Surgical Skill Evaluation by Force Data for Endoscopic Sinus Surgery Training System

Yasushi Yamauchi[1], Juli Yamashita[1], Osamu Morikawa[1], Ryoichi Hashimoto[1], Masaaki Mochimaru[1], Yukio Fukui[1,2], Hiroshi Uno[3], and Kazunori Yokoyama[4]

[1] National Institute of Advanced Industrial Science and Technology (AIST),
Tsukuba Central 6, 1-1-1, Higashi, 305-8566 Tsukuba, Japan
{y.yamauchi,yamashita-juli,morikawa.osamu,hashimoto-r,
m-mochimaru,y.fukui}@aist.go.jp
http://unit.aist.go.jp/humanbiomed/humed_en/indexE.htm
[2] Institute of Information Sciences and Electronics, University of Tsukuba,
1-1-1, Ten-noudai, 305-8573 Tsukuba, Japan
[3] KOKEN CO., LTD., 3-14-3, Mejiro, Toshima-ku, 171-0031 Tokyo, Japan
[4] ENT clinic@Tsukuba South Avenue, 2-18-2, Takezono, 305-0032 Tsukuba, Japan
ra4k-ykym@asahi-net.or.jp

Abstract. In most surgical training systems, task completion time and error ratio are common metrics of surgical skill. To avoid applying unnecessary and injurious force to the tissue, surgeons must know for themselves how much force they are exerting as they handle surgical tools. Our goal is to develop an endoscopic sinus surgery (ESS) training system that quantitatively evaluates the trainee's surgical skills. In this paper, we present an ESS training system with force sensors for surgical skill evaluation. Our experiment revealed that the integral of the force data can also be one of the useful metrics of surgical skill.

1   Introduction

Recently, endoscopic surgery has become widely used as a less invasive method than conventional means. In endoscopic surgery, a small opening must be made for the insertion of the endoscope and forceps, but deep lesions can be treated directly, i.e., without celiotomy. This provides marked advantages, such as less invasiveness and shorter hospital stays. On the other hand, it is difficult to interpret the positions of lesions on endoscopic images, due to the limited endoscopic visual field and the narrowness of the spaces around the legions; in addition the long forceps used in endoscopic surgery tend to restrict movement. It has been reported that these limitations can cause complications such as injury to tissues around the lesions. To solve these problems, new surgical devices and navigation systems supporting endoscopic surgery have been developed. However, the development of training devices to improve surgical skills has only just started, and new methods must be developed to quantitatively and objectively evaluate improvement in surgical skills.

A variety of systems have been developed for training in various types of surgical procedures. Most of the studies in surgical training based on virtual reality techniques [1-3] focus on helping the trainee learn a surgical procedure by simulation and imitation but do not mention surgical skill evaluation. These surgical 'simulation'

Fig. 1. Endoscopic sinus surgery training system. The endoscopic view *(lower right)* is also shown for the subject. The force output is shown on the computer display *(upper right)*.
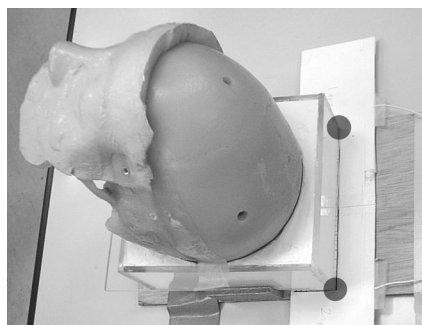
Fig. 2. Force sensors are fixed beneath the platform of the dummy (left and right circles). These sensors detect the forces perpendicular to the surface of the desk.

systems do not provide quantitative data about the trainee's technical progress, and should not be confused with surgical 'training' systems. The most popular metrics of surgical skill are task completion time and error ratio [4,5], which are also commonly used in surgical training systems [6-11]. However, these metrics are not adequate for evaluating the quality of surgical tasks, because a fast but rough technique can damage tissues. From the qualitative point of view, a few studies on surgical skills have added force analysis as a metric. O'Toole et al. [12] applied force analysis to a surgical training system for a suturing task.

We are developing a training system that employs force analysis [13] for endoscopic sinus surgery (ESS). As the surgeon's working space in the nasal cavities is far narrower than it is in laparoscopic surgery, avoiding excessive collision between surgical tools and tissue is an essential point in ESS training. A few ESS training systems have been proposed [3,9,10], but none have succeeded in quantifying collision force as a metric of surgical skill. In this study, we examine surgical force as well as task completion time as metrics of surgical skill.

2 ESS Training System

The ESS training system that we have developed (Fig. 1) consists of a head dummy, force sensors, endoscope, position sensor, video monitoring system, and computer systems.

The dummy is an accurate reproduction made of urethane resin and based on a mold of a cranial bone specimen. Since the dummy has a modular structure consisting of a skull module and left and right nasal cavity modules, the nasal cavity modules are interchangeable for future use. To make the nasal cavity module similar to an actual nasal cavity, the urethane bone was given soft tissues made of silicone and colored to look like nasal mucosa. The face of the dummy is made of a porous urethane epidermis. CT images of the dummy showed no excessive artifact, and the average CT value of the dummy was around 250HU. This indicates that our ESS training system is also applicable in the training of image-guided sinus surgery.

Two force sensors (PS-5KA, Kyowa Dengyo Inc.) are fixed beneath the platform of the dummy (Fig. 2), one on the left and the other on the right. The output of the force sensors is digitized (5 Hz), recorded by a computer (PC-9801RX, NEC Inc.) and shown in real time on the computer display. The output of the sensors is adjusted to 0 when no external force exists. If the output of both sensors is positive, it indicates that a force perpendicular to the surface of the desk is working. If one output is positive and the other is negative, a moment to rotate rightward or leftward is working.

The endoscopic system consists of a rigid forward-oblique endoscope at 30 degrees (Shinko Optics, 4mm in diameter), a video camera, and an illuminator. The endoscopic image is displayed on a 14-inch color CRT monitor (Sony Inc.). To determine the task completion time from the movement of the endoscope tip, a probe of an optical position sensor (Hybrid Polaris, Northern Digital Inc.) is fixed to the CCD unit of the endoscope.

The dummy, endoscope, and monitors are arranged so that they can be observed simultaneously on a single video frame. The conditions of the experiments are recorded with a camcorder (Sony Inc.).

3   Experiments

In the experiments, the subjects conducted two types of surgical tasks: 'evaluation tasks' and 'training tasks'. The evaluation tasks were performed in the left nasal cavity of the dummy, and the training tasks, which were similar to the evaluation tasks, were performed in the right nasal cavity. Each trainee's performance was evaluated in each evaluation task. Kinematical learning specific to the right nasal structure occurred by repeated training, but it was useless for the left nasal cavity, in which the evaluation tasks were performed, because the left and right cavities have different structures. Therefore, the evaluation of skills in the evaluation tasks was not affected by the kinematical learning that occurred in the training tasks.

3.1   Tasks

Figure 3 shows the two task types. The evaluation tasks were to confirm the ostium of the left frontal sinus by the forward-oblique endoscope at 30 degrees, and to take the tip of the forceps to the ostium. When the endoscope is introduced into the nasal cavity, the ostium can be spontaneously seen by direct insertion because the endoscope is usually inserted looking obliquely upward. Then the endoscope was removed, the forceps was inserted into the site slightly forward from the endoscope, and the tip of the forceps was taken to the ostium while the endoscope confirmed its position. This procedure was done once in each evaluation task period.

For the training tasks, the target was the ostium of the right maxillary sinus. Although the endoscope was inserted looking obliquely upward, it had to be rotated 90 degrees anticlockwise on the way to the target, because the ostium is situated on the side. The rotated endoscope was removed from the nasal cavity, the forceps was inserted into the site slightly forward from the endoscope, and the tip of the forceps was taken to the target while the endoscope confirmed its position. In each training task period, this procedure was repeated for five minutes.
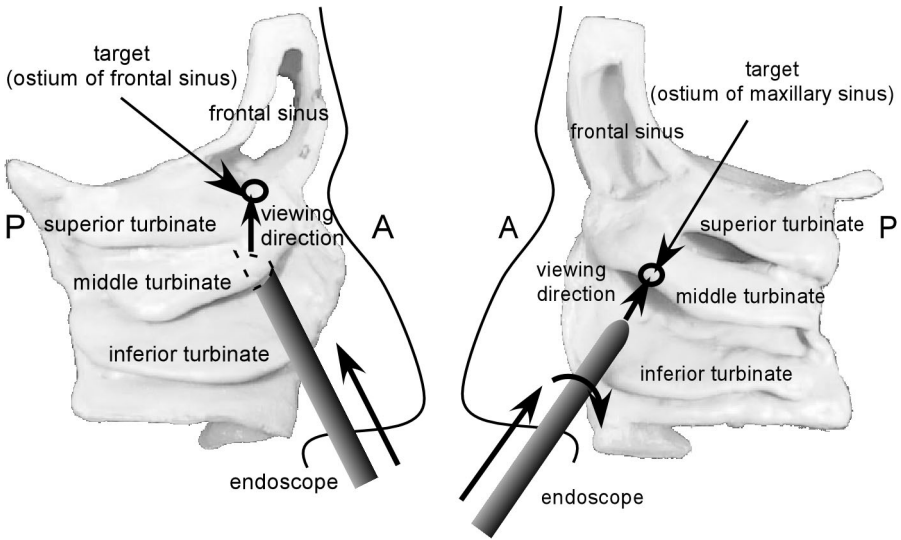
Fig. 3. The evaluation task and the training task. The target of the evaluation task is the ostium of the left frontal sinus *(left),* whereas the target of the training task is the ostium of the right maxillary sinus *(right).* In the training task, the endoscope should be rotated 90 degrees on the way.

3.2   Subjects

The subjects were three non-medical volunteers who had not used an endoscope and two resident physicians who had limited experience in endoscopic sinus surgery (fewer than 20 cases each). None of the five subjects had previously used the training system. Data from the evaluation tasks performed by an experienced otorhinologist who had performed endoscopic surgery in more than 200 cases were used as the expert reference data.

The output of the force sensors (5 Hz) and the position of the endoscope tip (10 Hz) were measured during the surgical and training tasks, and the camcorder synchronously recorded the conditions of the operation and the endoscopic images. The task completion time was defined as the period when the tip of the endoscope was in the nasal cavity, as measured by the position sensor.

3.3   Experiment Procedures

The experiment procedures for all subjects consisted of a pre-experiment learning session and five tasks as following steps. The total time of the experiment was around 1 hour per subject.

Pre-experiment learning: After outlines of the experiments were presented and the anatomical structures of the left and right nasal cavities were explained, the subjects studied the basic operation of the endoscope and forceps twice by watching afrom the videotape, and then actually performed the operation.
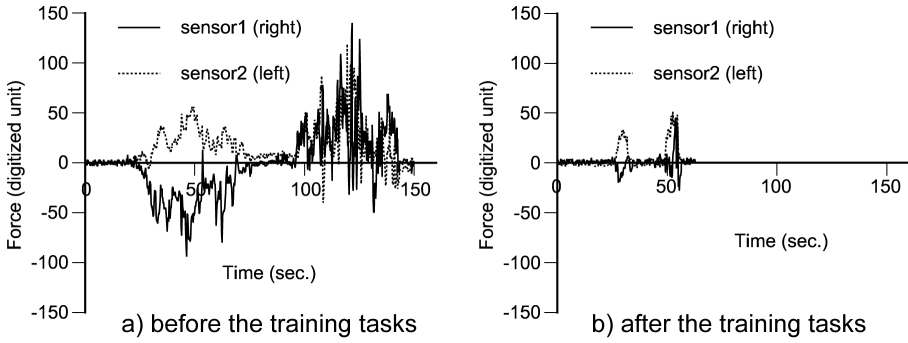
Fig. 4. Time evolutions of the output of force sensors of a resident. Before the training tasks (a), an evaluation task took over two minutes and excessive force occurred. After the training tasks (b), the task completion time was shorter and the force was lower than before.

First evaluation task: The subjects learned the evaluation tasks once by watching a videotape of the tasks being performed by an experienced ESS surgeon, then immediately performed the same tasks. Force and position data were recorded during the tasks.

First training task: Subsequently, the subjects learned the training tasks once by watching a videotape of this task as performed by the expert, then immediately performed the training tasks for five minutes. After the completion of the training tasks, his/her performance, as recorded by the video camera, was played back on the VCR and displayed on the image monitor. This was followed by a replay of the learning videotape about the training tasks. This was the feedback for the subjects.

Second evaluation task: This was the same as the first evaluation task.

Second training task: This was the same as the first training task.

**Third evaluation task:** This was the same as the first evaluation task.

4   Results

Fig. 4 shows the time evolutions of the output of the force sensors during a) the first and b) the third evaluation tasks performed by one of the residents. The unit of force is digitized voltage from an A/D converter.

As shown in Fig. 4a, strong and irregular forces and moments were detected. Pushing the dummy towards the platform with the forceps or endoscope caused strong positive forces, and twisting the dummy on the platform with the forceps or endoscope inserted into the nasal cavity caused strong negative forces. These forces were judged to be unnecessary, because the expert showed no such forces. In the third evaluation tasks (Fig. 4b), the forces were clearly lower than those in the first performance, indicating the training's effects on the force data.

To evaluate surgical skill, we need a kind of index for force exerted. We propose three indices: a) maximum, b) average and c) integral, based on the absolute values from the two sensors (Fig. 5). The expert's data normalize all the data in these figures. Clinically, the maximum shows the instantaneous strong force, the average indicates the average force on the tissue, and the integral shows the total force on the tissue

during the tasks. For all parameters, the data obtained from the expert were the lowest (1.0), indicating that the expert can totally reduce the unnecessary load on the tissue. Fig. 5d shows the subjects' task completion times. The trend is almost the same as that of the force data.
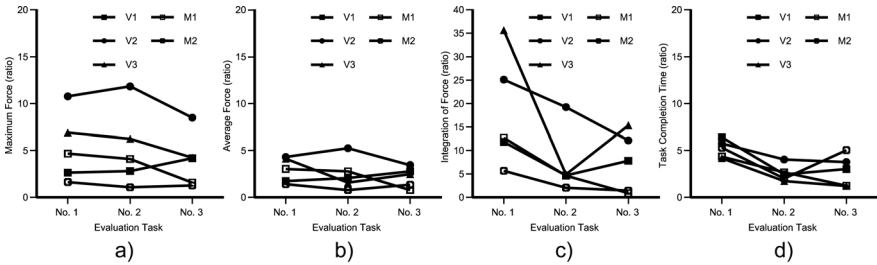
Fig. 5. Indices for the evaluation tasks: a) maximum force, b) average force, c) integration of the force, and d) completion time. V1, V2, and V3 are the non-medical volunteers, and M1 and M2 are the residents. All data are normalized by those of an expert.

5   Discussion

Of the three force indices, the integral was decreased significantly by the repeated performance of the surgery (non-parametric: Friedman's test ($p < 0.05$), Wilcoxson's coding rank test ($p < 0.05$)). The mean and maximum were reduced slightly, but the differences were not statistically significant. As the dimension of the integral is the multiplication of time and force, the characteristics of the integral can inherit those of both the task completion time and average force. Clinically, the integral reflects both the magnitude and duration of friction caused by the endoscope or forceps contacting tissues on the way to the lesions. Therefore, the integral is considered the most adequate force index of surgical skill in ESS.

The differences between the expert, the residents, and the non-medical subjects are clearly shown in Fig. 4c. The reference data obtained from the expert showed an excellent performance compared with any of the subjects. In the two residents, the initial performance was lower than that by the expert, but the integral steadily decreased with the completion of two training tasks, reaching a level close to that of the expert. On the other hand, in the three non-medical subjects, the integral decreased but did not finally reach the level of the expert. Within the limitations of the small number of subjects and the statistical analysis with the presumption of normal distribution, there was a significant difference in the integral of force between the residents and non-medical subjects (analysis of variance, $p < 0.05$).

Force indices are adequate to evaluate the interference between surgical tools and organs, and are applicable to the training of nasal, bronchoscopic, angioscopic, and gastrointestinal surgeries. Force evaluation is also valuable for tasks in microsurgery and robotic surgery, in which the surgeon does not obtain force-feedback directly. In the latter cases, the difference to an adequate force is the most meaningful metric [19]. The analysis of the surgical tool's position will also serve as a quantitative index

of surgical skill [15]. The position sensors in our system will also enable such analysis.

The tasks in this study are for short-term learning. But in real applications, training occurs in periods of several days to several months. It is necessary to evaluate the usefulness of the proposed training system in such long-term training, and also to study the skill learning system over a long period.

In most of the existing systems with surgical force measurement [12,14-18], force/torque sensors are attached to the surgical tools. When the objective of training is to avoid unnecessary tool-to-dummy collision, sensors can be attached to the dummy. Doing so, the systems accept any kind of existing surgical tools. It should also be added that the sensors attached to the dummy only detect tool-to-dummy collisions.

6   Conclusions

To aid endoscopic surgical training, we have developed a training system based on a dummy of the nasal cavities that is equipped with force sensors, conducted experiments using a training protocol for surgical operations, and evaluated the metrics for the improvement of surgical skills by repeated training. It was indicated that the integral of the absolute values of the force sensor output was an adequate physical index of surgical force. Additional studies are currently under way to precisely evaluate   surgical skill via a 6-DOF force sensor, and by a position sensor to analyze the motion of a surgical tool. Based on these studies, we hope to propose a new surgical training system and training protocols that would contribute to the training of endoscopic sinus surgery.

References

1. Kaufmann, C., Zakaluzny, S., Liu, A.: First Steps in Eliminating the Need for Animals and Cadavers in Advanced Trauma Life Support. In: Proc. of MICCAI 2000, Lecture Notes in Computer Science, Vol. 1935. Springer-Verlag (2000) 618-623
2. Gomes, M.P.S.F., Davies, B.L.: Computer-Assisted TURP Training and Monitoring. In: Proc. of MICCAI 2000, Lecture Notes in Computer Science, Vol. 1935. Springer-Verlag (2000) 669-677
3. Edmond, J.C.V., Heskamp, D., et al.: ENT Endoscopic Surgical Training Simulator. In: Proc. of MMVR, Studies in Health Technology and Informatics, Vol. 39. IOS Press (1997) 518-528
4. Rosser Jr, J.C., Rosser, L.E., Savalgi, R.S.: Objective evaluation of a laparoscopic surgical skill program for residents and senior surgeons. Arch. Surg. 133 (1998) 657-661
5. Hanna, G.B., Shimi, S.M., Cuschieri, A.: Task performance in endoscopic surgery is influenced by location of the image display. Ann. Surg. 227 (1998) 481-484
6. Jordan, J.A., Gallagher, A.G., McGuigan, J., McGlade. K., McClure, N.: A comparison between randomly alternating imaging, normal laparoscopic imaging, and virtual reality training in laparoscopic psychomotor skill acquisition. Am. J. Surg. 180 (2000) 208-211
7. Gallagher, A.G., McClure, N., McGuigan, J., Crothers, I., Browning, J.: Virtual reality training in laparoscopic surgery: a preliminary assessment of minimally invasive surgical trainer virtual reality (MIST VR). Endoscopy 31 (1999) 310-313

8. Derossis, A.M., Fried, G.M., Abrahamowicz, M., Sigman, H.H., Barkun, J.S., Meakins, J.L.: Development of a model for training and evaluation of laparoscopic skills. Am. J. Surg. 175 (1998) 482-487
9. Rudman, D.T.: Functional endoscopic sinus surgery training simulator. Laryngoscope 108 (1998) 1643-1647
10. Weghorst, S.: Validation of the Madigan ESS simulator. In: Proc. of MMVR, Studies in Health Technology and Informatics, Vol. 50. IOS Press (1998) 399-405
11. Yamashita, J., Yamauchi, Y., et al.: Real-Time 3D Model-Based Navigation System for Endoscopic Paranasal Sinus Surgery. IEEE Biomed. Eng. 46 (1999) 107-116
12. O'Toole, R.V., Playter, R.R., Krummel, T.M., Blank, W.C., Cornelius, N.H., Roberts, W.R., et al.: Measuring and developing suturing technique with a virtual reality surgical simulator. J. Am. Coll. Surg. 189 (1999) 114-127
13. Yamauchi, Y., Suzuki, M., et al.: A Training System for Endoscopic Sinus Surgery with Skill Evaluation. In: Proc. of CARS 2001, International Congress Series, Vol. 1230. Elsevier Science (2001) 1150
14. Rosen, J., Hannaford, B., MacFarlane, M.P., Sinanan, M.N.: Force controlled and teleoperated endoscopic grasper for minimally invasive surgery--experimental performance evaluation. IEEE Trans. Biomed. Eng. 46 (1999) 1212-1221
15. Emam, T.A., Hanna, G.B., Kimber, C., Dunkley, P., Cuschieri, A.: Effect of intracorporeal-extracorporeal instrument length ratio on endoscopic task performance and surgeon movements. Arch. Surg. 135 (2000) 62-65
16. Hanna, G.B., Shimi, S., Cuschieri, A.: Influence of direction of view, target-to-endoscope distance and manipulation angle on endoscopic knot tying. Br. J. Surg. 84 (1997) 1460-1464
17. Hanna, G.B., Drew, T., Clinch, P., Shimi, S., Dunkley, P., Hau, C., et al. Psychomotor skills for endoscopic manipulations: differing abilities between right and left-handed individuals. Ann. Surg. 225 (1997) 333-338
18. Rosen, J., Hannaford, B., Richards, C.G., Sinanan, M.N.: Markov modeling of minimally invasive surgery based on tool/tissue interaction and force/torque signatures for evaluating surgical skills. IEEE Trans. Biomed. Eng. 48 (2001) 579-591
19. Salcudean, S.E., Ku, S., Bell, G.: Performance measurement in scaled teleoperation for microsurgery. In: Proc. of CVRMed-MRCAS'97, Lecture Notes in Computer Science, Vol. 1205. Springer-Verlag (1997) 789-798