Autonomous image guided control of endoscopic orientation for OCT scanning

Fernando Gonzalez Herrera[1,2,*], Ameya Pore[3,4,*], Luca Sestini[1,5,*], Sujit Kumar Sahu[1,6,*], Guiqiu Liao[1,3,*], Philippe Zanne[1], Diego Dall'Alba[3], Albert Hernansanz[4], Benoit Rosa[1], Florent Nageotte[1], Michalina J Gora[7]

[1] *ICube, University of Strasbourg, Strasbourg 67000, France*

[2] *Robot-Assisted Surgery Group, Department of Mechanical Engineering, KU Leuven, Leuven 3001, Belgium*

[3] *Altair Robotics Laboratory, Department of Computer Science, University of Verona, Verona 37134, Italy*

[4] *Biomedical engineering, Universitat Politècnica de Catalunya, Barcelona 08028, Spain*

[5] *NearLab, Department of Electronics, Information and Bioengineering, Politecnico di Milano, Milan 20133, Italy*

[6] *The BioRobotics Institute and Department of Excellence in Robotics & AI, Scuola Superiore Sant'Anna, Pisa 56025, Italy*

[7] *Wyss Center for Bio and Neuroengineering, Campus Biotech, Geneva 1202, Switzerland*

gonzalezherrera@unistra.fr

INTRODUCTION

Minimally invasive gastrointestinal procedures such as Endoscopic Submucosal Dissection (ESD) are challenging to perform with standard flexible endoscopes [1]. Hence, robotic solutions have been proposed to improve dexterity and manipulation in intraluminal surgeries. Conventional systems such as Da Vinci (Intuitive Surgical) have been used for transanal access, but the usage of rigid instruments limits the application to operating in the vicinity of the natural access and makes manipulation in constrained areas highly challenging. Recently, STRAS robotic platform [1] has shown feasibility for intraluminal procedures in terms of safety and dissection speed with respect to manual systems.

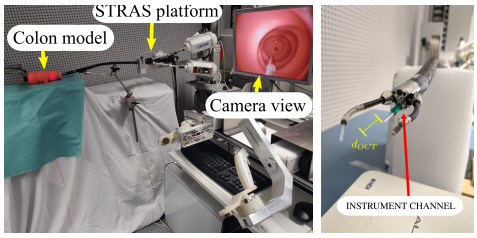
Optical Coherence Tomography (OCT) embedded in continuum robots offers minimally invasive scanning of internal tissues and organs with micrometre resolution and millimetre penetration depth. However, due to the limited depth perception, to achieve precise scanning, contact between the tissue should be maintained during scanning. Performing such tasks requires controlling several DoFs while relying on both the endoscopic camera and OCT images. This procedure has been proven to be difficult to realise by users, even in telemanipulation. In this context, automatic repositioning of the endoscope could allow deploying the OCT probe accurately and more easily.

The objective of this study is to develop autonomous endoscope positioning to facilitate OCT sensing in a phantom environment (Fig.1a).

MATERIALS AND METHODS

Autonomous positioning of the endoscope is carried out by developing an image-based controller. We describe the methodology and the OCT sensor data in this section.

System Overview: The STRAS robotic system consists of a main endoscope, which allows to house three working channels for instruments [1]. The main endoscope is equipped with a camera at the distal tip, a lighting system and, a channel for fluids such as air insufflation and water to cleanse the camera. The distal part of the endoscope can be deflected in two orthogonal directions, which are actuated by antagonist tendons. In total, the system has 10 degrees of freedom, 3 per arm (bending, rotation and translation) and 4 on the body (vertical and horizontal bending, rotation and translation). In our setup, a motorised OCT probe is inserted in the central instrument channel and extends 25mm out its distal tip (Fig.1b).



(a) (b)

Fig. 1. (a) STRAS robotic setup. (b) Motorised OCT probe.

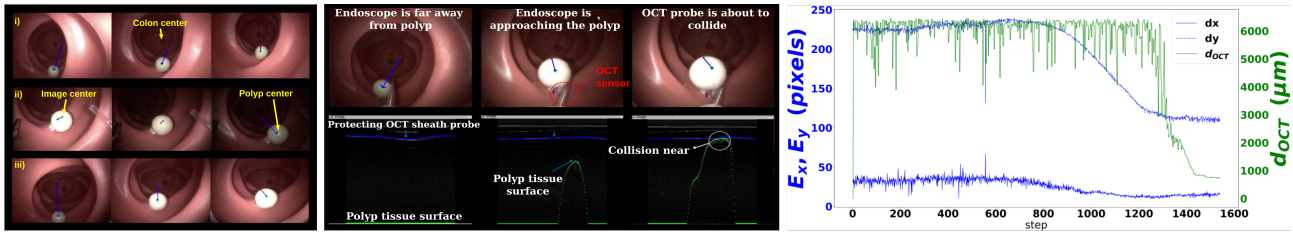
OCT: The probe comprises an external motor, an optical fibre, a distal ball lens, a casing and the associated electronic and optical elements. The ball lens deflects the light passing through the optical fibre, creating an orthogonal imaging beam to the probe. The reflected light is processed to obtain A-line scans with a perception range of 3.5mm outside the probe sheath. The high-speed rotation of the external motor enables the creation of radial images at 50Hz. In this study, we use planar radial B-scans images. The OCT image stream is firstly stabilised with a CNN-based method [2] and then segmented to calculate the distance and direction between the scanning centre and the surrounding tissue.

Polyp Detection: Polyps in the phantom are automatically detected using supervised deep learning techniques. The training set is created by capturing images through telemanipulating the robot. These images (720 × 576 pixels) are manually annotated by selecting the polyp area from the background; the 100 images collected are augmented to ensure robust training. A U-net [3] is implemented and trained in a supervised manner on the annotated image dataset. Using the output of the trained U-net, the centre of mass of the detected polyp (P_p) is estimated.

Endoscope Orientation: An image-guided position control is developed to reduce the distance error between the image centre P_c and the detected coordinates of polyp P_p. The controller outputs the desired joint position proportional to the error, $E = P_c - P_p$, on both coordinates $E = [E_x, E_y]$. With the horizontal bending q_{horz} reducing

*These authors contributed equally to this work.

This work was supported by the ATLAS project. This project has received funding from the European Union's Horizon 2020 research and innovation programme under the Marie Sklodowska-Curie grant agreement No 813782.



(a) (b) (c)

Fig. 2. a) Image guided control. Sequence of frames obtained from the monocular camera while aligning towards the polyp (i) in normal lighting conditions, (ii) in varying lighting conditions, (iii) Experiment two: translating and aligning towards the detected polyp. b) Top: Deployment of the OCT probe. Bottom: Data generated while tissue scanning. c) Plot of evolution of E_x, E_y and d_{OCT} for the second experiment performed.

the E_x component and, q_{vert} acting on E_y. When the error is below $\theta_1 = 60$ pixels and $\theta_2 = 100$ pixels, for E_x, E_y, the orientation control ends.

Endoscope translation: To ensure polyp-probe contact, a small motorised translation of $\sim 5cm$ is carried out at a constant speed. For simplification, the navigation is carried out at a constant speed. The OCT signal is used as feedback to terminate the translation when the tip is close to the polyp or colon wall, i.e. $d_{OCT} < \theta_3$, where d_{OCT} represents the distance measured by the OCT probe and $\theta_3 = 750\mu m$ is the distance threshold.

For the experiments performed, the STRAS robot was teleoperated by the user inside the colon model; the position allows polyp visibility. Then, after a set time, the autonomous technique takes control of the whole-DOF to perform the alignment.

RESULTS

We use the ascending colon part of the LM-107 Colonoscopy Simulator (KOKEN, Japan). The model has a white polyp that can be inserted in a predefined location within the selected section. Our aim is first to show robust polyp detection in varying lighting conditions; second to deploy an OCT probe for tissue scanning. For the first objective, we implement the image position control as described in the previous section, and evaluate detection in changing luminosity conditions. Note that the OCT probe is not employed in these experiments. For the second objective, we deploy the OCT probe and translate near the polyp to scan the tissue. We consider two experimental scenarios as follows (i) align first and then translate towards the polyp; (ii) align towards the detected polyp while translating.

The image position control and translation experiments are illustrated in Fig.2a. The distance from the centre of the image (red dot) to the centre of mass of the detected polyp (green dot) is shown in blue. Fig.2b (bottom) shows the detection of the polyp surface in the OCT image when the endoscope is approaching it, with the corresponding endoscopic image in the top line. The green line in this figure shows the polyp surface, while the blue shows the sheath of the OCT catheter probe. It can be observed that when the endoscope is not aligned with the polyp, both the line remains straight as the polyp remains outside the field of view of the OCT probe. However, while approaching towards polyp, a peak point occurs in the green line showing the presence of an object near the probe. The peak point continues growing more with endoscope advancement and finally touches the blue line when the probe is about to collide with the polyp. This way, the feedback from the OCT image avoids collision. In Fig. 2c, the results of the second experiment are shown. A 10 second (830 step) safety buffer is set for visual inspection at a 50Hz step rate. After this, pixel error E starts reducing; the d_{OCT} start decreasing after the 1300 step, in which the probe has come close enough to the polyp to provide a measurement with less error.

CONCLUSIONS AND DISCUSSION

The LM-107 colon model used has a 4cm diameter, making navigation with the arms outside the channels prone to collision. A model with a realistic colon diameter, from ~4-8 cm, would resolve the collision problems and allow for independent arm-body motions [4]. Increasing the complexity of the testbed would require a more robust control system, e.g., the one present in [5] where feedback from the endoscopic camera and OCT is used to actuate the translation and bending intelligently. The prior changes on the setup would require retraining the neural network to ensure polyp detection in real-time, given the background image modification. Polyp dimensions from the image can also serve as feedback on the selected thresholds θ_1, θ_2 to better suit the geometry of the polyp and OCT probe w.r.t. the camera frame; given dimensions of the polyp does not have any effect in our approach.

Another simplification considered in this work is the constant Jacobian matrix that maps the tendon displacement to camera displacement. This assumption can be valid for small orientation changes; future work will be directed towards estimating the Jacobian matrix online.

REFERENCES

[1] Nageotte Florent *et al.*, "Stras: A modular and flexible telemanipulated robotic device for intraluminal surgery," in *Handbook of Robotic and Image-Guided Surgery*. Elsevier, 2020, pp. 123–146.

[2] Liao Guiqiu *et al.*, "Data stream stabilization for optical coherence tomography volumetric scanning," *IEEE Transactions on Medical Robotics and Bionics*, 2021.

[3] Ronneberger Olaf *et al.*, "U-net: Convolutional networks for biomedical image segmentation," in *International Conference on Medical image computing and computer-assisted intervention*. Springer, 2015, pp. 234–241.

[4] Mora Caravaca Oscar *et al.*, "Oct image-guidance of needle injection for robotized flexible interventional endoscopy," in *Endoscopic Microscopy XVI*, vol. 11620. International Society for Optics and Photonics, 2021, p. 116200D.

[5] Zhang Zhongkai *et al.*, "Image-guided control of an endoscopic robot for oct path scanning," *IEEE Robotics and Automation Letters*, 2021.