



# Augmented reality-based autostereoscopic surgical visualization system for telesurgery

Tianqi Huang<sup>1</sup> · Ruiyang Li<sup>1</sup> · Yangxi Li<sup>1</sup> · Xinran Zhang<sup>1</sup> · Hongen Liao<sup>1</sup> 

Received: 11 January 2021 / Accepted: 15 July 2021 / Published online: 7 August 2021  
© CARS 2021

## Abstract

**Purpose** The visualization of remote surgical scenes is the key to realizing the remote operation of surgical robots. However, current non-endoscopic surgical robot systems lack an effective visualization tool to offer sufficient surgical scene information and depth perception.

**Methods** We propose a novel autostereoscopic surgical visualization system integrating 3D intraoperative scene reconstruction, autostereoscopic 3D display, and augmented reality-based image fusion. The preoperative organ structure and the intraoperative surface point cloud are obtained from medical imaging and the RGB-D camera, respectively, and aligned by an automatic marker-free intraoperative registration algorithm. After registration, preoperative meshes with precalculated illumination and intraoperative textured point cloud are blended in real time. Finally, the fused image is shown on a 3D autostereoscopic display device to achieve depth perception.

**Results** A prototype of the autostereoscopic surgical visualization system was built. The system had a horizontal image resolution of 1.31 mm, a vertical image resolution of 0.82 mm, an average rendering rate of 33.1 FPS, an average registration rate of 20.5 FPS, and average registration errors of approximately 3 mm. A telesurgical robot prototype based on 3D autostereoscopic display was built. The quantitative evaluation experiments showed that our system achieved similar operational accuracy ( $1.79 \pm 0.87$  mm) as the conventional system ( $1.95 \pm 0.71$  mm), while having advantages in terms of completion time (with 34.11% reduction) and path length (with 35.87% reduction). Post-experimental questionnaires indicated that the system was user-friendly for novices and experts.

**Conclusion** We propose a 3D surgical visualization system with augmented instruction and depth perception for telesurgery. The qualitative and quantitative evaluation results illustrate the accuracy and efficiency of the proposed system. Therefore, it shows great prospects in robotic surgery and telesurgery.

**Keywords** 3D interactive visualization · Augmented reality · Autostereoscopic display · Telesurgery · Percutaneous procedures

## Introduction

Telesurgery is an emerging surgical technique that allows surgeons to observe the remote surgical scene in real time from the surgical visualization system and to perform medical diagnosis and treatment without direct physical contact

with patients. This technology not only alleviates the current shortage of surgeons but also reduces complications and financial burden and eliminates risky long-distance travel [1]. The current research about telesurgery mainly focuses on applying robot-assisted endoscopic surgery [2]. The da Vinci Surgical System is the most widely used robot system for endoscopic telesurgery [3]. The visualization system of the da Vinci robot comprises a stereoscopic laparoscope to capture the surgical scene inside the patient and a binocular display device to provide surgeons with the 3D surgical scene.

In addition to endoscopic surgery, non-endoscopic percutaneous procedures play an important role in diagnosing and treating cancer, including tissue biopsy, brachytherapy,

---

Tianqi Huang and Ruiyang Li have contributed equally to this work.

---

✉ Hongen Liao  
liao@tsinghua.edu.cn

<sup>1</sup> Department of Biomedical Engineering, School of Medicine, Tsinghua University, Beijing, China

and radiofrequency ablation [4]. Researchers have recently begun to adopt surgical robots in needle-guided procedures [5]. However, a lack of usable visualization tools for teleoperated percutaneous interventions hinders its adoption in telesurgery.

To improve the efficiency and safety of telesurgery, a novel 3D visualization system, which can provide depth perception and augmented reality fusion, is indispensable. The current non-endoscopic telesurgery study utilizes a flat panel monitor to display three cross sections of intraoperative medical imaging and intraoperative scenes acquired from a monocular camera [6, 7]. This kind of display has insufficient depth cues, which can easily lead to misjudgment of spatial location by the surgeons. To provide depth perception in telesurgery, researchers use head-mounted binocular display [8] and glasses-based binocular display [9] to show 3D surgical scenes. The problem was the discomfort associated with wearing 3D glasses, as well as visual fatigue and dizziness [10, 11]. Compared with binocular display, autostereoscopic 3D display has the advantages of no additional equipment, support for multi-person observation, and lower visual fatigue [12, 13], which has preliminary applications in augmented reality surgical navigation [14, 15]. However, the autostereoscopic 3D display-based telesurgical visualization system has not been developed yet due to the lack of a suitable multi-viewpoint video capture system.

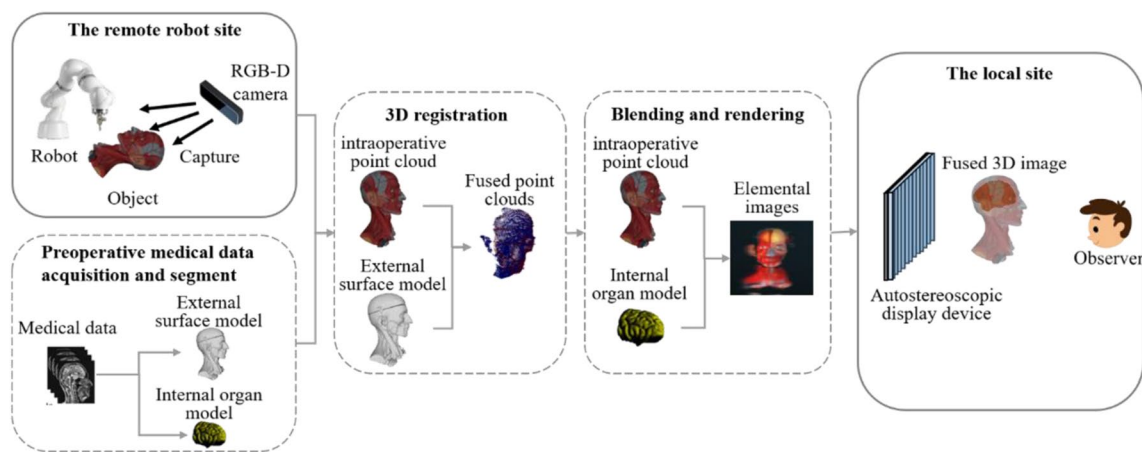
In this paper, we propose an autostereoscopic surgical visualization system with a 3D real-time fusion of the intraoperative scene and preoperative medical images for telesurgery. In the proposed system, we utilize a depth camera to capture intraoperative scenes, and an autostereoscopic display device to provide augmented surgical scenes with depth perception. To achieve real-time performance, we propose an algorithmic framework, including automatic marker-free intraoperative 3D registration and real-time augmented reality rendering method for 3D autostereoscopic display.

This system can collect intraoperative images in real-time and offer a 3D-fused image of the preoperative and intraoperative data on the autostereoscopic display device, which provides depth perception and supports multi-person observation. In this research, real-time performance is a crucial evaluation criterion and will be evaluated with phantom experiments. The proposed system can provide a real-time preoperative medical image fusion and 3D depth perception from multiple views, simultaneously displaying the intraoperative image information, the preoperative lesion information, and the preoperative plan information to help doctors perform complicated surgery more precisely and effectively. We believe that this technology has application prospects in robotic surgery and telesurgery.

## Materials and methods

### System overview

We propose an autostereoscopic surgical visualization system that can display preoperative medical models and intraoperative point cloud data on the 3D autostereoscopic display, thus helping doctors obtain sufficient medical information in robotic surgery telesurgery. The system consists of an RGB-D camera, a computer, and a 3D autostereoscopic display device, as shown in Fig. 1. The preoperative medical data are obtained from CT (computerized axial tomography scan), MRI (magnetic resonance imaging), ultrasound, or other imaging techniques and registered in the same coordinate system. The data are segmented to obtain an external surface model for subsequent intraoperative registration and internal organ models for image fusion. In the surgery, we use an RGB-D camera as the primary source of intraoperative data. Textured point cloud data are acquired by the RGB-D camera in real time and automatically fused with the



**Fig. 1** Workflow of the autostereoscopic surgical visualization system

preoperative models. Therefore, the preoperative medical models and the intraoperative point cloud data are simultaneously displayed on the 3D autostereoscopic display. The algorithm framework includes the real-time acquisition of textured point clouds, automatic marker-free intraoperative 3D registration, real-time fusion, and rendering of meshes and point clouds for autostereoscopic display.

### Automatic marker-free intraoperative 3D registration

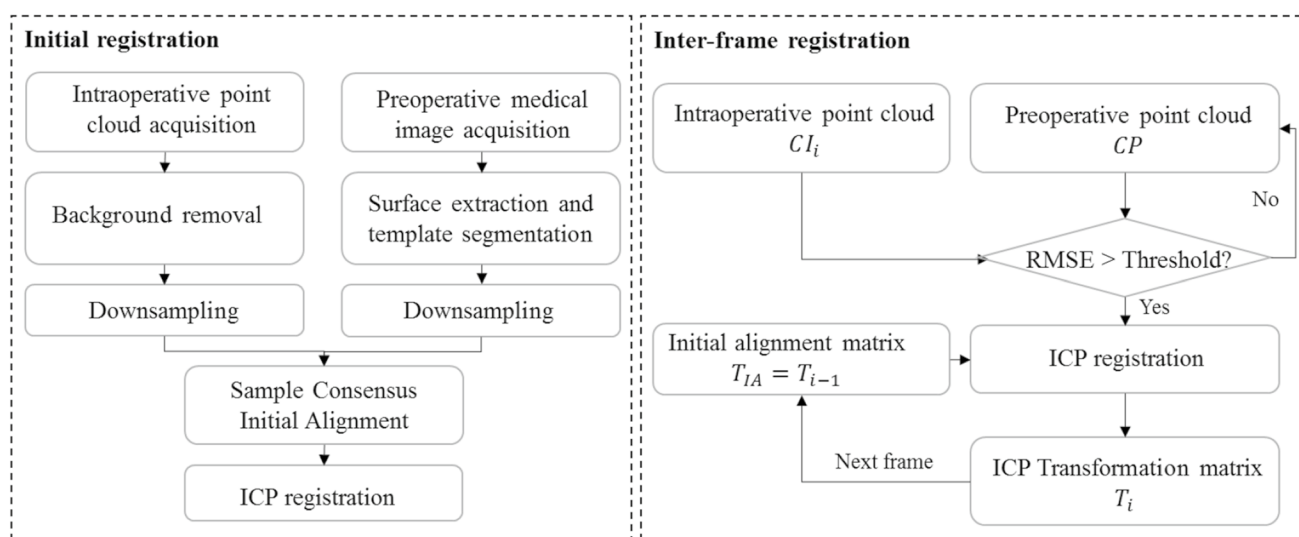
In order to fuse the preoperative and intraoperative data with high accuracy, we propose an automatic continuous marker-free 3D registration method, as shown in Fig. 2. Combining with the fusion and rendering method introduced in the above subsection, we can provide a real-time 3D scene of the surgical region, together with the anatomical structure inside the patient at the precise location.

The intraoperative registration workflow can be divided into two parts: the initial alignment and inter-frame registration. We first obtain preoperative medical data from MRI or CT in the initial alignment module, then segment it to obtain the external surface model for registration and the internal organ models for 3D display. After that, we remain the part of intraoperative and preoperative data in the ROI and downsample them to improve registration efficiency. To achieve initial alignment, we calculate the fast point feature histogram (FPFH) [16] of the point cloud to perform sample consensus initial alignment (SAC-IA) algorithm [17] as a coarse registration, then use iterative closest point (ICP) algorithm [18] for fine registration. Finally, we get the initial transformation matrix  $T_0$  from the preoperative coordinate system to the intraoperative coordinate system.

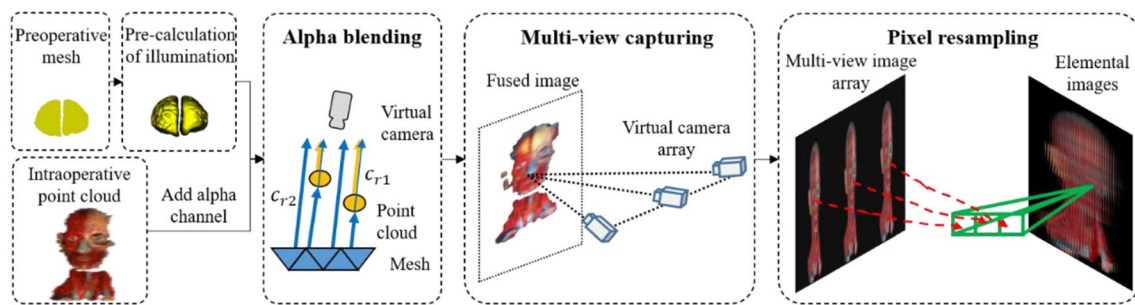
In the module of inter-frame registration, we regard the intraoperative point cloud ( $CI_i$ ) as the target and match the preoperative point cloud (CP) to it. The RMSE of the intraoperative point cloud and the preoperative point cloud is calculated in each frame, and the ICP algorithm is used for inter-frame alignment when the RMSE is greater than the threshold. By applying ICP for registration, we obtain the transform matrix  $T_i$  of the preoperative point cloud, which is continuously changed during 3D registration. To register the next frame, we use  $T_i$  as a matrix of initial alignment  $T_{IA}$  to calculate  $T_{i+1}$ .

### Real-time blending of meshes and point clouds for autostereoscopic display

We implement an efficient blending algorithm to simultaneously display intraoperative point cloud and preoperative model on an autostereoscopic display device. The rendering process is shown in Fig. 3. The first step is the data processing of the intraoperative point clouds and the preoperative meshes, which can be performed before the surgery. For the intraoperative point clouds, we set an additional alpha channel in the texture. We obtain the texture by precalculated illumination using the Phong reflection model [19] for the preoperative meshes and then update the model matrix according to the transformation matrix from the intraoperative 3D registration module. The second step is the blending part. We implement the alpha blending to fuse the intraoperative point clouds and the preoperative meshes. Then, we implement the high-performance shader and the lenticular display shader. These shaders are written in C for graphics (CG) language and implemented in GPU. The high-performance shader renders the intraoperative point clouds and



**Fig. 2** Process of the automatic marker-free intraoperative 3D registration



**Fig. 3** Real-time blending of meshes and point clouds for autostereoscopic display

the preoperative meshes by a virtual camera and calculates the rendered image after alpha blending. The lenticular display shader captures multiple view images and calculates the elemental images. With this method, the elemental images rendered can display a transparent fusion scene of the intraoperative point cloud data and the preoperative mesh model on the 3D autostereoscopic display.

In this algorithm, we use two methods to achieve the fusion of point clouds and meshes. The first one is to render point clouds with gaps. In the rendering procedure, the patient's outer surface is a point cloud, and the internal organs are meshes. The meshes can be seen from the gaps within the point cloud. The second way is to set the alpha value of the point cloud texture. In this way, the fusion effect is achieved by alpha blending. In computer graphics, the definition of alpha blending is the process of combining foreground and background images at a specific ratio to achieve partial or complete transparency [20]. During the rendering procedure, the transparency of the medical scene is influenced by alpha blending and the unfilled space of point clouds.

In the blending part, we achieve the fusion of the preoperative model and the intraoperative point cloud by rendering a semi-transparent point cloud. The alpha blending method is applied to achieve the transparency effect. For a virtual camera, the color of light can be calculated by the following formula:

$$\begin{cases} c_{r1} = c_{\text{point}} \times \alpha + c_{\text{mesh}} \times (1 - \alpha) \\ c_{r2} = c_{\text{mesh}} \times (1 - \alpha) \end{cases} \quad (1)$$

where  $c_{r1}$  represents the light color value through the point cloud,  $c_{r2}$  represents the light color value without passing through the point cloud, and  $\alpha$  represents the preset transparency. The transparency of the model is defined by the overall color intensity of the foreground with alpha blending and point cloud gaps. In the rendering of the point cloud, the diameter of a single point in the point cloud is generally set to be smaller than the point-to-point distance to avoid interfering with each other. At this point, the relationship

between the transparency of the displayed intraoperative point cloud ( $T_m$ ), the size of the single point, and the alpha value can be calculated by the following equation:

$$T_m = \alpha \times \frac{\pi r_p^2}{d_p^2} \quad (2)$$

where  $r_p$  is the radius size of a single point in the point cloud and  $d_p$  is the point-to-point distance. By the formula (2), the transparency of the displayed intraoperative point cloud is determined by the point cloud radius size  $r_p$ , and the preset transparency  $\alpha$ .

### Algorithm framework of the autostereoscopic surgical visualization system

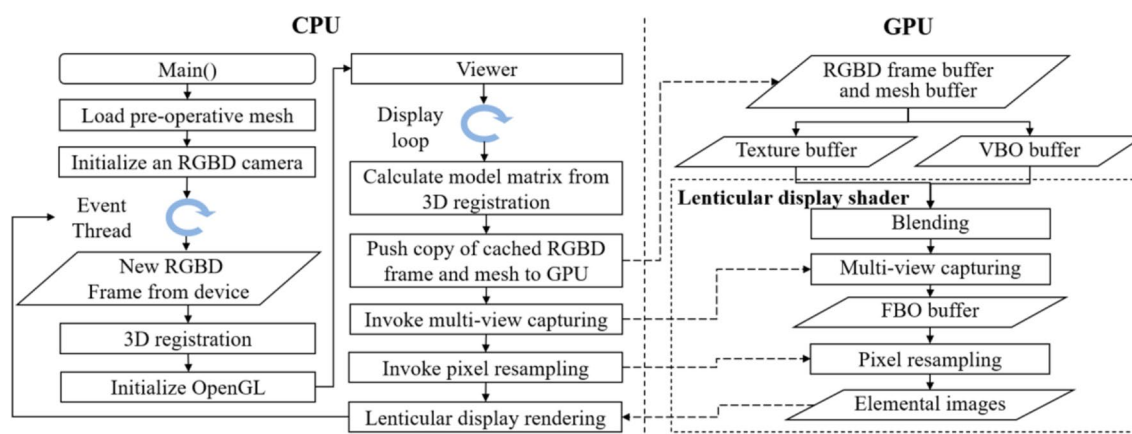
The processes of real-time 3D reconstruction, augmented reality-based image fusion, and 3D autostereoscopic display require a large amount of calculation. To improve the real-time performance of the autostereoscopic surgical visualization system, we build an OpenGL-based high-speed rendering framework and a personally configured lenticular display shader to integrate these processes and achieve a real-time workflow.

Figure 4 describes the process of the acquisition and the rendering on the CPU and GPU in detail. The CPU runs the main loop, including initialization of the RGB-D camera and OpenGL, a display loop, and an event loop for obtaining the RGB-D data. The GPU performed augmented reality fusion and image calculation of lenticular display rendering. The lenticular display rendering algorithm we used is based on the integral photography rendering algorithm. A more detailed description of the rendering algorithm can be found in previous papers [21, 22].

### Experiments and results

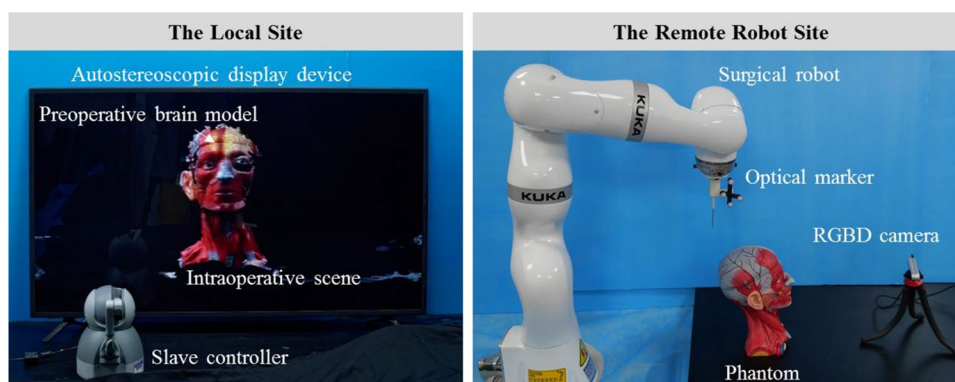
We performed experiments to verify the feasibility of the proposed method and evaluate the fusion accuracy and robustness of the prototype. A telesurgical robot prototype





**Fig. 4** The detailed program flow in CPU and GPU

**Fig. 5** The prototype of the autostereoscopic surgical visualization system



was built to validate the potential application in teleoperated percutaneous interventions.

### Prototype of the autostereoscopic surgical visualization system

We built a prototype of the autostereoscopic surgical visualization system. As shown in Fig. 5, we constructed a lenticular display device comprising a lenticular lens array (lens pitch 0.4438 mm, thickness 6.4 mm) and an LCD device (49.5 inches, 3840 × 2160 pixels). The LCD device is placed obliquely with a tilt angle of 9.5° to remove color moiré [23]. On the acquisition side, we used a tripod to fix the RGB-D camera (Intel, RealSense D435), which captured the textured point cloud from the human head model placed on the desktop. The camera capturing photos and videos in the experiment was the Canon EOS70D. The specifications of the computer hardware and the operating system are listed in Table 1.

**Table 1** Computer configuration used in the experiment

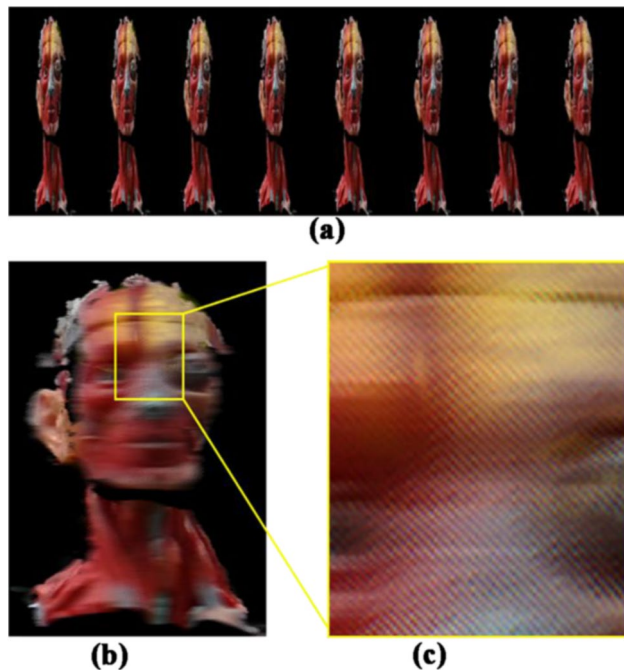
Environment parameter	Configuration
Computer model	Dell XPS 8930
CPU	Intel Core i7-8700K CPU @ 3.70 GHz
GPU	8 GB Nvidia GeForce GTX 1080
Memory	16 GB of RAM
Hard disk	512 GB solid-state drive
Operating system	Windows 10

### Point cloud fusion and 3D display

In the experiment, we used a human head model with internal brain structure models. We scanned the human head model as a preoperative image with CT. The slice thickness was 0.75 mm, and 417 axial slices with a matrix size of 512 × 512 pixels were obtained. We used the 3D Slicer software to segment the outer surface model and the frontal lobe model. The outer surface model was used for intraoperative registration with the point cloud data. We further extracted the face region of the outer surface model to reduce the

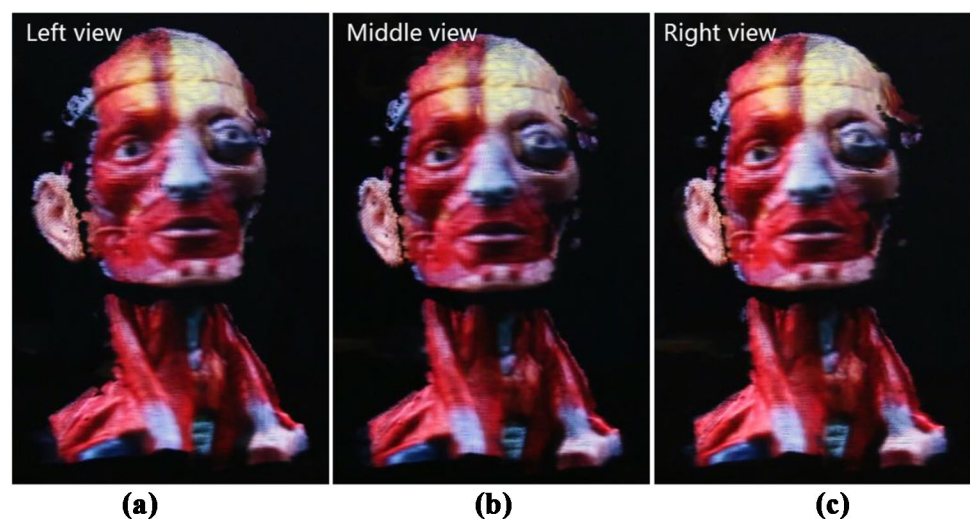
registration error. The frontal lobe model was used for fusion with the intraoperative point cloud data and displayed on the autostereoscopic display.

We rendered the point cloud data of the outer surface model using the rendering method described in the previous sections (Fig. 6). Figure 6a shows a multi-view image array captured by the virtual camera array in the 3D rendering algorithm. The elemental images and the enlarged ones from different viewpoints for autostereoscopic display are shown in Fig. 6b and Fig. 6c. By displaying the elemental



**Fig. 6** **a** The multi-view image array captured by the virtual camera array in the 3D rendering algorithm. **b** The elemental images are rendered by the 3D rendering algorithm. **c** The enlarged image shows sub-images from different viewpoints

**Fig. 7** The augmented reality fusion of the head model on the autostereoscopic display device from **a** left, **b** middle, and **c** right view



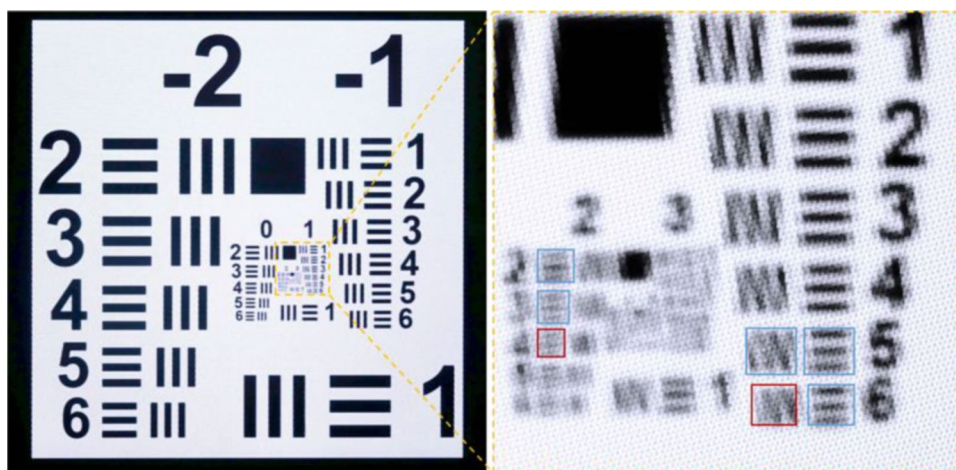
images on the autostereoscopic display device, the observer can obtain depth perception without special glasses. The fusion effects of the cloud point data and the frontal lobe model from different viewpoints are shown in Fig. 7 and Movie 1. The viewing angle is  $5.15^\circ$ . The limited viewing angle requires the observer to view at a distance of more than 2 m to avoid observing flipping images.

We used the 1951 USAF resolution test chart to evaluate image resolution of the built autostereoscopic display device. Figure 8 shows a captured image of the resolution test chart and an enlarged view of the center part. Rayleigh criterion is used to determine the minimum length of the resolved stripes. The details of the calculation can be found in [24]. The 1951 USAF resolution test chart was magnified 8.3 times on the display device. In Fig. 8, the red and blue rectangles, respectively, represented the unresolved and resolved stripes. The test elements 2–3 and 1–5 were the stripes with the minimum length, which can be resolved. According to the resolution stripe width table, the horizontal resolution of the 50-inch autostereoscopic display device was 1.31 mm, and the vertical resolution was 0.82 mm.

### Evaluation of real-time registration and time delay

We acquired the textured 3D point cloud of the head model in real time through an RGB-D camera. In the intraoperative registration process, we employed the SAC-IA and ICP algorithms in the Point Cloud Library (PCL). The PCL has optimized the performance of the ICP algorithm based on the octree data structure. We implemented the initial registration and inter-frame registration process with the C++ language and compiled them in release mode. Two parallel threads are set in our algorithm, including the inter-frame registration and the real-time rendering. The average frame rates of the inter-frame registration and the real-time rendering are 20.5 FPS and 33.1 FPS, respectively.

**Fig. 8** The resolution chart displayed on the autostereoscopic display device and the enlarged view of the center part. The red and blue rectangles, respectively, represented the unresolved and resolved stripes



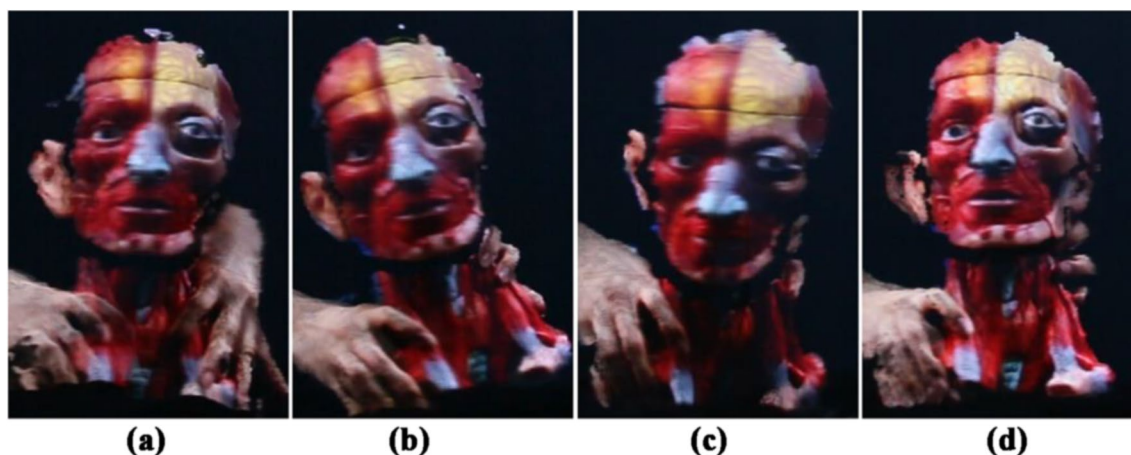
**Table 2** Latency evaluation with different settings

Frame rate of color and depth cameras/FPS	(30, 30)	(60, 60)	(60, 90)	(60, 30)
<b>Rendering thread</b>				
Camera capture/ms	153	72	58	111
Point cloud rendering/ms	5	6	5	10
Mesh rendering/ms	2	2	2	4
<b>Registration thread</b>				
Camera capture/ms	140	59	56	109
Registration/ms	20	22	19	35
Latency/ms	160	81	75	144

In addition to real-time performance, the time delay of the proposed system is also evaluated. Since there are two parallel threads, including registration and rendering, the total latency is the maximum delay time of the two threads. The tasks in each thread and the averaged delay time for each stage are shown in Table 2. The results illustrate that

the latency for camera capture has a primary proportion and can be shortened with a higher preset frame rate. However, there is a balance between the point cloud density and the camera's frame rate. For better visualization effect, the time delay of our setting is 144 ms, as the last column of Table 2 (resolutions of color and depth cameras for the last column are  $640 \times 480$  and  $1280 \times 720$ ; resolutions of color and depth cameras for the first three columns are  $640 \times 480$ ). With the developing performance of RGB-D cameras, the system can realize shorter latency.

We moved the human head model and evaluated registration accuracy and robustness during the motion. The experimental results are shown in Fig. 9 and Movie 2. We moved the head model in four motion states: keeping stationary, translating, rotating left and right, tilting forward and backward. The root-mean-square error (RMSE) was used to represent the accuracy of real-time registration. We repeated five trials for each motion state and recorded mean RMSEs and standard deviation values for each trial within 20 s,



**Fig. 9** The different movement of the head model on the autostereoscopic display device: **a** translation, **b** rotation, **c** tilt forward, and **d** tilt backward



as shown in Table 3. The average translation and rotation speeds were 16.11 mm/s and 3.46°/s during the trials. The results showed the repeatability of our system. The evaluated accuracy and robustness are acceptable for surgery.

### Surgical robot prototype for teleoperated percutaneous interventions

We built a surgical robot prototype for teleoperated percutaneous interventions based on the proposed 3D autostereoscopic surgical visualization system. The overall system was divided into two major parts: the local site and the remote robot site, as shown in Fig. 10a. The local site included an autostereoscopic display device and a robotic manipulator (Geomagic Touch), providing surgeons with a 3D fusion image of the preoperative data and the intraoperative data. The remote robot site comprised a surgical robot (KUKA LBR iiwa) to perform the surgical operation and an RGB-D camera to collect the surgical scene and transmit it to the local site. The puncture needle was mounted on the end of the robot with an optical marker attached. An optical

tracking system (NDI Polaris Vega) was used to track the posture of the puncture needle and the tip position. This allows the virtual needle to be displayed in a rendered surgical scene synchronized with the real needle and allows the surgeon to observe the needle tip when the real puncture needle enters the patient's body.

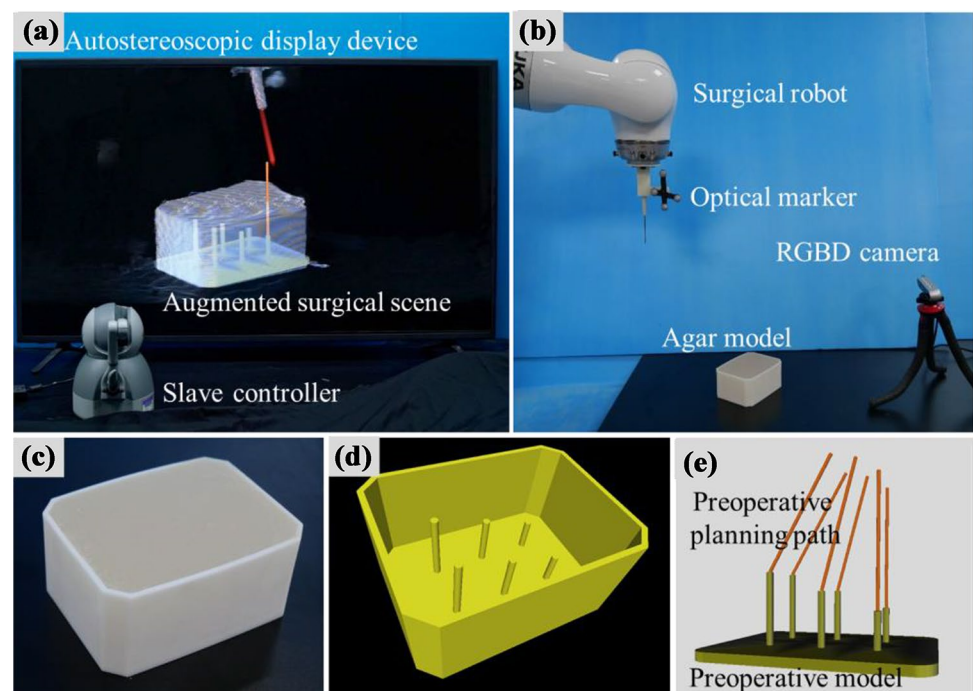
We fabricated agar phantoms (Fig. 10c) for quantitative evaluation experiments. Firstly, we fabricated a polymer box by 3D printing and added six cylinders with different heights inside with a radius of 2 mm and a distance of 1 cm, 2 cm, and 3 cm from the upper surface, respectively. The top of the cylinder was the target. We filled the box with agar solution (2.5% high gel strength agar, 1% flour, and 96.5% water.), dissolved it, and then cooled it in the refrigerator to get the agar phantoms. The agar phantoms were scanned by MRI and segmented to obtain preoperative models. In the navigation program, we set up planning paths for different target sites with different inclination angles of 0°, 15°, and 30°, respectively.

We designed a puncture navigation program for the autostereoscopic 3D display (Fig. 11a). The 3D autostereoscopic

**Table 3** The repetitive RMSEs (mm) of the intraoperative 3D registration in four motion states

Motion state	Trial No				
	1	2	3	4	5
Stationary	2.38 ± 0.17	2.38 ± 0.14	2.53 ± 0.13	2.53 ± 0.13	2.56 ± 0.23
Translating	3.12 ± 0.21	3.50 ± 0.20	3.13 ± 0.20	3.37 ± 0.35	3.27 ± 0.23
Rotating	4.01 ± 0.46	3.68 ± 0.43	3.97 ± 0.52	3.59 ± 0.40	3.30 ± 0.28
Tilting	3.48 ± 0.80	3.91 ± 0.81	3.12 ± 0.58	4.07 ± 0.66	3.00 ± 0.57

**Fig. 10** **a** The surgical robot prototype on the local site. **b** The surgical robot prototype on the remote robot site. **c** The agar phantom. **d** The 3D print file of the box with six columns for agar phantom. **e** The preoperative planning path with different insertion angles





display device showed the intraoperative acquired point cloud, the AR fused preoperative model, the preoperative planning path, and the virtual needle. In the navigation program, we added color cues to assist the surgeon with the puncture procedure. The initial virtual needle was red and turned blue when the virtual needle tip was less than 3 mm away from the planning path. When the virtual needle tip was less than 2 mm away from the target position, the virtual needle turned green. At this point, the needle was considered to have reached the surgical target in the program. Movie 3 showed the operator operating the master controller to allow the surgical robot to puncture to the target in the agar phantom.

### Quantitative evaluation experiment

In the quantitative evaluation experiment, we compare the proposed system with a traditional telesurgical puncture system. The traditional telesurgical puncture system uses the same hardware as the proposed system, except for a 32" 4k flat display device. The software was designed by modifying the surgical navigation software presented in the previous paper [25]. The flat panel display device provides three views of the preoperative model and the virtual needle as well as a live intraoperative image (captured by the color camera in the Realsense D435), as shown in Fig. 11b. In both systems, the registration of the preoperative and intraoperative data is achieved by the proposed marker-free intraoperative 3D registration method and the conditions for the virtual needle color change are the same.

Five users participated in the quantitative evaluation: three biomedical engineering Ph.D. students, one biomedical engineering postdoctoral fellow, and one orthopedic surgeon. Each user was guided by two telesurgery systems separately to perform the puncture operation. The puncture operation was performed with three cylinders of different heights as targets and was repeated five times for each target. The final needle tip position, needle tip trajectory, and operation time were recorded.

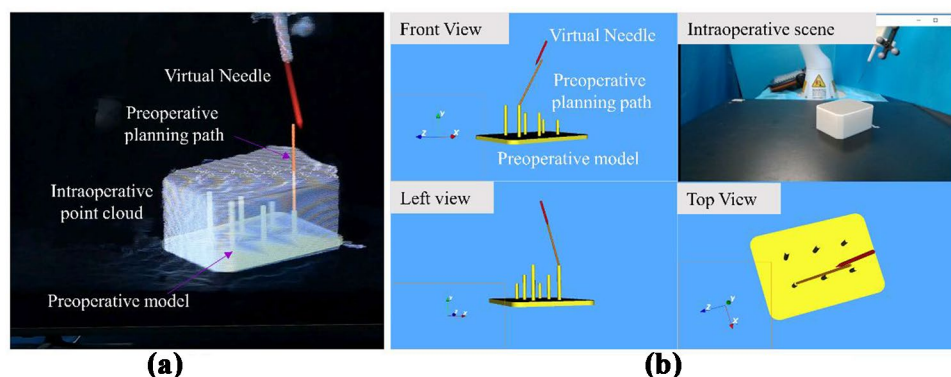
We evaluated the target positioning error (TPE), path length, and completion time for three targets in the two telesurgery systems. The target positioning error is defined by the distance between the needle tip position and the target position recorded in the coordinate system of the optical tracking system. The target position is obtained by the registration of the feature points. We used eight vertices on the agar box as feature points and collected the vertex positions with an optical marked probe by an optical tracking system, then registered them with the vertices of the design 3D model. After obtaining the registration matrix, we mapped the target coordinates of the design 3D model to the coordinate system of the optical tracking system with a registration error of about 0.5 mm.

We used a one-way ANOVA (Analysis of Variance) to examine the significant reduction of the target positioning error (TPE), the path length, and the completion time in the autostereoscopic telesurgery system and the traditional telesurgery system. Table 4 offers the detailed experimental results. The average TPE ( $1.79 \pm 0.87$  mm versus  $1.95 \pm 0.71$  mm), completion time ( $66.63 \pm 22.56$  s versus  $101.13 \pm 48.44$  s), and path length ( $297.03 \pm 122.26$  cm versus  $463.17 \pm 208.16$  cm) among all the participants and targets (totally 75 trials) for each system were calculated. The experimental results (Fig. 12) show significant differences ( $p < 0.05$ ) in completion time and path length between the two systems at most targets. The path length was shorter and the operation time was less in the proposed system by novices and the expert. But there is no significant difference in the target position errors. The experimental results showed that our system was able to achieve similar operational accuracy as the conventional system, while having advantages in terms of completion time (with 34.11% reduction) and path length (with 35.87% reduction).

### Post-experimental questionnaire

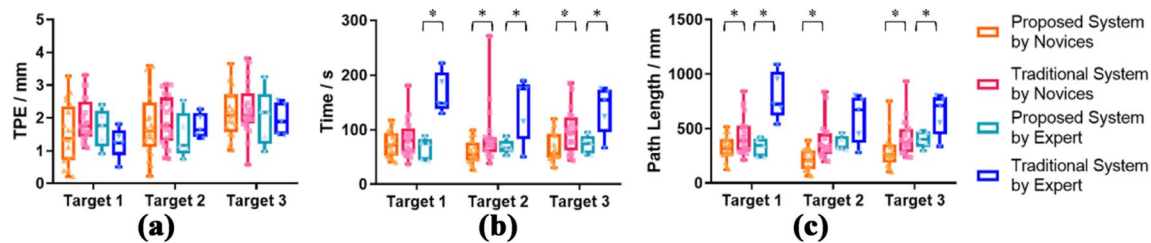
At the end of the experiment, the participants were asked to fill in a questionnaire inquiring about their experience of operating these two systems (Table 5). Each question

**Fig. 11** **a** The puncture navigation program for the autostereoscopic display device. **b** The traditional puncture navigation program for the flat panel display device



**Table 4** The target positioning error (TPE), path length, and completion time for each target in these two telesurgery systems

System	Autostereoscopic telesurgery system			Traditional telesurgery system		
Target	Target 1	Target 2	Target 3	Target 1	Target 2	Target 3
Novice						
TPE/mm	$1.49 \pm 0.89$	$1.79 \pm 0.97$	$2.16 \pm 0.72$	$1.92 \pm 0.64$	$1.91 \pm 0.73$	$2.25 \pm 0.78$
Completion time/s	$72.81 \pm 23.31$	$58.75 \pm 20.00$	$66.43 \pm 26.09$	$83.26 \pm 33.38$	$87.04 \pm 49.98$	$97.41 \pm 38.46$
Path length/cm	$323.94 \pm 104.00$	$219.71 \pm 92.39$	$296.55 \pm 149.47$	$439.37 \pm 173.45$	$439.57 \pm 173.45$	$412.62 \pm 163.86$
Expert						
TPE/mm	$1.69 \pm 0.52$	$1.45 \pm 0.63$	$2.00 \pm 0.78$	$1.25 \pm 0.43$	$1.75 \pm 0.33$	$1.96 \pm 0.44$
Completion time/s	$66.01 \pm 18.31$	$68.70 \pm 11.52$	$72.76 \pm 15.25$	$166.76 \pm 33.64$	$141.19 \pm 52.03$	$138.18 \pm 39.79$
Path length/cm	$329.31 \pm 77.06$	$362.46 \pm 59.58$	$402.91 \pm 67.79$	$801.11 \pm 195.52$	$596.02 \pm 197.64$	$635.32 \pm 171.81$

**Fig. 12** Boxplots indicate the target positioning error (TPE), the needle motion length, and the completion time for each target. Asterisks indicate statistical significance. The results show significant differ-

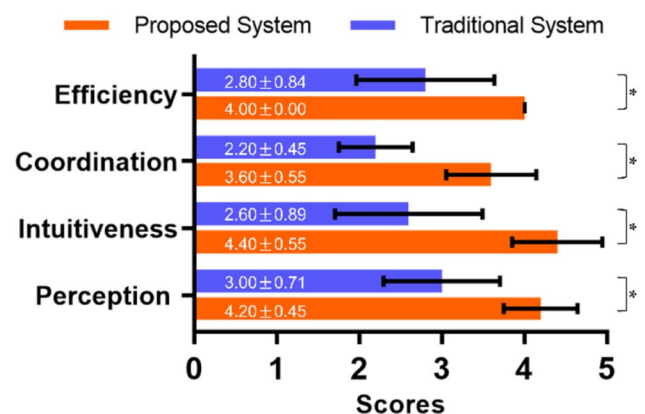
ences ( $p < 0.05$ ) in completion time and path length between the two systems at most targets

**Table 5** Questionnaire for assessing the usability of the system

Question	Evaluated Feature
1. The telesurgery system provides depth perception of the surgical scene	Depth perception
2. It is intuitive to judge the relative positions of the patient, the lesion and the virtual needle are intuitive through the telesurgery system	Intuitiveness
3. It is easy to complete operations on the robot controller based on navigation information	Hand–eye coordination
4. The telesurgery system helps improve the efficiency of remote surgery	Efficiency

was scored on a 5-point Likert scale, where 1 stands for “strongly disagree” and 5 for “strongly agree.”

Figure 13 shows the scores of the questionnaire. The proposed 3D autostereoscopic telesurgical system received higher scores on depth perception, intuitiveness, hand–eye coordination, and efficiency (with significant differences for all features by one-way ANOVA,  $p < 0.05$ ). After user experiments, both the novice group and the experts agreed that the proposed system was easier to use compared with the traditional telesurgical system based on 2D display. The expert indicated that although the 3D autostereoscopic telesurgical system is easier to use, the advantages are not obvious for percutaneous puncture procedures. In past clinical practice, physicians have become accustomed to using limited monocular depth cues on a 2D display for

**Fig. 13** Boxplots showing the scores of the post-experimental questionnaire. Asterisks indicate statistical significance. The results show significant differences ( $p < 0.05$ ) in all the features between the two systems

surgical operations, and it is not easy to promote 3D display-based systems in clinical practice.

## Discussion

A real-time 3D surgical visualization system based on the autostereoscopic display is proposed. In this work, we make two significant contributions: 1. Proposing a 3D surgical visualization system including real-time 3D reconstruction, augmented reality-based image fusion, and 3D autostereoscopic display; 2. Proposing a real-time blending algorithm for the fusion of textured point clouds and meshes on the autostereoscopic display device with significant depth perception.

In our algorithm, the average rendering frame rate is 33.1 FPS, and the average registration rate is 20.5 FPS. We parallelize the lenticular display rendering and marker-free intraoperative registration in the algorithm, fully utilizing the computing power by properly assigning tasks to the CPU and the GPU. Since the lenticular display rendering and RGB-D camera data acquisition processes are serialized, the current rendering frame rate is limited to the RGB-D camera's acquisition speed. Future advances in RGB-D cameras are expected to solve this problem by providing a high frame rate and high-resolution RGB-D information with better hardware performance. By employing a better RGB-D camera, we believe that the real-time performance of our algorithm will be sufficient to meet clinical requirements.

It should be noted that the surgeon should be cautious when performing a few surgeries requiring a high-precision surgical operation. We recommend a marker-based method to improve registration accuracy further in this system. The use of mark-based registration methods can significantly improve the accuracy and real-time performance of the intraoperative registration. However, the markers attach to patients' skin may cause additional infection problems.

Another issue to remind is the deformation of the RF needle and the soft tissue. We assume that the RF needle is rigid and does not deform or deflect during needle insertion. Based on this assumption, the position and posture of the needle after insertion into the body was estimated by tracking the markers on the back end of the RF. In addition, patient breathing and soft tissue deformation during RF needle insertion are also issues to be considered. In future studies, we will introduce intraoperative ultrasound into the system to achieve the integration of intraoperative ultrasound with the surgical scene to accurately display the deformation of the RF needle and soft tissues.

The TPE of the 3D autostereoscopic telesurgical system is around 2 mm, while the RMSE obtained by the intraoperative registration method is around 3 mm. In our method, the registration error between the intraoperative point cloud

and the preoperative segmentation model is quite large because the intraoperative point cloud reconstructed from the RGB-D camera has undulating surfaces. Another reason is that we used a complex human head model in the registration experiment, while a smooth agar phantom was used in the quantitative evaluation. The agar phantom may have a smaller registration error.

In quantitative evaluation experiments, the expert took longer completion times and longer paths compared to the novices. The reason was that the expert spent more time adjusting the needle's position and angle to achieve higher operational precision. In Fig. 12, the expert's mean value and variance of the TPE are smaller in both systems.

We have implemented a 3D autostereoscopic surgical visualization system that can display preoperative models and intraoperative scenes in real time. We built a preliminary prototype of the telesurgical robot system to demonstrate the potential of the 3D autostereoscopic surgical visualization system in telesurgery. Compared to binocular or head-mounted surgical visualization systems, the proposed system enables surgeons to see the 3D image and their hands simultaneously, thus ensuring better hand–eye coordination interaction. Moreover, the proposed system provides binocular parallax and motion parallax with multiple view-angle information, which improves surgeons' depth perception. Another major advantage of the 3D autostereoscopic surgical visualization system is that it can offer 3D images with accurate spatial information. In the future, we will further study in-situ interaction to achieve a more natural human–computer interaction, enabling surgeons to perform surgical operations on 3D images like a real surgery.

## Conclusions

In this paper, we propose a 3D surgical visualization system for telesurgery, which collects the intraoperative surgical scene and offers a 3D fusion image of the preoperative and intraoperative data on the autostereoscopic display device. To achieve the real-time performance of image fusion and 3D display, we propose an algorithmic framework including real-time augmented reality-based image fusion and 3D autostereoscopic display. We built a prototype system to validate our method. In the experiments, the horizontal resolution of the autostereoscopic display device was 1.31 mm, and the vertical resolution was 0.82 mm. The average rendering frame rate and the average registration rate were 33.1 FPS and 20.5 FPS. The average registration errors were approximately 3 mm when the phantom remained stationary or moved slowly. We performed both evaluation experiments and user experiments to prove that our system possesses relatively high spatial accuracy and strong robustness. Based on 3D surgical visualization, we built a telesurgical robot

system for teleoperated percutaneous interventions. The quantitative evaluation experiments showed that our system achieved similar operational accuracy ( $1.79 \pm 0.87$  mm) as the conventional system ( $1.95 \pm 0.71$  mm), while having advantages in terms of completion time (with 34.11% reduction) and path length (with 35.87% reduction). Post-experimental questionnaires indicated that the system was user-friendly for novices and experts.

**Supplementary Information** The online version contains supplementary material available at <https://doi.org/10.1007/s11548-021-02463-5>.

**Acknowledgements** The authors acknowledge supports from National Natural Science Foundation of China (82027807, 81771940), Beijing Municipal Natural Science Foundation (7212202), and Tsinghua University Initiative Scientific Research Program (20197020014).

**Author's contributions** TH and RL have contributed equally to this work. All authors read and approved the final manuscript.

**Funding** This work was supported in part by National Natural Science Foundation of China (82027807, 81771940), Beijing Municipal Natural Science Foundation (7212202), and Tsinghua University Initiative Scientific Research Program (20197020014).

**Data availability** Not applicable.

**Code availability** Not applicable.

## Declarations

**Competing interests** The authors declare no competing interests.

**Research involving human participants or animals** This article does not contain any studies with human participants or animals performed by any of the authors.

**Informed consent** For this study, formal consent is not required.

## References

1. Cazac C, Radu G (2014) Telesurgery—an efficient interdisciplinary approach used to improve the health care system. *J Med Life* 7(3):137
2. Peters BS, Armijo PR, Krause C, Choudhury SA, Oleynikov D (2018) Review of emerging surgical robotic technology. *Surg Endosc* 32(4):1636–1655
3. Freschi C, Ferrari V, Melfi F, Ferrari M, Mosca F, Cuschieri A (2013) Technical review of the da Vinci surgical telemanipulator. *Int J Med Robot Comput Assist Surg* 9(4):396–406
4. Vaida C, Al Hajjar N, Lazar V, Graur F, Burz A, Elisei R, Mois E, Pislă D (2019) Robotics in minimally invasive procedures: History, current trends and future challenges. In: 6th international conference on advancements of medicine and health care through technology, 17–20 October 2018, Cluj-Napoca, Romania. Springer, pp 267–273
5. Wen R, Tay WL, Nguyen BP, Chng CB, Chui CK (2014) Hand gesture guided robot-assisted surgery based on a direct augmented reality interface. *Comput Methods Programs Biomed* 116(2):68–80. <https://doi.org/10.1016/j.cmpb.2013.12.018>
6. Su H, Shang W, Li G, Patel N, Fischer GS (2017) An MRI-guided telesurgery system using a Fabry–Perot interferometry force sensor and a pneumatic haptic device. *Ann Biomed Eng* 45(8):1917–1928
7. Chevie J, Krupa A, Babel M (2019) Real-time teleoperation of flexible beveled-tip needle insertion using haptic force feedback and 3D ultrasound guidance. In: 2019 international conference on robotics and automation (ICRA), 20–24 May 2019, pp 2700–2706. <https://doi.org/10.1109/ICRA.2019.8794012>
8. Kihara K, Fujii Y, Masuda H, Saito K, Koga F, Matsuo Y, Numao N, Kojima K (2012) New three-dimensional head-mounted display system, TMDU-S-3D system, for minimally invasive surgery application: Procedures for gasless single-port radical nephrectomy. *Int J Urol* 19(9):886–889
9. Azizian M, Patel R (2011) Intraoperative 3D stereo visualization for image-guided cardiac ablation. In: Medical imaging 2011: visualization, image-guided procedures, and modeling. International Society for Optics and Photonics, p 79640F
10. Kong S-H, Oh B-M, Yoon H, Ahn HS, Lee H-J, Chung SG, Shiraishi N, Kitano S, Yang H-K (2010) Comparison of two- and three-dimensional camera systems in laparoscopic performance: a novel 3D system with one camera. *Surg Endosc* 24(5):1132–1143
11. Yamauchi Y, Shinohara K (2005) Effect of binocular stereopsis on surgical manipulation performance and fatigue when using a stereoscopic endoscope. *Stud Health Technol Inform* 111:611–614
12. Chen G, Huang T, Fan Z, Zhang X, Liao H (2019) A naked eye 3D display and interaction system for medical education and training. *J Biomed Inform* 100:103319
13. Urey H, Chellappan KV, Erden E, Surman P (2011) State of the art in stereoscopic and autostereoscopic displays. *Proc IEEE* 99(4):540–555
14. Ma L, Jiang W, Zhang B, Qu X, Ning G, Zhang X, Liao H (2019) Augmented reality surgical navigation with accurate CBCT-patient registration for dental implant placement. *Med Biol Eng Compu* 57(1):47–57
15. Wang J, Suenaga H, Hoshi K, Yang L, Kobayashi E, Sakuma I, Liao H (2014) Augmented reality navigation with automatic marker-free image registration using 3-D image overlay for dental surgery. *IEEE Trans Biomed Eng* 61(4):1295–1304
16. Rusu RB, Blodow N, Beetz M (2009) Fast point feature histograms (FPFH) for 3D registration. In: 2009 IEEE international conference on robotics and automation, pp 3212–3217
17. Rusu RB, Blodow N, Marton ZC, Beetz M (2008) Aligning point cloud views using persistent feature histograms. In: 2008 IEEE/RSJ international conference on intelligent robots and systems. IEEE, pp 3384–3391
18. Besl PJ, McKay ND (1992) Method for registration of 3-D shapes. In: Sensor fusion IV: control paradigms and data structures. International Society for Optics and Photonics, pp 586–606
19. Phong BT (1975) Illumination for computer generated pictures. *Commun ACM* 18(6):311–317
20. Baudisch P, Gutwin C (2004) Multiblending: displaying overlapping windows simultaneously without the drawbacks of alpha blending. In: Proceedings of the SIGCHI conference on Human factors in computing systems, pp 367–374
21. Jiao S, Wang X, Zhou M, Li W, Hong T, Nam D, Lee JH, Wu E, Wang H, Kim JY (2013) Multiple ray cluster rendering for interactive integral imaging system. *Opt Express* 21(8):10070–10086. <https://doi.org/10.1364/OE.21.010070>
22. Wang J, Suenaga H, Liao H, Hoshi K, Yang L, Kobayashi E, Sakuma I (2015) Real-time computer-generated integral imaging and 3D image calibration for augmented reality surgical



- navigation. *Comput Med Imaging Graph* 40:147–159. <https://doi.org/10.1016/j.compmedimag.2014.11.003>
23. Kim Y, Park G, Jung J-H, Kim J, Lee B (2009) Color moiré pattern simulation and analysis in three-dimensional integral imaging for finding the moiré-reduced tilted angle of a lens array. *Appl Opt* 48(11):2178–2187
24. Huang T, Han B, Zhang X, Liao H (2019) High-performance autostereoscopic display based on the lenticular tracking method. *Opt Express* 27(15):20421–20434
25. Chen F, Cui X, Han B, Liu J, Zhang X, Liao H (2021) Augmented reality navigation for minimally invasive knee surgery using enhanced arthroscopy. *Comput Methods Programs Biomed* 201:105952. <https://doi.org/10.1016/j.cmpb.2021.105952>

**Publisher's Note** Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.