

Intention Recognition For Gaze Controlled Robotic Minimally Invasive Laser Ablation

Gauthier Gras and Guang-Zhong Yang, *Fellow*, IEEE

Abstract—Eye tracking technology has shown promising results for allowing hands-free control of robotically-mounted cameras and tools. However existing systems present only limited capabilities in allowing the full range of camera motions in a safe, intuitive manner. This paper introduces a framework for the recognition of surgeon intention, allowing activation and control of the camera through natural gaze behaviour. The system is resistant to noise such as blinking, while allowing the surgeon to look away safely at any time. Furthermore, this paper presents a novel approach to control the translation of the camera along its optical axis using a combination of eye tracking and stereo reconstruction. Combining eye tracking and stereo reconstruction allows the system to determine which point in 3D space the user is fixating, enabling a translation of the camera to achieve the optimal viewing distance. In addition, the eye tracking information is used to perform automatic laser targeting for laser ablation. The desired target point of the laser, mounted on a separate robotic arm, is determined with the eye tracking thus removing the need to manually adjust the laser's target point before starting each new ablation. The calibration methodology used to obtain millimetre precision for the laser targeting without the aid of visual servoing is described. Finally, a user study validating the system is presented, showing clear improvement with median task times under half of those of a manually controlled robotic system.

I. INTRODUCTION

The use of minimally invasive surgery has been demonstrated to reduce patient trauma and pain, as well as increase post-operative recovery speed. However, the minimally invasive nature of laparoscopic surgery also introduces new challenges. Laparoscopic cameras typically have a limited field of view, implying that they must be frequently repositioned to allow for good visual coverage of the surgical scene. As such, it is difficult for a surgeon to perform a bimanual task without external assistance to move the camera. This task is usually performed by a surgical assistant via verbal communication. However studies have shown that communication failures between the surgeon and the assistant can be cognitively taxing as well as dangerous to the patient [1], [2].

The advent of robotic systems in the operating theatre have provided new tools to address this shortcoming. Different means of active robotic camera control have been explored, such as head-mounted tracking [3], verbal control [4], and manual control [5]. While these systems allow the surgeon to keep both hands free they present a relatively large footprint in an already crowded operating theatre, and require direct surgeon control to function.

Gauthier Gras and Guang-Zhong Yang are with the Hamlyn Centre for Robotic Surgery, Imperial College London, SW7 2AZ, London, UK (email: gauthier.gras12@imperial.ac.uk)

An attractive alternative is to use eye tracking data to control the movement of the camera. In [6] an articulated mechatronic laparoscope is controlled using a joint-based gaze method. This technique uses an eye-position based proportional controller to move one robot joint at a time. However this system does not provide task space control of the robot nor zooming functionality (translation of the camera along the optical axis of the camera). It also does not possess any built-in safety features, and would have required a footswitch to activate and de-activate the system. Another system proposed by [7] attempts to address these issues through the use of gaze gestures. The camera is mounted on a robotic arm and controlled in task space via a discrete velocity controller based on pre-defined sections of the screen. The machine-learned gaze gestures allow the user to activate and deactivate panning and zooming camera control modes. However gaze gestures require the user's eyes to actively perform a combination of predefined movements, which can feel unnatural and tiring. Furthermore, activating a gaze gesture while in a camera control mode can make the system move in undesired ways.

This paper presents a novel framework for intention recognition designed to take full advantage of natural gaze behaviour, without the need for artificial gestures. The proposed framework allows the user to activate and control a robot-mounted stereoscope, is resistant to noise such as blinking, and safely allows the surgeon to look away to talk to an assistant before resuming control of the camera at any point. Furthermore, the proposed framework takes advantage of the synergy between eye tracking and stereo reconstruction to obtain the 3D coordinates with regards to the robot base of the points fixated by the user. Accurate stereo reconstruction algorithms are computationally intensive, such that a compromise between accuracy and speed must typically be made for realtime applications. However, by using eye tracking to aim the stereo reconstruction around an area of an interest, accurate algorithms can be used with sufficient speed. This is in contrast to [8] and [9] in which 3D coordinates are obtained by assuming that the left and right eye pixel coordinates from binocular eye tracking represent valid feature pairs. The proposed system uses this synergy in a novel method to control the zoom of the camera so that optimal camera-to-tissue viewing distance is maintained. In addition, the proposed system takes further advantage of this synergy to perform automatic laser targeting with a second robotic arm for laser ablation. The laser calibration technique developed to achieve this with millimetre precision, without visual servoing, is also presented. Finally a user study

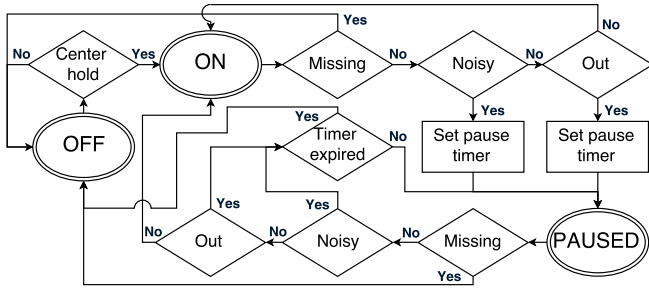


Fig. 1. Intention recognition state diagram.

involving the entire framework is presented, validating the usability of the system compared to a manually controlled robotic system.

II. GAZE CONTROLLER

The aim of the proposed gaze framework is to provide a tool capable of understanding whether the user is trying to interface with the system or not, and to provide the safest and least frustrating experience possible while interfacing with the system. Ideally the system should feel completely transparent to the user: it should activate as soon as the user wants to interact with it, disable safely if the user switches their attention away from the system, and be robust enough to disregard minor natural nuisances such as blinks or brief glances. While controlling the camera, the system should provide enough speed to prevent frustration when the surgeon wants to navigate, without sacrificing stability while performing a static task. The reason for placing such emphasis on the transparency of the system from the user perspective stems from the nature of eye data: while natural gaze behaviour is conducted everyday, artificial constraints as simple as forcing the user to fixate on a single point for a length of time feel distinctively unnatural and can become rapidly tiring. This section divides the eye tracking framework into three parts: the intention recognition is presented first, followed by the velocity controller used to direct the panning motion of the stereoscope, and finally the use of the stereo reconstruction to perform the zooming motion.

A. Intention Recognition

The intention recognition algorithm attempts to answer the following two questions:

- Is the user trying to interact with the system?
- Is the eye tracking data reliable enough for robotic control?

This question is processed using a state engine that updates with every cycle of the eye tracker. In this paper, a Tobii Pro X2-60 (Tobii Technology AB) with an update rate of 60 Hz was used. Every update cycle, the algorithm analyses multiple past sequences of eye data and settles in one of three possible states: ON, OFF, and PAUSED. The camera controller is active in the ON state and disabled in the OFF state. The PAUSED state is triggered when incoming data is deemed unreliable enough for the system to determine

whether the user is still trying to interact with it or not, or whether the data quality is no longer sufficient for robotic control. In this state the camera controller is temporarily suspended, without bringing the camera to a full stop, while the system waits for more data cycles to choose the subsequent state. The state diagram of the algorithm is displayed in Figure 1.

Circular buffers store all the raw data coming from the eye tracker for the last three seconds, more than required by the processing. Every transition in the state diagram requires the analysis of sequences of the stored data, to identify whether or not specific criteria are met. At the beginning of every update cycle a series of these checks is run, as shown in Figure 1. The results of these checks dictate which state the system achieves at the end of the update cycle. The different conditions and numerical values for the criteria are shown in table I. These values were obtained and fine-tuned using results from over 20 participants, over multiple iterations of the system. Each criterion possesses a tolerance threshold and sequence length. In order for a criterion to be met, the samples within the corresponding time-sequence have to satisfy the criterion's condition within the tolerance limit. For example, for the *center hold* criterion to met, no more than 15% of the samples from the previous 700 milliseconds should be outside a sphere centred in the middle of the screen with a diameter of 0.2 in normalised coordinates (normalised values range from 0 to 1 from one edge of the screen to the other).

The system starts in the OFF state and is activated by looking at the center of the screen. If the gaze is maintained in the center part of the screen long enough to validate the *center hold* criterion, and the data connection is stable, the system is activated. Once activated, the system verifies the stability of the eye data every cycle. The first check, named *missing*, is to ensure that the connection with the eye tracker is stable, and the received number of data points matches what is expected based on the update frequency. In practice this check will only come into effect when the tracking is experiencing critical difficulties. It should not occur even when the user disengages from the system, and

	Center Hold	Missing	Noisy	Out
Success Condition	Gaze point is inside centred sphere of normalised radius 0.2	Received samples match expected number of samples	Sample has one or both eyes visible	Gaze point is within screen margins
Threshold Tolerance	15%	15%	22%	90%, margins: 5%
Sequence Length	700 ms	300 ms	300 ms	300 ms

TABLE I

CRITERIA FOR THE EYE DATA ANALYSIS.

is only meant as a backup safety to quickly deactivate the system without passing through the PAUSED state in case something obstructs the eye tracker or the connection is lost. The next check verifies the level of noise on the eye tracking data. This refers to data obtained through a stable connection, but where the eyes are not found by the eye tracker. In this work, the eye data was considered valid if even only one eye was found. If the signal is considered noisy by failing the *noise* check, the system proceeds to the PAUSED state. Finally, the last check analyses whether the user is looking within the screen boundaries. A 5% hysteresis is allowed to accommodate for tracking imprecision around the edges of the screen, as well as a large tolerance limit. If either the *noise* or *out* check fails, the system moves to the PAUSED state. In that state, all three checks *missing*, *noise*, and *out* are performed every data cycle. If the system passes all checks before a certain time has elapsed, the system switches back to the ON state. If the *missing* test is failed, or if the timer expires, the system switches to the OFF state. The duration of the timer depends on the nature of the check that triggered the PAUSED state. If triggered by a failed *noise* check, the pause timer is 500 milliseconds. In the case of a failed *out* check, the timer is 300 milliseconds. The shorter value of the timer for the failed *out* check is due to the high velocities that can be generated from eye data at the edges of the screen, as explained in the following section.

The values presented here were optimised to reflect the intended capabilities of the system. With a large portion of the screen available to start the control, the system is easily activated by looking approximately at the center of the screen for just over half a second. In the event that the system is unintentionally activated, the noise and boundary checks will trigger a pause of the system within a few hundred milliseconds. Examples of the system behaviour in response to eye tracking signals are shown in Figure 2.

B. Panning Velocity Control

A remote centre of motion (RCM) was implemented to force the stereoscope to respect the trocar position. The subsequent panning motion around the RCM was implemented using an angular velocity controller. When the system is enabled (ON state), the eye data is used to generate two-dimensional vectors. The centre portion of the image is a zero velocity zone, any points within this zone will only slow down the camera. Points outside of this zone will generate a vector as shown in Figure 3.

The resulting vector is normalized according to screen dimensions, and scaled proportionally to the maximum allowable angular speed as shown in equation 1:

$$V_{2D} = \frac{Va_{max}}{(0.5 - R)} \cdot \left(V_t - R \cdot \frac{V_t}{norm(V_t)} \right), \quad (1)$$

where V_t is the vector obtained from Figure 3 in centred normalized screen coordinates (values ranging from -0.5 to 0.5, where 0 is the center of the screen), R is the radius of the zero velocity zone in the same coordinates, and Va_{max} is the maximum reachable angular velocity. The end result is

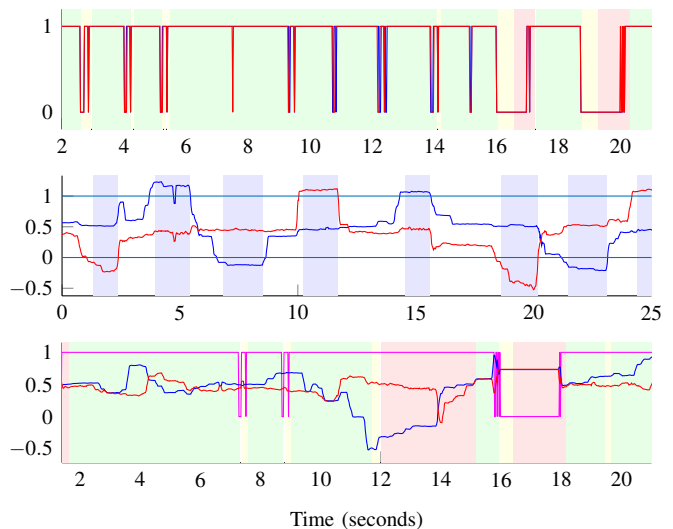


Fig. 2. Top: blue and red lines respectively represent whether the left and right eye are visible (value of 1) or not (value of 0). Background colors represent the state of the sytem: red = OFF, yellow = PAUSED, green = ON. Center: blue and red lines represent the normalized X and Y screen coordinates of the average gaze point (averaged between the left and right eyes, on-screen values are between 0 and 1, with 0.5 being the center of the screen). Background color represent whether the user is looking inside the screen limits, white = within limits / *Out* check succeeded, blue = outside limits / *Out* check failed. Bottom: blue and red lines same as center figure. Magenta line shows if either one of the two eyes is detected (values same as top figure). Background colors same as top figure.

transformed to a rotation vector with non-zero components along the camera X(pitch) and Y(yaw) axis. This rotation vector is then set as the desired velocity for a 3 degree of freedom velocity planner. Every control cycle of the robotic arm, the current rotation vector is obtained from the velocity planner and used to update the robot position and orientation accordingly.

This approach presents several advantages compared to [6] and [7]. The use of task space velocity control offers more intuitive movements than those of a joint by joint position

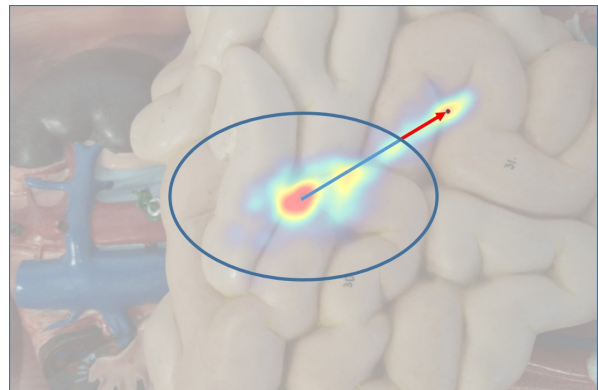


Fig. 3. Generation of velocity vectors based on eye data. The blue ellipse represents the zero velocity zone. The red arrow represents the two-dimensional vector used to generate desired angular velocities. The overlaid heatmap shows typical eye motions in this direction.

control based on image quadrants. However, unlike in [7] no discrete velocity zones are used, allowing for continuous transitions between desired velocities. The capacity to enforce desired angular accelerations also further increases the controllability of the system, such as by allowing faster braking when the system enters the OFF state.

Coupled with the intention recognition presented in section II-A, the velocity planner offers an intuitive and robust platform for the control of the stereoscope. Due to the generation of angular velocity and acceleration profiles, unintentional system activation will typically not result in any perceptible movement before the system pauses the control.

C. Eye Tracking and Stereo Reconstruction Synergy

In this paper, eye tracking is combined with stereo reconstruction to transform two-dimensional screen coordinates into corresponding 3 dimensional robot base coordinates. The proposed system uses the stereo reconstruction described in [10]. While the running time of this algorithm is typically over five seconds for high definition images, the input image size can be considerably reduced by considering only the area surrounding the gaze fixation point. Combined with image scaling, this allows the system to run the targeted stereo reconstruction algorithm at 500 ms per frame. This update rate is largely sufficient for the requirements of the proposed system, as the robot planning allows for smooth, non-jerky movements.

The information obtained from the stereo reconstruction is used to update a proportional controller adapting the translation of the scope along its optical axis (zoom). In the proposed system the controller tried to maintain a tissue-to-camera viewing distance of 60 millimetres. This distance was chosen based on the focal length set on the calibrated stereoscope during the studies.

III. SYSTEM SETUP AND CALIBRATION

In the proposed system, a stereoscope was mounted on a Kuka iiwa (KUKA Roboter GmbH) 7 degrees of freedom redundant robotic arm. The stereoscope was calibrated using standard camera calibration techniques, yielding a root mean square (RMS) reprojection error of under 0.2 pixels. A dual-quaternion hand-eye calibration method was implemented based on [11], and used to find the transformation between the camera frame of reference and the robot base frame. The RMS error for the hand-eye calibration was submillimetre. In order to simulate a surgical laser, a laser diode was placed inside a custom holder, mounted onto a second Kuka iiwa, and controlled via an Arduino Teensy 3.2. A picture of the complete system is shown in Figure 4.

The two robotic arms are mounted on optical tables screwed to each other and so are considered statically linked. The transformation between their respective bases was determined by screwing both of their end effectors rigidly to a metal board. The robots are then placed into a compliant mode and moved to cover as much of the workspace as possible. Let K_1 be the transformation from the robot base to the end effector for the first robot at a sampling time t , K_2 ,

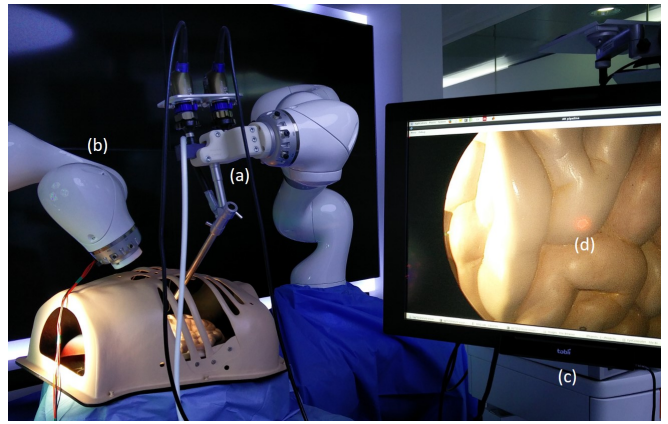


Fig. 4. System overview. (a) Robot 1 holding the stereoscope, (b) Robot 2 holding the laser diode, (c) Tobii X2-60, (d) View of the laser dot on the screen.

for the second robot, X the static transformation matrix from the end effector of the second robot to the end effector of the first robot, and L the transformation from the first robot base to the second robot base. It is then possible to write

$$K_{2,t}^{-1} \times K_{2,t} \times X = X \times K_{1,t}^{-1} \times K_{1,t}, \quad (2)$$

which is the standard formulation of the hand-eye calibration problem $AX = XB$. By solving for X using the hand eye calibration method mentioned previously, we can then solve for L in each of the previous measurements:

$$L = K_{1,t} \times X^{-1} \times K_{2,t}^{-1}. \quad (3)$$

Multiple different measurements of L can then be combined to minimize the overall error. The proposed system was calibrated with over 4000 measurements, and resulted in an RMS error of under 0.2 millimetres.

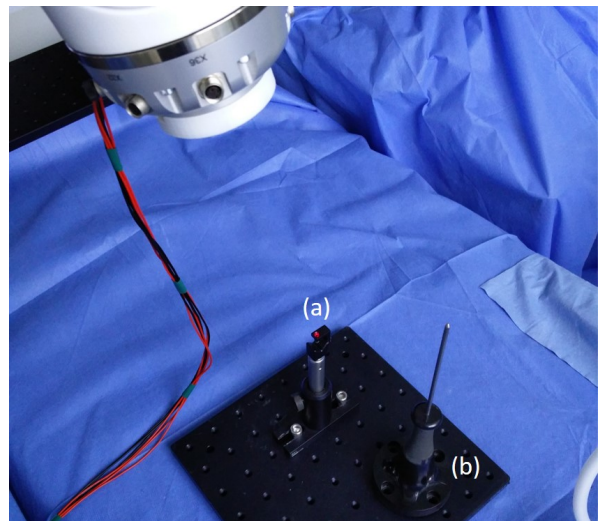


Fig. 5. Laser calibration setup. (a) Pivot point for the tool and the laser, (b) Pivoting tool used to determine the base frame coordinates of the pivot point.

The complete system was first tested by firing the laser held by the second robot at a 3D reconstructed point obtained from the stereoscope held by the first robot. It was initially thought that the laser could be assumed to fire straight along the end effector direction of the robotic arm. However, this assumption quickly proved to lead to a significant loss of accuracy when the laser was fired at a distance further than a few centimetres. This was due to imperfections in the laser manufacture, as well as in the mount holding the laser diode. As such, a method was designed to compensate for the fact that the laser did not fire precisely along the Z axis of the robot end effector. Figure 5 shows the setup used to calibrate the laser direction.

The calibration of the laser direction is a two step process. The initial step consists of calibrating a pointed tool attached to the second robot's end effector (figure 5b) to find the coordinates of the pivot point in the base frame of the robot (figure 5a). The second step consists of mounting the laser on the second robot and manually aligning it with the pivot point in multiple different poses. Let K_{2_t} be the transformation from the robot end effector to the robot base at sampling time t , X_{laser} the transformation from the robot end effector to the tip of the laser diode, P the position of the pivot point in the robot base frame, and V_t the translation from the laser tip to P . It then follows that:

$$V_t = X_{laser}^{-1} \times K_{2_t}^{-1} \times P. \quad (4)$$

The firing direction of the laser is then obtained by normalizing V_t and averaging it over multiple measurements. While taking into account the laser direction substantially improved the accuracy of the system, a remaining error of up to 5 millimetres in worst cases persisted. It was hypothesized that this error was due to the accumulation of calibration errors spread throughout the system. Indeed, while all individual calibration errors were kept low, the full kinematic chain of the system involved a total of six possible sources of error, not counting the robot encoders:

- The calibration of the stereo camera.
- The stereo reconstruction.
- The camera to robot 1 end effector hand-eye calibration.
- The robot 1 base frame to robot 2 base frame calibration.
- The laser mount hand-eye calibration.
- The laser firing direction.

One final calibration step was performed to attempt to model the residual error on the entire system. The same asymmetric circles calibration grid used for the camera calibration was placed in a fixed position. The circles were identified, triangulated, and their 3D coordinates brought into the base frame of the robot 2 using the stereoscope. Following this, the 3D coordinates of the points in the calibration grid were determined separately using the second robot holding the laser laser. This was performed by first aiming the laser at each of the four corners of the grid. Multiple measurements were made for every point, each resulting in the equation of a 3D line containing that point. The coordinates of each corner point on the calibration grid

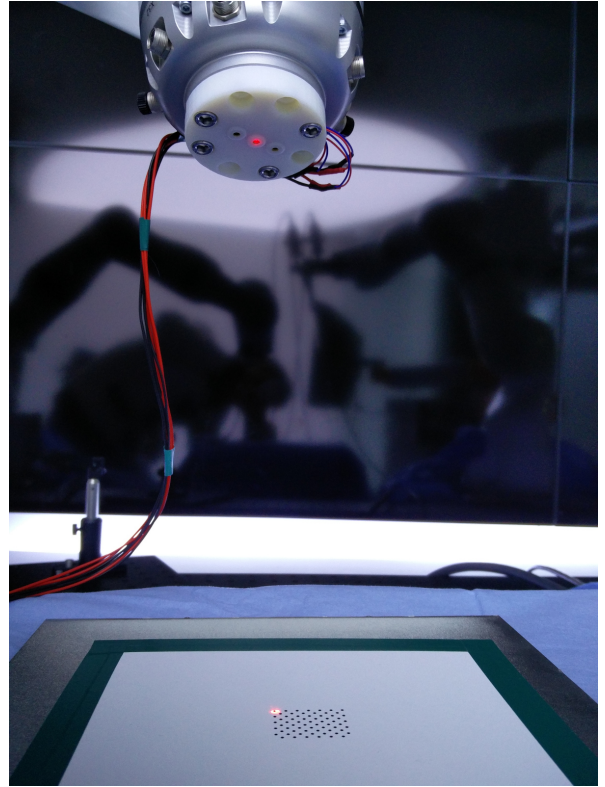


Fig. 6. Error compensation by projection of the laser on the calibration grid.

were then obtained by solving for the point of minimum distance between all of the corresponding lines. Figure 6 shows a picture of this compensation step.

Due to the manual adjustment of the laser on the points as well as the size of the laser dot itself, the laser pivot calibration can present submillimetre inaccuracies. The quality of these measurements were further improved by spreading the measurement error evenly across the four points. This was achieved by first finding the centre point P_c of the four points and then computing the vectors V_1 , V_2 , V_3 and V_4 from the P_c to each of the four corners. Figure 8 illustrates this process.

The vectors V_x and V_y along the directions of the grid are then found by summing up the corner vectors as follows:

$$V_x = \frac{V_3 + V_4 - V_1 - V_2}{2}, V_y = \frac{V_3 + V_2 - V_1 - V_4}{2}, \quad (5)$$

From there, 3D coordinates of all the points in the calibration grid in the robot 2 base frame can be rebuilt using P_c , V_x and V_y . The outcome of the previous sequence is two sets of points supposed to represent the same values. Using singular value decomposition, the transformation T_{error} can be computed that minimizes the difference between the two sets of points. Applying this transformation matrix to the system resulted in bringing the error of the entire system down to below two millimetres. This was a critical threshold to reach as the limiting accuracy factor was no longer the

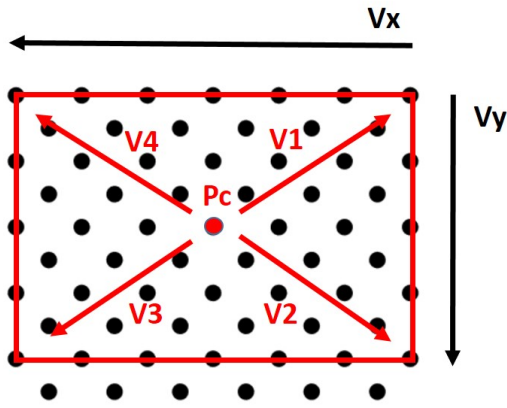


Fig. 8. Minimisation of the measurement error through reconstruction of the ideally spaced grid points.

system, but the eye tracking values.

IV. USER STUDY

In order to evaluate the system, an experiment was designed in which participants had to control both the stereoscope and the laser to mark four target points spread out in a phantom model at varying heights. The purpose of this study was to assess the usability of the system by determining whether it presented any improvements in terms of speed and ease of use over a manually controlled system. Participants had to bring each target point within the centre of the field of view of the stereoscope, and position the laser on top of the point. The task was performed twice for each participant, once with the eye-controlled system and once with manual control over the robotic system. The order of the tasks was randomized between participants. Manual control of the system was performed via keyboard input.

In the manual mode, arrow keys and the number pad of the keyboard were used to control the laser position and the stereoscope position, respectively. The movement of the

stereoscope controlled by the arrow keys was the panning motion around the RCM, although participants could ask to zoom in or out. In the event that participants achieved a tissue-to-camera distance of over 90 millimetres the system was zoomed in for them to prevent unrealistically high rotation angles around the RCM.

For the eye-controlled mode participants would perform an eye tracking calibration prior to the task. During the task they would fixate each target point until they felt they had reached a stable position, at which point they would activate the laser. If the laser was on target, the participants were free to continue to the next target. Otherwise the eye control would be paused and the user would adjust the laser on the target manually.

Participants were shown all the target points and trained once with each control mode before the experiment. Lastly, participants were made aware that they should attempt to minimize the amount of time the laser was on if possible. In a real case a laser pointer could be used to indicate where the laser is going to fire, but it would still be distracting for the surgeon to have to constantly keep the laser pointer on. A total of 8 participants, 5 male and 3 female, took part in the study.

The different metrics considered were:

- The relative time taken to reach each target point.
- The total task time.
- The total time with the laser activated.
- The tissue-to-camera distance.
- In eye control mode, the amount of time the system had to be disabled to manually adjust the laser.

Figures 7 and 9 display the results of the experiment. The relative time to reach each point shows significant improvement for all points except the last one. This can be explained by the fact that the two last points were close to each other such that it required only minimal movement from the stereoscope and the laser to reach it. Even then, the median time to reach that point is noticeably shorter for the eye controlled system. The overall task time shows a

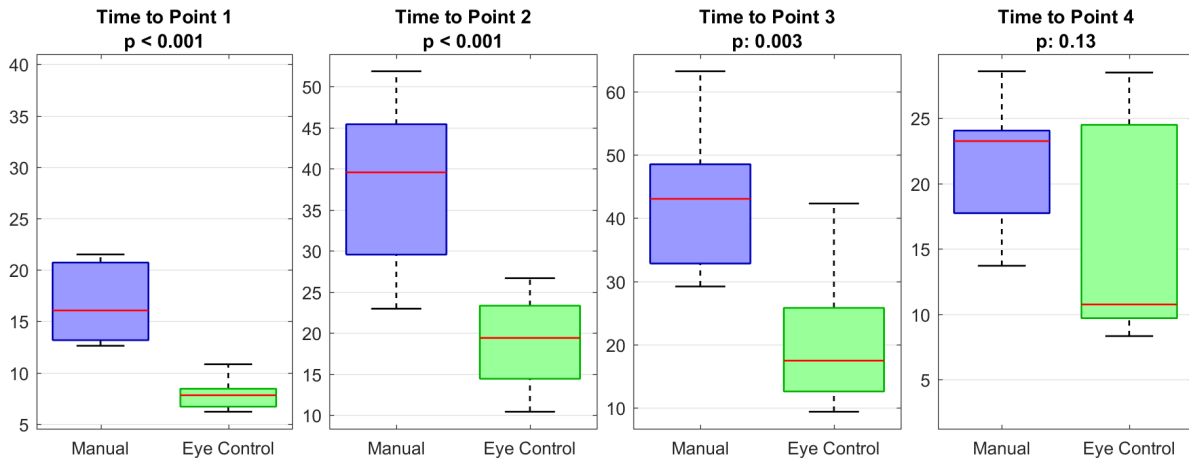


Fig. 7. Experimental results(N=8): Relative time to reach each target point(seconds).

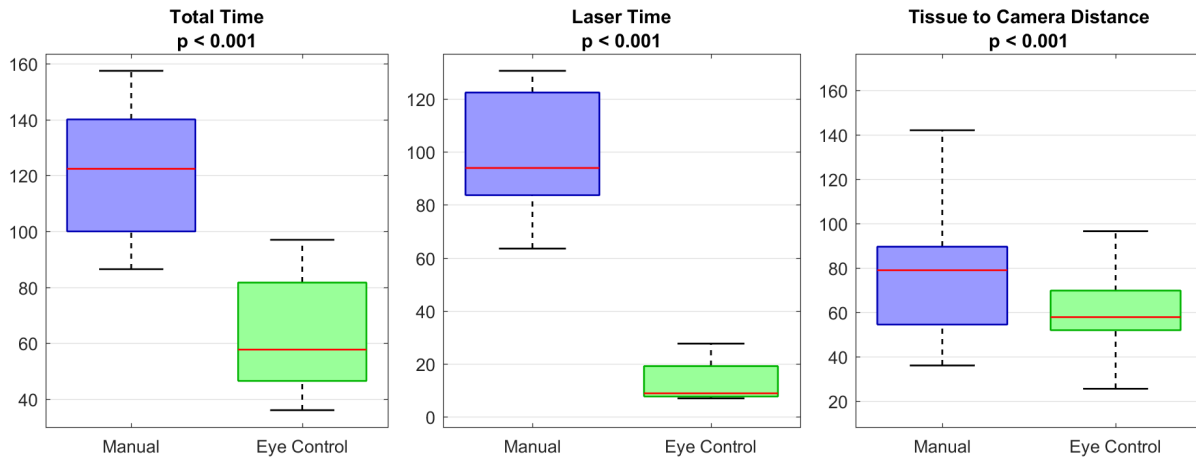


Fig. 9. Experimental results(N=8): Total task time(seconds), total laser activation time(seconds), and tissue-to-camera distance(millimetres, 30000 data points).

clear improvement in usability with median task times over twice as long for the manually controlled system than for the eye-controlled system. The total laser activation time is unsurprisingly significantly higher with the manual system, as it is difficult for participants to assess where the laser is going without visual feedback. The tissue-to-camera distances show that the stereo-guided zoom controller managed to stabilize the system around the desired 60 millimetres, as well as showing that the manually controlled system typically overshoot that distance by 33%. Lastly, the mean time across all participants that the eye controller had to be turned off to manually position the laser dot on the target was 2.87 seconds. This error can be mostly attributed to the quality of the eye tracking calibration prior to the experiment.

V. CONCLUSION

This paper presents a novel framework for intention recognition in eye tracking applications. By using only natural gaze behaviour, users are capable of activating the proposed system without any artificial gestures. Furthermore, the system is robust to noise, can recover from brief interruptions, and possesses redundant layers of built-in safety to prevent it from unintentionally moving in an uncontrolled manner. Additionally, the proposed system makes full use of the synergy between eye tracking and stereo reconstruction to provide an automated zoom functionality to the stereoscope, as well as automatic laser targeting. The results obtained from the user study conducted clearly demonstrate the potential of this approach, and the speed gains that can be made using eye tracking. As shown, eye tracking data lends itself well to use with other sources of information. Future work could improve on the presented system by taking into consideration other information available in the surgical environment, such as the position of the instruments to improve the stereoscope and laser motions. The user interface could also benefit from information regarding the surgical context to fine-tune the behaviour of the system for different parts of the procedure.

ACKNOWLEDGMENTS

The authors would like to thank Petros Giataganas for his assistance in building the system, and Lin Zhang for his help with the stereo reconstruction.

REFERENCES

- [1] Zheng, B., Rieder, E., Cassera, M., Martinec, D., Lee, G., Panton, O. N., et al. (2012b). Quantifying mental workloads of surgeons performing natural orifice transluminal endoscopic surgery (NOTES) procedures. *Surgical endoscopy*, 26(5), 1352-1358.
- [2] Nduka, C. C., Super, P. A., Monson, J. R., Darzi, A. W. (1994). Cause and prevention of electrosurgical injuries in laparoscopy. *Journal of the American College of Surgeons*, 179(2), 161-170.
- [3] Kommu, S., Rimington, P., Anderson, C., Rane, A. (2007). Initial experience with the EndoAssist camera-holding robot in laparoscopic urological surgery. *Journal of Robotic Surgery*, 1(2), 133-137.
- [4] Kraft, B. M., Jager, C., Kraft, K., Leibl, B. J., Bittner, R. (2004). The AESOP robot system in laparoscopic surgery: Increased risk or advantage for surgeon and patient? *Surgical Endoscopy And Other Interventional Techniques*, 18(8), 1216-1223.
- [5] Sriskandarajah, K., Gillen, S., Marco, A. D., Sodergren, M., James, D. R. C., Clark, J., et al. (2012). Robotic Tele-Manipulating Devices for Laparoscopy Improve Surgical Performance in Simulated Porcine Laparoscopic Cholecystectomies on the ELITE Simulator. *Proceedings of the 2012 Hamlyn Symposium on Medical Robotics*, 36-37.
- [6] Noonan, D. P., Mylonas, G. P., Shang, J., Payne, C. J., Darzi, A., Yang, G.-Z. (2010). Gaze Contingent Control for an Articulated Mechatronic Laparoscope. Paper presented at the IEEE International Conference on Biomedical Robotics and Biomechanics.
- [7] Fujii, K., Salerno, A., Sriskandarajah, K., Kwok, K.-W., Shetty, K., Yang, G.-Z. (2013, 3-7 Nov. 2013). Gaze contingent cartesian control of a robotic arm for laparoscopic surgery. Paper presented at the IEEE/RSJ International Conference on Intelligent Robots and Systems.
- [8] Mylonas, G. P., Darzi, A., Yang, G.-Z. (2006) Gaze-contingent control for minimally invasive robotic surgery. *Computer Aided Surgery* 11, no. 5: 256-266.
- [9] Noonan, D. P., Mylonas, G. P., Darzi, A., Yang, G.-Z. (2008) Gaze contingent articulated robot control for robot assisted minimally invasive surgery. In *Intelligent Robots and Systems, 2008. IROS 2008. IEEE/RSJ International Conference on*, pp. 1186-1191. IEEE.
- [10] Yamaguchi, K., McAllester, D., Urtasun, R. (2014) Efficient joint segmentation, occlusion labeling, stereo and flow estimation. In *Computer Vision-ECCV 2014*, pp. 756-771. Springer International Publishing.
- [11] Malti, A., Barreto, J. P. (2010) Robust hand-eye calibration for computer aided medical endoscopy. In *Robotics and Automation (ICRA), 2010 IEEE International Conference on*, pp. 5543-5549. IEEE.