

Real-Time Visual Tracking of the Surgeon's Face for Laparoscopic Surgery

Atsushi Nishikawa¹, Toshinori Hosoi¹, Kengo Koara¹, Daiji Negoro¹, Ayae Hikita¹, Shuichi Asano¹, Fumio Miyazaki¹, Mitsugu Sekimoto², Yasuhiro Miyake², Masayoshi Yasui², and Morito Monden²

¹ Department of Systems and Human Science,
Graduate School of Engineering Science, Osaka University
1-3 Machikaneyama-cho, Toyonaka City, Osaka 560-8531, Japan

² Department of Surgery and Clinical Oncology,
Osaka University Graduate School of Medicine
2-2 Yamadaoka, Suita City, Osaka 565-0871, Japan

Abstract. We have developed an image-based human-machine interface that tracks the surgeon's face robustly in real-time(30Hz) and does not require to use any body-contacted sensing devices. Based on this face tracker we have developed a new robotic laparoscope positioning system for solo surgery. Our system completely frees the surgeon's hands and feet from the laparoscope guiding task. To evaluate the performance of the proposed system and its applicability to clinical use, an *in vivo* experiment was carried out in which a surgeon used the system to perform a laparoscopic cholecystectomy on a pig.

1 Introduction

In current laparoscopic surgery, a camera assistant usually holds the laparoscope for the surgeon and positions the scope according to the surgeon's instructions. This method of operation is frustrating and inefficient for the surgeon because commands are often interpreted and executed erroneously by the assistant. The views may be suboptimal and unstable because the scope is sometimes aimed incorrectly and vibrates due to the assistant hand trembling. The introduction of robotic technologies –development of robotic laparoscope positioning systems to replace the human assistant– is a major step towards the solution of this problem and the user(surgeon)-friendly design of human-machine interface that controls the laparoscope positioner plays an important role in this step.

Almost all the laparoscope positioning systems proposed so far only have the human-machine interface requiring the use of the surgeon's hand and/or foot such as instrument-mounted buttons/joystick, foot pedal and the like[1,2,3]. This type of interface, however, seems to be sometimes uneasy to use because the surgeon already uses his/her hands/feet to control a variety of surgical tools. To solve this problem, several researchers have tried to introduce "voice" controller interface based on the use of advanced voice recognition systems[4,5,6]. It seems an effective approach because the verbal instructions are natural for human and

the use of neither hands nor feet is required in controlling the laparoscope. However, it has originally some limitations such as reduced accuracy in positioning, long reaction time and erratic movements in a noisy environment. We believe that the surgeon's head controllable laparoscope is the best solution because nonverbal instructions such as facial gestures are more intuitive and faster than verbal instructions (typically voice commands) and have the potential ability to represent not only the direction of scope motion but also the motion degrees such as velocity so that the laparoscope positioning accuracy may be improved. Several laparoscope manipulators with the head controller interface have been developed [7,8,9]. Such systems, however, require not only head movements but also simultaneous control of a foot/knee switch. Furthermore, the surgeon must wear a head-mounted sensing devices such as headband and gyro sensor, which are stressful for surgeon.

We have developed an image-based human-machine interface that tracks the surgeon's face robustly in real-time (30Hz) and does not require to use any body-contacted sensing devices. Based on this face tracker we have developed a new robotic laparoscope positioning system for solo surgery. Our system completely frees the surgeon's hands and feet from the laparoscope guiding task.

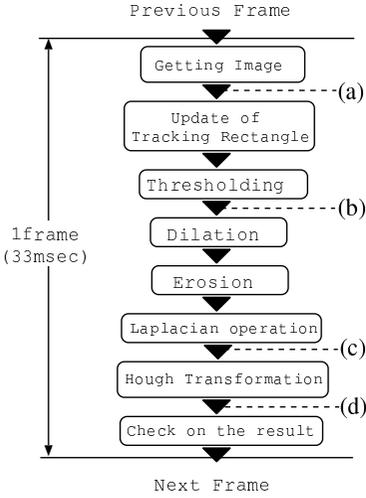
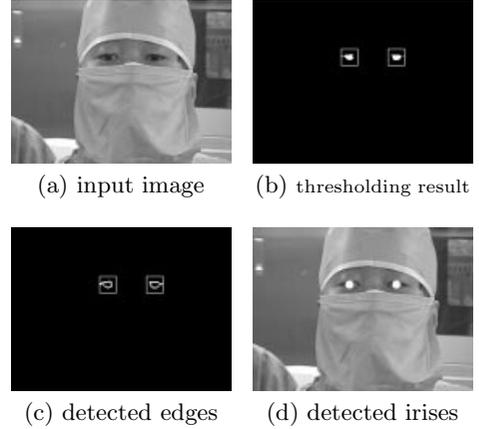
2 Real-Time Face Tracking

In operation, the laparoscopic surgeon usually stands in front of the TV monitor on which the scope image is displayed and gazes at a surgical point of interest on the screen. Therefore, if a surveillance camera is placed just over the TV monitor, the surgeon's face can be observed almost all the time. However, as the surgeon wears a gown, a cap and a mask, almost all the face features such as mouth, nose, and hair do not appear in the surveillance image. We at first turn our attention to the surgeon's irises. The main reason is the following: in our case, not only it is ensured that both left and right irises always appear in the image but also it is easier to detect them than in the general case of face image processing.

2.1 Iris-Based Method

Basically the iris has a low intensity and a circular shape. Thus we do not use color image but gray-value one. The surgeon's face image is taken by a CCD camera and its intensity information is digitized into the memory of the host computer through a video capturing device at a frequency of 30Hz. The video capturing process is done in parallel to the following image processing so that the surgeon's irises are also detected and tracked at a rate of 30Hz. The outline of our algorithm for detecting the surgeon's irises and an example of the detection process are respectively shown in Fig. 1 and Fig. 2.

First of all, based on the previous frame information two rectangle-shaped windows (respectively for tracking the left/right irises) are set on the image. The following processes run only within the rectangles and not the whole image. The

**Fig. 1.** Iris detection algorithm**Fig. 2.** An example of iris detection process

widow images are thresholded at a level, driving lower intensity pixels to 1 and higher ones to zero. For example, see Fig. 2(a)(original) and (b) (result). (in the latter figure, the white region indicates the extracted low-intensity pixels and the rectangle indicates the tracking window.) The gaps are then bridged through dilation-erosion process. After that, the resulting image is convolved with the Laplacian edge operator so that the edge image is yielded such as Fig. 2(c). Finally the Hough transformation technique is used to detect circular shapes as the surgeon's irises. This result is shown in Fig. 2(d). The circular region detected by the Hough technique are overlaid on the original image. Fig. 3 shows ten images taken from a tracking sequence. In addition to the tracking rectangles, the estimates of the surgeon's irises are marked with circular boundaries.

2.2 Marker-Based Method

The iris-based method requires "careful" selection of the thresholding/Hough transformation parameters according to illumination conditions or individual variations in visible size of irises(a part of them is occluded by the eyelid). To cope with this problem, as the alternative, we also developed a marker-based method for face tracking. The marker we use is black(for non-color image processing) and made of low-reflectance materials, and has a long-narrow rectangle shape. It is attached on the surgical cap in advance (see Fig. 3). The flow of marker-based tracking process is the following: (1) the black region is at first extracted by simple thresholding, (2) the conventional labeling algorithm is executed on the thresholded image so that the maximal region is selected as the marker region,



Fig. 3. Real-time tracking of the surgeon’s face (iris-based method)

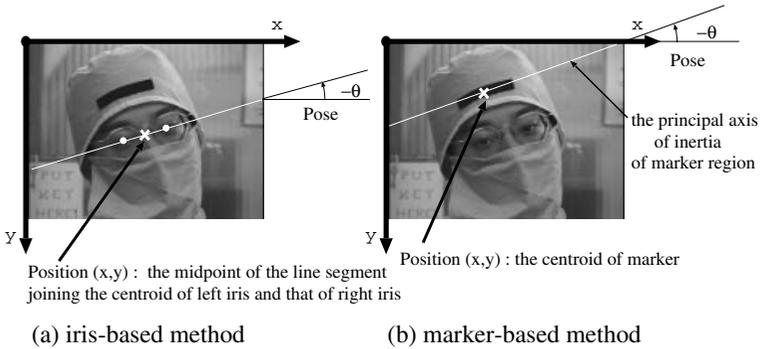


Fig. 4. Estimating the position and the pose of the surgeon’s face

(3) the centroid of the marker and its principal axis of inertia are calculated for robust and rapid tracking at the next frame. The marker-based face tracker can also work at a frequency of 30Hz.

2.3 Estimation of the Position and the Pose of the Surgeon’s Face

We have already assumed that a CCD camera can be placed just over the TV monitor on which the laparoscope image is displayed. In standard laparoscopic surgery such as laparoscopic cholecystectomy, we can also assume *distance-constant* and *fronto-parallel* interaction, i.e., that the surgeon’s face remains almost parallel to the TV monitor screen and the distance between the surgeon and the screen is almost constant during the whole interaction time. In this case, the position and the pose of the surgeon’s face can be easily estimated in real-time from the result of the image processing described above (see Fig. 4). Notice that during such kind of interaction the face DOF(degrees of freedom) are reduced from six to three –namely, a translation 2 dimensional vector (x, y) and a rotation θ in the face plane.

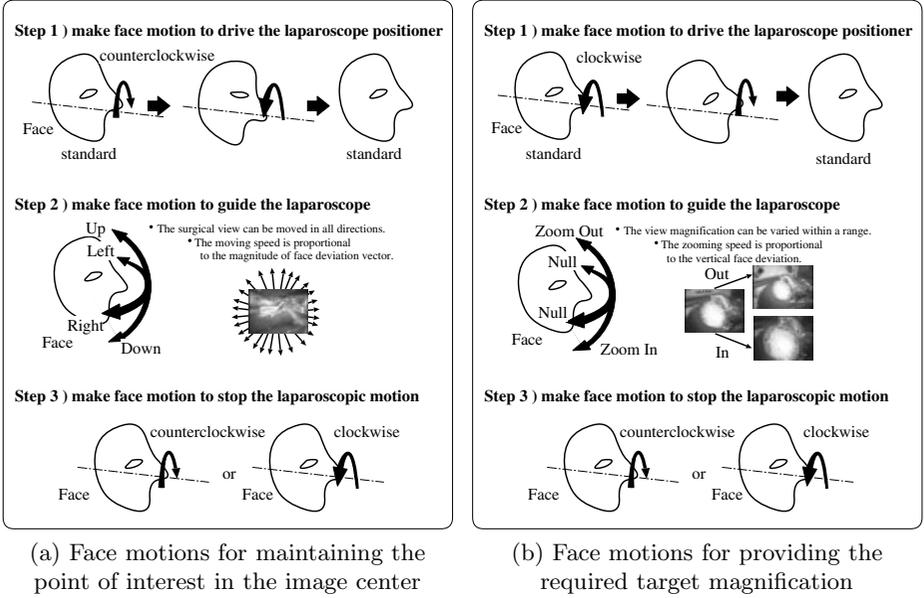


Fig. 5. Positioning the laparoscope by making face motion

3 Application to Laparoscopic Surgery

The proposed method was applied to laparoscopic cholecystectomy on a pig.

3.1 Face Gestures

At first, we worked out a method for positioning the laparoscope only by face motion. It consists of the following three steps (see Fig. 5).

1. *Make face motion to drive the laparoscope positioner*

In this step, the following three consecutive face motions are required. (1) put the position and pose of the face in standard, (2) roll the face counterclockwise/clockwise, (3) return the face “precisely” to the standard position/pose. The counterclockwise rotation is selected when the surgeon wants to guide the laparoscope for maintaining the surgical point of interest in the center of the video frame, while the clockwise motion is made when he/she wants to guide it for providing the required target magnification. Notice that the surgeon cannot make these consecutive motions unconsciously.

2. *Make face motion to guide the laparoscope*

Once coming into the step, as shown in Fig. 5, the face motion is represented as a vector from the standard position, and the direction and the magnitude of the vector are respectively transformed into the direction and the velocity of the laparoscopic motion, according to the surgeon’s requirement.

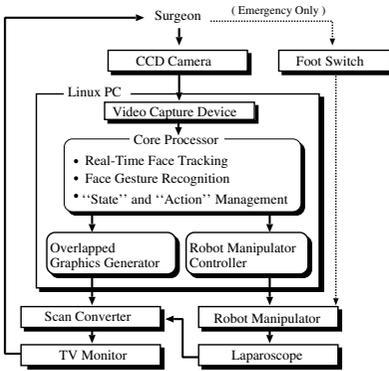


Fig. 6. System configuration

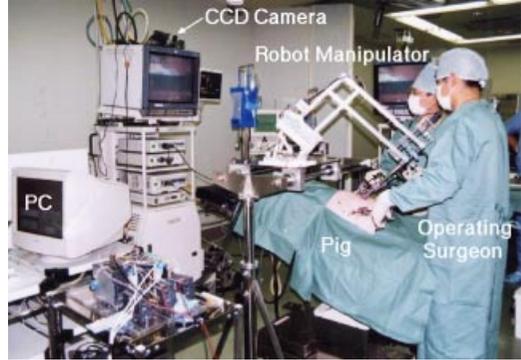


Fig. 7. *In vivo* experiment

3. Make face motion to stop the laparoscopic motion

In this final step, all the surgeon has to do is to roll the face. Note that this action is very easy to make.

Notice that these gestures can be recognized based on the time sequence of the estimated position and pose of the surgeon's face.

3.2 Face Tracker-Based Laparoscope Positioning System

Based on the face tracker and the gesture recognizer, we have developed a novel robotic laparoscope positioning system for solo surgery. The system configuration is shown in Fig. 6. Our laparoscope positioning system mainly consists of a CCD camera, a video capture board, a general PC (Intel Pentium III, 600MHz, OS: Linux), a robot manipulator that holds a laparoscope, a scan converter for superimposing some graphics on the scope image, and a foot switch (in an emergency only). The core system in the PC can detect and track the surgeon's face features (either left/right irises or a marker) in real-time (30Hz) from a sequence of video images taken through the CCD camera and estimates the position and the pose of the surgeon's face from the image processing result and then recognizes the face gestures. According to the motion recognition result, the control command is sent to the laparoscope manipulator.

3.3 Laparoscopic Cholecystectomy on a Pig

To evaluate the applicability of our system to clinical use, an *in vivo* experiment was carried out in which a surgeon used the system to perform a laparoscopic cholecystectomy on a pig (see Fig. 7). Instead of human camera assistant, the system was applied to all of the procedure after trocar insertion. As a result, the whole operative procedure was successfully and safely completed with our system. No one used the emergency foot switch for shutting down the system. The

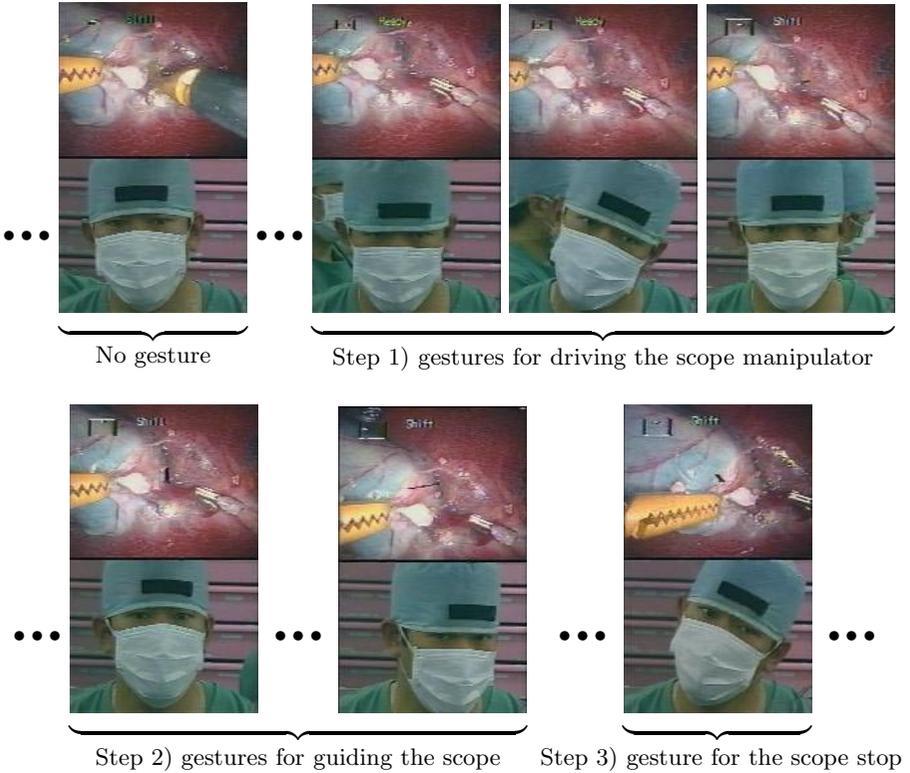


Fig. 8. A scene in the *in vivo* laparoscopic cholecystectomy experiment

number of the lens cleaning was also zero. The operating time from trocar insertion till the removal of the gallbladder inclusive was about 44 minutes. In this experiment, no case has been found in which the robot obstructed the surgeon's work and neither particular incidents nor technical problems had occurred. Fig. 8 shows a scene of the surgeon's facial motions in the experiment (upper part: the scope image which the surgeon looked at, lower part: the surgeon's face image from the surveillance camera, each pair was taken at the same time).

The number of times in operation that the surgeon made face motions to the system to drive the laparoscope manipulator (i.e., step 1 of Fig. 5) was 97 times, which was broken down into 40 times for maintaining the point of interest in the image center, 50 times for providing the required target magnification, and 7 times for being not recognized by the system. No case was found in which the system mistook any other motion such as unintentional action of the operating surgeon and/or another surgeon who looked at the experiment. (e.g., see Fig. 8. In the 2nd-4th images, a surgeon was walking behind the operating surgeon.) The number of times that the surgeon made face gestures to stop the laparoscope motion (step 3 of Fig. 5) was 90 times and these were all recognized completely.

We got many positive comments such as fast reaction time, high positioning accuracy, and easy and intuitive camera guidance (for step 2 of Fig. 5), from the surgeons who performed or looked at the experiment. The operating surgeon, however, also made a negative comment that after the experiment he felt a little fatigue in the cervix from a lot of rolling face motions.

4 Conclusion

We have developed a new robotic laparoscope positioning system for solo surgery based on the real-time face tracking technique. In an *in vivo* experiment, our system succeeded in freeing the surgeon's hands and feet from the laparoscope guiding task while achieving safety, rapid reaction, and high positioning accuracy. Now we are studying a new method for guiding the laparoscope with the aim of reducing not only mental stress but also "physical" one such as cervical fatigue.

References

1. Sackier, J.M., Wang, Y.: Robotically Assisted Laparoscopic Surgery. From Concept to Development. *Surg Endosc* **8** (1) (Jan 1994) 63–66
2. Hurteau, R., DeSantis, S., Begin, E., Gagner, M.: Laparoscopic Surgery Assisted by a Robotic Cameraman: Concept and Experimental Results. In: Proc. 1994 IEEE Int Conf Robotic Automat, San Diego, California, USA, May 1994, pp. 2286–2289
3. Taylor, R.H., Funda, J., Eldridge, B., Gomory, S., Gruben, K., LaRose, D., Talamini, M., Kavoussi, L., Anderson, J.: A Telerobotic Assistant for Laparoscopic Surgery. *IEEE Eng Med Biol Mag* **14** (3) (May-Jun 1995) 279–288
4. Allaf, M.E., Jackman, S.V., Schulam, P.G., Cadeddu, J.A., Lee, B.R., Moore, R.G., Kavoussi, L.R.: Laparoscopic Visual Field. Voice vs Foot Pedal Interfaces for Control of the AESOP Robot. *Surg Endosc* **12** (12) (Dec 1998) 1415–1418
5. Buess, G.F., Arezzo, A., Schurr, M.O., Ulmer, F., Fisher, H., Gumb, L., Testa, T., Nobman, C.: A New Remote-Controlled Endoscope Positioning System for Endoscopic Solo Surgery. the FIPS Endoarm. *Surg Endosc* **14** (4) (Apr 2000) 395–399
6. Munoz, V.F., Vara-Thorbeck, C., DeGabriel, J.G., Lozano, J.F., Sanchez-Badajoz, E., Garcia-Cerezo, A., Toscano, R., Jimenez-Garrido, A.: A Medical Robotic Assistant for Minimally Invasive Surgery. In: Proc. 2000 IEEE Int Conf Robotic Automat, San Francisco, California, USA, Apr 2000, pp. 2901–2906
7. Finlay, P.A., Ornstein, M.H.: Controlling the Movement of a Surgical Laparoscope. EndoSista, with Four Degrees of Freedom, Operates in Concert with Surgeon's Intuitive Head Motions. *IEEE Eng Med Biol Mag* **14** (3) (May-Jun 1995) 289–291
8. Dowler, N.J., Holland, S.R.J.: The Evolutionary Design of an Endoscopic Telemannipulator. *IEEE Robotic Automat Mag* **3** (4) (Dec 1996) 38–45
9. Kobayashi, E., Masamune, K., Sakuma, I., Dohi, T., Hashimoto, D.: A New Safe Laparoscopic Manipulator System with a Five-Bar Linkage Mechanism and an Optical Zoom. *Comput Aided Surg* **4** (4) (1999) 182–192