2019 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)
Macau, China, November 4-8, 2019

A Handheld Master Controller for Robot-Assisted Microsurgery

Dandan Zhang, *Student Member, IEEE*, Yao Guo, *Member, IEEE*, Junhong Chen,
Jindong Liu, *Senior Member, IEEE* and Guang-Zhong Yang, *Fellow, IEEE*

***Abstract*— Accurate master-slave control is important for Robot-Assisted Microsurgery (RAMS). This paper presents a handheld master controller for the operation and training of RAMS. A 9-axis Inertial Measure Unit (IMU) and a micro camera are utilized to form the sensing system for the handheld controller. A new hybrid marker pattern is designed to achieve reliable visual tracking, which integrated QR codes, Aruco markers, and chessboard vertices. Real-time multi-sensor fusion is implemented to further improve the tracking accuracy. The proposed handheld controller has been verified on an in-house microsurgical robot to assess its usability and robustness. User studies were conducted based on a trajectory following task, which indicated that the proposed handheld controller had comparable performance with the Phantom Omni, demonstrating its potential applications in microsurgical robot control and training.**

I. Introduction

One of the main challenges faced by surgeons for microsurgery, such as vitreo-retinal and cochlear implant surgeries, is the capability to perform complex tasks with a high level of dexterity, accuracy, and safety. It has been shown that the perceptual accuracy of a naked human eye is within 10s of micrometers, while the human hand can only reach a precision of about 100 micrometers under optimal conditions [1]. This natural limitation makes the manipulating at microscale difficult even for expert surgeons.

In the last decades, RAMS is rapidly growing, bringing significant benefits to help realize the full potential of microsurgery in terms of dexterity, accuracy and safety [2], [3]. The master-slave paradigm has been extensively applied to the existing platforms, with the surgeon manipulating the master controller to control slave manipulators. Various manipulator designs for microsurgical tasks have been presented in [4].

Among the existing systems, grounded robotic masters for microsurgery have seen promising commercial opportunities and clinical success, due to their appreciated advantages of stability and reliability. Generally, surgeons' hands are tracked by wire and pulleys or mechanical linkages to generate commands for remote control of a slave robot. However, these available systems usually have a large footprint, high cost, and long setup time, thus preventing their widespread application.

With the advantages of being compact and lightweight, handheld master controllers are attractive alternatives for their grounded counterparts [5]. Without mechanical linkages, the physical footprint can be significantly reduced. Besides, intuitive ungrounded design can assist the surgeons to manipulate the slave robot through teleoperation in a more natural way without motion constraints. Therefore, it is beneficial for reducing extensive surgical training before the adoption of robotic microsurgery [6]. More importantly, for clinical translation, handheld master controllers can be executed as a user interface at the patient's bedside without introducing a physical distance between surgeons and patients [7], which enhances the potential safety benefits.

For handheld controllers, capturing the surgeons' motion in a precise manner plays a critical role in generating adequate commands for microsurgical robot remote control. In the past decades, various sensing technologies have been explored to track surgeons' hand movements. Acoustic systems utilize the time of flight and intensity of audio signals to affirm the marker locations, while electromagnetic systems employ the magnetic field intensity [8]. These two modalities are convenient to use but are sensitive to noises and external disturbances, which reduces their reliability and hamper their potential applications for RAMS.

Inertial Measurement Unit (IMU) is now common for motion tracking [9]. The advantages of IMUs include high sampling rate and being invariant to the environment [9], though sensor drift and accumulative errors are their inherent limitations.

With recent advances in computer vision, depth cameras can provide promising tracking of hands and fingers [10]. Existing products include a hand-tracking system, namely KinectTM-based 3Gear. However, these systems incorporating depth sensors are typical with large size and need to be carefully placed in a specific location, which limits their flexibility [11]. Optitrack and Vicon are two commercial products for motion tracking. However, they are not suitable for clinical applications, since the resolutions of both systems may not be high enough to track the gestures for microsurgical operation.

On the other hand, various optical markers, such as chessboard vertices [12], QR code, and Aruco marker [13], are developed to achieve high precision tracking for the moving camera. However, the existing marker patterns have their limitations while being utilized in the handheld controller for RAMS.

Comprehensive comparisons of motion tracking systems based on different modalities were summarized in [14]. Considering that different modalities have pros and cons, a hybrid tracking method is beneficial for combining the advantages

This work is supported by the Hamlyn Centre. All authors are with the Hamlyn Centre for Robotic Surgery, Imperial College London, London, United Kingdom. G.-Z. Yang is also with the Institute of Medical Robotics, Shanghai Jiao Tong University, China. Corresponding author email: d.zhang17@imperial.ac.uk

of different methods as well as compensating their drawbacks [15]. Inertial sensors may suffer from drift when estimating position, but they can deal with fast motion acquisition. In spite of the low sampling rate of the visual sensors, they have precise tracking results of position with well-designed markers. Moreover, the visual cues are necessary to deal with the drift problems of the IMU. Therefore, hybrid motion tracking methods can have many advantages.

In [16], inertial data and 2D locations of optical markers in the stereo camera images were fused to track a handheld device. However, the complexity of the system setup may limit its adoption. The optical tracking system is a lab-based design, which is difficult to be fitted in a general clinical operation room. Occlusion may limit the consistency of the optical tracking, and thus hamper the safety for microsurgical operation.

Using an on-board camera is more effective and practical than placing extra tracking devices such as a depth camera or a stereo camera at a specific place. The additional setup procedures are not necessary while the influence originated from the illumination effects can be reduced. Therefore, an on-board micro camera is considered, which ensures the system is easy-to-use.

In this paper, we develop a novel handheld master controller for RAMS, which consists of a 9-axis IMU and a micro stereo camera. The rest of the paper is organized as follows. Section II introduces the proposed hybrid marker pattern design and a multi-sensor fusion framework for the motion tracking system, while the evaluation of the motion tracking results is performed. Section III presents the system integration of the microsurgical robot and the handheld controller. User studies are conducted in Section IV. Conclusions are drawn in Section V.

II. Methodology

In this section, the hardware of the proposed handheld controller will be briefly introduced. A novel hybrid marker pattern is designed to realize precise marker-based motion tracking for the handheld controller, with the illustration of sensor fusion algorithm.

A. Handheld Master Controller

The handheld controller incorporates a miniature stereo camera (Precision Robotics, UK) and a 9-axis IMU sensor (Xsens, Netherland). The overview of the handheld controller can be found in Fig. 1. After fusing the pose estimation information from the micro camera and the orientation estimation results provided by the IMU, the full 6-DoF tracking can be obtained.

B. Pose Estimation from IMU

A 9-axis IMU combines 3-axis magnetometer with inertial sensors. Acceleration information can be measured by the 3-axis accelerometer, while angular rate information is obtained by the 3-axis gyroscope based on the Coriolis acceleration effect [17]. The 3-axis magnetometer measures the magnetic field strength in a given direction, allowing the north direction to be found. In this way, the yaw information can be gained.

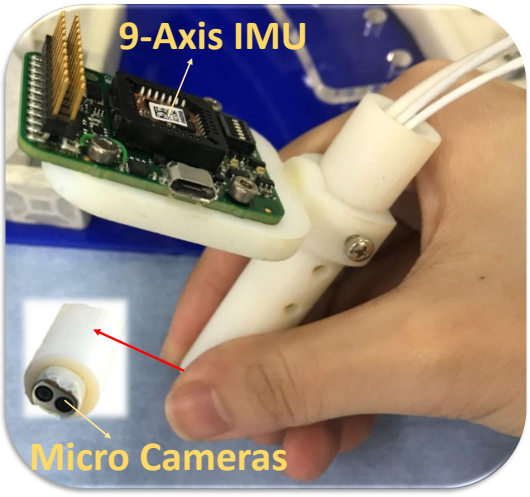

Fig. 1. Details of the handheld controller.

3-axis acceleration and 3-axis magnetic field messages are used to estimate the static orientation. The accelerometer measurements can be directly computed into Euler angles using trigonometry, while the magnetometer coupled with the accelerometer can effectively calculate the heading angle by measuring the magnetic field. Angular velocity is the rate of change of angular displacement, as measured by the 3-axis gyroscope. Since the gyroscope sensor outputs angular rates, the measurement can be simply integrated over a period of time (the sampling rate) and summed up to obtain the absolute angle. In this way, the dynamic update of the orientation estimation can be achieved.

XKF3i is a frequently used fusion algorithm for the 9-axis IMU, which fuses the Information from inertial and magnetic sensors and combines their advantages to obtain accurate orientation estimation [18].

Although the position value can be obtained by the double integral of the acceleration data provided by the IMU, the accuracy of the position estimate of inertial sensors degrades with time due to the random errors.

C. Hybrid QR-Aruco-Chessboard Pattern Design

Marker-based visual tracking method is much more accurate and reliable than the markless ones. Considering that safety and accuracy are fundamental for microsurgery, marker-based visual tracking is utilized.

A hybrid QR-Aruco-Chessboard (QAC) pattern is designed as a computer-vision-friendly 2D pattern that has enough salient features for 6 DoFs pose estimation (see Fig. 2), which combines the advantages of QR codes [19] [20], Aruco markers [13] and chessboard vertices [12] together. Three tracking modules are used to provide real-time pose estimation information concurrently, while the information is merged along with the orientation estimation results to provide robust 6D pose estimation.

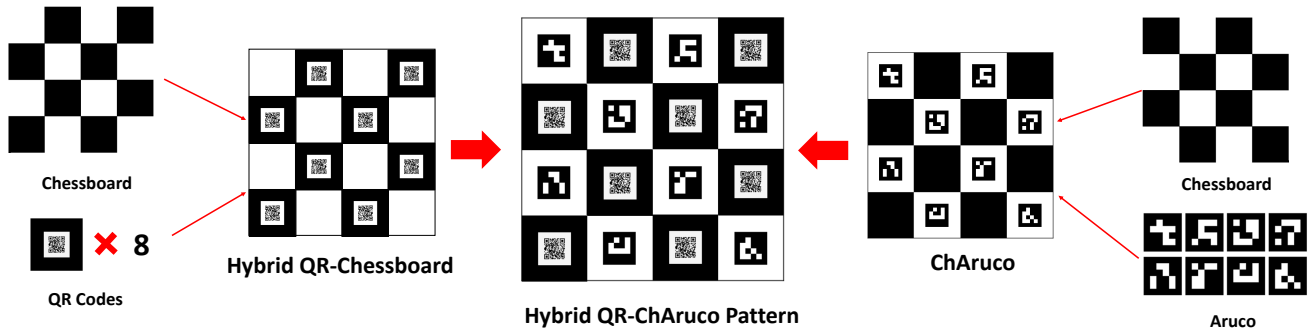

Fig. 2. A hybrid QR-Aruco-Chessboard (QAC) pattern for marker-based visual tracking.

D. Pose Estimation from Chessboard

In order to accelerate the speed of real-time pose estimation, the chessboard is limited to a 3×3 non-symmetrical pattern. Hence, the chessboard vertices can be detected quickly. For each pixel in the filtered image, a ring of 11 pixels around the pixel is sampled at a constant radius with equal angular spacing.

Given the tracked projections on an image of a calibrated camera and the corresponding spatial positions in the marker's local coordinate frame, the pose estimation based on 2D-3D correspondences can be realized by a Perspective-n-Point (PnP) method [21] [22].

However, when the camera is closed to the pattern, the number of tracked chessboard vertices will not be enough to realize pose estimation via PnP method.

E. Pose Estimation from Multi-Aruco Markers

Aruco markers [23] and their derivatives are frequently used in augmented reality and robotics. Eight different Aruco markers with ID range from 1 to 8 are used to provide tracking information. This can ensure a wider range of tracking. Therefore, a multi-Aruco module is utilized to compensate for the limitation of the chessboard vertices based tracking, when the micro-camera approaches the hybrid QAC pattern.

However, the precision of the multi-Aruco tracking module will degrade when the micro-camera is far from the plane of the markers.

F. Pose Estimation from chAruco Pattern

To achieve higher tracking accuracy, chAruco pattern can be generated by merging Aruco markers and the chessboard information as part of the hybrid QAC pattern. The pose estimation accuracy from the chAruco is higher than using only chessboard information or pure Aruco markers. However, the successful estimation depends on whether more than five of the Aruco markers are within the view of the camera.

Therefore, the ChAruco only plays an important role in providing accurate tracking information while enough markers are within the camera's view.

G. Pose Estimation from QR Code

Another pose information can be obtained based on the QR-code detection, where the pose estimation process is achieved from the location of the four QR-code corners. The 2D coordinates are extracted from the image. The transform relationship between the 2D image and the pose in the 3D space can be obtained thanks to camera intrinsic parameters. Eight QR-codes are merged to the chessboard to form the final hybrid QAC pattern, as shown in Fig. 2.

One advantage of the QR-code is that the tracking precision is high, due to more feature information is available than the Aruco markers. In practice, due to the varying light conditions or fast movement of the handheld controller, failures of QR-codes detection may occur. The successful rate of QR codes detection will decrease when the distance between the camera and the hybrid QR-Chessboard pattern is too large.

H. Multi-Sensor Fusion Framework Construction

The calibration process, using Zhang's calibration method [24], is conducted first to determine the intrinsic camera parameters and lens distortion coefficients. Gaussian filter is applied to the greyscale image to remove speckles and noises before sending the real-time images for pose estimation.

The visual tracking framework is constructed for pose estimation via three major modules, i.e. i) multi-Aruco module, ii) chAruco module, iii) QR-codes module. Considering that the position estimation from IMU data suffers from drift problem, while the orientation estimation from single Aruco is unstable, the 3D orientation information from the IMU and the 3D position information from the multi-Aruco module are combined to form a stable 6D pose vector.

As for the chAruco module and the QR-code module, the 6D pose information can be fully utilized. Fig. 3 shows the available micro-camera view for online visual data processing.

A status monitor is utilized to select the suitable information for multi-sensor fusion. Visual tracking state vector $\boldsymbol{V}(t) = [a, c, q]$ indicates the status of the tracking situation of the multi-Aruco, chAruco and QR-code at time step t. If the value of $q/c/a$ is equal to one, the tracking result of the corresponding mode is reliable. Suppose that there are

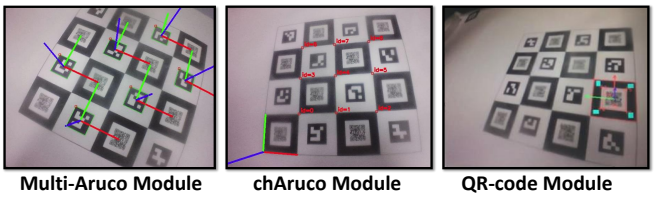

Fig. 3. Camera tracking results of the three modules in the micro-camera view.

k Aruco markers in total, $\boldsymbol{V_a^j}(t)(j = 1, 2, ...k)$ identifies whether the Aruco marker with the ID value of j_{th} is within the view of the camera or not.

$\boldsymbol{Q_q}(t)$, $\boldsymbol{Q_c}(t)$ represent the pose estimation results from the QR-code module and the chAruco module respectively at time step t, $\boldsymbol{P_*}(t)(* = q, c, a)$ represents the position while $\boldsymbol{O_*}(t)(* = q, c, a)$ represent the orientation.

$$\boldsymbol{P_a}(t) = \frac{\sum_{j=1}^{j=k} \boldsymbol{P_a^j(t)} * \boldsymbol{V_a^j}(t)}{\sum_{j=1}^{j=k} \boldsymbol{V_a^j}(t)}, \tag{1}$$

where $\boldsymbol{P_a^j(t)}$ $(j = 1, 2, ...k)$ represents the position estimation result based on the Multi-Aruco tracking with ID value of j.

Suppose that $\tilde{Q}(t) = [\tilde{P}(t), \tilde{O}(t)]$ is the final pose estimation result calculated by the sensor fusion framework at time step t. Position $\tilde{P}(t)$ can be estimated as follows.

$$\boldsymbol{\tilde{P}}(t) = \frac{(q(t) \times \boldsymbol{P_q}(t) + c(t) \times \boldsymbol{P_c}(t) + a(t) \times \boldsymbol{P_a}(t))}{(q(t) + c(t) + a(t))} \tag{2}$$

$\boldsymbol{O_i}(t)$ represents the orientation estimation result based on IMU. The visual tracking framework is applied to every image in the sequence, by which the motion tracking of the handheld controller can be realized. Orientation $\tilde{O}(t)$ can be estimated as follows.

$$\boldsymbol{\tilde{O}(t)} = \frac{(q(t) \times \boldsymbol{O_q}(t) + c(t) \times \boldsymbol{O_c}(t) + \boldsymbol{O_i}(t))}{(q(t) + c(t) + 1)} \tag{3}$$

The pose increment value $\delta\boldsymbol{\tilde{Q}}(t)$ will be obtained and scaled down, which will be sent to the microsurgical robot as command to realize pose incremental control. The algorithm of the multi-sensor fusion is shown as Algorithm 1.

I. Tracking Performance Analysis

Reliable visual tracking is fundamental for the usability of the handheld controller. The detection rate and the motion tracking accuracy are utilized to analyze the performance of the motion tracking technique.

The tracking status of the three modules can be monitored simultaneously. Suppose that N frames were obtained online, n_i is the number of frames that the marker detection can successfully provide pose estimation information. The detection rate β can be defined as $\beta = n_i/N$. Different visual tracking modules have different detection ranges, depends on their inherent properties.

Algorithm 1 Multi-Sensor Fusion

Input: $\boldsymbol{Q_q}(t)$, $\boldsymbol{Q_c}(t)$,$\boldsymbol{O_i}(t)$, $\boldsymbol{\tilde{Q}}(t-1)$
$\boldsymbol{V}(t) = [q(t), c(t), a(t)]$,
$\boldsymbol{V_a^j}(t)$, $\boldsymbol{V_a^j}(t-1)(j = 1, 2, ...8)$
Output: Pose Increment Value $\delta\boldsymbol{\tilde{Q}}(t)$

1: Use Extended Kalman Filter (EKF) to filter the original data.
2: Estimate $\boldsymbol{P_a(t)}$ based on Equation (1).
3: Estimate $\boldsymbol{\tilde{P}(t)}$ based on Equation (2).
4: Estimate $\boldsymbol{\tilde{O}(t)}$ based on Equation (3).
5: Obtain the pose vector $\boldsymbol{\tilde{Q}}(t) = [\boldsymbol{\tilde{P}}(t), \boldsymbol{\tilde{O}}(t)]$
6: Use EKF to filter $\boldsymbol{\tilde{Q}}$.
7: Obtain the pose increment value $\delta\boldsymbol{\tilde{Q}}(t) = \boldsymbol{\tilde{Q}}(t) - \boldsymbol{\tilde{Q}}(t-1)$.

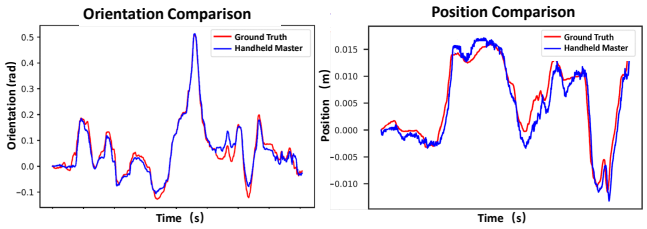

Fig. 4. Orientation and position tracking results comparison between the ground truth data and the handheld master controller.

The experiments were divided into three groups based on different distances from the pattern to the camera: near (20-100 mm), middle (100-180 mm) and far (180-260 mm). For each group, we repeated three trials and recorded the data of the tracking results. The quantitative results, as shown in Table I, confirms that the tracking system for the proposed handheld master is superior to the single tracking module by achieving nearly 100% detection rates in various range.

TABLE I
DETECTION RATE OF DIFFERENT MODES

	Near (20-100mm)	**Middle** (100-180mm)	**Far** (180-260mm)
Chessboard	0%	70%	96%
Multi-Aruco	100%	100%	97%
QR-Codes	98%	82%	76%
chAruco	0%	80%	98%
Hybrid Marker	100%	100%	100%

The handheld controller can be easily mounted with the pen-like controller of the Phantom Omni, which is a reliable grounded system. Thus, the motion tracking data can be recorded simultaneously, while the data obtained from the Phantom Omni is utilized as the ground truth data. Fig. 4 shows the orientation and position tracking results comparison between the ground truth data and the handheld master controller. The results show that there is almost no significant differences between the grounded and the ungrounded master controller.

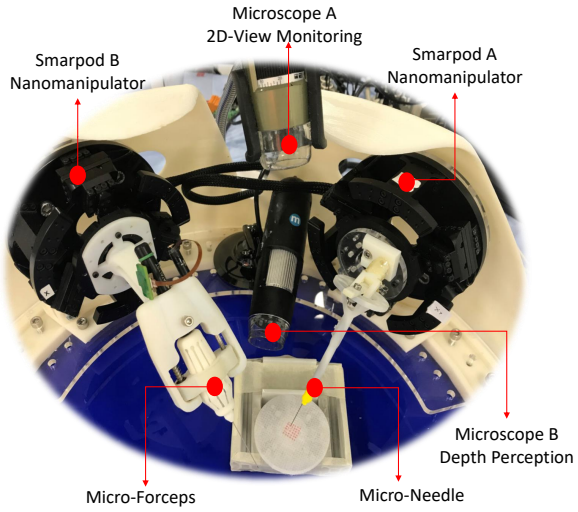

Fig. 5. Overview of the in-house microsurgical robot as the experimental platform.

III. System Overview

In this section, the hardware platform and the software architecture of the in-house microsurgical robot are introduced.

A. Experimental Platform

An in-house microsurgical robot is utilized as the research platform for the validation of the effectiveness of the handheld master controller. It is a teleoperated tethered micromanipulation research platform for RAMS. A pair of 6 Degrees of Freedom (DoFs) positioning stage, named SmarPod (SmarAct, German), are utilized as the slave robot (see Fig. 5), where microsurgical tools are mounted, including a micro-forcep and a micro-needle. The hexapod-like SmarPod is a parallel-kinematic system, which is considerably more compact than common serial kinematics.

As demonstrated in Fig. 5, two digital microscopes (A: Dino Lite, B: Maplin n43HH) provide top and side views, with adjustable magnification ranged from 20 to 50. Microscope A is used for 2D-View monitoring, while microscope B is employed for providing depth perception of visual cues. A foot pedal (Philips LFH2310) is utilized to provide the "engage" and "clutch" information during the remote control.

The original controller for the SmarPod (Nanomanipulator) is non-intuitive, which requires the operator to directly control each actuator with buttons or joystick-like interfaces. This solution influences the control efficiency and fails to provide their users with meaningful insights into the micromanipulation [25]. A handheld master controller for teleoperation, which is portable and does not occupy too much space, will be more suitable for both microsurgical training and clinical applications.

B. Software Architecture

The software for the whole system is constructed with the Robotic Operating System (ROS) framework on a Linux Host PC. The states of the robots and the images obtained from the microscopes and all the other sensor information can be transmitted to the host PC. They are subscribed as ROS messages (topics) for high-level processing. A Smarpod-ROS Bridge is built up to link the low-level kinematic control of the SmarPod with the ROS interface.

The final control commands generated by master robots can be published to ROS topics and control the microsurgical robot via the Smarpod-ROS Bridge to set its end-effector at the targeted pose in the Cartesian Space.

For master-slave mapping, incremental control is utilized. The real-time increment value for the pose of the handheld controller is scaled down by a fixed motion-scaling ratio τ.

IV. User Studies

A user study was designed to examine the usability of the handheld master controller in the general trajectory following maneuvers based on a predefined protocol. The proposed handheld controller is compared with Phantom Omni, a grounded master device, with results demonstrating the relative merits in terms of task completion time, average control speed and Root-Mean-Square-Error (RMSE).

A. Experimental Setup and Task Description

Fig. 6 (a) shows the experimental setup. The whole system includes the microsurgical robot as a slave, the proposed handheld controller as a master device, the hybrid QAC pattern for motion tracking, and an experimental platform for user studies. The top and side overview of the experimental platform are shown in Figs. 6 (b) and (c) respectively. A trajectory following task was designed based on a predefined protocol.

The subjects were asked to trace the boundary of a square based on the top-to-down 2D view (see Figs. 6 (d)). The square was drawn on a piece of textile and the subject was asked to place the micro-needle tip as close as to the predefined points in the order of A-B-C-D-A (see Fig. 6 (d)). Another view is used to monitor the distance between the tooltip to the textile to ensure that the micro-needle tip is within a reasonable distance to the textile plane (see Fig. 6 (e)).

B. Participants

Preliminary user studies were conducted with six novice users (1 female and 5 males; mean age = 27.3±3.4) to evaluate the usability of the handheld controller. All the subjects had five minutes to get familiarized with the task before the actual experiment. They were asked to go through the whole procedure twice to get accustomed to the experimental protocols.

During the user studies, the subjects were asked to perform the same task for three to five trials using the Phantom Omni and the handheld master respectively. A total of 36 trails were recorded for statistical analysis.

C. Performance Evaluation

Fig. 7 illustrates the representative trace trajectory of the microsurgical tooltip locations in blue colour of the same subject when using the handheld master controller and the

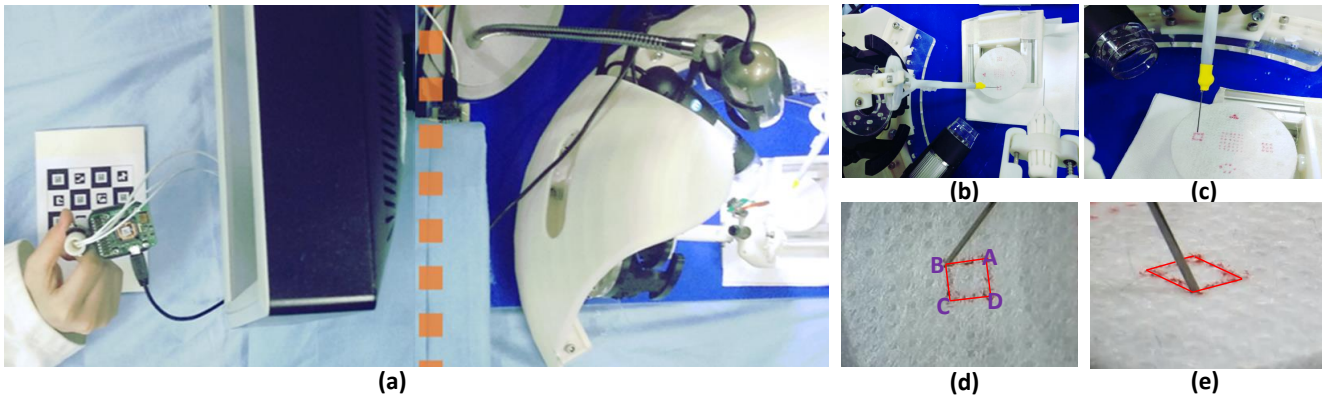

Fig. 6. The user study design based on a trajectory following task. (a) Experimental setup overview. (b) Top-to-down view of the experimental platform. (c)Side view of the experimental platform. (d) Top-to-down view of the pre-defined trajectory.(e) Side view of the experimental platform.

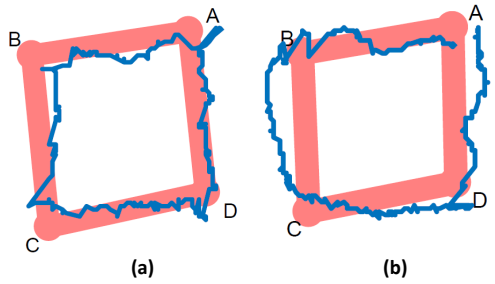

Fig. 7. The trajectory following task results. (a) Results of the trajectory following of Subject A using the handheld controller. (c) Results of the trajectory following of Subject A using the Phantom Omni.

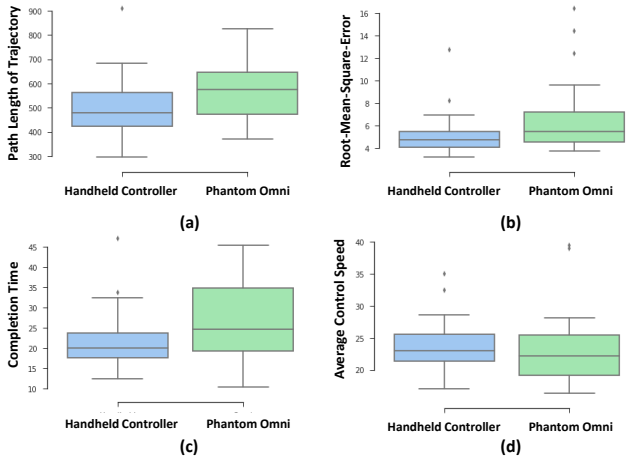

Fig. 8. Results for comparison in the form of box plot. Comparisons in term of (a) P_i; (b) R; (c)T_c; (d) A_s .

Phantom Omni. The target trajectory is represented by thick light red lines.

It can be seen that with the use of the handheld controller, all users managed to achieve more precise and smooth motion trajectories. The trajectories were more chaotic when using the Phantom Omni, with a lower tracing accuracy.

For quantitative analysis, four evaluation metrics are used to evaluate the performance of the subjects during teleoperation, including the path length of the trajectory, the Root Mean Square Error (RMSE), the task completion time, and the average control speed.

- **Path Length of the Trajectory** P_i: The total length of the trajectory generated by the micro-surgical tooltip when conducting a trial.
- **Root Mean Square Error** R: RMSE is the square root of the variance of the trajectory, known as the standard error.
- **Completion Time** T_c: The time used to finish each trial.
- **Average Control Speed** A_s : The average control speed for a single trial.

D. User Study Results and Analysis

Table II shows the average results of the user studies and corresponding standard deviation among all the trials for the four evaluation metrics. As the subjects involved have different levels of expertise, unsurprisingly these are reflected in the performance variation.

The results for each master controller using different evaluation metrics are plotted in a boxplot form, as shown in Fig. 8. The overall results indicated that the microsurgical robot can achieve a better track following the motion of the handheld controller.

Normality tests (Shapiro-Wilk test) at 0.05 significance level were performed to check if the metrics satisfy the normal distribution assumption. The results indicated that except for P_i, all the metrics reveal non-parametric nature. Wilcoxon signed-rank tests are therefore conducted for non-parametric statistical comparison between variables A_s, T_c and R. T-test was used for p-value calculation for P_i. The mean value with standard deviation value and the p-value results can be seen in Table II.

Based on the results mentioned above, it can be seen that the handheld controller has better performance than the Phantom Omni in terms of RMSE, slave robot trajectory,

TABLE II
USER STUDIES RESULTS

	Phantom Omni	Handheld Controller	p-value
P_i	130.8 ± 37.1	500.2 ± 128.8	0.0506
$T_c(s)$	26.4 ± 9.4	21.9 ± 7.7	0.1137
A_s	11.9 ± 6.13	15.9 ± 4.3	0.4654
R	6.8 ± 3.4	5.2 ± 2.0	0.0636

average speed and task completion time.

Although Phantom Omni is a mature commercial product, the placement of the Phantom Omni during user study is not optimized. The mechanical constraints may influence the users' performance. As for the handheld controller, ergonomics design can be taken into consideration in the future prototype.

V. CONCLUSIONS

In this paper, we designed a handheld master controller for RAMS. To achieve high precision motion tracking performance of the controller, a 9-axis IMU and a micro camera are integrated into the controller. For realizing reliable visual tracking, a new marker pattern (hybrid QAC) is designed by effectively combining QR codes, Aruco markers and chessboard vertices. To fuse the information from both the IMU and the micro camera, a multi-sensor fusion framework was utilized to further improve the tracking accuracy for the handheld controller.

The proposed handheld system was verified on an in-house microsurgical robot to demonstrate its usability and robustness. Comparison experiments were conducted between the commercial grounded controller (Phantom Omni) and the handheld controller. User study results indicated that the proposed handheld controller has high control efficiency during teleoperation, the performance of which is comparable to the Phantom Omni. Results demonstrate that the handheld controller facilitates the training of basic microsurgery operational skills, which enables users to become familiarized with the hand-motor axis control, leading to potential future clinical use.

VI. ACKNOWLEDGEMENTS

The authors would like to acknowledge all the subjects for taking part in the user studies and would like to thank Dr. Yang Hu for valuable discussion.

REFERENCES

[1] C. Duval and J. Jones, "Assessment of the amplitude of oscillations associated with high-frequency components of physiological tremor: impact of loading and signal differentiation," *Experimental brain research*, vol. 163, no. 2, pp. 261–266, 2005.

[2] G.-Z. Yang, J. Bellingham, P. E. Dupont, P. Fischer, L. Floridi, R. Full, N. Jacobstein, V. Kumar, M. McNutt, R. Merrifield *et al.*, "The grand challenges of science robotics," *Sci. Robot.*, vol. 3, no. 14, p. eaar7650, 2018.

[3] V. Vitiello, S.-L. Lee, T. P. Cundy, and G.-Z. Yang, "Emerging robotic platforms for minimally invasive surgery," *IEEE reviews in biomedical engineering*, vol. 6, pp. 111–126, 2013.

[4] H. Das, H. Zak, J. Johnson, J. Crouch, and D. Frambach, "Evaluation of a telerobotic system to assist surgeons in microsurgery," *Computer Aided Surgery*, vol. 4, no. 1, pp. 15–25, 2015.

[5] C. J. Payne and G.-Z. Yang, "Hand-held medical robots," *Annals of biomedical engineering*, vol. 42, no. 8, pp. 1594–1605, 2014.

[6] A. N. Sridhar, T. P. Briggs, J. D. Kelly, and S. Nathan, "Training in robotic surgeryan overview," *Current urology reports*, vol. 18, no. 8, p. 58, 2017.

[7] A. Simorov, R. S. Otte, C. M. Kopietz, and D. Oleynikov, "Review of surgical robotics user interface: what is the best way to control robotic surgery?" *Surgical endoscopy*, vol. 26, no. 8, pp. 2117–2125, 2012.

[8] S. Yabukami, H. Kikuchi, M. Yamaguchi, K. Arai, K. Takahashi, A. Itagaki, and N. Wako, "Motion capture system of magnetic markers using three-axial magnetic field sensor," *IEEE transactions on magnetics*, vol. 36, no. 5, pp. 3646–3648, 2000.

[9] M. K. Kim, K. Ryu, Y. Oh, S.-R. Oh, and K. Kim, "Implementation of real-time motion and force capturing system for tele-manipulation based on semg signals and imu motion data," in *Robotics and Automation (ICRA), 2014 IEEE International Conference on*. IEEE, 2014, pp. 5658–5664.

[10] V. Frati and D. Prattichizzo, "Using kinect for hand tracking and rendering in wearable haptics," in *World Haptics Conference (WHC), 2011 IEEE*. IEEE, 2011, pp. 317–321.

[11] T. B. Moeslund and E. Granum, "A survey of computer vision-based human motion capture," *Computer vision and image understanding*, vol. 81, no. 3, pp. 231–268, 2001.

[12] C. Shu, A. Brunton, and M. Fiala, "A topological approach to finding grids in calibration patterns," *Machine Vision and Applications*, vol. 21, no. 6, pp. 949–957, 2010.

[13] R. M. Salinas, "Aruco: a minimal library for augmented reality applications based on opencv," 2012.

[14] M. Field, D. A. Stirling, F. Naghdy, and Z. Pan, "Motion capture in robotics review," 2009.

[15] Y. Tao, H. Hu, and H. Zhou, "Integration of vision and inertial sensors for 3d arm motion tracking in home-based rehabilitation," *The International Journal of Robotics Research*, vol. 26, no. 6, pp. 607–624, 2007.

[16] A. Tobergte, M. Pomarlan, G. Passig, and G. Hirzinger, "An approach to ulta-tightly coupled data fusion for handheld input devices in robotic surgery," in *Robotics and Automation (ICRA), 2011 IEEE International Conference on*. IEEE, 2011, pp. 2424–2430.

[17] Z. Fu, G. Zhang, Y. Lin, Y. Liu, and J. Tan, "Calibration and compensation of inertial sensor errors in portable applicationsa review," in *Control (CONTROL), 2016 UKACC 11th International Conference on*. IEEE, 2016, pp. 1–4.

[18] A. Jouybari, A. Ardalan, and M. Rezvani, "Experimental comparison between mahoney and complementary sensor fusion algorithm for attitude determination by raw sensor data of xsens imu on buoy," *The International Archives of the Photogrammetry, Remote sensing and Spatial Information Sciences*, vol. 42, pp. 497–502, 2017.

[19] T. Anezaki, K. Eimon, S. Tansuriyavong, and Y. Yagi, "Development of a human-tracking robot using qr code recognition," in *2011 17th Korea-Japan Joint Workshop on Frontiers of Computer Vision (FCV)*. IEEE, 2011, pp. 1–6.

[20] É. Marchand, F. Spindler, and F. Chaumette, "Visp for visual servoing: a generic software platform with a wide class of robot control skills," *IEEE Robotics & Automation Magazine*, vol. 12, no. 4, pp. 40–52, 2005.

[21] S. Dinc, F. Fahimi, and R. Aygun, "Mirage: an o (n) time analytical solution to 3d camera pose estimation with multi-camera support," *Robotica*, vol. 35, no. 12, pp. 2278–2296, 2017.

[22] Y. Zheng, Y. Kuang, S. Sugimoto, K. Astrom, and M. Okutomi, "Revisiting the pnp problem: A fast, general and optimal solution," in *Proceedings of the IEEE International Conference on Computer Vision*, 2013, pp. 2344–2351.

[23] S. Garrido-Jurado, R. Muñoz-Salinas, F. J. Madrid-Cuevas, and M. J. Marín-Jiménez, "Automatic generation and detection of highly reliable fiducial markers under occlusion," *Pattern Recognition*, vol. 47, no. 6, pp. 2280–2292, 2014.

[24] Z. Zhang, "A flexible new technique for camera calibration," *IEEE Transactions on pattern analysis and machine intelligence*, vol. 22, 2000.

[25] S. Sakr, T. Daunizeau, D. Reversat, S. Régnier, and S. Haliyo, "An ungrounded master device for tele-microassembly," in *2018 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*. IEEE, 2018, pp. 1–9.