A Gaze-contingent Robotized Flexible Endoscope

Alexandros A. Kogkas, Benjamin Glover, Nisha Patel, Ara Darzi, George P. Mylonas, *Member, IEEE*

I. Introduction

Flexible endoscopy is a common medical technique used for diagnosis in the upper and lower gastrointestinal (GI) tract. The flexible endoscope consists of a flexible tube, up to two working channels for insertion of flexible instruments, a camera and a light source at the distal end (tip). Bending the tip left-right and up-down, is achieved by rotating two dials with one hand, whilst the other hand is used for rotating, inserting and retracting the endoscope shaft. Often, more than one operator is needed for effective manipulation. As a result, the conventional control of a flexible endoscope creates challenges, such as poor ergonomics, inefficient handling of the endoscopic tip and suboptimal collaboration [1].

Therefore, the development of an intuitive interface to overcome these challenges is deemed necessary. Robotization of the existing flexible endoscopes has been attempted in previous studies [2], [3]. In [4] we proposed a gaze-controlled robotized system, which allows hands-free control of the endoscopic view in an intuitive fashion, using the natural gaze of the user to steer the endoscope tip. Evaluation with novice users showed improved performance and preference over traditional hand control. Limitations of this system include the lack of motorization of the rotation, insertion and retraction of the shaft of the endoscope and the restricted positioning of the user in front of a screen, due to the use of screen-based eye-tracking sensors.

In this study, we propose a fully motorized gaze-controlled system for non-restricting, free-view flexible endoscopy. The feasibility and comparison against traditional hand control are assessed. Simulated examination of the upper GI tract is performed by novice users.

II. System Overview

The system presented here allows a user to remotely control the endoscope movements without handling the device, while expert endoscopist studies art currently ongoing. A flexible gastroscope is attached to an articulated robotic arm, mounted onto a rail and placed on top of a surgical table (Fig. 1). The dials used to control the distal tip steering are motorized using two 3D printed gears and two motors, controlled by a gaze-contingent closed loop velocity controller. Gaze on the screen is estimated based on a 3D gaze reconstruction

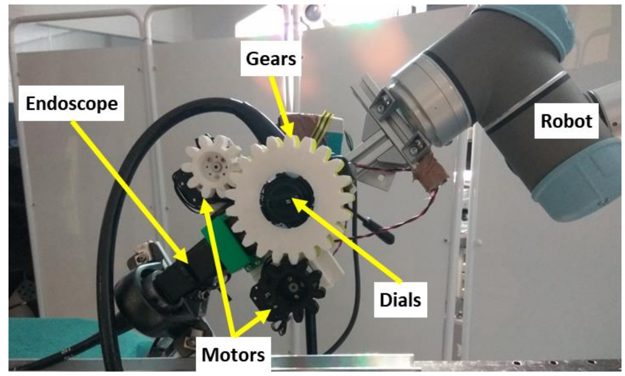

Fig. 1: The robotized flexible endoscope mounted on the UR5.

framework we developed [5] with the synergy of conventional wearable eye-tracking glasses, a motion capture system and fixed in space RGB-D cameras for real-time 3D reconstruction of the environment. The robot controls the endoscope shaft rotation and insertion/retraction. The former is controlled with the head sideways rotations and the latter with a joystick handle, which also allows features such as system pause and automatic retroflexion. Audio feedback is provided when the user enables rotation, pauses the system or enables retroflexion.

III. Materials and Methods

Free-view gaze interaction with the screen is achieved with the real-time framework presented in [5]. Its core functionality is to provide the user's 3D point of regard in space. It relies on estimating the pose of the eye-tracker scene camera in the world frame and tracing the gaze ray provided on the head frame of reference, onto the 3D reconstructed space. For the work presented here, the head pose is estimated with the employment of the OptiTrack motion capture system. Spherical markers are mounted on the eye-tracker to form an asymmetric rigid body and allow OptiTrack to provide its unique 6 degrees of freedom (DOF) pose in space. The rigid transformations between the rigid body – eye-tracker scene camera and OptiTrack – world coordinate systems are calibrated. By using the hybrid macro/micro-scale model presented in [6], the mode (macro or micro) and the 2D screen fixation (for micro mode) is provided as output.

For gaze control, the PID controller presented in [4] is employed, which relies on the distance between the fixation point on the screen and the center of the screen.

The rotation of the endoscope is achieved by rotating the end-effector of the robot with a constant speed. It is initiated with the rotation of the user's head on the eye-tracker's scene camera's z-axis, above a predefined angular threshold. The

*Research supported by the NIHR Imperial Biomedical Research Centre (BRC).

Alexandros A. Kogkas and George P. Mylonas are with the HARMS Lab, Department of Surgery and Cancer, Imperial College London, UK (e-mail: a.kogkas15@imperial.ac.uk).

Benjamin Glover and Nisha Patel are with the Department of Gastroenterology, Imperial College London, UK.

Ara Darzi is with the Department of Surgery and Cancer, Imperial College London, UK.

head orientation reference is defined at the beginning of the experiment.

Insertion and retraction of the endoscope is implemented with the linear movement of the robot with constant velocity. It is triggered by a joystick (up/down respectively) connected to an Arduino Uno which streams data to the system PC.

Retroflexion of the distal tip is activated by holding the joystick to the right for 1 s. The system is paused (motors and robot remain idle) by pushing the joystick handle and is enabled by pushing the joystick again. No action is performed when user's gaze is outside the screen.

For eye-tracking, the SMI (SensoMotoric Instruments GmbH) glasses are used. For RGB-D sensing, the Microsoft Kinect v2 is used and for head pose tracking the OptiTrack motion capture system with four Prime 13 cameras. The robot arm is a UR5 by Universal Robots. Two Dynamixel RX-24F motors were employed for the motorization of the endoscope. For the endoscopic task view, a 42″ LG screen with 1920x1080 px resolution is used. The endoscope is a flexible gastroscope (Karl Storz 13801 PKS).

IV. Experimental Setup

The evaluation task simulates a diagnostic gastroscopy using gaze or hand (conventional) control. The experimental setup involves a head phantom, a silicon tube to simulate the oesophagus and a stomach phantom filled with 10 numbered targets (Fig. 2b,c). Seven subjects, 6 males and 1 female, between the ages of 26 and 35 with normal uncorrected vision, were invited to take part in the study. After informed consent, the subjects were taken through the experimental setup, starting either with the gaze- or hand-control setup, in randomized order, and given time to familiarize themselves with it. Each subject was then instructed to intubate the oesophagus, locate and fit the targets in ascending order within a circle drawn at the center of the screen (Fig. 2c).

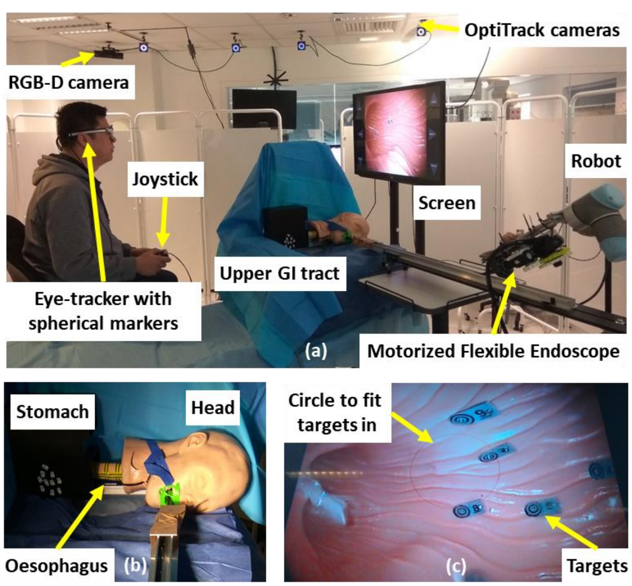

Fig. 2: (a) The experimental setup. (b) The upper GI tract phantom and (c) the camera view of the targets.

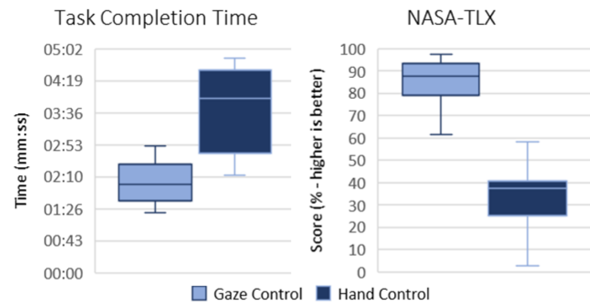

Fig. 3: Results of task completion time (left) and NASA-TLX (right) for both gaze- and hand-control modalities.

V. Results

All the subjects were able to independently intubate the oesophagus and accurately locate ten targets placed in the stomach using the gaze-controlled system and the traditional hand control. The average time to complete the task with gaze was 2:04±0:30 minutes vs 03:40±01:00 with hand control. Significant difference between the two modalities at the $p<.05$ level $[F(1, 14) = 16.32, p = 0.001]$ is found. NASA-TLX questionnaires show preference in gaze control (84±12%) over hand control (33±17%) and significant difference at the $p<.05$ level $[F(1, 12) = 41.55, p<.001]$ (Fig. 3).

VI. Discussion

A fully robotized gaze-contingent flexible endoscope has been presented, which allows touchless control of a flexible endoscope in a free-viewing fashion. Testing with non-endoscopist subjects in a simulated diagnostic gastroscopy (upper GI) showed that gaze controlled endoscopy is a feasible concept. It allows ergonomic, user-friendly and intuitive control whilst maintaining the benefits of a flexible endoscope. Further experiments involving expert endoscopists are ongoing.

References

[1] C. G. L. Cao and P. Milgram, "Disorientation in minimal access surgery: A case study," in *Proceedings of the Human Factors and Ergonomics Society Annual Meeting*, 2000, vol. 44, no. 26, pp. 169–172.

[2] K. Kume, T. Kuroki, T. Sugihara, and M. Shinngai, "Development of a novel endoscopic manipulation system: The Endoscopic operation robot," *World J. Gastrointest. Endosc.*, vol. 3, no. 7, p. 145, 2011.

[3] H. J. M. Pullens, N. van der Stap, E. D. Rozeboom, M. P. Schwartz, F. van der Heijden, M. G. H. van Oijen, P. D. Siersema, and I. A. M. J. Broeders, "Colonoscopy with robotic steering and automated lumen centralization: a feasibility study in a colon model," *Endoscopy*, vol. 48, no. 03, pp. 286–290, 2016.

[4] T. J. C. O. Vrielink, J. G.-B. Puyal, A. Kogkas, A. Darzi, and G. Mylonas, "Intuitive Gaze-Control of a Robotized Flexible Endoscope," in *2018 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*, 2018, pp. 1776–1782.

[5] A. A. Kogkas, A. Darzi, and G. P. Mylonas, "Gaze-contingent perceptually enabled interactions in the operating theatre," *Int. J. Comput. Assist. Radiol. Surg.*, pp. 1–10, 2017.

[6] A. A. Kogkas, M. H. Sodergren, A. Darzi, and G. P. Mylonas, "Macro- and Micro-Scale 3D Gaze Tracking in the Operating Theatre," in *The Hamlyn Symposium on Medical Robotics*, 2016, p. 100.