FAce MOUSe: A Novel Human–Machine Interface for Controlling the Position of a Laparoscope

Atsushi Nishikawa, Toshinori Hosoi, Kengo Koara, Daiji Negoro, Ayae Hikita, Shuichi Asano, Haruhiko Kakutani, Fumio Miyazaki, *Member, IEEE*, Mitsugu Sekimoto, Masayoshi Yasui, Yasuhiro Miyake, Shuji Takiguchi, and Morito Monden

***Abstract*—Robotic laparoscope positioners are now expected as assisting devices for solo surgery among endoscopic surgeons. In such robotic systems, the human–machine (surgeon–robot) interface is of paramount importance because it is the means by which the surgeon communicates with and controls the robotic camera assistant. We have designed a novel human–machine interface, called "FAce MOUSe," for controlling the position of a laparoscope. The proposed human interface is an image-based system which tracks the surgeon's facial motions robustly in real time and does not require the use of any body-contact devices, such as head-mounted sensing devices. The surgeon can easily and precisely control the motion of the laparoscope by simply making the appropriate face gesture, without hand or foot switches or voice input. Based on the FAce MOUSe interface, we have developed a new robotic laparoscope positioning system for solo surgery. Our system allows nonintrusive, nonverbal, hands off and feet off laparoscope operations, which seem more convenient for the surgeon. To evaluate the performance of the proposed system and its applicability in clinical use, we set up an *in vivo* experiment, in which the surgeon used the system to perform a laparoscopic cholecystectomy on a pig.**

***Index Terms*—Face gesture, human–machine interface, laparoscopic surgery, robotic camera assistant, solo surgery.**

Manuscript received June 21, 2002; revised February 2, 2003. This paper was recommended for publication by Editor R. Taylor upon evaluation of the reviewers' comments. This work was supported in part by the Mikiya Science and Technology Foundation. This paper was presented in part at the 15th International Congress and Exhibition on Computer Assisted Radiology and Surgery, Berlin, Germany, June 2001, and in part at the 4th International Conference on Medical Image Computing and Computer-Assisted Intervention, Utrecht, The Netherlands, October 2001.

A. Nishikawa and F. Miyazaki are with the Department of Mechanical Science and Bioengineering, Graduate School of Engineering Science, Osaka University, Toyonaka 560-8531, Japan (e-mail: atsushi@me.es.osaka-u.ac.jp; miyazaki@me.es.osaka-u.ac.jp).

T. Hosoi is with the Multimedia Research Laboratories, NEC Corporation, Kawasaki 216-8555, Japan (e-mail: t-hosoi@ap.jp.nec.com).

K. Koara is with the Nagoya Works, Mitsubishi Electric Corporation, Nagoya 461-8670, Japan (e-mail: koara@kai.mei.melco.co.jp).

D. Negoro is with the Social Network and Infrastructure Systems Company, Toshiba Corporation, Kawasaki 212-8501, Japan (e-mail: daiji.negoro@toshiba.co.jp).

A. Hikita is with the FITEL Network Laboratory, Furukawa Electric Corporation, Hiratsuka 254-0016, Japan (e-mail: hikita@inf.furukawa.co.jp).

S. Asano is with Denso Corporation, Kariya 448-8661, Japan.

H. Kakutani is with the Manufacturing Engineering Center, Mitsubishi Electric Corporation, Hyogo 661-8661, Japan (e-mail: kakutani.haruhiko@mdl.melco.co.jp).

M. Sekimoto, M. Yasui, S. Takiguchi, and M. Monden are with the Department of Surgery and Clinical Oncology, Osaka University Graduate School of Medicine, Suita 565-0871, Japan (e-mail: sekimoto@surg2.med.osaka-u.ac.jp; yasui@surg2.med.osaka-u.ac.jp; taky@surg2.med.osaka-u.ac.jp; monden@surg2.med.osaka-u.ac.jp).

Y. Miyake is with the Division of Hematology and Oncology, Department of Internal Medicine, University of Michigan, Ann Arbor, MI 48109 USA (e-mail: ymiyake@umich.edu).

Digital Object Identifier 10.1109/TRA.2003.817093

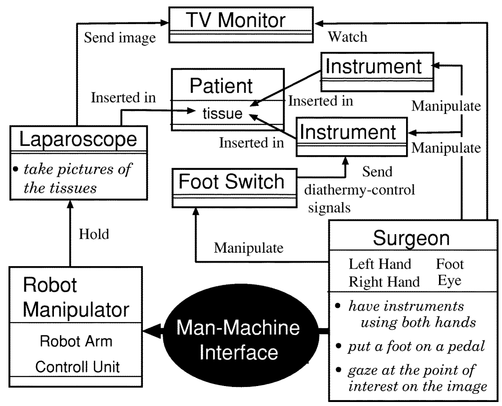


Fig. 1. Diagram of action flow in endoscopic solo surgery.

I. INTRODUCTION

In CURRENT laparoscopic surgery, the vision of the operating surgeon usually depends on the camera assistant responsible for guiding the laparoscope. The assistant 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. Also, the views may be suboptimal and unstable because the scope is sometimes aimed incorrectly and vibrates due to the assistant's hand tremors. The introduction of robotic technologies, specifically, the development of robotic laparoscope positioning systems to replace the human assistant, is a major step toward the solution of this problem. Laparoscope positioning systems will enable endoscopic solo surgery, in which the surgeon carries out all surgical work alone, without the support of the human camera assistant.

Fig. 1 illustrates the action flow in an endoscopic solo surgery. As shown in this figure, the operating surgeon manipulates a variety of surgical instruments using both hands and one foot, while watching the TV monitor screen displaying the image of the patient's body from the laparoscope. A robotic manipulator, instead of a camera assistant, holds the laparoscope. The surgeon is relying on a human–machine interface to carry out the positioning of the laparoscope manipulator. Thus, a user-friendly (i.e., surgeon-friendly) design of the human–machine interface that controls the laparoscope positioner plays an important role in the realization of solo surgery.

Several robotic laparoscope positioning systems have been devised in the last ten years. The first product, named AESOP, was released to the marketplace in 1994 [1] by Computer

Motion, Inc., Goleta, CA, USA. The human–machine interface of most laparoscope positioners proposed in the early years included a joystick (a teleoperation type [2] or instrument-mounted type [3]) or foot pedal [4], and each required the use of the surgeon's hand and/or foot. These types of interfaces, however, seem generally difficult to use because surgeons already use their hands and/or feet to control a variety of surgical tools. To solve this problem, several researchers have introduced a voice-activated system based on the verbal aspect of human speech [4]–[8]. In these systems, the surgeon's voice is at first turned into words or sentences using a voice-recognition engine, and the appropriate actions (as manipulator control commands) are then generated from the recognized texts. Initially, the voice-activated system seemed to be an effective approach because verbal instructions are natural for humans and neither hands nor feet are required to control the laparoscope. However, this system does have some inherent limitations, such as reduced accuracy in positioning, long reaction times, and erratic movements in a noisy environment. We believe that a motion-based laparoscope controller, using the movement of the surgeon's head, is the best solution because nonverbal instructions such as face gestures are more intuitive and faster than verbal instructions. Also, because these gestures have the potential ability to represent not only the direction of scope motion but also the degree of motion, such as velocity, laparoscope positioning accuracy may be improved. Several laparoscope manipulators with a head navigation interface have previously been developed [9]–[13]. Such systems, however, failed to fully utilize the nonverbal features of facial motion. These systems were limited to detecting dominant head gestures, which only served as discrete (verbal) commands, and required not only head movements but also simultaneous control of an additional footswitch. Furthermore, the surgeon had to wear head-mounted sensing devices, such as a headband and gyro sensor, which were stressful for the surgeon.

To make the most of the advantages of nonverbal and noncontact instructions, we have designed a novel human–machine interface, called "FAce MOUSe," for controlling the laparoscope positioner. This proposed human interface is an image-based system which tracks the surgeon's facial motion robustly in real time and does not require the use of body-contact sensing devices. Using the FAce MOUSe interface, we have developed a new robotic laparoscope positioning system for solo surgery. Our system, which seems to be the most convenient for the surgeon, allows nonverbal, hands off and feet off laparoscope operations.

In Section II, we will explain the concept of FAce MOUSe in greater detail.

II. Concept of FAce MOUSe

First of all, we will summarize the degrees of freedom (DOFs) of the laparoscope and the surgeon's face. Then, we will discuss the correlation between laparoscopic and facial motions, and describe the FAce MOUSe concept.

A. Laparoscope DOFs

Due to the constraints imposed by operating through the trocar point, laparoscope movements are kinematically restricted to four DOFs: the first and second DOFs for pivoting about the trocar insertion point, i.e., pan and tilt motions, which produce a translation of the patient's body image on the monitor screen; a third DOF for insertion and retraction along the longitudinal axis, which produces a magnification or reduction of the visualized laparoscopic field; and another DOF for spinning about the insertion axis, which produces a rotation of the visualized image around the image center. The surgeon, however, usually needs to avoid the rotation of the visualized image on the monitor screen during the operation because this demands additional mental effort [14] that is not needed. The camera-holding robot thus requires only three DOFs-two rotation angles (*pan* and *tilt*) and a translation along the camera optical axis (*insertion* and *retraction*), as shown in Fig. 2.

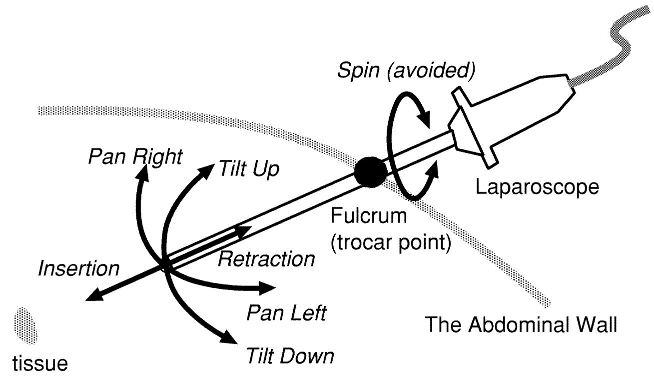


Fig. 2. Laparoscope DOFs.

B. DOFs for the Surgeon's Face

In standard laparoscopic surgery, such as a laparoscopic cholecystectomy, the surgeon usually stands in front of the TV monitor displaying the scope image, and observes the surgical point of interest on the screen. Therefore, we can assume *fronto-parallel* and *distance-constant* interaction, i.e., 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 entire interaction time. Notice that during this kind of interaction, the DOFs for the surgeon's face are reduced from six to three, namely, a translation two-dimensional (2-D) vector (x, y) and a rotation θ in the face plane [see Fig. 3(a) and (b)]. When the motion of the face is small, the translation vector can be approximately replaced with pitch and yaw motions [Fig. 3(c)], while the scalar θ corresponds to the roll motion angle.

C. Correspondence Between Laparoscope Motions and Face Motions in Previous Work

Kobayashi *et al.* [9] related the face-roll, face-pitch, and face-yaw motions to scope-zoom, scope-tilt, and scope-pan motions, respectively. Furthermore, they developed another command method [10], in which the determination of the scope-pan and scope-tilt angles is based on the time of the face-roll motion while the scope-zoom is controlled by the face-pitch motions (in this method, the face-yaw motion is not used for positioning the scope). In their system, the surgeon can drive the laparoscope manipulator by moving his/her head in

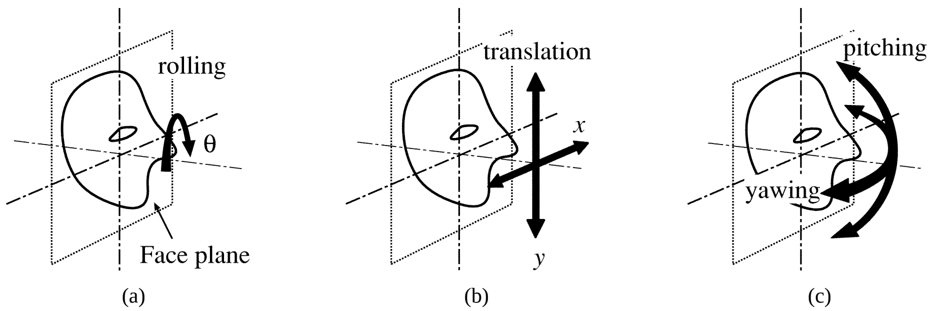


Fig. 3. Face DOFs. (a) A rotation in the face plane (rolling). (b) A translation in the face plane. (c) Pitching and yawing motions.

the appropriate direction, followed by turning on an additional knee switch to confirm the motion request.

In the endoscope-holding robot, EndoSista [11], commercially available from Armstrong Healthcare Ltd., High Wycombe, U.K., a face-rolling motion is used as a toggle switch to change the significance of the pitch motion [12]. At first, a nod motion (face pitching) indicates the camera tilt and a side-to-side head motion (face yawing) indicates the camera pan. After the surgeon makes a roll motion, a nod up indicates the camera retraction (image reduction) and a nod down indicates the camera insertion (image magnification). The surgeon can control the motion of the laparoscope by making such head gestures, in conjunction with a foot switch to avoid the execution of inappropriate input. In the latest version of EndoAssist [13], the foot switch is also used as a trigger instead of rolling the face, which changes the significance of the face-pitch motion.

D. Discussion

Face gestures are nonverbal and seem suitable for the precise control of scope motion. Furthermore, face gestures do not require the use of the hands or feet, and so they are convenient for the operating surgeon. The previous face motion-based laparoscope positioners, however, did not fully utilize the advantages of face gestures.

- They did not allow hands-free and feet-free laparoscopic operations. The surgeon was required not only to make the face gesture but also to control an additional foot or knee switch, which could confuse the surgeon.
- Although face gestures are nonverbal, typical head gestures were related only to several discrete commands (verbal texts) such as “zoom/tilt/pan the scope,” which are inherently the same as voice-activated systems. While this approach is reasonable for discrete, high-level tasks, it is inappropriate for continuous low-level controls, such as the fine adjustment of laparoscopic positions.

To make the most of the merits of face gestures, we worked out a new laparoscope control scheme: FAce MOUSe.

E. FAce MOUSe

Fig. 4 illustrates the FAce MOUSe control scheme. As shown in Fig. 4(a), the state of the robotic laparoscope positioning system can be broken down into the following three states [17].

1) *SHIFT state: guiding the laparoscope for maintaining the surgical point of interest in the center of the video frame*. This state corresponds to pan and tilt camera functions.
2) *ZOOM state: guiding the laparoscope for providing the required target magnification*. This state corresponds to insertion and retraction camera functions.
3) *STILL state: keeping the laparoscope still*. This state means the surgical view is to be fixed.

All surgery work time can be classified into any of these states from the viewpoint of laparoscope operation.

Let us consider how to control not only the laparoscope motion itself but also the transition between these states by making face gestures only. We refer to the face motion for the transition state as the *Trigger Action*, and that for guiding the laparoscope as the *Guiding Action* [see Fig. 4(a)]. The method for positioning the laparoscope through face motions is summarized as follows.

Step 1) Make a Trigger Action to Change the STILL State to the SHIFT/ZOOM State: The thick solid lines in Fig. 4(a) correspond to this step. To complete the transition, the following three consecutive face motions are required [also see Fig. 4(b) and (c)]: 1) put the position and pose of the face in the standard position and pose; 2) roll the face counterclockwise (for *SHIFT*) or clockwise (for *ZOOM*); and 3) return the face “precisely” to the standard position and pose. Note that the surgeon cannot make this action unconsciously.

Step 2) Make a Guiding Action in the SHIFT/ZOOM State: The thick broken lines in Fig. 4(a) correspond to this step. Once the system comes into the *SHIFT* or *ZOOM* state, the face translation is represented as a vector from the standard position, and the direction and magnitude of the vector are, respectively, transformed into the direction and velocity of the laparoscope motion. As shown in Fig. 4(b), the face intuitively shifts parallel to the scope image plane, corresponding to the identical pan and tilt movements of the laparoscope when the state is *SHIFT*. On the other hand, when the state is *ZOOM*, the up and down movements of the face correspond to the zoom-out and zoom-in movements, respectively, of the laparoscope [see Fig. 4(c)].

Step 3) Make a Trigger Action to Change the SHIFT/ZOOM State to the STILL State: The thin solid lines in Fig. 4(a) correspond to this step. All the surgeon has to do is roll the face [see Fig. 4(b) and (c)]. As soon as the rolling motion is detected, the laparoscope motion stops and the state returns to *STILL*. Note that this action is very easy to do.

This control scheme is analogous to that of a computer mouse device (with two buttons), as shown in Fig. 5. Accordingly, this is why we called our scheme FAce MOUSe. *Guiding Action*, in which the 2-D translation in the face plane is dominant, corresponds to the mouse body movement on the mouse pad plane.

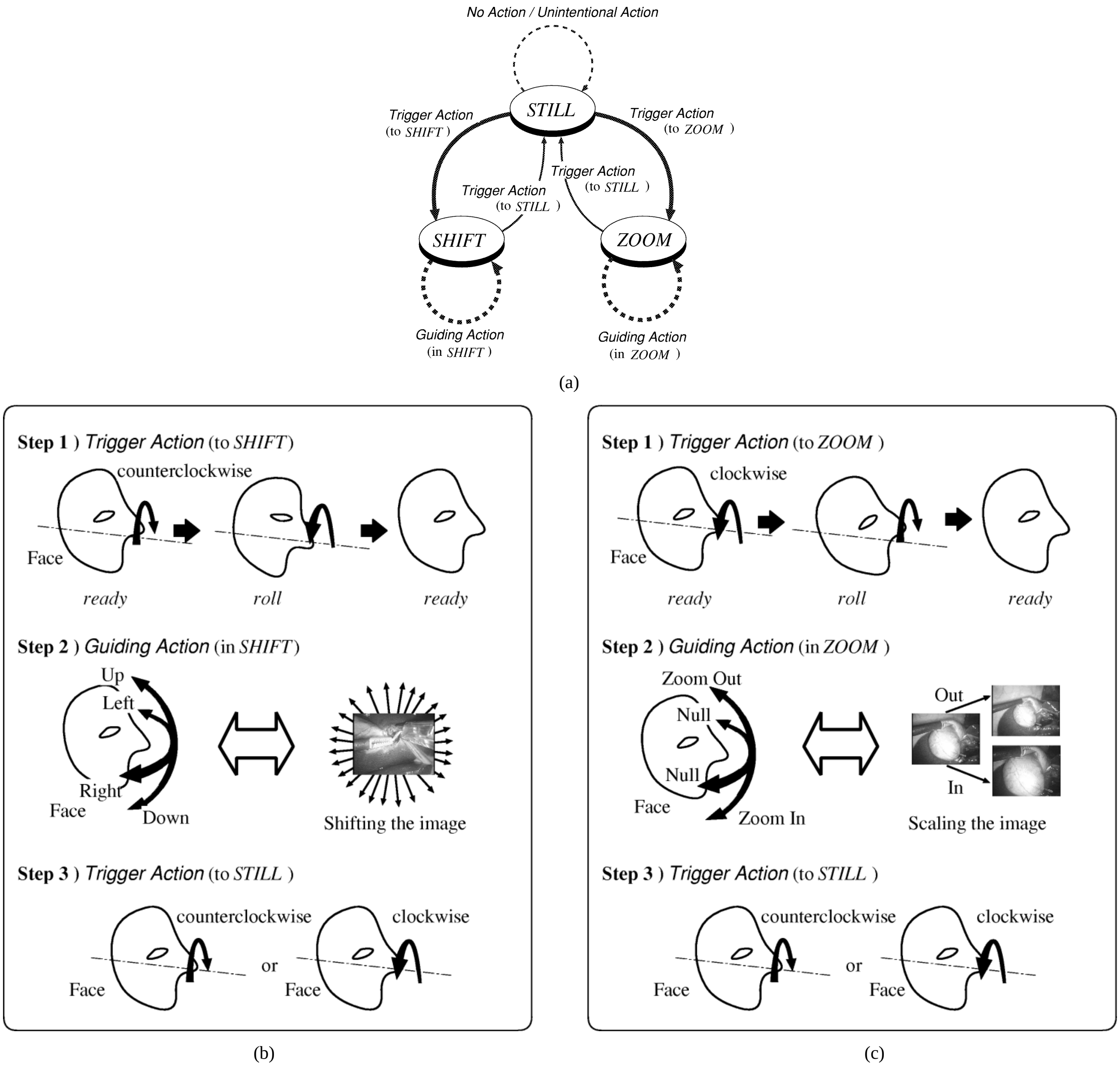


Fig. 4. FAce MOUSe control scheme. (a) State transition diagram. (b) Face motions for pan/tilt camera functions. (c) Face motions for scope insertion/retraction.

Trigger Action, in which the 1-D rotation in the face plane is dominant, corresponds to mouse button operations such as click (i.e., press and release the button). The correspondence between the *Guiding* and *Trigger Actions* and mouse device operation is summarized in Table I.

Our FAce MOUSe control scheme is based on the nonverbal aspect of human gestures and provides the surgeon with a means of total hands-off and feet-off laparoscope operations, while also achieving rapid reaction and high positioning accuracy. To maintain high levels of "safety" during surgery, however, we must note the following two points.

- Without exception, unintentional movements, which could be misunderstood as *Trigger Action* (to *SHIFT*) or *Trigger Action* (to *ZOOM*) in the *STILL* state, should be avoided.
- *Trigger Action* (to *STILL*) should be definitely and immediately recognized in the *SHIFT/ZOOM* state.

We performed a successful FAce MOUSe implementation by paying great attention to these points.

III. Implementation

A. FAce MOUSe System Overview

We designed a novel human–machine interface for controlling the laparoscope with the above method. The system configuration and overview are shown in Fig. 6. Our laparoscope positioning system, the FAce MOUSe system, consists primarily of a charge-coupled device (CCD) camera placed just over the TV monitor, an all-purpose PC (CPU: Intel Pentium III, 600MHz, OS: Vine Linux 2.0) with a video-capturing device, a robot manipulator that holds the laparoscope, a scan converter for superimposing graphics on the scope image, and a foot switch [see Fig. 6(a) and (b)]. For face tracking, we used a commercially available zoom camera, Sony EVI-D30, whose

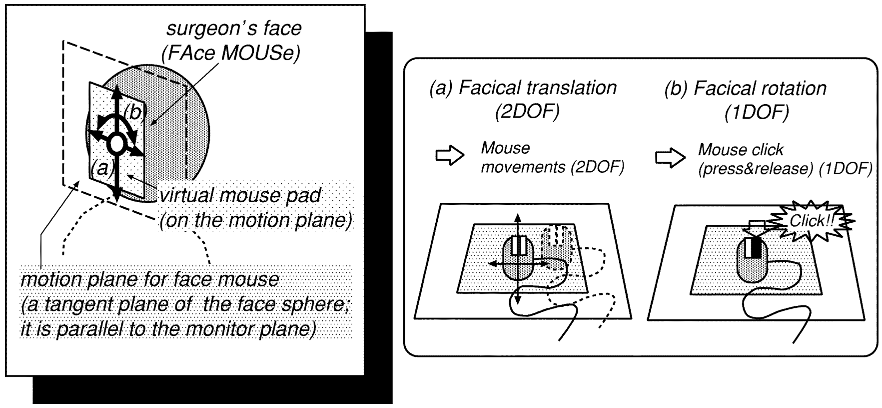
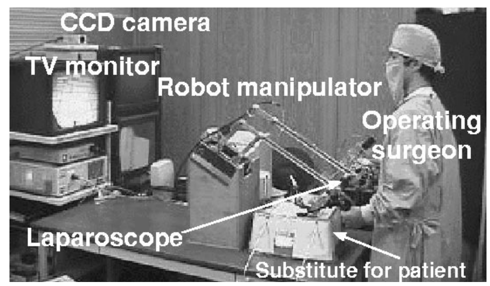


Fig. 5. Analogy between facial motion and mouse operation.

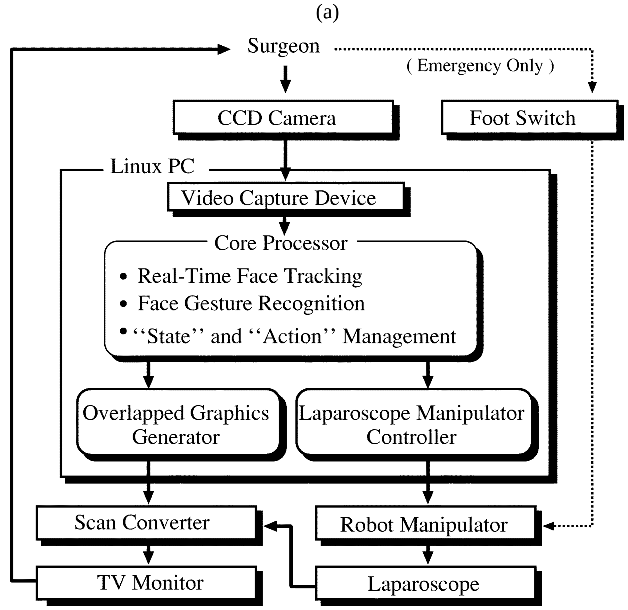
TABLE I
CORRESPONDENCE BETWEEN GUIDING/TRIGGER ACTIONS AND MOUSE DEVICE OPERATIONS

FAce MOUSe	Computer mouse device
Trigger Action (to *SHIFT*)	Click the left button
Trigger Action (to *ZOOM*)	Click the right button
Trigger Action (to *STILL*)	Press the left/right button
Guiding Action	Slide the mouse body on the pad

pan/tilt/zoom can be controlled manually using an infrared (IR) remote commander. A systematic calibration process is not necessary; all one needs to do before using the FAce MOUSe interface during surgery is to use the remote commander to roughly adjust the camera parameters so that the surgeon's face is near the center of the video image and the projection size is appropriate. Once the camera is set in the appropriate position and pose, the core system in the PC can detect and track the surgeon's facial features in real time (30 Hz) from a sequence of video images captured through the CCD camera. Assuming that the surgeon's face is moving on a virtual plane parallel to the TV monitor screen (that is, the surgeon's face has only three DOFs), the system estimates the position and pose of the surgeon's face in real time from the image-processing result and then recognizes the surgeon's facial gestures (i.e., the *Trigger* and *Guiding Actions*). According to the state of the system and the gestural action recognition result, the control command is sent to the laparoscope manipulator. The kinematic mechanism of this manipulator was inspired by and is similar to that used in the LARS surgical robot designed by Taylor *et al.* [3]. As shown in Fig. 7, our manipulator has a planar double-parallelogram mechanism (two DOFs) and a ball-screw mechanism (one DOF). The planar double-parallelogram mechanism allows movements only around an arc whose center is the point of trocar insertion, which corresponds to the pan and tilt scope motions and produces a translation of the patient's body image on the monitor screen, with respect to the surgeon's reference work frame. The ball-screw mechanism enables translation along the longitudinal axis of the laparoscope, which corresponds to the zoom-in and zoom-out motions and produces a magnification or reduction of the visualized laparoscopic field. The system state and image-processing results are also superimposed graphically on the laparoscope image using a scan converter (as feedback information for the surgeon), and



(a)



(b)

Fig. 6. FAce MOUSe system configuration. (a) System overview. (b) Hardware components.

the resulting image is displayed on the TV monitor. In the current version, an emergency foot switch for stopping a scope movement is also provided, but there has been no need to use it.

B. Software Components Overview

Fig. 8 shows the overview of the FAce MOUSe software system. The following three processes run on the main computer: FAce MOUSe core process (CORE), overlapped graphics generator (OGG), and laparoscope manipulator controller (LMC).

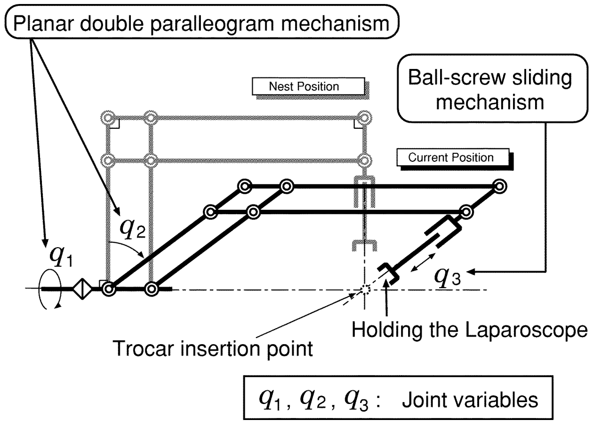


Fig. 7. Mechanism of laparoscope manipulator.

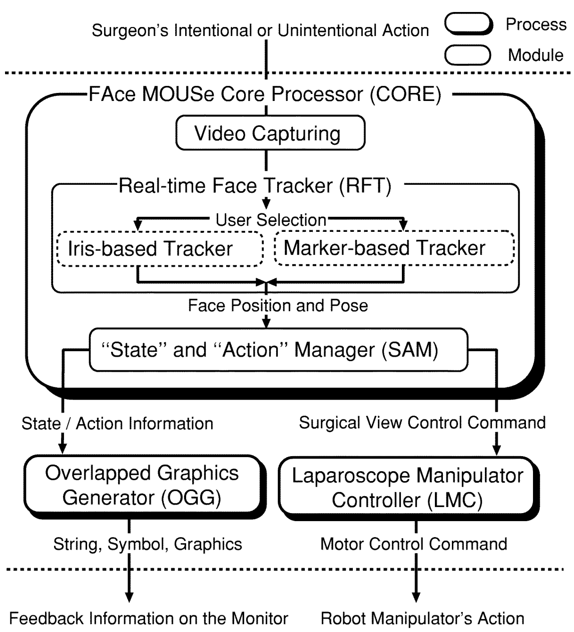


Fig. 8. FAce MOUSe software components.

The CORE process mainly consists of the following two software modules: real-time face tracker (RFT) and state and action manager (SAM). The RFT module detects and tracks the surgeon's facial features in real time from the image sequence captured through the surgeon-surveillance CCD camera. The SAM module manages the state of the laparoscope manipulator and the surgeon's facial actions, based on the time sequence of the position and pose coming from the RFT module. As a result, the CORE process generates the surgical view control command and outputs it to the LMC process while sending the state and action information to the OGG process.

The OGG process outputs feedback information for the surgeon graphically on the PC display, and the resulting image (graphics) is synthesized with the laparoscopic image signals using the scan converter.

The LMC process converts the surgical view control commands to motor control commands and then drives the robot manipulator.

In Sections III-C–E, we will explain each software component in more detail.

C. FAce MOUSe Core Process (CORE)

1) Real-Time Face Tracker (RFT): During surgery, the laparoscopic surgeon usually wears a gown, cap, and mask, so that almost all facial features, such as the mouth, nose, and hair do not appear in the surveillance image. Taking this into account, we developed two robust, RFTs: an iris-based tracker and a marker-based tracker. The iris-based tracker extracts, at most, two black circular shapes, which are the surgeon's irises, from the surveillance image using simple thresholding and the Hough transformation technique. The marker-based tracker uses simple thresholding and a conventional labeling algorithm to detect and track a black rectangular marker attached to the surgical cap in advance. Although the iris-based method requires careful selection of the thresholding and Hough transformation parameters according to illumination conditions or individual variations of the visible size of the irises (the irises are partially occluded by the eyelid), there is no difference in the tracking performance between the two face trackers. Both trackers can work at a frequency of 30 Hz. The details of these face-tracking methods can be found in [18]. Assuming *distance-constant and fronto-parallel interaction* (described in Section II-B), the system estimates the position and pose of the surgeon's face in real time from the image-processing result. When using the iris-based tracker, the midpoint of the line segment joining the centroids of the left and right irises is known as the position of FAce MOUSe, and the angle between this line segment and the horizontal axis of the image (i.e., the x axis) is known as the pose of FAce MOUSe [see Fig. 9(a)]. When using the marker-based tracker, the centroid of the marker and the angle between the principal axis of inertia of the marker region and the image x axis are regarded as the position and pose of FAce MOUSe, respectively [see Fig. 9(b)]. As shown in Fig. 8, the surgeon can select either tracker in advance as the FAce MOUSe tracker (in the current version, we cannot use both trackers simultaneously).

In actual practice, it could be difficult sometimes to set the TV monitor just in front of the operating surgeon due to the presence of other surgical equipment in the operating room. Furthermore, because we did not require a precise setting of the face-tracking camera, affine distortion of the face plane, the image plane and the TV monitor plane could result. Nevertheless, our system works well because we utilize the relative displacement (deviation) of FAce MOUSe, not the absolute coordinates, and the surgeon can always monitor the relative change of the detected face position and pose by checking the indicator superimposed on the TV monitor (even if distortion exists, the surgeon can flexibly adjust his/her face motion). For details, see Sections III-C.2 and D.

2) State and Action Manager (SAM): The software module SAM recognizes the surgeon's face gestures based on the time sequence of the position and pose of FAce MOUSe, and controls the state of the laparoscope manipulator. We will first define the three virtual mouse pads and two special states of FAce MOUSe.

Three virtual mouse pads (see Fig. 5): We denote the three rectangular regions, which correspond to *virtual* mouse pads superimposed on the face motion plane, as D_{ready}, D_{trigger}, and D_{guiding} ($D_{\text{ready}} \subset D_{\text{trigger}} \subset D_{\text{guiding}}$). D_{ready} indicates the standard position of FAce MOUSe. D_{trigger} and D_{guiding} are

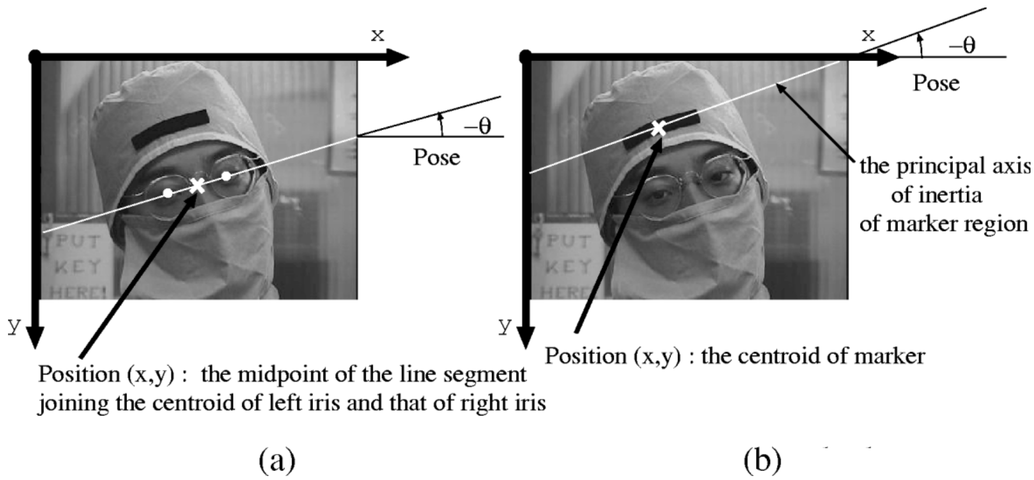


Fig. 9. Estimating the position and the pose of the surgeon's face. (a) Iris-based method. (b) Marker-based method.

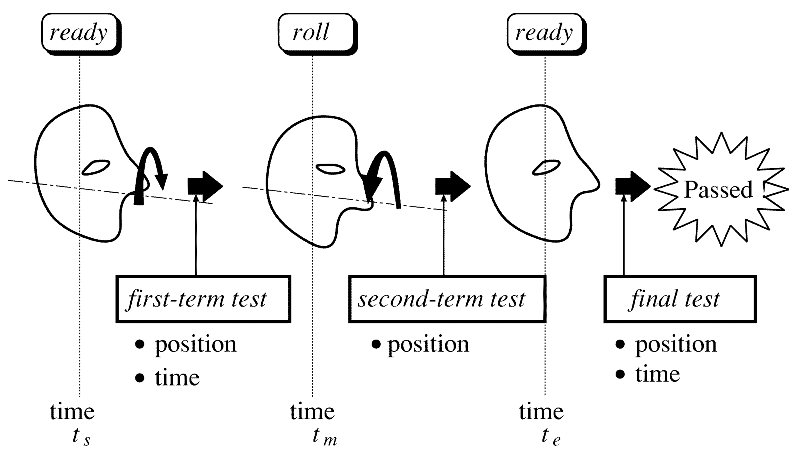


Fig. 10. Recognition process of *Trigger Action*.

the virtual mouse pads for *Trigger Action* and *Guiding Action*, respectively. As an example, the *Guiding Action* must be made within the region D_{guiding}.

Two special states of FAce MOUSe: Let $(x(t), y(t))$ and $\theta(t)$ be the position and pose, respectively, of FAce MOUSe coming from the RFT module at time t. For convenience, we define the following special states of FAce MOUSe.

- *ready* state:
  The state of FAce MOUSe is referred to as "*ready* at time t" if both $(x(t), y(t)) \in \mathcal{D}_{\text{ready}}$ and $|\theta(t)| < \theta_{\text{flat}}$.
- *roll* state:
  The state of FAce MOUSe is referred to as "*roll* at time t" if both $(x(t), y(t)) \in \mathcal{D}_{\text{trigger}}$ and $|\theta(t)| \geq \theta_{\text{roll}}$.

θ_{flat} and θ_{roll} indicate the thresholds for restricting the pose of FAce MOUSe and $0 \leq \theta_{\text{flat}} \leq \theta_{\text{roll}}$ (in practice, $\theta_{\text{flat}} = 2°$, $\theta_{\text{roll}} = 12°$).

The process in the SAM module varies with the state of the laparoscope manipulator, as follows.

a) *The case where the state of the laparoscope manipulator is STILL.*

The system regularly checks whether the state of FAce MOUSe changes to *ready*. The following special process is executed just after the *ready* event occurs (at most, this process occurs during the 5 s period $T_{\max}$, which will be defined below). Let t_s be the time when the *ready* transition event has occurred. If the FAce MOUSe state changes to *roll* at time $t = t_m (> t_s)$ while passing the *first-term test*, and if it returns to *ready* at time $t = t_e (> t_m)$ while passing the *second-term test*, and if the *final test* is also passed, then the face action during time $[t_s, t_e]$ is regarded as a *Trigger Action* (see Fig. 10), and the state of the laparoscope manipulator is changed according to the rolling motion angle at time $t = t_m$: the state is changed from *STILL* to *SHIFT* if $\theta(t_m) > 0$, or *STILL* to *ZOOM* if $\theta(t_m) < 0$. The details of the above three-term tests are the following.

- The *first-term test* (from *ready* to *roll*)
  The following conditions must be satisfied simultaneously:

$$(x(t), y(t)) \in \mathcal{D}_{\text{trigger}} \text{ during } t = [t_s, t_m) \tag{1}$$

$$t_m - t_s < T_{\text{stay}}. \tag{2}$$

- The *second-term test* (from *roll* to the second *ready*)
  The following condition must be satisfied:

$$(x(t), y(t)) \in \mathcal{D}_{\text{trigger}} \text{ during } t = [t_m, t_e). \tag{3}$$

- The *final test*.

  The following conditions must be satisfied simultaneously:

$$|x(t_e) - x(t_s)| \le \varepsilon_x \tag{4}$$

$$|y(t_e) - y(t_s)| \le \varepsilon_y \tag{5}$$

$$t_e - t_s < T_{\max} \tag{6}$$

  where ε_x and ε_y indicate predetermined positive constants. T_{stay} and $T_{\max}$ are the thresholds for restricting the time for action, and $T_{\text{stay}} < T_{\max}$ (in practice, $T_{\text{stay}} = 2$ s, $T_{\max} = 5$ s).

b) *The case where the state of the laparoscope manipulator is SHIFT/ZOOM.*

If either the condition

$$(x(t), y(t)) \notin \mathcal{D}_{\text{guiding}} \tag{7}$$

or the condition

$$|\theta(t)| \ge \theta_{\text{roll}} \tag{8}$$

is satisfied, then the system considers the surgeon's face action as a *Trigger Action* and changes the state of the laparoscope manipulator to *STILL* while sending the 3-D vector $\mathbf{0} = (0,0,0)^T$ (indicating the stop of laparoscope motion) to the laparoscope manipulator controller (LMC). Otherwise, the system regards the surgeon's facial gestures as a *Guiding Action* and calculates the face deviation vector $\boldsymbol{x}(t)$ as

$$\boldsymbol{x}(t) = (x(t) - x_{\text{origin}}, y(t) - y_{\text{origin}})^T \tag{9}$$

where $(x_{\text{origin}}, y_{\text{origin}})$ indicates the position of FAce MOUSe just after the state of the laparoscope manipulator changes from *STILL* to *SHIFT/ZOOM*. Then, the system outputs the deviation vector $\boldsymbol{x}(t)$ to the overlapped graphics generator (OGG) as one portion of the feedback information for the surgeon, while sending the following 3-D vector to the LMC process as the surgical view control command:

$$\boldsymbol{v}(t) = \begin{cases} (\boldsymbol{x}(t)^T, 0)^T, & \text{if the state is } SHIFT \\ \left(0, 0, \boldsymbol{x}(t)^T \begin{pmatrix} 0 \\ 1 \end{pmatrix}\right)^T, & \text{if the state is } ZOOM. \end{cases} \tag{10}$$

The first and second elements of the command vector $\boldsymbol{v}(t)$ correspond to the horizontal and vertical components of the view shifting velocity, while the third element indicates the view scaling velocity along the optical axis.

3) Consideration of Safety: Our implementation takes surgery safety into account. Basically, *Trigger Action* (to *SHIFT/ZOOM*), representing a left/right click of the mouse, is made by changing the state of FAce MOUSe from *ready* to *roll* and by returning to *ready*, as shown in Fig. 10. Notice that the *ready* state is natural for the surgeon. Therefore, FAce MOUSe may often change to the *ready* state by the surgeon's unintentional movements. But, in such a case, it would be very rare to pass the *first-term test* because the *roll* state requires the surgeon to make unusual movements. On the other hand, when the surgeon feels fatigue or doubt, the angle of the head may indicate a slight leaning. In the current implementation, the system cannot recognize such fatigue or doubt signs. As a result, the system may regard this natural leaning of the surgeon's head as a *roll* gesture. Even in this case, it is very difficult to pass all three tests. Especially, such nonmanipulative gestures do not seem to satisfy the final conditions (4)–(6), which require the surgeon to make the fast and fine positioning of FAce MOUSe.

Compared with *Trigger Action* (to *SHIFT/ZOOM*), it is much easier to make *Trigger Action* (to *STILL*) in the *SHIFT/ZOOM* state. All the surgeon has to do is satisfy (7) or (8). This provides the means for a rapid stop of laparoscope movements. Although unintentional movement may infrequently satisfy these conditions, a high level of safety can be maintained because this type of misunderstanding only results in stopping of the scope movement.

D. Overlapped Graphics Generator (OGG)

The OGG process assists the surgeon in making the appropriate face gesture by transforming the following information, which comes from the SAM module in the CORE process, into strings, or symbols, or graphics.

1) The state of the laparoscope manipulator as a string. According to the state of manipulator, one of three strings ("Still," "Shift," or "Zoom") is superimposed at the top center of the laparoscope image.
2) The state of FAce MOUSe as a string.

   In the current version, when the state of the laparoscope manipulator is *STILL* and the state of FAce MOUSe is *ready*, the string "Ready" is displayed at the top-center position, instead of the string "Still". Furthermore, the string "Stillhold" is output at the same position when the time constraint (2) or (6) is violated.
3) The position and pose of FAce MOUSe and the virtual mouse pad as graphics.

   The FAce MOUSe position (x, y) and pose θ and the virtual mouse pad (either $\mathcal{D}_{\text{ready}}$, $\mathcal{D}_{\text{trigger}}$, or $\mathcal{D}_{\text{guiding}}$), which are graphically represented in Fig. 11, are displayed at the top-left corner of the monitor screen. The kind of virtual mouse pad is automatically selected according to the state of the laparoscope manipulator.
4) The face deviation vector $\boldsymbol{x}(t)$ as a symbol.

   The face deviation vector $\boldsymbol{x}(t)$ is drawn as a vector at the center of the laparoscope image when the state of the laparoscope manipulator is *SHIFT* or *ZOOM*.

Examples of OGG outputs overlaid on the laparoscopic image are shown in Fig. 12.

E. Laparoscope Manipulator Controller (LMC)

The LMC process transforms the view control command vector $\boldsymbol{v}(t) = (v_1, v_2, v_3)^T$ (10) into the motor control command $\dot{\boldsymbol{q}}(t) = (\dot{q}_1, \dot{q}_2, \dot{q}_3)^T$ as follows:

$$\text{For } i = 1,2,3\text{: } \dot{q}_i = \begin{cases} K_i v_i, & \text{if } v_i > \epsilon_i \\ 0, & \text{if } v_i \le \epsilon_i \end{cases} \tag{11}$$

where $q_i (i = 1,2,3)$ indicates the joint variable of the 3-DOF laparoscope manipulator shown in Fig. 7. K_i indicates a proportional constant and ϵ_i is an infinitesimal value. By filtering using ϵ_i, the sway caused by unintentional body movements or

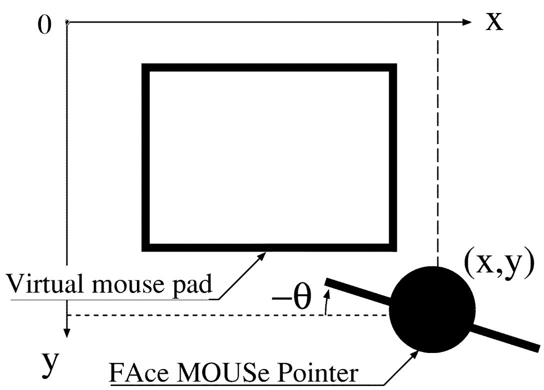
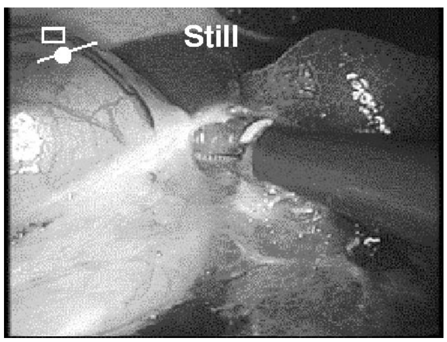
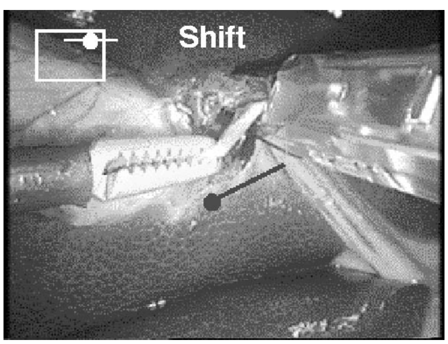


Fig. 11. Graphical representation of FAce MOUSe and virtual mouse pad.



(a)



(b)

Fig. 12. Examples of feedback information overlaid on the laparoscopic image.

image noise can be eliminated, and also the surgeon can easily stabilize the motion of the laparoscope.

Our FAce MOUSe system is applicable to other types of laparoscope manipulators or other medical robots by simply rewriting (11) according to the robotic kinematics. Please note that, in our case, this transformation equation is very simple due to the remote-center-of-motion mechanism in which the motions of all axes are kinematically decoupled at the insertion point.

IV. Experiment

To evaluate the performance of our system, an experiment was first conducted under laboratory conditions using a conventional laparoscopic surgical training box and standard laparoscopic equipment (all from Olympus Optical Co. Ltd, Tokyo, Japan). Next, an *in vivo* experiment was carried out, in which

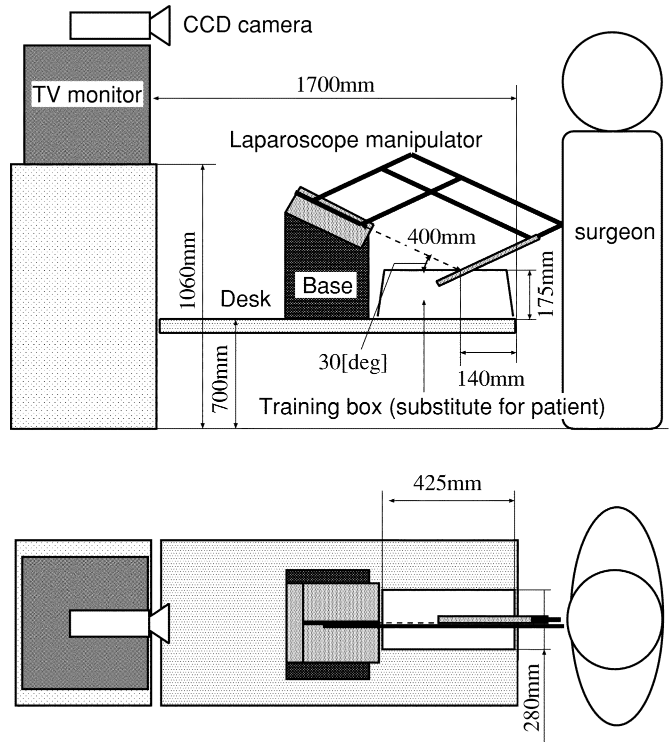


Fig. 13. Setup of laboratory experiment.

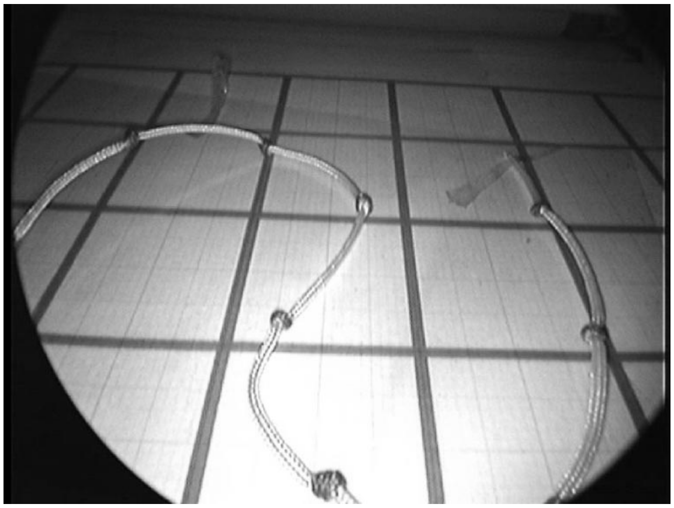
a surgeon used our system in a conventional laparoscopic operating room environment and a regular set of instruments (also all from Olympus Optical Co. Ltd) to perform a laparoscopic cholecystectomy on a pig.

A. Laboratory Experiment

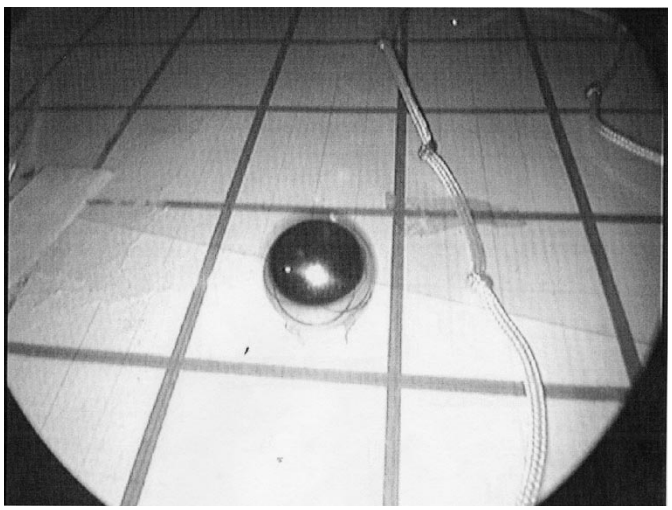
The following two tasks were used to evaluate the basic performance of our system: 1) directing the scope along a 25-cm rope (for evaluating scope movements in the *SHIFT* state) and 2) magnification of a 10-mm ball (for evaluating scope motion in the *ZOOM* state). Note that these tasks did not require a surgical operation or surgical instruments, such as forceps and scalpel. This experiment involved three subjects (named A, B, C). Subject A was familiar with the FAce MOUSe interface but the other two subjects had never used the system before. A schematic representation of the experimental setup is shown in Fig. 13 (see Fig. 6(a) also).

1) Task 1: The 25-cm rope was placed in the training box so that it formed a shape like a reversed "S" [see Fig. 14(a)]. The rope contained 10 knots at 20-mm intervals. The visual field of the camera was displayed on a 400×300-mm monitor, and a 40-mm circular mark was put on the screen. The three subjects were asked to direct the scope from one end of the rope to the other, such that the 10 knots passed through the circular mark in order and, at the end, the rope tip was at the center of the mark. A referee measured the time taken to complete the task with a stopwatch and recorded the distance between the center of the mark and the end point of the rope on the display screen, as a measure of positioning accuracy. This procedure was repeated for 10 consecutive trials per subject.

The times for task completion and the final positioning errors for each trial are plotted in Fig. 15(a) and (c), where the three



(a)



(b)

Fig. 14. Examples of laparoscopic images taken at the laboratory experiment. (a) A 25-cm rope with 10 knots. (b) A 10-mm metallic ball.

curves correspond to the three subjects A, B, and C. The overall mean time for task completion was 45 s. The improvement in speed due to learning between trial 1 and trial 10 was about 42% of the initial trial time (the mean time for trial 1 was 62 s; for trial 10, it was 36 s). The mean and maximum positioning errors were 4.8 mm and 12 mm, respectively, which corresponded to 0.97% and 2.4% of the diagonal length of the monitor screen.

2) Task 2: A 10-mm metallic ball was placed in the training box [see Fig. 14(b)] and initially the laparoscope was set so that the display size of the ball was 30 mm. Also, a 100-mm circular mark was put on the monitor. The three subjects were asked to magnify the ball so that its size was consistent with the size of the mark. A referee measured the task completion time and recorded the final display size of the ball as a measure of accuracy. This was repeated for 10 consecutive trials per subject.

The task completion times and the accuracy measures at each trial are plotted in Fig. 15(b) and (d). The learning curve for Task 2 is almost flat, while the Task 1 curve does not level off within ten trials. This happened because Task 2 was simpler than Task 1, in which the subject had to make a variety of face motions. The overall mean time for task completion was 18 s. The mean and maximum positioning (magnification) errors were 1.3 and 5 mm, respectively (i.e., 1.3% and 5% of the target size).

3) Summary of Results: In all of the tests, we did not find any case in which our system was misguided. All subjects performed both tasks correctly and nonstop for all trials. These experimental results demonstrate the effectiveness of FAce MOUSe, such as easy camera guidance and high positioning accuracy. Notice especially that it is very difficult to use the conventional head navigation systems or voice-activated systems to perform nonstop pan and tilt scope operations along a complex, curved path, as used in Task 1. Although one of the subjects was familiar with our system, there were no differences in task time and error among the three subjects.

B. *In Vivo Experiment*

In the above laboratory experiment, the use of surgical instruments was not included in the assessment tasks. In using our system to perform a real operation, however, the surgeon would have to make face gestures while also precisely controlling surgical tools with his/her hands and/or foot. Fatigue caused by much longer operation times (than those for Task 1 and Task 2) may have a negative influence on face motions. Furthermore, the surgeon may sometimes move his/her face in an extreme and unintentional manner because of tool extraction or insertion, or conversation with another person in the operating room. Even in these cases, the system should maintain a high level of safety. To evaluate the applicability of our system to clinical use, an *in vivo* laparoscopic cholecystectomy was performed on a pig.

1) Experimental Setup: A schematic representation and photograph of the *in vivo* experimental setup are shown in Fig. 16. Four insertion holes were made on the abdominal wall of the pig prior to the surgical operation. In the Fig. 16 schematic representation, L and R indicate the trocar points for instruments used by the surgeon, C indicates the insertion point for the laparoscope, and A indicates the insertion position of additional instruments. After trocar insertion, the laparoscope manipulator was mounted on a holder hanging over the surgical table, on which the pig was already in place. The laparoscope manipulator was positioned precisely so that its remote rotation center was consistent with trocar point C (see Fig. 7). The setup time for the manipulator was about 20 min. Instead of a human camera assistant, the system was used for the entire procedure until the removal of the gallbladder.

In addition to the operating surgeon and robotic camera assistant, another surgeon took part in the laparoscopic cholecystectomy experiment as an assistant. He undertook the responsibility of lifting up the liver, which was hanging over the gallbladder to be removed. He performed this task only by supporting the liver with an additional instrument inserted through trocar point A. Incidentally, this task could be performed without an assistant surgeon by introducing a passive instrument holder (see [15] for an example). Therefore, the operating surgeon was in a situation similar to solo surgery.

2) Results: The entire operative procedure was successfully and safely completed with our system. No one used the emergency footswitch for shutting down the system. The number of lens cleanings was also zero. The operating time inclusive from trocar insertion until the removal of the gallbladder was about

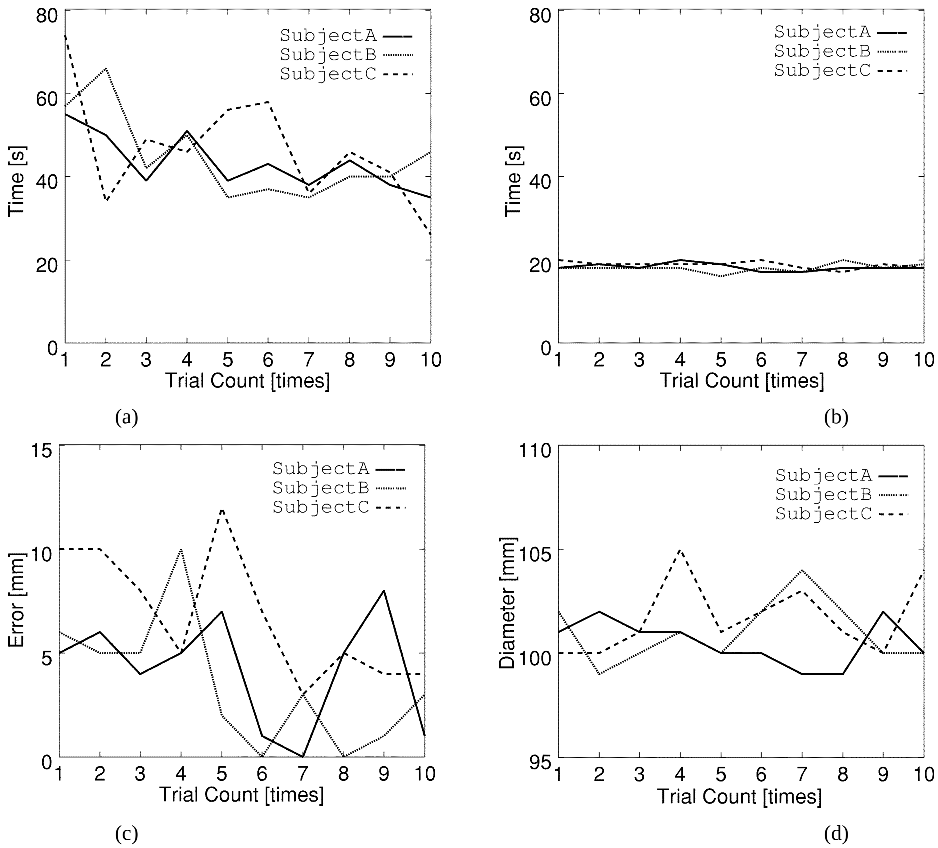


Fig. 15. Results of laboratory experiment. (a) Time to complete Task 1. (b) Time to complete Task 2. (c) Positioning accuracy in Task 1. (d) Positioning accuracy in Task 2.

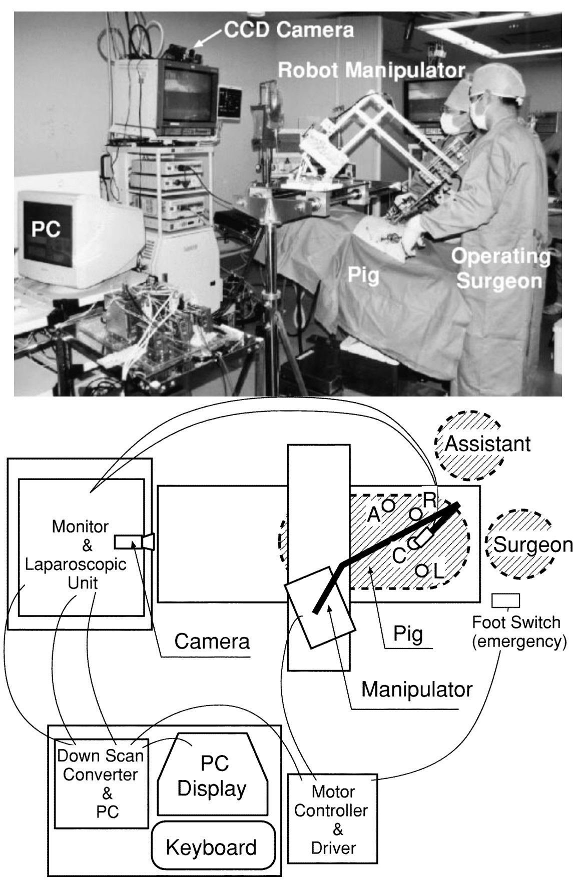


Fig. 16. Setup of *in vivo* experiment.

44 min (2642 s), which was broken down into 2113 s (80.0%) for the *STILL* state, 311 s (11.8%) for the *SHIFT* state, and 218 s (8.2%) for the *ZOOM* state. The number of state transitions for *STILL* → *SHIFT* and *STILL* → *ZOOM*, and vice versa, were 40 and 50, respectively. In this experiment, the robot never obstructed the surgeon's work, and worrisome incidents and technical problems did not occur. Fig. 17 shows scenes of the surgeon's facial motions in the experiment. The upper part of Fig. 17 consists of the scope images which the surgeon looked at; the lower part consists of the images of the surgeon's face from the surveillance camera. Each pair of images was taken at the same time.

The number of times during the operation that the surgeon made a *Trigger Action* to the system to drive the laparoscope manipulator (i.e., Step 1 of Fig. 4) was 97 times, which breaks down into 90 times for being recognized correctly by the system and seven times for not being recognized correctly. Although 252 *ready* transitions (including the intentional 97 transitions) were observed in the *STILL* state, the system never mistook any other motion of the operating surgeon or any other surgeon who was present during the experiment. (e.g., see the second, third, and fourth images of Fig. 17, which show a surgeon walking behind the operating surgeon.) This result demonstrates that our face-gesture recognition algorithm in the SAM worked well.

The number of times that the surgeon made face gestures to stop the laparoscope motion (i.e., Step 3 of Fig. 4) was 90 and these were all correctly recognized. We received many positive comments, such as fast reaction time, high positioning accuracy,

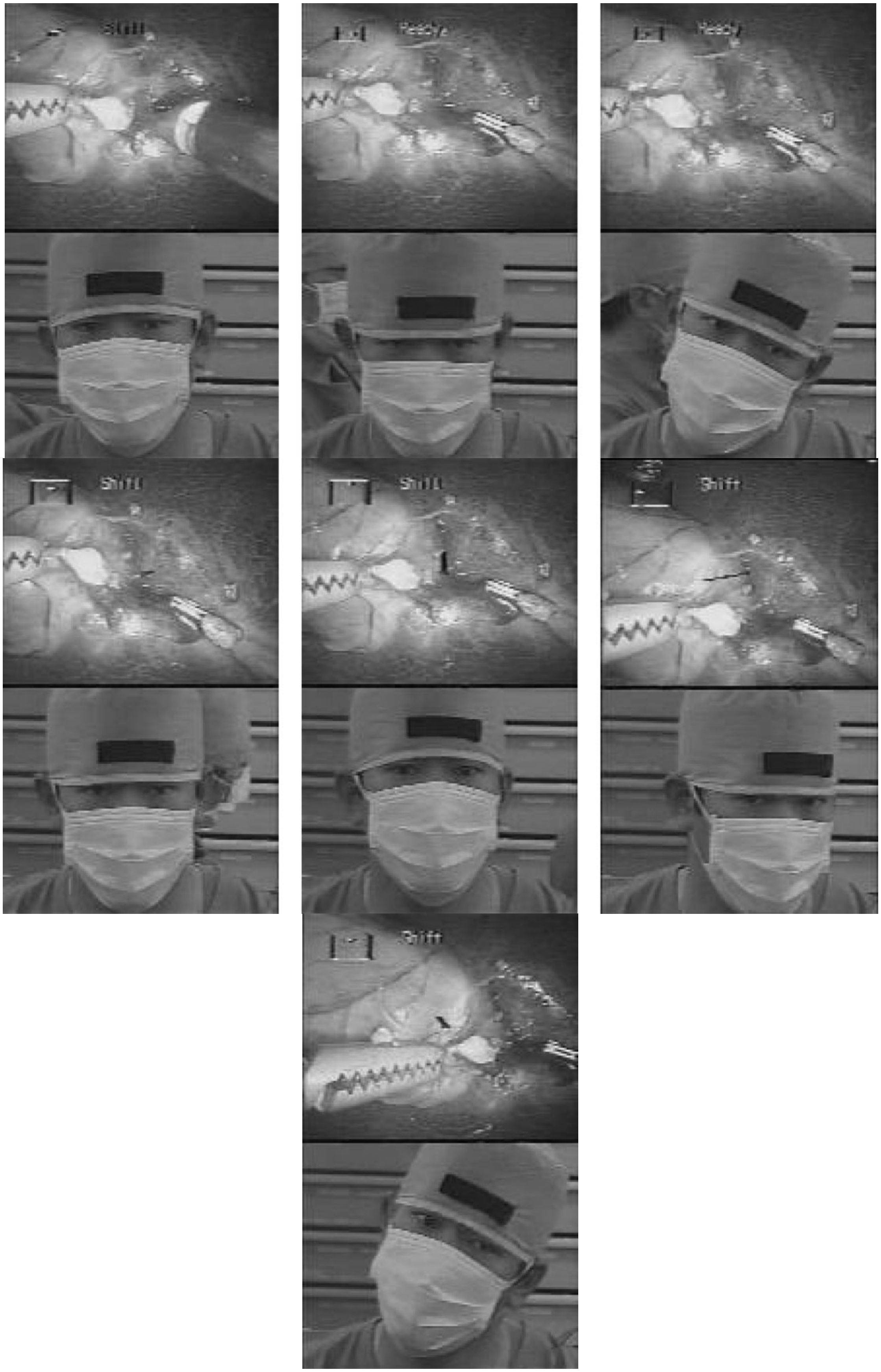


Fig. 17. An *in vivo* laparoscopic cholecystectomy experiment using the FAce MOUSe system.

and easy and intuitive camera guidance (for Step 2 of Fig. 4), from the surgeons who performed or looked in during the experiment. The operating surgeon, however, did make the negative comment that after the experiment he felt a little fatigue in his neck from a lot of rolling face motions. In fact, he made rolling gestures 187 times $(= 97 + 90)$, that is, 4.2 times/min during the operation.

V. Comparative Study With a Solo Surgery Simulator

To test the advantage of the proposed system compared to other existing human–machine interfaces, a comparative experiment was conducted using a laparoscopic solo surgery simulator. The reason for using the simulator, not the real system, is because we wanted reproducibility and objectivity in our comparative evaluation.

A. Solo Surgery Simulator

Our solo surgery simulator consists primarily of a main computer (CPU: AMD Athlon 1.4 GHz, OS: Windows 98) and a commercially available instrument simulation device, VLI (Immersion Corporation, San Jose, CA, USA), which tracks the 3-D motion of a pair of laparoscopic surgical instruments. The computer simulates a laparoscope, patient body and body tissue, a camera-holding robot manipulator, and two surgical instruments, and generates virtual laparoscopic images using the Open GL graphics library. This simulator has several standard communication capabilities, such as RS232 serial ports so that a "real" human–machine interface such as FAce MOUSe can be directly connected with the simulator instead of a "real" laparoscope manipulator. The details of our simulator can be found in [20] and [21].

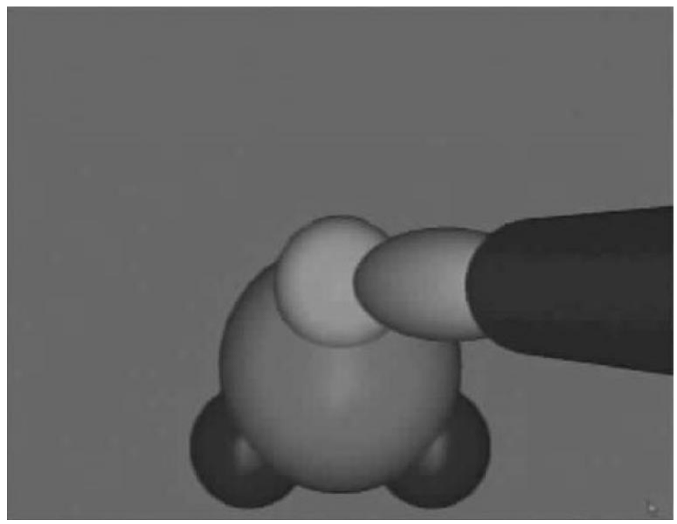
B. Voice Control Interface

We developed a voice-activated system as an alternative laparoscope control interface [21]. Our voice control system has two modes of movement (continuous and discontinuous), similar to that of the commercial voice-activated laparoscope-holding robot, AESOP [4]. Continuous movement is activated with a single Japanese word direction: "HIDARI" ("left"), "MIGI"("right"), "UE" ("up"), "SHITA" ("down"), "KAKUDAI" ("magnification"), "SHUKUSYOU" ("reduction"). The robot would move the laparoscope in the desired direction until the command "MARU" ["period" (full stop)] is voiced. The discontinuous mode uses the same commands as before, preceded by the adverb "SUKOSHI" ("a little"), and the robot would move the camera toward the corresponding direction by a predetermined short distance.

C. Assessment Task

The assessment task was designed to simulate the process of a "dissection of the gallbladder from the liver," which is one of the basic and frequent operations in a laparoscopic cholecystectomy. During this process, the surgeon applies a cutting tool to the dissection location in the gallbladder bed (and repeats this same action until the dissection is complete). Fig. 18 shows the comparison between a real laparoscopic image taken during the actual process of the dissection of the gallbladder from the liver, and the corresponding image generated by our simulator. The large sphere and three small spheres correspond to the gallbladder and three dissection locations on the gallbladder surface, respectively. The basic graphical model and instrument operation portion of the assessment task were inspired by those of the commercially available, computer-based surgical trainer, MIST VR [16]. Because the camera movement process was not included in the MIST VR tasks, we worked out a camera-control induction mechanism for this assessment of laparoscope control interfaces.

(a)



(b)

Fig. 18. Screenshots of "dissection of gallbladder from liver." (a) A real laparoscopic image taken during an actual cholecystectomy operation. (b) A virtual laparoscopic image generated by our simulator.

Camera-control induction mechanism: Let d_{scope} be the distance between the tip of the laparoscope and the center of gravity of the target sphere. Also, let θ_{scope} be the angle between the direction of the longitudinal axis of the laparoscope and the direction from the tip of the scope to the centroid of the target. In our simulator, the target under consideration can be touched or grasped if, and only if, it is captured with the appropriate magnification in the video center, that is, $d_{\mathrm{scope}} < T_d$ and $\theta_{\mathrm{scope}} < T_\theta$, where T_d and T_θ indicate predetermined thresholds. This inevitably produces regular camerawork without exception.

The state of the target sphere is represented by its color, which serves as feedback information for the operator. For example, the operator can tell whether the above condition concerning

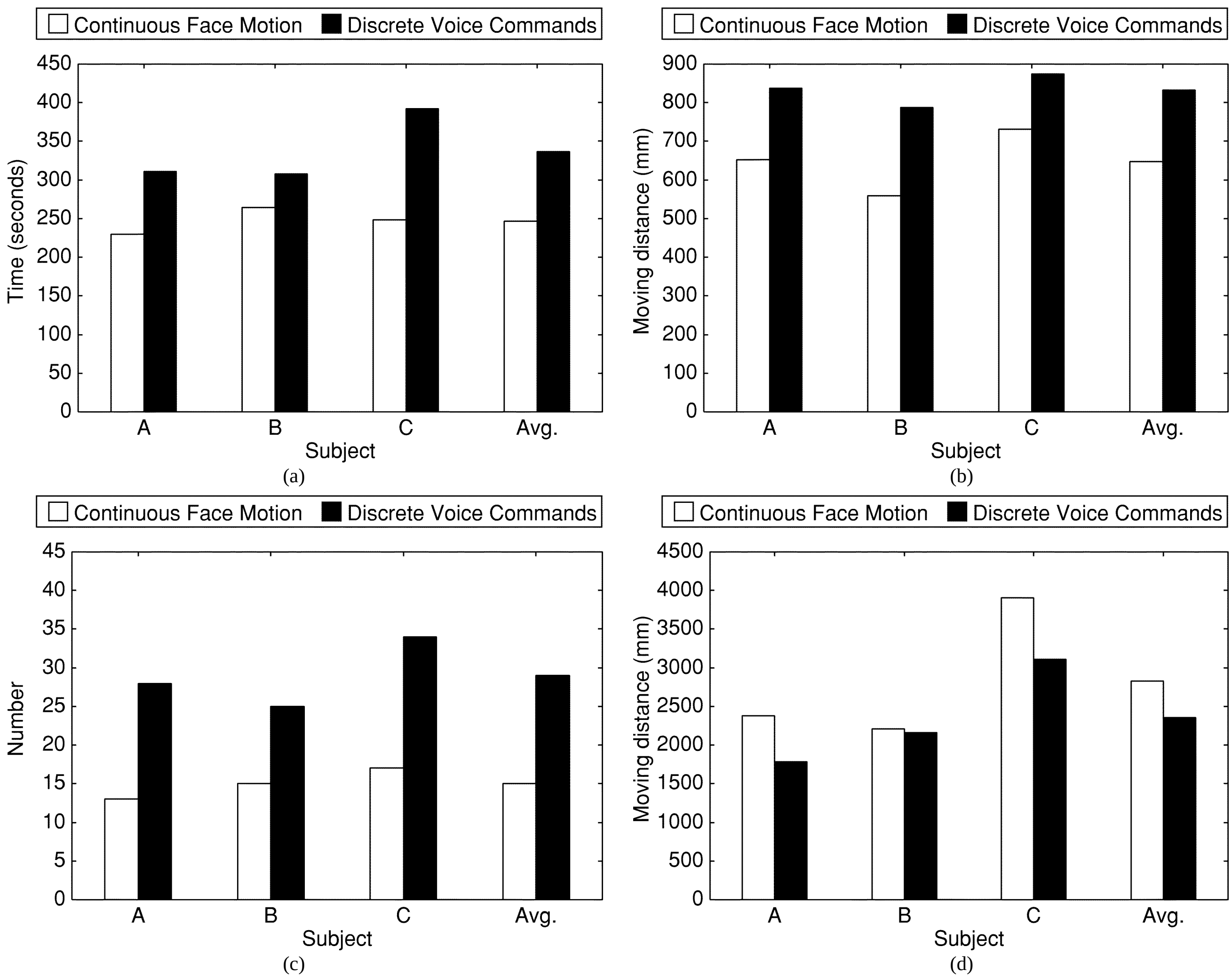


Fig. 19. Results of comparison between face and voice control interfaces. (a) Task completion time. (b) Camera moving distance. (c) The number of commands. (d) Instrument moving distance.

d_{scope} and θ_{scope} is satisfied by confirming the color of the target (e.g., light blue indicates the target has been appropriately captured).

D. Experimental Setup

To validate the advantage of using a continuous face position to guide the scope, the FAce MOUSe interface was compared to the voice control interface using the solo surgery simulator. For this experiment, the environment of Fig. 13 was conscientiously simulated on the computer and the assessment task described above was applied. This experiment involved three subjects. During the task execution, the coordinates of the tip point of the laparoscope, the coordinates of the tip point of the surgical tool, and the type of camera control commands were sequentially logged in a text file. After completion, the task beginning time and termination time were also logged in the file. Then, the number of camera control commands, camera moving distance, instrument moving distance, and task completion time were analyzed from the log file. In the experiment, the radius of the gallbladder sphere was 15 mm and that of the dissection sphere was 7.5 mm. The parameters for task management were as follows: $T_d = 125$ mm and $T_\theta = 5^\circ$.

E. Results and Discussion

Experimental results are illustrated in Fig. 19(a)–(d). In this figure, the white bar indicates data for the face tracking-based system while the black bar indicates the result by voice control.

- For all subjects, both the task completion time and the laparoscope moving distance were shorter for the face motion-based interface compared to the voice control interface [Fig. 19(a) and (b)]. This is probably because the FAce MOUSe system allows the operator, while guiding the scope, to make continuous fine adjustments of not only the direction of scope motion but also its motion velocity; using the continuous face motions, the operator can guide the laparoscope more efficiently and quickly, and thus the scope can take the shorter route to the target. Consequently, the time to complete the task can also be reduced.
- For all subjects, the number of camera guidance commands using the face control was smaller than the number using the voice control interface [Fig. 19(c)]. In the dissection task, the operator needs to position the laparoscope with high accuracy to remove the small target spheres. In this experiment, all the subjects had to give the same voice command many times over for fine adjustment of

the scope position. To the contrary, they all could keep the command losses to almost a minimum when using FAce MOUSe. Please note that there is little difference among the individuals in the face command number.

- For all subjects, the instrument moving distance was longer for the face interface compared with the voice interface [Fig. 19(d)]. The differences between them seems mainly due to the result of the accumulation of slight instrument movements accompanying quick face motions, such as rolling gestures.

The results of Fig. 19(a)–(c) demonstrate the effectiveness of our novel approach, as well as the main advantage of continuous, nonverbal human–machine interactions. On the other hand, another result [Fig. 19(d)] suggests that the face motion may have had a negative influence on precise surgical actions. Nevertheless, the task completion time with the face interface was shorter, and the *in vivo* laparoscopic experiment was successfully and safely performed on a pig with our system, as described above. Thus, the influence of face motion on instrument operation may not be a serious problem in the more simple manipulations performed during a laparoscopic cholecystectomy, but it may lead to major restrictions for the surgeon in operations where more complex surgical techniques are required. A more detailed analysis and investigation of the influence of facial motions will be one of our future studies.

VI. Conclusion

We have developed a new robotic laparoscope positioning system for solo surgery based on a real-time, face-tracking technique. In both *ex vivo* and *in vivo* experiments, 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. In a comparative experiment using a solo surgery simulator, the face control interface was more efficient in laparoscopic camera functions than voice control, even under the stress of controlling the instruments. The major advantages of our approach are summarized as follows: 1) *easy to use*: the system allows complete *hands-off* and *feet-off* laparoscope operations, which makes the surgeon more comfortable; 2) *rapid reaction*: nonverbal instructions by face gesture commands are more intuitive and faster than verbal instructions (typically voice commands). In practice, our system can make a frame-rate (30 Hz) response to the surgeon's face commands by realizing real-time image processing; 3) *natural communication*: the proposed system does not require the surgeon to put on a microphone or use head-mounted (body-contact) sensing devices. It is natural and generally nonstressful for the surgeon; and 4) *precise operation*: the face motion can represent not only the direction of scope motion but also the degree of motion, such as velocity, so that the laparoscope positioning accuracy can be improved.

Now we are studying an improved method for guiding the laparoscope based on the visual tracking of both the surgeon's face and the surgical instruments [19], with the goal of reducing not only mental stress but also physical stress, such as neck fatigue. A more exhaustive comparative study between our system and other human–machine interfaces is also important and on-going [20], [21].

References

[1] J. M. Sackier and Y. Wang, "Robotically assisted laparoscopic surgery from concept to development," *Surgical Endoscopy*, vol. 8, no. 1, pp. 63–66, Jan. 1994.

[2] R. Hurteau, S. DeSantis, E. Begin, and M. Gagner, "Laparoscopic surgery assisted by a robotic cameraman: Concept and experimental results," in *Proc. IEEE Int. Conf. Robotics and Automation*, San Diego, CA, May 1994, pp. 2286–2289.

[3] R. H. Taylor, J. Funda, B. Eldridge, S. Gomory, K. Gruben, D. LaRose, M. Talamini, L. Kavoussi, and J. Anderson, "A telerobotic assistant for laparoscopic surgery," *IEEE Eng. Med. Biol. Mag.*, vol. 14, pp. 279–288, May-June 1995.

[4] M. E. Allaf, S. V. Jackman, P. G. Schulam, J. A. Cadeddu, B. R. Lee, R. G. Moore, and L. R. Kavoussi, "Laparoscopic visual field. voice vs. foot pedal interfaces for control of the AESOP robot," *Surgical Endoscopy*, vol. 12, no. 12, pp. 1415–1418, Dec. 1998.

[5] G. F. Buess, A. Arezzo, M. O. Schurr, F. Ulmer, H. Fisher, L. Gumb, T. Testa, and C. Nobman, "A new remote-controlled endoscope positioning system for endoscopic solo surgery the FIPS endoarm," *Surgical Endoscopy*, vol. 14, no. 4, pp. 395–399, Apr. 2000.

[6] T. Kimura, Y. Umehara, and S. Matsumoto, "Laparoscopic cholecystectomy performed by a single surgeon using a visual field tracking camera," *Surgical Endoscopy*, vol. 14, pp. 825–829, Sept. 2000.

[7] V. F. Murioz, C. Vara-Thorbeck, J. G. DeGabriel, J. F. Lozano, E. Sanchez-Badajoz, A. Garcia-Cerezo, R. Toscano, and A. Jimenez-Garrido, "A medical robotic assistant for minimally invasive surgery," in *Proc. IEEE Int. Conf. Robotics and Automation*, San Francisco, CA, Apr. 2000, pp. 2901–2906.

[8] C. Vara-Thorbeck, V. F. Murioz, R. Toscano, J. Gomez, J. Fernandez, M. Felices, and A. Garcia-Cerezo, "A new robotic endoscope manipulator a preliminary trial to evaluate the performance of a voice-operated industrial robot and a human assistant in several simulated and real endoscopic operations," *Surgical Endoscopy*, vol. 15, pp. 924–927, Sept. 2001.

[9] E. Kobayashi, K. Masamune, I. Sakuma, T. Dohi, and D. Hashimoto, "A new safe laparoscopic manipulator system with a five-bar linkage mechanism and an optical zoom," *Comput.-Aided Surgery*, vol. 4, no. 4, pp. 182–192, 1999.

[10] E. Kobayashi, K. Masamune, I. Sakuma, T. Dohi, K. Shinohara, and D. Hashimoto, "Safe and simple man-machine interface for a laparoscope manipulator system comparison of command input methods," *J. Jpn. Soc. Comput.-Aided Surgery*, vol. 3, no. 2, pp. 71–78, 2001.

[11] P. A. Finlay and M. H. Ornstein, "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.*, vol. 14, no. 3, pp. 289–291, May-June 1995.

[12] N. J. Dowler and S. R. J. Holland, "The evolutionary design of an endoscopic telemanipulator," *IEEE Robot. Automat. Mag.*, vol. 3, pp. 38–45, Dec. 1996.

[13] P. A. Finlay, "A robotic camera holder for laparoscopy," in *Proc. 10th Int. Conf. Advanced Robotics, Workshop 2 on Medical Robotics*, Budapest, Hungary, Aug. 2001, pp. 129–132.

[14] A. Casals, J. Amat, and E. Laporte, "Automatic guidance of an assistant robot in laparoscopic surgery," in *Proc. IEEE Int. Conf. Robotics and Automation*, Minneapolis, MN, Apr. 1996, pp. 895–900.

[15] M. O. Schurr, A. Arezzo, B. Neisisu, H. Riminsland, H. U. Hilzinger, J. Dorn, K. Roth, and G. F. Buess, "Trocar and instrument positioning system TISKA an assist device for endoscopic solo surgery," *Surgical Endoscopy*, vol. 13, pp. 528–531, May 1999.

[16] M. S. Wilson, A. Middlebrook, C. Sutton, R. Stone, and R. F. McCloy, "MIST VR: A virtual reality trainer for laparoscopic surgery assesses performance," *Ann. Roy. Coll. Surgeons of England*, vol. 79, pp. 403–404, 1997.

[17] A. Nishikawa, T. Hosoi, K. Koara, A. Hikita, D. Negoro, S. Asano, F. Miyazaki, M. Sekimoto, Y. Miyake, M. Yasui, and M. Monden, "A laparoscope positioning system with the surgeon's face image-based human-machine interface," in *Proc. CARS2001 –Computer Assisted Radiology and Surgery, 15th Int. Congr. Exhibit.*, Berlin, Germany, June 27–30, 2001, pp. 165–170.

[18] A. Nishikawa, T. Hosoi, K. Koara, D. Negoro, A. Hikita, S. Asano, F. Miyazaki, M. Sekimoto, Y. Miyake, M. Yasui, and M. Monden, "Real-time visual tracking of the surgeon's face for laparoscopic surgery ," in *Proc. Medical Image Computing and Computer-Assisted Intervention, MICCAI 2001*, vol. 2208, Lecture Notes in Computer Science , Utrecht, The Netherlands, Oct. 2001, pp. 9–16.

[19] A. Nishikawa, S. Asano, F. Miyazaki, M. Sekimoto, M. Yasui, and M. Monden, “A laparoscope positioning system based on the real-time visual tracking of surgeon’s face and surgical instruments,” in *Proc. CARS2002 –Computer Assisted Radiology and Surgery, 16th Int. Congr. Exhibit.*, Paris, France, June 26–29, 2002, pp. 1090–1090.

[20] A. Nishikawa, D. Negoro, H. Kakutani, F. Miyazaki, M. Sekimoto, M. Yasui, S. Takiguchi, and M. Monden, “Using an endoscopic solo surgery simulator for quantitative evaluation of human-machine interface in robotic camera positioning systems,” in *Proc. Medical Image Computing and Computer-Assisted Intervention, MICCAI 2002*, vol. 2488, Lecture Notes in Computer Science , Tokyo, Japan, Sept. 2002, pp. 1–8.

[21] ——, “Design of an endoscopic solo surgery simulator for quantitative evaluation of human-machine interface in robotic camera positioning systems,” *Adv. Robot.*, vol. 17, pp. 503–521, Sept. 2003.

Atsushi Nishikawa received the B.S., M.S. and Ph.D. degrees in mechanical engineering from Osaka University, Osaka, Japan, in 1990, 1992, and 1995, respectively.

In 1995 and 1996, he was a Visiting Researcher at the Institute for Robotics and Intelligent Systems, University of Southern California, Los Angeles. From 1994 to 1996, he was a Research Fellow of the Japan Society for the Promotion of Science. He is currently a Research Associate of the Department of Mechanical Science and Bioengineering, Graduate School of Engineering Science, Osaka University. His research interests are robot vision, human interfaces and computer-assisted surgery.

Dr. Nishikawa is a member of the Robotics Society of Japan, the Institute of Electronics, Information and Communication Engineers, the Information Processing Society of Japan, and the Japan Society of Computer Aided Surgery.

Toshinori Hosoi received the B.S. and M.S. degrees in mechanical engineering from Osaka University, Osaka, Japan, in 1999 and 2001, respectively.

He is currently with Multimedia Research Laboratories, NEC Corporation, Kanagawa, Japan.

Kengo Koara received the B.S., M.S. and Ph.D. degrees in mechanical engineering from Osaka University, Osaka, Japan, in 1997, 1998, and 2001, respectively.

He is now with Mitsubishi Electric Corporation, Aichi, Japan. His research interests include computer vision, human interfaces and robotics.

Dr. Koara is a member of the Robotics Society of Japan and the Institute of Electronics, Information and Communication Engineers.

Daiji Negoro received the B.S. and M.S. degrees in mechanical engineering from Osaka University, Osaka, Japan, in 2000 and 2002, respectively.

He has joined Toshiba Corporation, Kanagawa, Japan.

Ayae Hikita received the B.S. and M.S. degrees in mechanical engineering from Osaka University, Osaka, Japan, in 2000 and 2002, respectively.

She has joined Furukawa Electric Corporation, Kanagawa, Japan.

Shuichi Asano received the B.S. and M.S. degrees in mechanical engineering from Osaka University, Osaka, Japan, in 2001 and 2003, respectively.

He is now with Denso Corporation, Aichi, Japan.

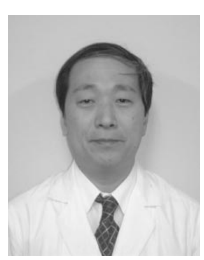
Haruhiko Kakutani received the B.S. degree in mechanical engineering from Osaka University, Osaka, Japan, in 2002.

He is now with Mitsubishi Electric Corporation, Hyogo, Japan.

Fumio Miyazaki (M’88) received the B.S., M.S., and Ph.D. degrees in mechanical engineering from Osaka University, Osaka, Japan, in 1975, 1977, and 1982, respectively.

In 1979, he served as a Research Associate, an Associate Professor in 1986, and a Professor in 1991, all with Osaka University. Currently, he is a Professor of the Department of Mechanical Science and Bioengineering, Graduate School of Engineering Science, Osaka University. In 1987 and 1988, he was a Guest Professor at the University of California, Santa Barbara. His research interests are intelligent robots and human skill.

Dr. Miyazaki is a member of the Robotics Society of Japan, the Institute of Systems, Control and Information Engineers, and the Society of Instrument and Control Engineers.

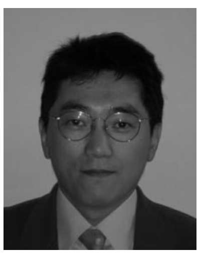


Mitsugu Sekimoto received the M.D. and Ph.D. degrees from Osaka University Medical School, Osaka, Japan, in 1981 and 1988, respectively.

He then joined Yao City Hospital, Osaka, Japan, and was a Director of Surgery at the hospital in 1991. From 1991 to 1993, he was a Director of Surgery at Nagayoshi General Hospital, Osaka, Japan. From 1993 to 1995, he was a Director of Surgery at Osaka Teishin Hospital, Osaka, Japan. In 1995, he served as an Assistant Professor and a Director of the Division of Endoscopic Surgery, Department of Surgery II, Osaka University Medical School. Since 1999, he has been an Associate Professor, a Director of the Division of Endoscopic Surgery, and a Director of the Division of Colorectal Surgery, Department of Surgery and Clinical Oncology, Osaka University Graduate School of Medicine.

Masayoshi Yasui received the M.D. degree from Wakayama Medical College, Wakayama, Japan, in 1995.

From 1995 to 1998, he joined the Department of Surgery, Osaka National Hospital, Osaka, Japan. From 1998 to 2000, he was with Kaiduka City Hospital, Osaka, Japan. He is currently a Postgraduate Student at the Division of Coloproctology, Department of Surgery and Clinical Oncology, Osaka University Graduate School of Medicine, Osaka, Japan, and engaged in basic research and laparoscopic surgery.



Yasuhiro Miyake received the M.D. degree from Osaka University Medical School, Osaka, Japan, in 1993.

From 1994 to 1997, he was a Senior Resident of Surgery at Kawanishi City Hospital, Hyogo, Japan. From 1997 to 2001, he was a Senior Resident of Surgery at the Department of Surgery and Clinical Oncology, Osaka University Medical School, Osaka, Japan. Currently, he is a Research Fellow at the Division of Hematology and Oncology, Department of Internal Medicine, University of Michigan, Ann Arbor.

Shuji Takiguchi received the M.D. degree from Osaka University Medical School, Osaka, Japan, in 1990.

From 1991 to 1997, he was a Senior Resident of Surgery at Osaka Teishin Hospital, Osaka, Japan. In 1997, he served as an Assistant Professor of the Department of Endoscopic Surgery, Osaka University Medical School, and a Director of Surgery at Osaka Teishin Hospital. Currently, he is an Assistant Professor of the Department of Surgery and Clinical Oncology, Osaka University Graduate School of Medicine.

Morito Monden received the M.D. and Ph.D. degrees from Osaka University Medical School, Osaka, Japan, in 1970 and 1979, respectively.

In 1979, he served as a Lecturer at the Department of Surgery II, Osaka University Medical School; an Assistant Professor in 1987, an Associate Professor and Vice Chairman of the Department in 1990, a Professor and the Chairman in 1994. Currently, he is a Professor and Chairman of the Department of Surgery and Clinical Oncology, Osaka University Graduate School of Medicine. From 1979 to 1981, he was a Visiting Fellow at the Department of Surgery, Memorial Sloan-Kettering Cancer Center, New York, NY. In 1984, he was a Visiting Fellow at the Department of Surgery, University of Pittsburgh Medical Center, Pittsburgh, PA.

Dr. Monden is a member of the International College of Surgeons, the International Surgical Society, the Transplantation Society, the International Gastro-Surgical Club, American Association for the Study of Liver Disease, American Association of Cancer Research, American Gastroenterological Association, American Society of Clinical Oncology, New York Academy of Sciences, the Japan Surgical Society, the Japanese Society of Gastroenterological Surgery, the Japanese Society for Clinical Surgery, the Japan Society for Transplantation, the Japanese Society for Vascular Surgery, the Japanese Cancer Association, the Japan Society of Clinical Oncology, the Japanese Society for Hepatology, and the Japanese Society of Gastroenterology.