Attention-Aware Robotic Laparoscope for Human-Robot Cooperative Surgery

Songpo Li, Jiucai Zhang, Member, IEEE, Linting Xue, Fernando J. Kim, MD, FACS, Xiaoli Zhang, Member, IEEE

ABSTRACT—Current laparoscopic surgeries suffer from inconvenient, limited visualization due to the small incisions. It normally needs an assistant to hold and readjust a laparoscope to view the surgical set when performing procedures. This assistant can be either a human or a robotic laparoscope holder. However, at the current state, both of them only passively follow the control commands from the surgeon (The human assistant relies on the surgeon’s verbal communications and the robotic holder is controlled by the voice or joystick/buttons), whenever the laparoscope needs to be readjusted. The surgeon’s explicit intervention in laparoscope readjustment might bring extra mental and physical burdens to him/her. In this paper, the work of granting attention-awareness to a robotic laparoscope holder is presented. The robot can actively observe the surgeon’s viewing attention by tracking the surgeon’s eye movements and automatically adjust the laparoscope’s view correspondingly. Our experimental results demonstrated that this system effectively released the surgeon from the intervention of the laparoscope, which could eliminate the visualization barriers of laparoscopic surgeries, reduce the learning curve in laparoscopic surgeries, and reduce the operation time and cost.

I. INTRODUCTION

In the last several decades, laparoscopic surgery has been widely adopted in modern medical practice. The surgeon uses slender instruments to perform the surgery through small incisions in the patient’s skin, watching the internal images of the patient’s abdomen projected from a laparoscope to a monitor. Laparoscopic surgery can significantly reduce patients’ tissue trauma, post-operative pain, recovery time and hospitalization. Consequently, it has become the favorite option of patients. However, performing such a surgical procedure is challenging to the surgical team. The limited range of motion and dexterity of the surgical instruments, as well as the poor depth perception and full-crum effect, bring more mental burden and work load to the surgeon, in comparison with open surgeries. On top of that the surgeon needs a third hand to hold and position the laparoscope during the surgery when his/her two hands are occupied with surgical instruments. Typically this third hand is from a human assistant or a robotic laparoscope holder.

However, the surgeon faces various challenges in communicating with a human assistant or controlling a robotic holder during the surgeries. Typically, the laparoscope is positioned and held by a human assistant. The surgeon needs to verbally communicate with the assistant to move the laparoscope to a desired viewpoint. The human assistant’s experience and understanding of the surgical task and surgeon can significantly affect the outcome of the surgery. In light of the fact that the human assistant is normally watching the surgical field on the monitor from a different location relative to the surgeon due to the compactness of the work area, it can be difficult for the assistant to fully understand the surgeon’s intention and correctly move the laparoscope to focus on the target area. In addition, the human factors of the assistant like tremors and fatigue may also affect the surgical outcome [1], especially in the long surgical procedures.

Robotic laparoscope holders have been introduced into laparoscopic surgeries since 1994, to replace the human assistant. Those robotic holders allow the surgeon to directly control the laparoscope with specially designed control interfaces. The first robot approved by the FDA for clinical use in the abdomen was the automated laparoscope system for optimal positioning (AESOP) (Computer Motion, Santa Barbara, CA, USA) [2], and the advanced version can be controlled through pre-calibrated voice commands [3]. Another laparoscope holding and positioning system, EndoAssist from Armstrong Healthcare Ltd., can be controlled by the surgeon’s head movements [4], [5]. In this system, an infrared (IR) emitter is mounted on the surgeon’s head, and the surgeon uses the head movement to indicate the direction (e.g. up, down, left, and right) where the laparoscope should move. A foot clutch is adopted for the surgeon to switch the IR control mode on and off. Freehand from Prosurgics [6], [7] is another robotic laparoscope system with a compact design, which has the same gesture control interface as EndoAssist. In the da Vinci surgical robot system, the surgeon can control the robotic arm that is holding the laparoscope using a joystick and clutch [8]. Inherently, these robotic laparoscope holders offer stable and tremor-free motions and feedback videos, and do not have the problem of human factors such as fatigue and emotion [9-12].

Manuscript received September 15, 2013.

S. Li is a Ph. D. student in the Mechanical Engineering Department at Colorado School of Mines, Golden, CO 80401 USA (e-mail: soli@mymail.mines.edu).

J. Zhang is a Lead System Engineer at General Electric, Schenectady, NY 12345 USA (e-mail: jiangjiewei@gmail.com).

L. Xue is a Master student in the Mechanical Engineering Department at Wilkes University, Wilkes-Barre, PA 18776 USA (e-mail: linting.xue@wilkes.edu).

F. Kim is an Associate Professor at the University of Colorado Denver:
X. Zhang is an Assistant Professor of Mechanical Engineering in the Mechanical Engineering Department at Colorado School of Mines, Golden, CO 80401 USA (corresponding author, phone: 303-384-2343; fax: 303-273-3602; e-mail: xilizhang@mines.edu).
Though robotic laparoscope holders bring benefits into the operations by replacing the human assistant, their control interface encounters various challenges in practice. For example, the voice control may distract other surgical staff in the operating room. Moreover, it limits the communication between the surgeon and other surgical personnel, as the voice recognition may pick up wrong commands from the conversation. In [13], the efficiency and safety of the laparoscope holding and positioning tasks were compared between human assistants and a robotic laparoscope holder (AESOP) with a voice control interface. The results showed that the voice controlled robotic holder was as safe but less efficient due to the limitation in its ease-of-use, when compared to a practiced and attentive human laparoscope-holding assistant by the lack of ease-of-use.

Head motion plus foot clutch require the surgeon’s multi-laparoscope-holding assistant by the lack of ease-of-use. The operation. The laparoscope, which significantly affects the continuity of an operation. When the surgical instruments to manipulate the laparoscope correspondingly. Through granting the ability of attention-awareness to the robotic assistant system, the robot can actively respond to the surgeon’s attention by his/her attention is on the surgical field. In that way, the robot can actively collaborate with the surgeon to better perform the surgery, rather than only act as a passive commands receiver and executant. To accomplish this, the robot needs to observe the surgeon’s gaze movements on the surgical field, and analyze these data to find out the surgeon’s viewing attention. When the robot senses the surgeon’s attention is stably located at a new view interest, the robot automatically adjusts the laparoscope’s position to focus on it. We hope this framework can completely release the surgeon from the mental and physical burden of the intervention of the laparoscope.

In this paper, we present an attention-aware interaction framework for a robotic laparoscope. With this framework, the robot can actively collaborate with the surgeon to better perform the surgery, rather than only act as a passive commands receiver and executant. To accomplish this, the robot needs to observe the surgeon’s gaze movements on the surgical field, and analyze these data to find out the surgeon’s viewing attention. When the robot senses the surgeon’s attention is stably located at a new view interest, the robot automatically adjusts the laparoscope’s position to focus on it. We hope this framework can completely release the surgeon from the mental and physical burden of the intervention of the laparoscope.

II. SYSTEM DESIGN AND INTEGRATION

A. Robotic Holder--CoBRASurge

CoBRASurge (Compact Bevel-gear Robot for Advanced Surgery) [22], [23] is used as the robotic laparoscope holder, shown in the Fig. 1 & 2. It is based on a spherical bevel gear mechanism consisting of three gear pairs and six turning pairs. CoBRASurge creates a mechanically constrained remote center of motion (RCM), which contains three rotational degrees of freedom (DOFs) and one translational DOF passing through it. The rotation center would coincide with the surgical entry port during the surgery. A laparoscope or surgical tool can be fitted into the articulated mechanism using a collar. Such a robot can be employed for guiding surgical tools or positioning a laparoscope. CoBRASurge can produce a cone workspace with a 60° vertex angle and the tip of the workspace will be located at the incision port. This workspace can satisfy the typical workspace required in laparoscopic surgeries [24].

![Fig.1 CoBRASurge as a robotic assistant for laparoscopic surgery](image-url)
A new coordinate system (Frame 8) is attached on the tip of the laparoscope. A series of transformations are performed to relate the new frame to the robot’s base frame, which is shown as Equation 1.

\[
\begin{align*}
\hat{b}T &= \hat{b}T \cdot \hat{a}T \cdot \hat{a}T \cdot \hat{a}T
\end{align*}
\]  

(1)

where the involved sub-system transformations are:

- \( \hat{b}T \): the tip of laparoscope to the sleeve linkage 4 of the robot for the laparoscope housing.
- \( \hat{a}T \): the sleeve linkage to the robot linkage 3.
- \( \hat{a}T \): the robot linkage 3 to the robot linkage 2.
- \( \hat{a}T \): the robot linkage 2 to the robot base link 1.

These sub-system transformations are formulated using the Denavit-Hartenberg method [25]. The new position of the laparoscope’s tip is firstly given as \([x, y, z]^T\) referring to Frame 8. From the above transformation, this new position with respect to the robot’s base is calculated as:

\[
\begin{align*}
[x', y', z']^T &= \hat{b}T \cdot [x, y, z]^T
\end{align*}
\]  

(2)

In our previous work, CoBRASurge was controlled by a joystick Phantom Omni (SensAble), which was able to provide three DOFs of position input in Cartesian space, three DOFs of rotation input. For controlling CoBRASurge, the joystick was programmed in such a way that the tip of the tool was mapped to the endpoint of the joystick stylus in order to achieve a feeling similar to open surgery. LabVIEW (national Instruments) was used to create a flexible, modular user interface on a laptop. This robot system has been verified in multiple porcine tests controlled by a surgical staff holding the joystick across the operating room. The robot provided good stability and smooth motion while holding and maneuvering a laparoscope for viewing the liver, spleen, and small intestine [26], [27]. Its compact design saved large working space around the operating table for the surgeon. It has also been used to guide a robotically actuated grasper in tissue manipulation [28]. Although the feasibility of the robot as a camera assistant was justified and it was controlled remotely by a surgical staff to save the space around the patient table, the joystick control interface still relies on the surgical/assistant’s participation.

**B. S2 Eye Tracker**

The eye tracking system, S2 Eye Tracker, from Mirametrix is used to track where the surgeon is looking at on the monitor. The point, where the surgeon is looking at, is referred as the gaze point. S2 is a video-based remote eye tracking system which allows a certain amount of head movement within a volume of 25×11×30 cm³. It can report the gaze data at 60Hz with an accuracy of 0.5° - 1° and draft <0.3°. An advanced calibration is in need before it can be used to track the eyes and estimate the gaze points. S2 reports the coordinates of gaze points in percentages of the tracking window, and by default it takes the full screen as the tracking window. Raw gaze data is analyzed to obtain a stable view attention before being transmitted to the laparoscope control software.

**B.1 Calibration**

The performance of the eye tracker system greatly depends on the initial calibration, which maps the eye movements to the gaze positions on the screen. The user stands or sits in front of the monitor and face to it. At the same time, the eye tracker is mounted above the monitor where it can successfully track both eyes when the user looks at discretionary positions on the screen, even the four corners. In the calibration, the user is asked to stare 9 points on the screen and each point shrinks from a circle to a point which can help the user focus on it. At the end, the system will evaluate the calibration performance and show the result to the user. The calibration configuration and one calibration result are shown in Fig. 3. The green points are the gaze points estimated from the left eye, while the red ones are from the right eye. The purple points are the average of the
left gaze and right gaze, and taken as the final gaze of both eyes. The average error is 26.8, presenting a very good calibration. At the same time, another purple point is used to indicate the user’s current gaze point. Such that, the user can sense the calibration result intuitively.

B. 2 Attention Determination

In general, the natural eye movements can be divided into fixation and saccade. A fixation is a set of the continuous visual gaze points on a single location and saccade is the fast movements of eyes, which typically can be considered as the movements between fixations. When the surgeon maintains the visual gaze on a single location, it means the object on this location is attractive to the surgeon’s viewing attention. Such that, the surgeon’s attention can be sensed by tracking gaze fixations.

To find fixations, the raw gaze data are firstly refined to filter the noises, which are caused by the involuntary ocular microtremor. This ocular microtremor is a natural eye movement even when eyes are apparently still. A modified rolling average method was used to smooth the raw gaze data and filter the noises. In addition, this algorithm can also filter the saccades from the raw gaze data. When the gaze remains at a small area for more than a time period $T_p$, we consider that a fixation is detected. The time period $T_p$ is a predefined cutting off value that could be different for applications and individual persons. Generally, a fixation lasts between 100-1000 ms, with the majority being between 200-500 ms, which is mainly dependent on the quality of information being processed and current cognitive load [29]. With a small $T_p$, the system response will be fast; however it leads to a poor stability and accuracy. On the other hand, a large $T_p$ makes the system slow in response and the surgeon has to wait. For the safety concern in surgical applications, the $T_p$ is initially defined slightly longer than 1000 ms. The surgeon can adjust the value of $T_p$ according to his/her preference.

C. System Integration and Control

The entire system of the attention-aware robotic laparoscope holder is illustrated as Fig. 4. The surgeon performs the surgery using two surgical instruments, while watching the surgical set from a monitor. The gaze of the surgeon is tracked by the eye tracking system, which is mounted above the monitor. The CoBRASurge holds the laparoscope and adjusts the position according to the detected surgeon’s attention. An elliptic surviving area is defined at the screen’s center, used as the boundary to differentiate between “current view is good” and “current view has expired”. The small block indicates the surgeon’s current gaze position. In practical applications, the surviving area and current gaze point are not displayed to the surgeon to avoid potential distractions. All the devices take the processor as the core and connection center as shown in Fig. 5, which is also an abstract expression of Fig. 4.

The robotic laparoscope system keeps tracking the surgeon’s attention by analyzing the gaze data. When a new stable viewing interest (fixation) outside the surviving area is detected, it means the surgeon has lost the interest in the current center view. And then the robot is triggered to adjust the position of laparoscope to focus on the new interest area. Otherwise, the robot holds the laparoscope at the current position and continually tracks the surgeon’s attention. In this way, the laparoscope will always focus on the surgeon’s attention, without his/her explicit request.

When a new reported fixation falls out the surviving area, the vector $[x_p, y_p, z_p]^T$ from the image center to this fixation is converted to a motion vector $[x, y, z]^T$ of the laparoscope’s
tip using a predefined mapping relation. This mapping relation is predefined with the consideration of camera coefficients and the current insertion depth of the laparoscope. With this mapping, the deviation from the new attention point to the current view center can be directly interpreted to the motion commands of the robot. That means, the surgeon does not need to gain any knowledge of robotic motion/control and his/her viewing attention can drive the laparoscope tip intuitively and directly.

III. EXPERIMENTAL SETUP

To validate the effectiveness of this novel attention-aware interaction framework, a set of tests were conducted on the proposed system. Six subjects were involved in the task trials and none had experience in eye tracking or gaze control before the tests. They were asked to view four targets inside of a human dummy in four individual trials, using this robotic laparoscope system. One trial was successfully completed if the laparoscope moved the viewing center on each target respectively. At first, the gaze tracking concept and the interaction framework were introduced to each participant and ten minutes were given to them for familiarizing themselves with our system. In each trial, the robot started from a home position and returned to the home after completion. The $T_p$ was about 1s in the tests.

Comparisons were carried out between the joystick control mode and the proposed attention-aware interaction framework. Each participant was asked to do the same visualization task using both interfaces. The motion trajectories of the laparoscope were recorded for later analysis.

IV. RESULTS

The results indicate that even in the familiarization stage when the user were first brought into contact with the attention-aware robotic laparoscope system, all participants were able to successfully complete the visualization tasks. The responses from the user were positive. Fig. 6 validates the rolling average method for the data noise filtering and fixation determination, as well as the inherent relationship between gaze fixations and viewing attentions. These data were collected from one participant, when he was viewing the four targets following the order T2-T1-T4-T3 within 6.9s. At the top, it is the raw data plot and at the bottom it is the refined gaze data with fixations. As shown, scattered gaze points, when the gaze focus is transferring from one object to the next, are eliminated. The gaze points on each target got concentrated after the rolling average filtering. The fixations are marked as the green stars. 7 fixations were determined during 6.9 seconds.

Figure 7 shows a comparison of the laparoscope trajectories under the attention-aware interaction model (smooth and short trajectory) and joystick control (jerky and long trajectory). In attention-aware model, robot directly guided the laparoscope to the field of view when it sensed the subject’s attention had deviated from the current viewing center. In addition, participants reported that they could naturally interact with the field of view without feeling the existence of the laparoscope. While in joystick control, the participant’s hand needed to coordinate with his/her eyes to adjust the field of view back and forth to gradually approach the target.

V. CONCLUSION AND FUTURE WORK

In this paper, we introduced a novel attention-aware interaction framework for robotic laparoscope systems to assist the surgeon to better perform surgeries. In this framework, the robot can track surgeon’s gaze points and
analyze them to find the surgeon’s attention on the surgical field of view. More importantly, the robot can automatically adjust the laparoscope to focus on the surgeon’s viewing attention. This system enables the robot to actively observe and cooperate with the surgeon, rather than passively follow commands from the surgeon to move the laparoscope. Its effectiveness has been validated in our experimental tests. We hope that this interaction system could release the surgeon from the intervention of the laparoscope.

In the future, the surgeon would be given a larger workspace, so that in practice the position of surgeon will not be limited and the setup will be more feasible for the operating room. New calibration strategy will be drawn into, therefore when the surgeon leaves and returns he/she will not need recalibration. The fixation time period would be explored to adaptively accommodate individual surgeon’s character.

REFERENCES