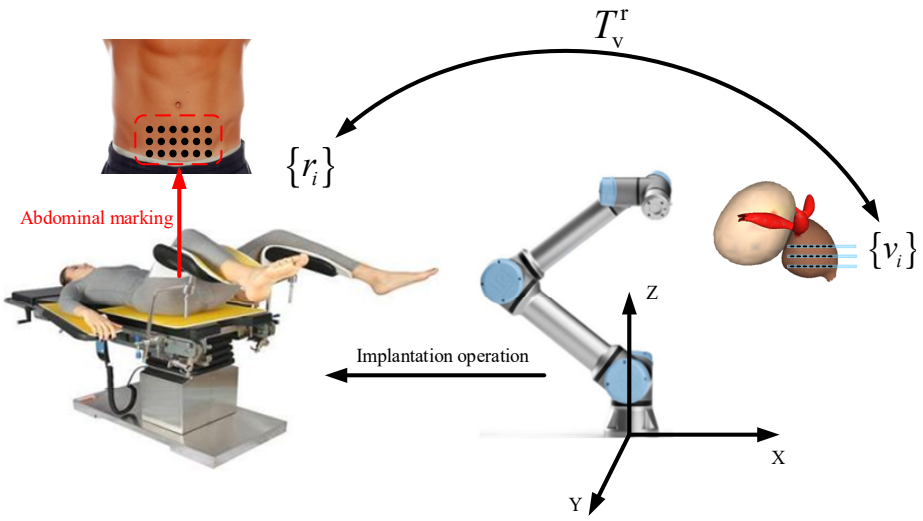


Figure 12. Schematic diagram of real-time optimization based on three-dimensional registration algorithm.

The transformation matrix T_v^r from $\{r_i\}$ to $\{v_i\}$ is solved based on $\{v_i\}$. The motion control parameters of the robot are updated by this matrix, which can ensure that the information from preoperative planning to actual control during surgery is consistent. The process of solving the transformation matrix is actually a three-dimensional registration process.

During the operation, the patient's breathing and slight trembling will reduce the accuracy of the operation. These influencing factors are microscopic changes, and mainly reflected in the change of the positional relationship between the patient and the robot in the translation direction. The augmented reality visualization device will update the transformation matrix T_v^r between the virtual world coordinate system and the real-world coordinate system in real time through sensors. The translation parameters in the transformation matrix T_v^r can be extracted in real time during the operation, and the positional relationship between the patient and the robot in the translation direction and the motion control parameters can be continuously corrected.

The motion accuracy of UR robot can reach 0.1 mm, while the registration accuracy of 3D registration algorithm is lower than 3 mm. Therefore, through the 3D registration technology, the corrected theoretical accuracy error is within an acceptable range, and the non-real-time problems caused by voice control can be initially solved.

2.3.3. Voice control instruction denoising

In the voice control system of prostate seed implantation, the collected voice signals are complex and contain various noises. Spectral subtraction is often used to eliminate broadband noise, which cannot meet actual needs. If the BP neural network is used to denoise complex time-varying noisy voice, the denoise result cannot guarantee convergence or fall into local optimum due to the imperfection of the training model. Based on the above problems and the actual demand of voice control system, a fusion algorithm based on spectral subtraction and BP neural network is proposed.

After the signal preprocessing and spectral subtraction denoising process are completed, the time domain signal is a noisy signal because it contains the music noise generated by spectral subtraction denoising. The BP neural network is used for secondary processing to remove music noise. For periodic noise in the system, the low-pass filter is used for hardware denoising in anti-aliasing filtering during preprocessing. In the subsequent denoising process, the periodic noise is further removed by spectral subtraction and the BP neural network, to obtain the final denoised signal. The denoising process is as follows.

1). The preprocessed voice signal is processed by three operations: Noise estimation, amplitude quadratic solution, phase angle calculation, and retention. The average energy estimation of the leading voiceless segment (noisy segment) of the voice signal is shown in Eq (13).

$$E(k) = \frac{1}{fnT}\sum_{i=1}^{fnT}|y_i(k)|^2, \tag{13}$$

where $E(k)$ is the average energy, fnT is the number of frames corresponding to the preamble voiceless segment, $y_i(k)$ is the preprocessed voice signal, and $|y_i(k)|^2$ is the amplitude quadratic solution of the voice signal. The phase angle of voice signal is calculated as shown in Eq (14).

$$y_i^a(k) = arctan\left[\frac{Imy_i(k)}{Rey_i(k)}\right], \tag{14}$$

where $y_i^a(k)$ is the phase angle of the voice signal. $Imy_i(k)$ is the imaginary part of complex number of $y_i(k)$, and $Rey_i(k)$ is the real part of complex number of $y_i(k)$.

2). As shown in Eq (15), the average energy estimated value $E(k)$ and the amplitude quadratic solution $|y_i(k)|^2$ are processed by spectral subtraction, so as to obtain the amplitude quadratic solution of the denoised voice signal.

$$\left|y_i^{'}(k)\right|^2 = \begin{cases} |y_i(k)|^2 - a \times E(k) & |y_i(k)|^2 \geq a \times E(k) \\ |y_i(k)|^2 - b \times E(k) & |y_i(k)|^2 < a \times E(k) \end{cases}, \tag{15}$$

where $\left|y_i^{'}(k)\right|^2$ represents the amplitude quadratic value of the denoised signal. Over-reduction factor a and gain compensation factor b can be used to reduce the amplitude of broadband peak, and remove

part of broadband noise. The setting range of over-reduction factor a should satisfy $a \geq 1$, and the setting range of gain compensation factor b should satisfy $0 < b \leq 1$.

3). Finally, combining the result of Eq (16) with the inverse fast Fourier transform (IFFT) technique, the voice signal in time domain after spectral subtraction denoising is obtained.

$$y_i'(k) = \left|y_i'(k)\right| exp(j \times y_i^a(k)) \tag{16}$$

In denoising, the error between the denoised voice and the learning model is calculated. The optimal voice signal denoising is completed by combining the iterative adjustment of weight parameters and offset values. Based on the characteristics and requirements of voice denoising, the structure of the BP neural network model is shown in Figure 13.

The BP neural network model structure shown in Figure 13 includes input layer, hidden layer and output layer. The hidden layer consists of hidden layer neurons, external offset value, summation processing and training function. There are ten neurons in the hidden layer from *H1* to *H10*, and the external offset value b is 1 by default. The trainglm function is used as the training function f. The maximum training number is 600, the error accuracy is 0.001, the momentum parameter is 0.8, and the learning rate is 0.002.

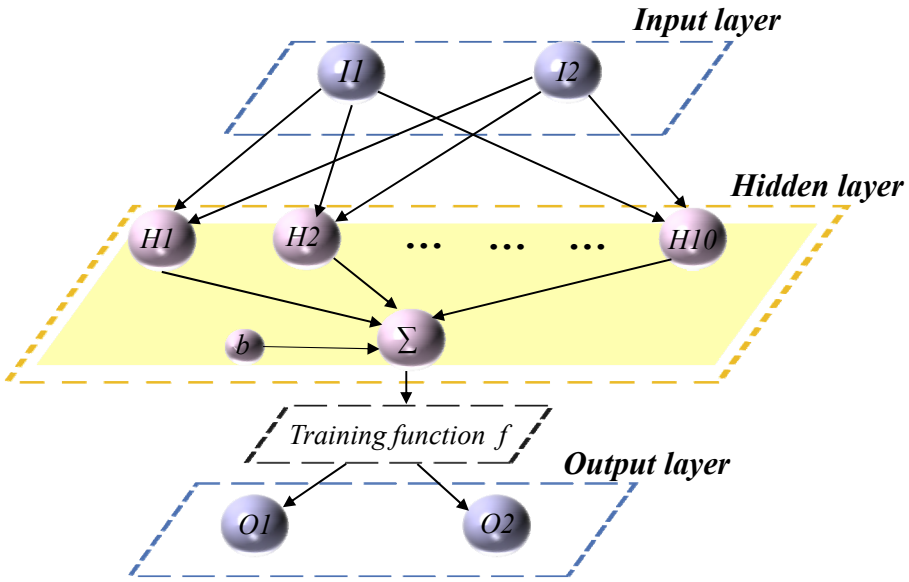


Figure 13. Model structure of BP neural network.

According to the above BP neural network model structure and parameter settings, the voice signal denoising program is written. The noisy voice signal is input into the BP neural network algorithm program. According to the preset parameters of the program and the learned model, the voice signal is denoised to obtain the denoised digital signal in the time domain.

Structure of the BP neural network model is shown in Figure 13.

3. Results

3.1. Three-dimensional registration experiment

3.1.1. Accuracy experiment of three-dimensional registration algorithm

In the augmented reality system, the spatial coordinate error of the point cloud data is within $30 \times 30 \times 30$ mm^3, considering the displacement of marker points caused by a patients' breathing. A three-dimensional point cloud set P containing fifty data points is randomly generated, and a homogeneous matrix T for spatial transformation is set. Combining P and T, the new three-dimensional point cloud set P_i is obtained. The data of three-dimensional point cloud set P will fluctuate randomly within $30 \times 30 \times 30$ mm^3, and the SVD algorithm and improved algorithm are used to solve the three-dimensional registration matrices T_1 and T_2 of P and P_i, respectively. Using three-dimensional registration matrices T_1, T_2, and P_i, new three-dimensional point cloud sets Q_1 and Q_2 are obtained, and the Euclidean distances between Q_1 and P and between Q_2 and P are calculated. The experimental results are shown in Figure 14(a)–(c).

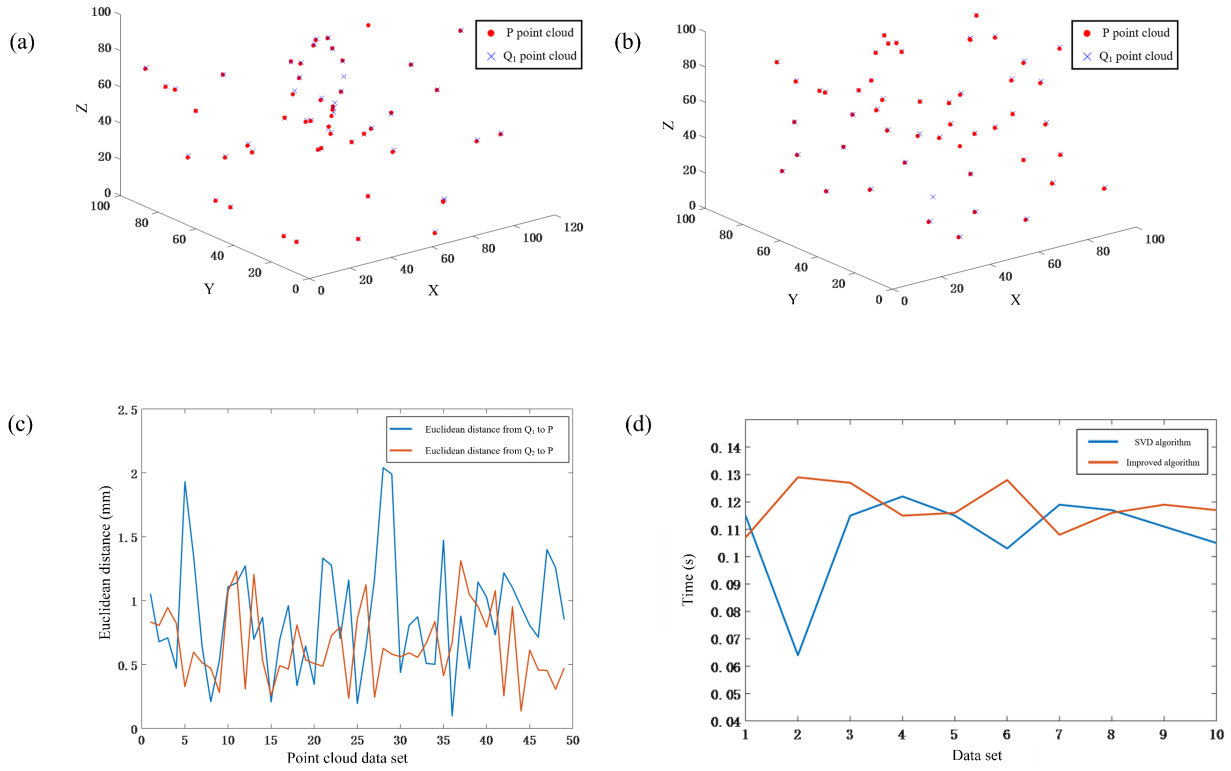


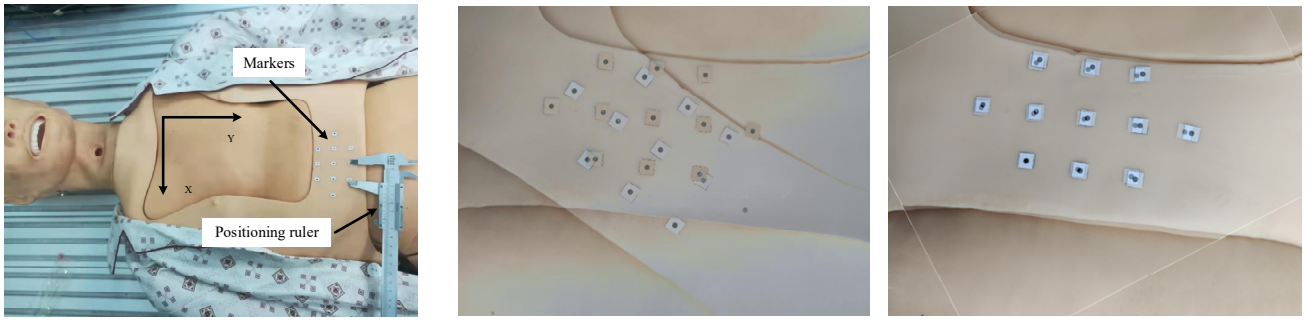
Figure 14. (a) Three-dimensional registration effect of SVD algorithm based on non-theoretical data, (b) three-dimensional registration effect of improved algorithm based on non-theoretical data, (c) three-dimensional registration algorithm accuracy based on non-theoretical data, and (d) delay of three-dimensional registration algorithm: Blue line represents 3D registration delay time based on SVD algorithm, and orange line represents 3D registration delay time based on improved algorithm.

3.1.2. Real-time experiment of three-dimensional registration algorithm

After improving the 3D registration algorithm, the 3D registration process has changed, so the 3D registration delay time is different under different algorithms. The real-time experiments of SVD algorithm and improved algorithm are carried out, and ten groups of time data under each algorithm are collected. Among them, the iterative condition of the improved algorithm is set as the minimum value to ensure its convergence, and the numerical value is accurate to one decimal place. Therefore, the set iteration condition is that the maximum distance between the corresponding points is less than 0.5 mm. The experimental results are shown in Figure 14(d).

3.1.3. Repetitive experiment of three-dimensional registration algorithm

As shown in Figure 15, the markers are pasted on the surface of the mannequin, and then three groups of three-dimensional registration experiments are conducted for the SVD algorithm and improved algorithm. In the experiment, the z-axis data in the depth direction are all replaced by zero.



(a) (b)

Figure 15. (a) The markers are pasted on the surface of the mannequin, and (b) comparison diagram before and after three-dimensional registration based on improved algorithm.

Table 1. Average error of three-dimensional registration experiment.

Experiment number	Average error					
	Six markers		Nine markers		Eleven markers	
	SVD algorithm	Improved algorithm	SVD algorithm	Improved algorithm	SVD algorithm	Improved algorithm
1	3.85	3.41	3.55	3.11	3.12	2.64
2	3.71	3.35	3.41	3.05	3.20	2.73
3	3.87	3.40	3.37	2.96	3.31	2.81
4	3.97	3.42	3.31	3.23	3.54	2.98
5	3.90	3.38	3.47	3.15	3.25	2.76

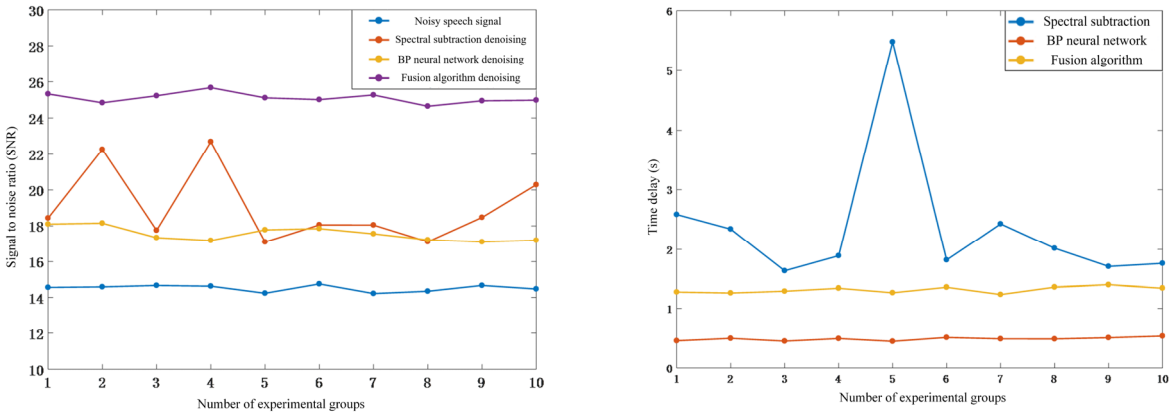
The position data of the markers need to be divided into ideal data and actual data. The ideal data are equivalent to the position of the three-dimensional virtual model in the virtual world coordinate system. The actual data are determined by obtaining the pixel information of the markers in the image. There will be certain errors between the actual data and the ideal data due to the actual pasting position, shooting distortion and coordinate information extraction method. To reduce the influence of errors on

the experimental results, aiming at the same group of ideal data, five groups of actual data were collected for repeated experiments. Then, the SVD algorithm and improved algorithm are used to calculate the average error of each group of actual data. There are three groups of ideal data corresponding to six, nine and eleven markers. The experimental results are shown in Table 1.

3.2. Voice experiments

3.2.1. Voice signal-to-noise ratio experiment

The pure voice signal was recorded in a quiet environment. Ten segments of voice signals were recorded as noisy voice signals in the actual environment of robot movement. The noisy voice signals are denoised by spectral subtraction, BP neural network and fusion algorithm, so as to obtain ten groups of denoised voice signals. The signal-to-noise ratio of noisy voice signals and denoised voice signals is shown in Figure 16(a).



(a) (b)

Figure 16. (a) Voice signal-to-noise ratio, (b) quantitative analysis of voice denoising delay using different denoising algorithms.

3.2.2. Denoising delay experiment

The delay experiment of the recorded noisy voice signal is carried out. Quantitative analysis of the voice denoising delay using different denoising algorithms is made, as shown in Figure 16(b). Figure 16(b) shows that the denoising delay of spectral subtraction is the lowest, and that of BP neural network is the highest. The average delay of spectral subtraction is 0.814 s, that of the BP neural network is 2.368 s, and that of the fusion algorithm is 1.314 s.

3.2.3. Voice denoising waveform diagram experiment

The voice signal waveform diagram experiment was carried out to observe the distribution and distortion of the waveform. Figure 17 shows the waveform comparison diagrams. Figure 18 shows the voice denoising waveform diagram based on fusion algorithm.

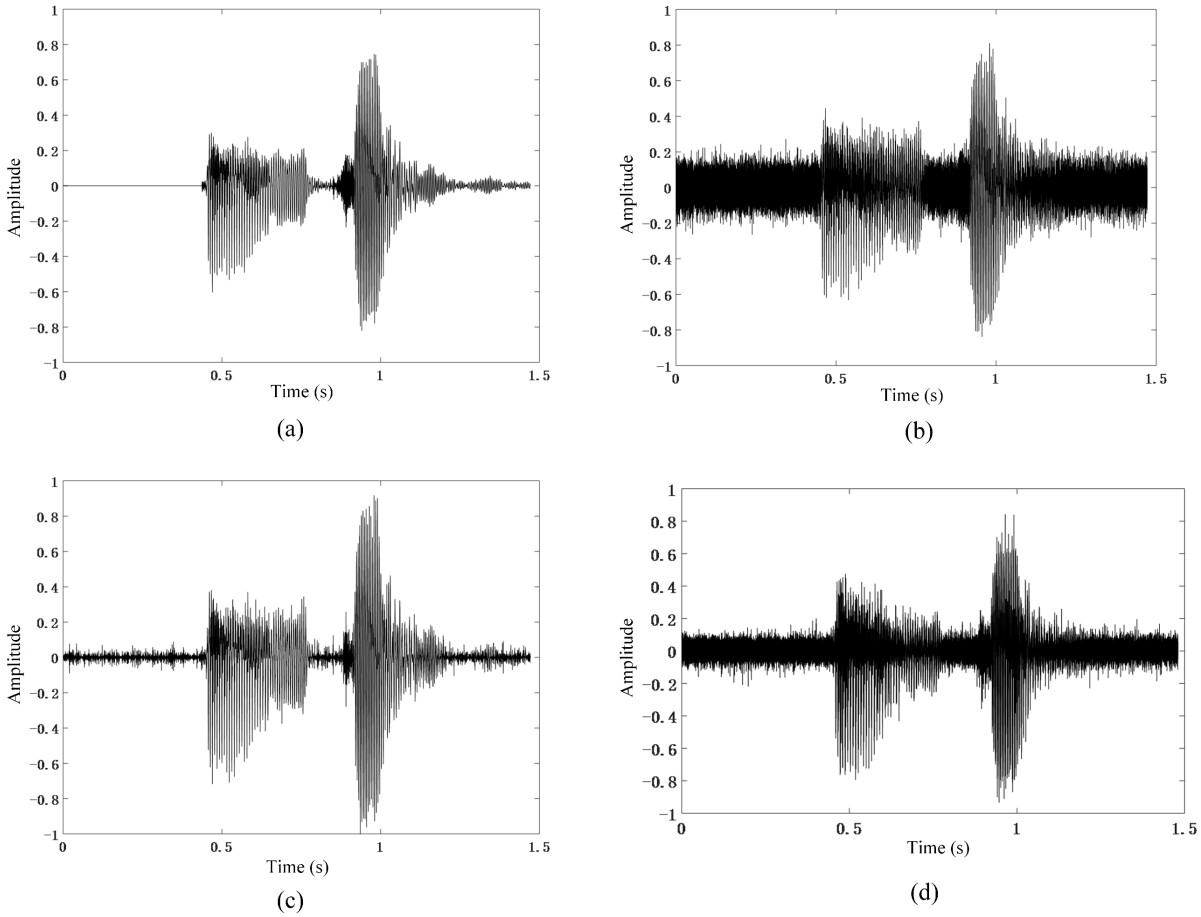


Figure 17. Waveform comparison diagrams: (a) pure voice waveform, (b) noisy voice waveform, (c) voice waveform denoised by spectral subtraction, and (d) voice waveform denoised by BP neural network.

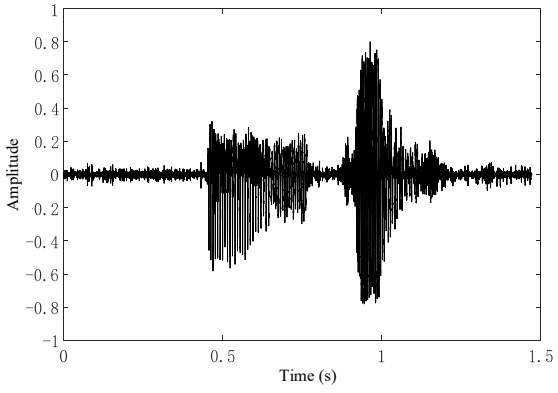
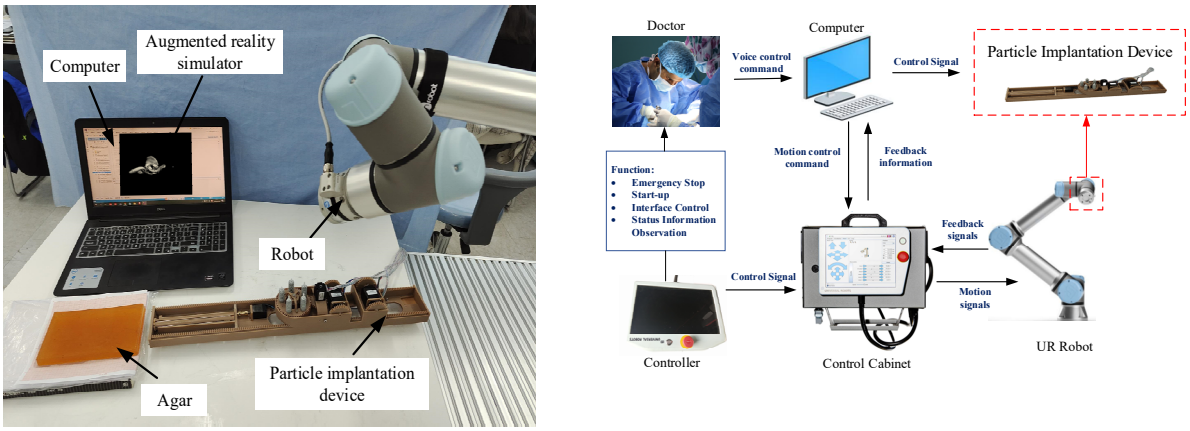


Figure 18. Voice denoising waveform diagram based on fusion algorithm.

3.2.4. Voice recognition experiment

To evaluate the fusion algorithm in voice denoising more comprehensively, the voice recognition

experiment was carried out based on the experimental platform. The experimental platform is shown in Figure 19. The computer is connected to the robot control cabinet. The seed implantation device is installed at the end of the robot, and keeps communication with the computer through the motion control card. The agar is used to imitate human tissues. According to the preset control scheme, five groups of voice control instructions are selected, and the robot motion control program corresponding to the instructions is written. In the experiment, the robot voice control is carried out under the conditions of original voice, spectral subtraction denoising, BP neural network denoising and fusion algorithm denoising. Each group of voice control instructions should be subjected to 40 groups of experiments under different voice processing conditions, and then the voice recognition degree under different processing conditions can be obtained by calculating the proportion that the robot can complete the corresponding actions according to the voice control instructions, as shown in Table 2.



(a) (b)

Figure 19. (a) Voice recognition experiment platform, and (b) voice control system of prostate seed implantation robot.

Table 2. Experimental results of voice recognition.

Voice control instruction	Original voice	BP neural network denoising	Spectral subtraction denoising	Fusion algorithm denoising
AA3	80%	82.5%	85%	95%
DC4	77.5%	80%	87.5%	92.5%
BE2	82.5%	82.5%	85%	95%
Up	82.5%	85%	90%	92.5%
Stop	75%	75%	77.5%	87.5%

3.2.5. Response speed experiment of integrated system

The augmented reality application saved on the computer is opened by the HoloLens 2 simulator to complete the response speed experiment of the integrated system. Five different voice control instructions were sent in turn. The time from voice control instruction to Hololens 2 simulator and UR5e robot is recorded, and five pairs of time data are obtained, as shown in Table 3.

Table 3. Response time of two systems under different control instructions.

Voice control instruction	1	2	3	4	5
Augmented reality system (s)	0.253	0.261	0.249	0.257	0.264
Robot control system (s)	1.47	1.53	1.44	1.49	1.51

4. Discussion

To verify the effectiveness of the improved 3D registration algorithm, accuracy experiments, real-time experiments, and repetitive experiments are carried out. During the simulation process, due to the error accumulation effect of the algorithm, the calculation result of one point cloud data among the 50 registered point cloud data has a big deviation. Figure 14(a),(b) shows that the deviation does not affect the overall three-dimensional registration effect. Therefore, the data can be removed as a singular value in the subsequent Euclidean distance calculation. After removing singular values, the results in Figure 14(c) shows that the overall accuracy of the improved algorithm is better than that of the SVD algorithm, and the data fluctuation of the improved algorithm is more stable than that of the SVD algorithm. Figure 14(d) shows that the real-time performance of the SVD algorithm is lower than that of the improved algorithm when processing actual data. Therefore, in practical applications, the real-time performance of the improved algorithm in the system is better than that of the SVD algorithm. The data in Table 1 show that the average error of the improved algorithm is smaller than that of the SVD algorithm, as shown in Figure 20. Compared with the accuracy experiment, the average error has increased, which is mainly due to the introduction of errors. However, the improved algorithm is better than the SVD algorithm in 3D registration.

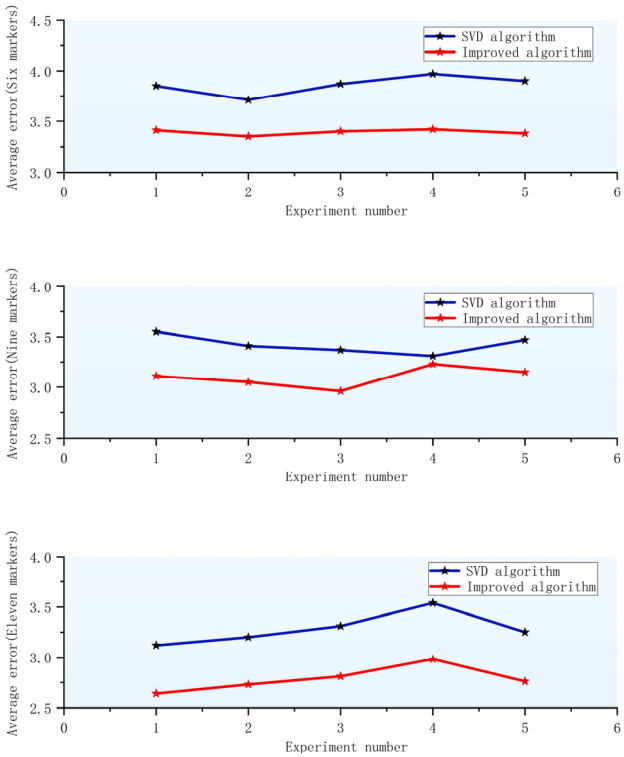
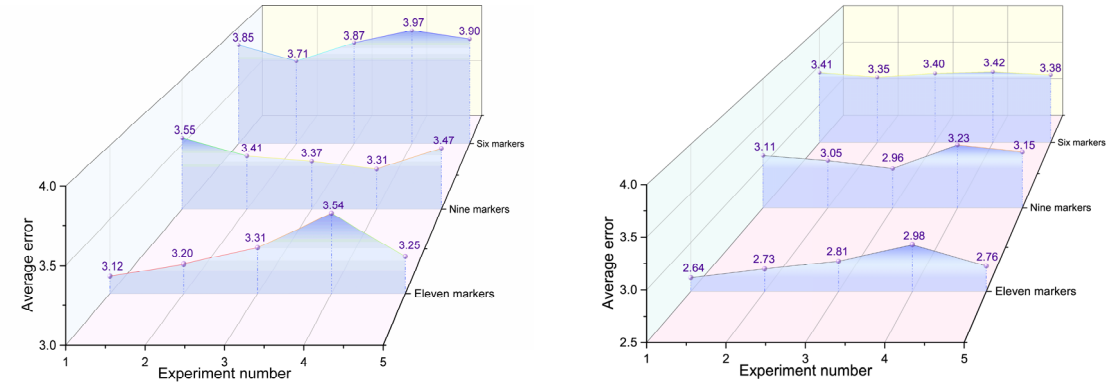
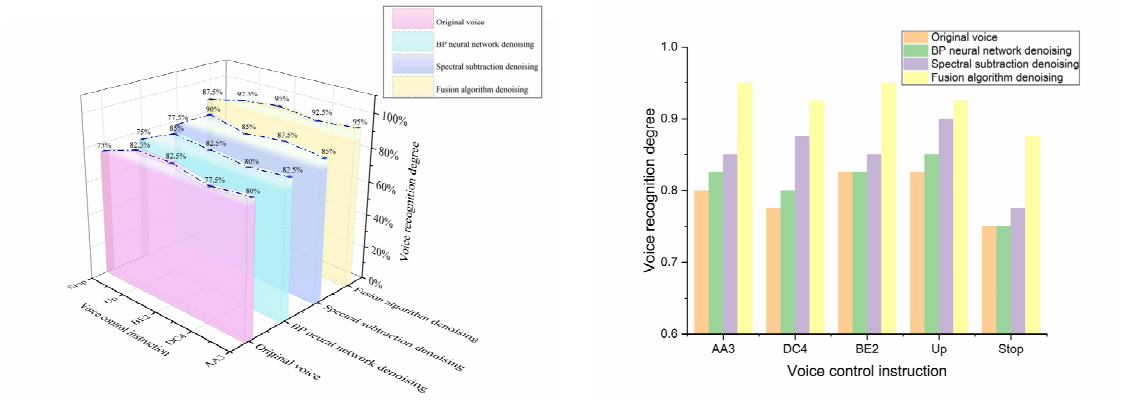


Figure 20. Comparison of average error of 3D registration based on SVD algorithm and improved algorithm.



(a) (b)

Figure 21. (a) Average error of 3D registration of different marker points under SVD algorithm, and (b) average error of 3D registration of different marker points under improved algorithm.



(a) (b)

Figure 22. (a) Comparison of voice recognition degree of voice control instructions based on different algorithms, and (b) comparison of voice recognition degree of algorithms based on different voice control instructions.

As shown in Figure 21, the more labeled points are pasted, the smaller the average error of data is. The comparison diagram before and after three-dimensional registration based on improved algorithm for 11 markers is shown in Figure 15(b). Experimental results show that the improved algorithm can meet the requirements of augmented reality system for visualization accuracy. Therefore, considering the above factors and clinical needs comprehensively, the improved algorithm can meet the visualization accuracy requirements of augmented reality system for prostate seed implantation. Three-dimensional registration technology is used to optimize the non-real-time problems in the voice control system to ensure the accuracy of prostate seed implantation. The signal-to-noise ratio curves shown in Figure 16(a) indicate that the denoising effect of spectral subtraction fluctuates greatly. The denoising effect of the BP neural network and fusion algorithm is stable. Considering the difference in the signal-to-noise ratio between the data, the denoising effect of the fusion algorithm is approximately twice that of the spectral subtraction and three times that of the BP neural network. Figure 16(b) shows that the average delay of spectral subtraction is 0.814 s, that of the BP neural network is 2.368 s, and

that of the fusion algorithm is 1.314 s. To ensure the denoising effect of the BP neural network, it is necessary to set more iterations and higher learning accuracy; thus, it takes more time to learn and approximate the model. The fusion algorithm combines the advantages of spectral subtraction and BP neural network. The fusion algorithm is less affected by the number of iterations and learning accuracy, so the delay of the fusion algorithm is lower than that of the BP neural network. However, the denoising process of fusion algorithm is more complicated than that of spectral subtraction, so the delay of the fusion algorithm is greater than that of spectral subtraction. By comparing the voice waveforms before and after denoising in Figures 17 and 18, it can be seen that the denoising effect of spectral subtraction is better than that of the BP neural network. However, compared with the pure voice waveform, there is music noise, and the signal distortion is large. Compared with fusion algorithm, the music noise generated by the fusion algorithm is smoother, the noise amplitude is reduced, and the distortion of voice signals is smaller. By analyzing the experimental results of the voice recognition degree shown in Table 2, we can know that the voice recognition degree of the system is the highest under the voice denoising of the fusion algorithm, as shown in Figure 22. Therefore, the experiment also proves that the use of the fusion algorithm optimizes and enhances the voice denoising function of the integrated system. By analyzing the data in Table 3, the response time of the augmented reality system is approximately 0.25 s, and that of robot control system is approximately 1.5 s. We mainly consider the response time of the robot system. Therefore, the overall response time of the integrated system is approximately 1.5 s. The experimental results show that the fusion algorithm can effectively improve the reliability of the voice control system and ensure the control effect during the process of system operation. The response speed of the system can meet the requirements of prostate seed implantation surgery.

5. Conclusions

In this paper, we propose a prostate seed implantation robot system based on AR and voice control. The median filtering algorithm is used to denoise prostate MR images, the local threshold segmentation method is used to segment the features of medical images, and the surface rendering technology is used to reconstruct the three-dimensional virtual model of the prostate and its surrounding organs and tissues. Combined with holographic application, the augmented reality system and scene display of prostate seed implantation are built. According to the three-dimensional registration process, an improved SVD algorithm based on ICP is proposed, and the effectiveness of the improved algorithm is verified by accuracy experiment, real-time experiment and repetitive experiment. The voice control system is built for the prostate seed implantation robot. The non-real-time problem in voice control is analyzed, and the control process is optimized in combination with three-dimensional registration technology. A fusion algorithm based on spectral subtraction and BP neural network is proposed for voice denoising. The reliability of the fusion algorithm is verified through voice signal-to-noise ratio experiment, denoising delay experiment, voice denoising waveform diagram experiment, and voice recognition experiment. Finally, the response speed experiment of the integrated system shows that the overall response time of the system is approximately 1.5 s. The experiments prove that the responsiveness of the voice control system of the prostate seed implantation robot in the augmented reality environment can meet the surgical requirements.

According to the needs of surgery, augmented reality and voice control technology are used to provide a more friendly visualization system and control system for prostate seed implantation surgery.

The system not only ensures the robot motion control based on voice commands, but also provides the corresponding augmented reality visualization information for doctors. The combination of augmented reality and voice control provides more convenient and natural non-contact control for the surgical process, and greatly improves the intelligence, visualization and information of the surgical process. The limitation of this paper is that the system and method have only carried out relevant experiments and tests, and have not really been used in clinic. The next step is to further improve the system functions and conduct clinical trials. In the future, we will continue to conduct in-depth research on the following aspects.

1). The reconstruction method of the 3D virtual model should be further optimized to improve the reconstruction accuracy. Appropriate image processing methods are adopted for human soft tissues to solve the problem of unclear feature boundary transition and improve the accuracy of feature extraction.

2). The real-time registration speed of the 3D registration algorithm should be further optimized. The rigid registration process of the 3D registration algorithm should be improved to solve the problem of soft tissue deformation and improve the accuracy and real-time performance of the 3D registration algorithm.

3). More voice commands and visualization functions will be developed, so that the prostate seed implantation robot system can have better human-computer interaction ability, and is more intelligent and more multimodal.

Use of AI tools declaration

The authors declare they have not used Artificial Intelligence (AI) tools in the creation of this article.

Author contributions

Xinran Zhang: Formal analysis, Software, Validation, Writing–original draft, Writing–review & editing; Yongde Zhang: Conceptualization, Method-ology; Jianzhi Yang: Formal analysis, Software; Haiyan Du: Writing – review & editing.

Acknowledgments

This research was funded by the National Natural Science Foundation of China, grant number 52275015, the Reserve Leader Funding Project of Leading Talent Echelon of Heilongjiang Province of China, grant number 2501050628, and the Science and Technology Innovation Team Project of Foshan City of China, grant number 2018IT100302.

Conflict of interest

The authors declare that there are no conflicts of interest.

References

1. S. Lim, C. Jun, D. Chang, D. Petrisor, M. Han, D. Stoianovici, Robotic transrectal ultrasound guided prostate biopsy, *IEEE Trans. Biomed. Eng.*, **66** (2019), 2527–2537. https://doi.org/10.1109/TBME.2019.2891240
2. M. R. Tangel, A. R. Rastinehad, Advances in prostate cancer imaging, *F1000Research*, **7** (2018), 1337. https://doi.org/10.12688/f1000research.14498.1
3. F. J. Siepel, B. Maris, M. K. Welleweerd, V. Groenhuis, P. Fiorini, S. Stramigioli, Needle and biopsy robots: A review, *Curr. Rob. Rep.*, **2** (2021), 73–84. https://doi.org/10.1007/s43154-020-00042-1
4. J. Michael, D. Morton, D. Batchelar, M. Hilts, J. Crook, A. Fenster, Development of a 3D ultrasound guidance system for permanent breast seed implantation, *Med. Phys*., **45** (2018), 3481–3495. https://doi.org/10.1002/mp.12990
5. Y. Chen, Q. Wang, H. Chen, X. Song, H. Tang, M. Tian, An overview of augmented reality technology, *J. Phys.: Conf. Ser.*, **1237** (2019), 022082. https://doi.org/10.1088/1742-6596/1237/2/022082
6. Z. Makhataeva, H. A. Varol, Augmented reality for robotics: A review, *Robotics*, **9** (2020), 21. https://doi.org/10.3390/robotics9020021
7. L. Qian, J. Y. Wu, S. P. DiMaio, N. Navab, P. Kazanzides, A review of augmented reality in robotic-assisted surgery, *IEEE Trans. Med. Rob. Bionics*, **2** (2019), 1–16. https://doi.org/10.1109/TMRB.2019.2957061
8. H. Younes, J. Troccaz, S. Voros, Machine learning and registration for automatic seed localization in 3D US images for prostate brachytherapy, *Med. Phys*., **48** (2021), 1144–1156. https://doi.org/10.1002/mp.14628
9. C. Rossa, J. Carriere, M. Khadem, R. Sloboda, N. Usmani, M. Tavakoli, An ultrasound-guided mechatronics-assisted system for semi-automated seed implantation and tracking in prostate brachytherapy, *Brain and Cognit. Intell. Control Rob.*, (2022), 21–46. https://doi.org/10.1201/9781003050315
10. Y. Zhang, Z. Lu, C. Wang, C. Liu, Y. Wang, Voice control dual arm robot based on ROS system, in *2018 IEEE International Conference on Intelligence and Safety for Robotics (ISR)*, IEEE, (2018), 232–237. https://doi.org/10.1109/IISR.2018.8535942
11. B. Li, L. Yuan, C. Wang, Y. Guo, Structural design and analysis of pneumatic prostate seed implantation robot applied in magnetic resonance imaging environment, *Int. J. Med. Rob. Comput. Assisted Surg.*, **18** (2020), e2457. https://doi.org/10.1002/rcs.2457
12. G. Fichtinger, J. P. Fiene, C. W. Kennedy, G. Kronreif , I. Iordachita, D. Y. Song, et al., Robotic assistance for ultrasound-guided prostate brachytherapy, *Med. Image Anal*., **12** (2008), 535–545. https://doi.org/10.1007/978-3-540-75757-3_15
13. M. Djohossou, A. B. Halima, A. Valérie, J. Bert, D. Visvikis, Design and kinematics of a comanipulated robot dedicated to prostate brachytherapy, *Robotica*, **39** (2021), 468–482. https://doi.org/10.1017/S026357472000051X
14. A. B. Halima, J. Bert, J. F. Clément, D. Visvikis, Development of a 6 degrees of freedom prostate brachytherapy robot with integrated gravity compensation system, in *2021 International Symposium on Medical Robotics (ISMR)*, IEEE, (2021), 1–7. https://doi.org/10.1109/ISMR48346.2021.9661571

15. B. Wang, Y. Liang, D. Xu, Y. Zhang, Y. Xu, Design of a seed implantation robot with counterbalance and soft tissue stabilization mechanism for prostate cancer brachytherapy, *Int. J. Adv. Rob. Syst*., **18** (2021), 17298814211040687. https://doi.org/10.1177/17298814211040687
16. S. Chen, B. Gonenc, M. Li, D. Y. Song, E. C. Burdette, I. Iordachita, et al., Needle release mechanism enabling multiple insertions with an ultrasound-guided prostate brachytherapy robot, in *39th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC)*, IEEE, (2017), 4339–4342. https://doi.org/10.1109/EMBC.2017.8037816
17. S. Jiang, Y. Yang, Z. Yang, Z. Zhang, S. Liu, Design and experiments of ultrasound image-guided multi-dof robot system for brachytherapy, *Trans. Tianjin Univ.*, **23** (2017), 479–487. https://doi.org/10.1007/s12209-017-0067-9
18. X. Dai, Y. Zhang, J. Jiang, B. Li, S. Zuo, Design of a transrectal ultrasonic guided prostate low dose rate brachytherapy robot, *Mech. Sci*., **13** (2022), 399–409. https://doi.org/10.5194/ms-13-399-2022
19. M. Bakouri, M. Alsehaimi, H. F. Ismail, K. Alshareef, A. Ganoun, A. Alqahtani, et al., Steering a robotic wheelchair based on voice recognition system using convolutional neural networks, *Electronics*, **11** (2022), 168. https://doi.org/10.3390/electronics11010168
20. P. Tran, S. Jeong, F. Lyu, K. Herrin, S. Bhatia, D. Elliott, et al., FLEXotendon Glove-III: Voice-controlled soft robotic hand exoskeleton with novel fabrication method and admittance grasping control, *IEEE/ASME Trans. Mechatron.*, **27** (2022), 3920–3931. https://doi.org/10.1109/TMECH.2022.3148032
21. E. Watanabe, M. Satoh, T. Konno, M. Hirai, T. Yamaguchi, The trans-visible navigator: A see-through neuronavigation system using augmented reality, *World Neurosurg*., **87** (2016), 399–405. https://doi.org/10.1016/j.wneu.2015.11.084
22. D. Cohen, E. Mayer, D. Chen, A. Anstee, J. Vale, G. Z. Yang, et al., Augmented reality image guidance in minimally invasive prostatectomy, in *Prostate Cancer Imaging*, *Computer-Aided Diagnosis*, *Prognosis, and Intervention*, Springer, (2010), 101–110. https://doi.org/10.1007/978-3-642-15989-3_12
23. T. Yamamoto, N. Abolhassani, S. Jung, A. M. Okamura, T. N. Judkins, Augmented reality and haptic interfaces for robot-assisted surgery, *Int. J. Med. Rob. Comput.*, **8** (2012), 45–56. https://doi.org/10.1002/rcs.421
24. T. Song, C. Yang, O. Dianat, E. Azimi, Endodontic guided treatment using augmented reality on a head-mounted display system, *Healthcare Technol. Lett.*, **5** (2018), 201–207. https://doi.org/10.1049/htl.2018.5062
25. F. Gîrbacia, R. Boboc, B. Gherman, T. Gîrbacia, D. Pîsla, Planning of needle insertion for robotic-assisted prostate biopsy in augmented reality using RGB-D camera, in *26th International Conference on Robotics in Alpe-Adria Danube Region (RAAD)*, Springer, (2017), 515–522. https://doi.org/10.1007/978-3-319-49058-8_56
26. D. Lee, H. W. Yu, S. Kim, J. Yoon, K. Lee, Y. J. Chai, et al., Vision-based tracking system for augmented reality to localize recurrent laryngeal nerve during robotic thyroid surgery, *Sci. Rep.*, **10** (2020), 8437. https://doi.org/10.1038/s41598-020-65439-6
27. B. Xu, Z. Yang, S. Jiang, Z. Zhou, B. Jiang, S. Yin, Design and validation of a spinal surgical navigation system based on spatial augmented reality, *Spine*, **45** (2020), E1627–E1633. https://doi.org/10.1097/BRS.0000000000003666

28. G. Samei, K. Tsang, C. Kesch, J. Lobo, S. Hor, O. Mohareri, et al., A partial augmented reality system with live ultrasound and registered preoperative MRI for guiding robot-assisted radical prostatectomy, *Med. Image Anal.*, **60** (2019), 101588. https://doi.org/10.1016/j.media.2019.101588
29. R. Schiavina, L. Bianchi, S. Lodi, L. Cercenelli, F. Chessa, B. Bortolani, et al., Real-time augmented reality three-dimensional guided robotic radical prostatectomy: Preliminary experience and evaluation of the impact on surgical planning, *Eur. Urol. Focus*, **7** (2020), 1260–1267. https://doi.org/10.1016/j.euf.2020.08.004
30. H. Reichenspurner, R. J. Damiano, M. Mack, D. H. Boehm, H. Gulbins, C. Detter, et al., Use of the voice-controlled and computer-assisted surgical system ZEUS for endoscopic coronary artery bypass grafting, *J. Thorac. Cardiovasc. Surv.*, **118** (1999), 11–16. https://doi.org/10.1016/S0022-5223(99)70134-0
31. K. Zinchenko, C. Y. Wu, K. T. Song, A study on speech recognition control for a surgical robot, *IEEE Trans. Ind. Inf.*, **13** (2016), 607–615. https://doi.org/10.1109/TII.2016.2625818
32. K. Gundogdu, S. Bayrakdar, I. Yucedag, Developing and modeling of voice control system for prosthetic robot arm in medical systems, *J. King Saud Univ. Comput. Inf. Sci.*, **30** (2018), 198–205. https://doi.org/10.1016/j.jksuci.2017.04.005
33. M. F. Ruzaij, S. Neubert, N. Stoll, K. Thurow, Hybrid voice controller for intelligent wheelchair and rehabilitation robot using voice recognition and embedded technologies, *J. Adv. Comput. Intell.*, **20** (2016), 615–622. https://doi.org/10.20965/jaciii.2016.p0615
34. S. K. Pramanik, Z. A. Onik, N. Anam, M. M. Ullah, A. Saiful, S. Sultana, A voice controlled robot for continuous patient assistance, in *2016 International Conference on Medical Engineering, Health Informatics and Technology (MediTec)*, IEEE, (2016), 1–4. https://doi.org/10.1109/MEDITEC.2016.7835366
35. R. Matarneh, S. Maksymova, O. Zeleniy, V. Lyashenko, Voice control for flexible medicine robot, *Int. J. Comput. Trends Technol.*, **55** (2018), 1–5. https://doi.org/10.14445/22312803/IJCTT-V56P101
36. T. S. Newman, H. Yi, A survey of the marching cubes algorithm, *Comput. Graphics*, **30** (2006), 854–879. https://doi.org/10.1016/j.cag.2006.07.021
37. R. T. Azuma, A survey of augmented reality, *Presence Teleoperators Virtual Environ.*, **6** (1997), 355–385. https://doi.org/10.1162/pres.1997.6.4.355

AIMS Press