MUTHUKRISHNAN, R., KANNAN, S., PRABHU, R., ZHAO, Y., BHOWMICK, P. and HASAN, M.J. 2023. Tracking and estimation of surgical instrument position and angle in surgical robot using vision system. *In Proceedings of the 2023 Network, multimedia and information technology international conference (NMITCON 2023), 1-2 September 2023, Bengaluru, India*. Piscataway: IEEE [online], article number 10275983. Available from: <u>https://doi.org/10.1109/NMITCON58196.2023.10275983</u>

Tracking and estimation of surgical instrument position and angle in surgical robot using vision system.

MUTHUKRISHNAN, R., KANNAN, S., PRABHU, R., ZHAO, Y., BHOWMICK, P. and HASAN, M.J.

2023

© 2023 IEEE. Personal use of this material is permitted. Permission from IEEE must be obtained for all other uses, in any current or future media, including reprinting/republishing this material for advertising or promotional purposes, creating new collective works, for resale or redistribution to servers or lists, or reuse of any copyrighted component of this work in other work.



This document was downloaded from https://openair.rgu.ac.uk SEE TERMS OF USE IN BOX ABOVE

Tracking and Estimation of Surgical Instrument Position and Angle in Surgical Robot using Vision System

¹Ramkumar Muthukrishnan School of Engineering Robert Gordon University Aberdeen, UK r.muthukrishnan@rgu.ac.uk

School of Engineering Robert Gordon University Aberdeen, UK s.kannan1@rgu.ac.uk

²Somasundar Kannan

⁵Parijat Bhowmick Electronics and Electrical Engineering Department Indian Institute of Technology Guwahati, India parijat.bhowmick@iitg.ac.in

Abstract-A da Vinci robot endoscopic-camera gives surgeons a magnified 2D view of the operating area, but additional time is required to detect and estimate the location of the surgicalinstrument during an operation. The main focus and novelty of this research is to develop a new virtual coloured marker-based tracking algorithm for estimating the posture and orientation of the instrument. Initially, the developed algorithm begins by determining the coloured area of the instrument as referencecontour. Followed by a new Virtual-Rotating Bounding Rectangle (V-RBR) created over the reference-contour by meeting the minimum area of contour criteria. Additionally, a new Virtual Dynamic Multi-line Crossbar and a Virtual Static Graph (VDMC-VSG) was constructed to trace the movement of V-RBR, which helps to estimate the pose and angle of the targeted instrument in 2D during observations. V-RBR is considered as virtual coloured marker, it avoids ambient illumination-related difficulties. The proposed approach performed excellently in Gazebo-simulation and the overall accuracy is 91.3 % obtained by comparing with Robot Operating System (ROS)-based Transform measuring system, which uses robot kinematics.

Index Terms—Image Processing, Robot Operating System, Gazebo simulation, Surgical robotics

I. INTRODUCTION

Robotics is at the heart of current medical engineering. For example, robotic surgery has been one of the most significant technological developments in medical realm during the past 20 years. Similar to this, it is projected that in the next few years, the field of Robotic-Assisted Minimally Invasive Surgery (RAMIS) will grow greatly. The exceedingly complex

Surgery (RAMIS) will grow greatly. The exceedingly complex medical discipline known as (RAMIS) includes eye-hand coordination difficulties, restricted mobility, and a restricted field of vision especially during laparoscopic surgery. However, there are still many challenges. Suturing is one of the most difficult tasks that surgeons felt during minimally ³Radhakrishna Prabhu School of Engineering Robert Gordon University Aberdeen, UK r.prabhu@rgu.ac.uk

> ⁶ Md Junayed Hasan National Subsea Centre Robert Gordon University Aberdeen, UK j.hasan@rgu.ac.uk

invasive operations. An autonomous surgical robot must be able to monitor and localise the tool during suturing. But inaccuracies in tool trajectories occurs by robot or camera miscalibration. Thus, it might cause tool slips and incorrectly guide the tool, potentially leading to operation failures. Poor tool trajectory, particularly when a tool manipulation, could result in additional organ damage or leads to fail the suture operation. Related work for this issue is discussed in section II.

⁴Yafan Zhao

School of Engineering

Robert Gordon University

Aberdeen, UK

y.zhao@rgu.ac.uk

The remaining parts of the paper are structured as follows: section III presents the comprehensive development of the proposed method, which is divided into three subsections and accompanied by corresponding experiments. The analysis of the experimental results is provided in section IV, while section V serves as the conclusion of the paper.

II. RELATED WORKS

The suturing technique cannot be carried out autonomously unless the surgical needle is identified and monitored throughout the task. Therefore, many approaches for monitoring surgical tool has been put forth during the past few decades. To begin with the early stage, robot kinematics, computer vision, and Bayesian state prediction were utilised to develop a particle filter to track a surgical needle through stereo endoscopic images [1]. Later that, a method was developed to register the position of needle with respect to the needle holder while accurately locating the instrument, based on how an endoscopic camera is used [2].In addition, refined and coarse detection modules were combined to develop a convolutional neural network model for surgical instrument identification in minimally invasive surgery videos [3]. A reinforcement learning-controlled single- phase surgical tool identification framework was constructed based on a sample adaptive approach, as this process maintains higher

precision and constant speed [4]. Thereafter, a faster R-CNNbased modified Anchoring Network was prepared to identify the instrument during the key-hole operation [5]. A geometric object descriptor-based multiple tool tracing architecture was introduced for limited surgical tool data sets [6].

Moreover, Kalman filter was suggested to combine sensor data with robot kinematics in order to identify and track needles [7]. Recently, a modified TernausNet-16 network model was proposed to track surgical tools automatically [8]. Thereafter, a fine tuned InceptionResnetV2 model was developed for automatic instrument detection during operations [9]. An improved pattered-based marker was designed to trace a circular-type surgical instrument [10]. Through this research study, it has been concluded that a simple approach is required, which is detailed in section III.

III. PROPOSED METHOD

In this work, a simplified, robust approach is presented to trace and estimate the posture and angle of the instrument in 2D during suturing tasks. The main contribution of this work is developing an algorithm framework with;

- Virtually created a Virtual-Rotating Bounding Rectangle (V-RBR) over the instrument, which can be considered as a Virtual Marker (VM).
- Additionally, a Virtual Dynamic Multi-line Crossbar with a Virtual Static Graph (VDMC-VSG) constructed virtually, which tracks the instrument in 2D view.

A detailed explanation is given in three sub sections about the algorithm flowchart shown in Fig. 1.

A. First stage of processing

When the node (algorithm) starts, it initiates to extract each frame separately from live video data, which is obtained from camera in Gazebo simulator. During the start of the processing phase, each frame is converted from RGB to HSV. Merging the hues of red, green, and blue leading to the creation of the RGB colour model. HSV is the result of combining hue, saturation, and value, or brightness. Hue, which is described as an angle with a dimension of $0 - 360^{\circ}$, symbolises the nature and characteristics of a hue. Saturation is the technical term for colour intensity. The colour's contrast, often known as brightness or a representation of value, means that moving from RGB to HSV frames enables better processing to identify instruments in diverse lighting conditions, shadows, etc. The thresholding strategy was applied to select the tool's colour from the HSV frame in order to construct a binary frame. The targeted objects are shown in white with a black backdrop scene.

B. Second stage of processing

Thereafter, coming to the next processing phase, binary frames are helpful for measuring the first contours of the targeted object since they only capture the tool's area and keep them in a first loop for later processing. If the instrument region is more than 90 pixels within the first loop then V-RBR is created over the instrument after obtaining centroid coordinate for the instrument. The centroid coordinates for the instrument, from the centre of mass, X_c and Y_c , has been determined in (1) and (2);

$$X_c^1 = \frac{\sum_{i=1}^{K_1} X_i^1}{K_1} \tag{1}$$

$$Y_c^1 = \frac{\sum_{i=1}^{K_1} Y_i^1}{K_1}$$
(2)

where K_i is the total number of pixels that match the instrument, and X_i , Y_i are the location value of i^{th} pixel in x, ydirection on Screen Plane. Technically, V-RBR is considered a static, non-scale-changeable Virtual Marker (VM), which helps to avoid zooming in and out issues. Then it moves on to the further stage of processing after constructing a VM over the tool. Here, it follows the actions that were taken in the initial stage of processing, including transforming the RGB to HSV frame, choosing the VM's colour from the HSV frame, obtaining binary frames using the threshold approach, and grabbing the VM by treating it as a second contour to keep in the second loop for final phase of processing.

C. Final stage of processing

Final processing phase begins with calibrating the VM position if the VM area is greater than 100 pixels. An 8-bit RGB camera with a 640 x 480 resolution was used in this research. Due to this, the position of the VM within the Screen Plane always appears in pixels and the VM coordinates are obtained using (3) and (4);

$$X_c^2 = \frac{\sum_{i=1}^{K_2} X_i^2}{K_2}$$
(3)

$$Y_c^2 = \frac{\sum_{i=1}^{K_2} Y_i^2}{K_2}$$
(4)

After obtaining centroid coordinates for VM, pixel-tocentimetre conversion is determined as in (5) and (6):

$$S_{sw} = \frac{1 - VM_{sw(pixel)} * 2.54}{100} = 0.0245cm \tag{5}$$

$$S_{sh} = \frac{1 - VM_{sh_{(pixel)}} * 2.54}{100} = 0.0245cm \tag{6}$$

The height and width of the screen are represented, respectively, by S_{sh} and S_{sw} . The locations of the VM on the tool in the screen's height and width are represented by VM_{sh} and VM_{sw} , respectively. The Screen Plane (SP) to the Global Plane (GP) x and y axis coordinates has been calibrated using (7) and (8);

$$S_{sw_{(SP-GP)}} = -(0.0036 * (S_{sw}^2)) + (0.7938 * (S_{sw})) - 0.0403$$
(7)
$$S_{sh_{(SP-GP)}} = -(0.0068 * (S_{sh}^2)) + (0.0395 * (S_{sh})) - 0.4439$$
(8)

The VM on the instrument's position is calibrated to the GP's x and y-axis coordinates, which are represented by the $S_{sw_{(SP-GP)}}$ and $S_{sh_{(SP-GP)}}$, respectively. Following this, a vertical and horizontal dynamic line were developed virtually



Fig. 1. Algorithm flow chart of proposed approach with three stage of processing for developing the Virtual-Rotating Bounding Rectangle and Virtual Dynamic Multi-line Crossbar with a Virtual Static Graph.

to track the posture of the VM on the instrument in 2D. Additionally, a Virtual-Fit Line (V-FL) was created to estimate the angle of the instrument through VM, which was also converted from radian to degree as follows;

$$\hat{U}_{V-FL} = \frac{\vec{V}_{V-FL}}{|M_{V-FL}|} \tag{9}$$

$$\hat{U}_{V-FL} = [X_{V-FL}, Y_{V-FL}]$$
 (10)

$$y_{axis} = [0, 1] \tag{11}$$

$$DP = [y_{axis}] \cdot [\hat{U}_{V-FL}] \tag{12}$$

$$R_{SP} = \cos^{-1} * (DP) = Radian \tag{13}$$

$$R_{SP} = \left(roll_{SP} * \frac{(180)}{\pi}\right) = Degree \tag{14}$$

The following characters (V_{V-FL} , $|M_{V-FL}|$, \hat{U}_{V-FL} , DP, R_{SP}) are representation of vector's direction, vector's magnitude of displacement, unit vector for V-FL, dot product and rotational angle value of instrument in SP. Finally, a Static Virtual Graph (SVG) is constructed, which visually represents the x and y axes directions. After processing those three phases, followed by an illustration of the produced algorithm's result, as shown in Fig. 3. Here, the combination of virtual horizontal, vertical line and V-Fl is represented as Virtual Dynamic Multi-line Crossbar. This VDMC-VSG aids to trace and predicts the precise position and angle of tool in 2 Dimension. Thus, tasks performed by the surgeon during a surgery are reduced.

D. Experiment

A customised surgical robot model with one Endoscopic Camera Manipulator (ECM) and Patient Side Manipulator (PSM) was created and imported to the Gazebo simulation platform with the setup environment depicted in Fig. 2 to assess the effectiveness of the proposed approach. Dilation and erosion blur filters have been used to assess the area of contour size under various depth ranges of the surgical tools in dynamic pictures, as indicated in the Table. I.

 TABLE I

 Testing the contour range under different blur scenes and depth ranges

Type of Blur	Zoom range (cm)	Area of contour
Erosion	140	91
Erosion	120	98.5
Erosion	70	125.5
Dilation	140	91
Dilation	120	98
Dilation	70	126.5

This test shows that for each blur filter, the resulting area of contour size is not constant at various depth ranges of surgical tools. This is a crucial justification for building a non-scale-changeable VM over the surgical tool. It acquires uniform, constant pixel values. Finally, using this different approach helps to monitor the instrument in 2D and estimate its pose and angle. The whole procedure is described in the preceding section, and the results are displayed in Fig. 3.

IV. VALIDATION AND RESULT

To validate the proposed work, it is required to conduct a comparative analysis between proposed work with existing research work for validation purpose. However, the given



Fig. 2. Customised surgical robot model in simulation environment with setup.



Fig. 3. Visual output of Virtual Dynamic Multi-line Crossbar with a Virtual Static Graph and Virtual-Rotating Bounding Rectangle.

commands to the surgical robot for motion trajectory is unique. Thus, the obtained data from the proposed work is completely different from existing research work. For this significant reason forward kinematics were employed to verify the proposed work. Here, the position and angle of each joint are determined via the Denavit-Hartenberg (DH) method. A DH chain diagram is first created (see Fig. 6) using a set of coordinate frames that are positioned on all joints of the customised surgical robot, as seen in Figs. 4 and 5. As a result, it is simple to acquire and use DH parametric values in the DH table for PSM, depicted in Table. II. Link length, link twist, link offset and joint angle are denoted as these DH parameters a_{i-1} , α_{i-1} , d_i and θ_i . Additionally, the acronyms TF, BF, and RF stand for Target Frame, Base Frame, and Reference Frame, respectively.

Matrix values were obtained for all joints of the PSM from the DH parametric values in the DH table. The next step is to execute Homogenous Transformation Matrices (HTM), which are done by employing cross-multiplication on all joints of the matrices from the target frame (surgical tool) to the reference



Fig. 4. PSM with coordinate frame. Fig. 5. ECM with coordinate frame.



Fig. 6. DH-chain diagram for customised surgical robot from base frame to target frame and reference frame.

TABLE IIDH-TABLE FROM RF TO TF

Link	Joint-Type	a_i	α_i	d_i	θ_i
$RF \rightarrow BF$	Fixed	0	0	0	0
$BF \to PSM_1$	Prismatic	l_1	0	d_1	0
:	•	:	:	:	÷
$PSM_{12} \rightarrow TF$	Rotation	0	$\frac{\pi}{2}$	0	θ_{13}

frame (camera), as shown below. Additionally, ECM joints are kept in a consistent position to achieve a steady camera viewpoint. Consequently, it is not required to acquire matrices for all joints in the ECM.

$$T_{PSM_1}^{BF} = \begin{bmatrix} 1 & 0 & 0 & l_1 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & d_1 \\ 0 & 0 & 0 & 1 \end{bmatrix}$$
(16)

$$T_{TF}^{PSM_{12}} = \begin{bmatrix} c\theta_{TF} & 0 & -s\theta_{TF} & 0\\ s\theta_{TF} & 0 & c\theta_{TF} & 0\\ 0 & 1 & 0 & 0\\ 0 & 0 & 0 & 1 \end{bmatrix}$$
(17)

where $s\theta_i$ and $c\theta_i$, respectively, stand in for $sin\theta_i$ and $cos\theta_i$, correspondingly. Additionally, l_i stands for the link length, while d_i symbolises the prismatic junction. The transformation between TF and RF is performed as follows (18);

$$T_{TF}^{RF} = T_{BF}^{RF}(q_0) * T_{PSM_1}^{BF}(q_1) * \dots * T_{TF}^{PSM_{12}}(q_{13})$$
(18)

The joint q_i (revolute or prismatic joint) is concerned with the joint i (i = 1, 2, ..., n), and the result of the cross multiplication is (19);

$$T_{TF}^{RF} = \begin{bmatrix} R_{11} & R_{12} & R_{13} & p_x \\ R_{21} & R_{22} & R_{23} & p_y \\ R_{31} & R_{32} & R_{33} & p_z \\ 0 & 0 & 0 & 1 \end{bmatrix}$$
(19)

where $k = \{1, 2, 3\}$ and $j = \{1, 2, 3\}$, and R_{kjs} , indicates the rotational components of the final outcome matrix. P_x , P_y , and P_z stand for location vector elements in the x, y, and z axes, respectively. The values from the suggested approach SP are given in centimetres. However, results from the GP are in metres and were obtained by using forward kinematics. These values has been calibrated and translated from metres to centimetres as follows in (20) and (21);

$$p_x(centimeter) = \frac{0.099 - p_x(meter)}{100}$$
(20)

$$p_y(centimeter) = \frac{0.043 - p_y(meter)}{100}$$
(21)

Thereafter, these rotational components, R_{23} and R_{33} , are utilised to obtain rotational values, which are part of Euler angles. Additionally, the rotational values of the respective joints in GP are calibrated to match the rotational values in SP, which are all determined as follows;

$$R_{GP} = tan^{-1} * \left(\frac{R_{32}}{R_{33}}\right) = Radian \tag{22}$$

$$R_{GP} = R_{GP} * \left(\frac{180}{\pi}\right) = Degree \tag{23}$$

$$R_{GP-SP} = 0.0812 * (R_{GP})^2 + 2.8828 * (R_{GP}) + 20.89$$
(24)

where $R_{(GP-SP)}$ denotes the calibration of rotation value from GP to SP and R represents the rotation value.

Furthermore, there is a chance for errors since calculating the entire robot's kinematics, whether manually or by coding, is a very laborious operation. The Robot Operating System (ROS) provided a module called Transform Measuring System (TF) to address this problem. Selecting the reference and target frames enables the operation of inverse and forward kinematics for any type of robot, regardless of the number of joints. The suggested approach needs this kind of absolute data to be validated. As shown in Figs. 7, 8 and 9 the efficacy of



Fig. 7. Comparison between absolute and estimation value for X-axis direction.



Fig. 8. Comparison between absolute and estimation value for Y-axis direction.

the suggested approach is verified with absolute data. The blue plot shows the estimated value utilising the suggested approach, while the red plot shows the absolute value, which is determined by conducting forward kinematics.

To obtain some additional data like velocity and acceleration, single and double differentiation are performed from displacement values (w.r.t time in seconds) in both SP and GP, which are determined as follows;

$$v^{SP} = \frac{d(S_{sw_{(SP-GP)}})}{dt} \tag{25}$$

$$v^{SP} = \frac{d(S_{sh_{(SP-GP)}})}{dt}$$
(26)

$$v^{GP} = \frac{d(p_x)}{dt} \tag{27}$$



Fig. 9. Comparison between absolute and estimation value for rotational angle.

$$v_y^{GP} = \frac{d(p_y)}{dt} \tag{28}$$

$$a_x^{SP} = \frac{d^2(S_{sw_{(SP-GP)}})}{dt^2}$$
(29)

$$a_y^{SP} = \frac{d^2(S_{sh_{(SP-GP)}})}{dt^2}$$
(30)

$$a_x^{GP} = \frac{d^2(p_x)}{dt^2} \tag{31}$$

$$a_y^{GP} = \frac{d^2(p_y)}{dt^2} \tag{32}$$

The following characteristics are $(v_x^{SP}, v_y^{SP}, v_x^{GP}, v_y^{GP}, a_x^{SP}, a_y^{SP}, a_x^{SP}, a_x^{SP},$

(

For evaluating the accuracy of proposed approach performance, the coefficient of determination (R^2) is used (33);

$$R^{2} = \left(1 - \frac{\sum_{i=1}^{n} (f_{i} - y_{i})^{2}}{\sum_{i=1}^{n} (y_{i} - y')^{2}}\right) * (100)$$
(33)

where y' is the average (mean) absolute value, f_i represents the estimated value, y_i represents the absolute value, and nsignifies the number of values.

Accuracy of each element (displacement, velocity, acceleration and rotation angle) and the overall performance accuracy are summarized in Table. III. In this research, the complete work is conducted on a virtual machine-based Ubuntu 18.04

 TABLE III

 Accuracy validation between estimated and absolute values

List of position direction and angle	Accuracy range (percentage)
Xaxis	93.86
Yax is	80.92
Rotational(angle)	99.32
$Overall_{accuracy}$	91.36

version with a 6-core processor and 16 GB of RAM. For simulation, Gazebo 9.0 with ROS 1 were used in this research, with Python 2.7 as the programming language.

V. CONCLUSION

In robotic surgery, real-time observation of surgical instrument tracking is quite difficult task. Surgery is getting increasingly challenging for surgeons, especially when it comes to needle tracking. Therefore, a new visual system-based approach is developed in order to track and predict the needle's orientation and location from the camera point of view with the help of the developed V-RBR as Virtual Marker. The data obtained from the proposed approach demonstrates the potential to train automated surgical robots to perform procedures with reduced errors, thereby alleviating challenges faced by surgeons. However, the proposed approach may encounter challenges in real-world scenarios. The future objective of this research is to develop a more dependable model utilizing neural networks to address the issues inherent in real-life situations.

REFERENCES

- R. A. Gondokaryono and A. Agrawal, "An approach to modeling closedloop kinematic chain mechanisms, applied to simulations of the da vinci surgical system," *Acta Polytechnica Hungarica*, vol. 16, no. 8, 2019.
- [2] H. Peng, X. Yang, Y.-H. Su, and B. Hannaford, "Real-time data driven precision estimator for raven-ii surgical robot end effector position," in 2020 IEEE International Conference on Robotics and Automation (ICRA), pp. 350–356, IEEE, 2020.
- [3] P. Shi, Z. Zhao, S. Hu, and F. Chang, "Real-time surgical tool detection in minimally invasive surgery based on attention-guided convolutional neural network," *IEEE Access*, vol. 8, pp. 228853–228862, 2020.
- [4] G. Wang and S. Wang, "Surgical tools detection based on training sample adaptation in laparoscopic videos," *IEEE Access*, vol. 8, pp. 181723– 181732, 2020.
- [5] B. Zhang, S. Wang, L. Dong, and P. Chen, "Surgical tools detection based on modulated anchoring network in laparoscopic videos," *IEEE Access*, vol. 8, pp. 23748–23758, 2020.
- [6] M. Robu, A. Kadkhodamohammadi, I. Luengo, and D. Stoyanov, "Towards real-time multiple surgical tool tracking," *Computer Methods in Biomechanics and Biomedical Engineering: Imaging & Visualization*, vol. 9, no. 3, pp. 279–285, 2021.
- [7] I. Rivas-Blanco, C. J. Pérez-Del-Pulgar, I. García-Morales, and V. F. Muñoz, "A review on deep learning in minimally invasive surgery," *IEEE Access*, vol. 9, pp. 48658–48678, 2021.
- [8] T. Cheng, W. Li, W. Y. Ng, Y. Huang, J. Li, C. S. H. Ng, P. W. Y. Chiu, and Z. Li, "Deep learning assisted robotic magnetic anchored and guided endoscope for real-time instrument tracking," *IEEE Robotics and Automation Letters*, vol. 6, no. 2, pp. 3979–3986, 2021.
- [9] J. Jaafari, S. Douzi, K. Douzi, and B. Hssina, "Towards more efficient cnn-based surgical tools classification using transfer learning," *Journal* of Big Data, vol. 8, pp. 1–15, 2021.
- [10] J. Cartucho, C. Wang, B. Huang, D. S. Elson, A. Darzi, and S. Giannarou, "An enhanced marker pattern that achieves improved accuracy in surgical tool tracking," *Computer Methods in Biomechanics and Biomedical Engineering: Imaging & Visualization*, vol. 10, no. 4, pp. 400–408, 2022.