The Role of Tactile Feedback in Laparoscopic Surgery

MV Ottermo, MS [1], M Øvstedal, BS [1], T Langø, PhD [2], Ø Stavdahl, PhD [1], Y Yavuz, MD [3], TA Johansen, PhD [1], R Mårvik, MD, PhD [3]

[1] Norwegian University of Science and Technology, Department of Engineering Cybernetics, Trondheim, Norway

[2] SINTEF Health Research, Department of Medical Technology, Trondheim, Norway

[3] National Center for Advanced Laparoscopic Surgery, St. Olav's Hospital, Trondheim, Norway

Corresponding author:
Maria V. Ottermo, NTNU, Department of Engineering Cybernetics
Odd Bragstads Plass 2D , 7491 Trondheim, Norway
Phone: +47 73 59 02 43, Fax: +47 73 59 43 99, Email: ottermo@itk.ntnu.no

Contributions:

MV Ottermo designed equipment, designed and conducted experiments, performed analysis of experimental data and wrote manuscript. M Øvstedal designed and conducted experiments and performed analysis of experimental data. T Langø helped with the planning and design of experiments. Ø Stavdahl helped with planning and design of equipment and experiment.
Y Yavuz contributed to the design of the experiments. TA Johansen has worked on the design of equipment and experiment. R Mårvik worked on design and planning of experiments.

Acknowledgements:

The research is funded by the Research Council of Norway, grant 147830/320. The authors would like to acknowledge all the participants who made this study possible. We would also like to thank the National Center for Advanced Laparoscopic Surgery at St. Olav's hospital for providing the necessary equipment to conduct this study.

Abstract

Two experiments aiming at comparing palpation with gloved fingers, conventional laparoscopic instruments and a laparoscopic instrument with a sensor array attached to its end effector are described. The sensor array provides the surgeon with visually presented tactile information. 15 subjects were asked to discriminate hardness and size of objects (rubber balls hidden in pig's intestine) with the 3 palpation methods. The experiments showed that gloved fingers are better at differentiating hardness and size compared to conventional laparoscopic instruments and the instrument with sensor. There was no significant difference between conventional instruments and the instrument with sensor. This indicates that visual presentation may not be an ideal way of presenting tactile information. It also indicates that the presence of the array does not make the task more difficult.

Keywords: Tactile feedback, palpation, laparoscopic surgery, tactile sensor array.

1 Introduction

In laparoscopic surgery the operation is performed with instruments and viewing equipment inserted into the body through small incisions created by the surgeon. This method has many advantages, including minimization of surgical trauma and damage to healthy tissue. However, laparoscopic surgery requires specialized dexterity even beyond that needed for open surgery. Decreased tactile feedback, different eye-hand coordination and translation of a two-dimensional video image into a three-dimensional working area are just some of the obstacles when performing laparoscopic surgery. The lack of the tactile feedback limits the surgeon's abilities to palpate internal organs, a technique actively used for locating tumors, gallstones and abnormalities in the tissue during open surgery. Based on this problem it is important to identify which information is lost in laparoscopic surgery compared to open surgery. The future goal is to develop a surgical instrument which serves as an extension of the surgeon's fingers (Ottermo, Stavdahl and Johansen, 2004), where a sensor array attached to the instrument's end effector measures contact forces between the array and the tissue. This tactile information is then sent to the surgeon's fingers to provide him with a feeling of the shape or hardness of the tissue under investigation.

In this study the tactile information is presented visually on a monitor. Sensory substitution is a mature research area, among others in the world of prosthetics, and in this field it was long assumed that the human mind has great powers of adaption that enables it to replace afferent signals by user's vision or other senses. For amputees, however, this assumption will often influence the ability to focus on other tasks, for instance indulging in intellectual activity at any but the trivial level (Simpson, 1973). At present, the concept of extended physiological proprioception (E.P.P.) introduced by David Simpson is considered to be more effective. In E.P.P. a control system is constructed such that the movement at a joint in for instance an artificial arm is made to correspond to the movement of, for example, one of the joints in the shoulder girdle. The information from the natural joint that corresponds to the angle of the appropriate joint in an artificial arm will then be sent to the central nervous system (Simpson, 1974). It is therefore unlikely that visual presentation of the tactile information is the optimal solution to the reduced tactile feedback in laparoscopic surgery, although it may contribute to the ability of making correct diagnoses.

Some previous studies describing haptic feedback in laparoscopic surgery have been conducted. The word haptic refers to something of or relating to the sense of touch. This term is further divided into two categories: the kinesthetic sense and the tactile sense. Kinesthetic information is often called force feedback and is related to the movement or force in muscles and joints, while tactile sense information includes the sensation of shapes and textures (McLaughlin, Hespanha and Sukhatme, 2002). The latter requires a sensor array to measure several contact points at the same time. Bholat, Haluck, Murray et al. (1999) conducted a single-blinded study where participants were asked to identify objects by direct palpation, conventional (surgical) instruments and laparoscopic instruments. From this study they stated that haptic feedback is altered, but not eliminated when using laparoscopic instruments. Bicchi, Canepa, Rossi et al. (1996) used a position sensor and a force sensor attached to a conventional instrument to measure the force acting on the tissue from the instrument and the angle deflection of the end effector. These measurements were used to find the viscous elasticity of objects, and the information was presented graphically (position versus force) to the user. Experiments showed that subjects could discriminate between objects of different materials using this instrument. den Boer, Herder, Sjoerdsma et al. (1999) examined the sensitivity of laparoscopic dissectors when touching a simulated arterial pulse. The results showed that feedback quality was significantly better with reusable dissectors than for disposable dissectors, but that the overall sensitivity of instruments was low compared to bare fingers. Tholey, Desai and Castellanos (2005) developed a prototype of a laparoscopic grasper with force feedback and demonstrated that an operator could easily differentiate between tissue samples of varying stiffness. The same results were found in MacFarlane, Rosen, Hannaford et al. (1999) and indicated that a force feedback device is significantly better than a standard Babcock grasper at rating tissue compliance, but not as successful as a gloved hand. Wagner, Stylopoulos and Howe (2002) showed that the absence of force feedback increased the average force magnitude applied to the tissue by at least 50%, and that the number of errors that damage tissue increased by more than a factor of 3. Other attempts of providing tactile feedback have also been done. Jackman, Jarzemski, Listopadzki et al. (1999) compared the EndoHand, a laparoscopic three-fingered hand, with standard laparoscopic instrumentation. They found a significant promise in the ability to perform sophisticated manipulation of objects, although it fell short in both dexterity and tactile feedback.

The experiments described in this paper aim at comparing the sensitivity of a laparoscopic grasper and gloved fingers, and investigating whether additional information about the contact forces between the instrument and the internal tissue contributes to a better understanding of the properties of the tissue. The main difference from earlier experiments is that the tests are done in a more realistic environment known to the laparoscopic surgeon. Instead of manipulating the objects directly, they are hidden in pig's intestine. In this way we simulate a situation with resemblance to real surgery. In addition, the sensor array used differs from the sensors used for force feedback in (Tholey et al., 2005, MacFarlane et al., 1999) in that our system measures and displays spatially distributed parameters of contact (not single point, as in the mentioned force feedback devices).

2 Materials and Method

2.1 Tactile sensor array

The tactile sensor array we use is a PPS TactArray (Pressure Profile Systems Inc.) developed for measuring the tactile pressure distribution between objects in direct physical contact (see Figures 1 and 2). It consists of a two-dimensional array (15 x 4) of pressure sensing capacitive elements in a thin, continuous sheet, and the total size of the array is 3.5 cm x 1 cm, which is exactly the size of a plate fixed to the grasper jaw. Each element of the sensor array is 2 mm x 2 mm, and since it is built using strips of capacitive material there is no spacing between the elements, resulting in a 2 mm resolution. Included in the system is software for acquiring, visualizing and storing data. Our system has been custom made to fit on to a reusable laparoscopic grasper (see Figures 1 and 2), and the active part of the sensor array (a total of 60 sensing elements) covers most of the grasper jaw (3.0 cm x 0.8 cm). The maximum pressure range of the sensor array is 0-7 N/cm^2. When sensory information is available, it is presented visually to the surgeon on a screen (see Figure 3). The surface of the sensor array is smooth, while conventional graspers usually have a toothed surface. Sedaghati, Dargahi and Singh (2004) have attempted to make a toothed sensor array surface, and depending on the application it is an important parameter to consider for future instruments.

2.2 Silicone rubber balls

Palpation is used actively for many discrimination tasks, but we chose to focus on identification of tumors. We developed several artificial tumor samples using a silicone rubber. This is a two-component silicone rubber, where the amount of component B determines how soft the rubber becomes. The recommended compositions are 9:1, 9:2, 9:3 and 9:4, where the hardest (9:1) is 25 shore A. For the hardness tests we used the 9:2, 9:3 and 9:4 compositions. Additionally, we used a glass ball (marble) and a foam rubber ball to provide extreme points (see Figure 6). The choice of tumor samples was based on a pilot where experienced surgeons were asked to distinguish between the samples with their fingers. The soft silicone (9:2) has a hardness that resembles a lymph node or the intestine wall, i.e. it is not as hard as a tumor. The medium silicone (9:2) has a hardness close to that of a benign tumor, for instance in prostate or lymph node and lipoma. The hard silicone (9:4) resembles a malign tumor, while the glass ball is as hard as bone, gallstones etc. The balls used in the hardness experiments all have a diameter of 1.5 cm. For the size experiments we used the following 5 diameters: 0.5 cm, 0.9 cm, 1.3 cm, 1.8 cm and 2.4 cm (see Figure 6), and all these balls had the same silicone rubber composition (9:2).

2.3 Experimental design

We were interested in evaluating the information needed, the information present and the information lost during laparoscopic surgery. In addition, we wanted to evaluate the new instrument with sensor array attached to see if the technology complicates the tasks. To do this, the following points of interest were listed:

- How well can the surgeon discriminate size and hardness using his fingers (F), state of the art laparoscopic instruments (LI) and a laparoscopic instrument with tactile sensor (LIS), respectively?
- Is the attention drawn from the original task when LIS is used, and does the visually presented tactile information introduce any additional benefits or problems to the surgeon?

As our target goal was to evaluate the loss of information in laparoscopic surgery, we chose to reproduce this environment as closely as possible. Therefore, all objects to be identified were hidden in pig's intestine and placed in a simulator for laparoscopic training (Figure 4).

To avoid the problem of the surgeon not being able to locate the objects at all, we separated the intestine into pockets, where each pocket contained one silicone ball (see Figures 5 and 6). The subjects were given visual feedback on a video screen by use of standard endoscopic viewing equipment (Olympus A5294A) inside the simulator (see Figures 7 and 8).

For the first experiment the subject was presented with 5 silicone balls, all with varying hardness. The subject was asked to rank the 5 balls from softest to hardest, using either his fingers, conventional laparoscopic instruments or the laparoscopic instrument with sensor (the order of which instrument to use first, second and last was randomized between subjects). The subjects were allowed to feel the different samples as many times as necessary. After ranking the 5 samples, the process was repeated (with the same instrument) for a total of three trials. In order to prevent memorization, each trial was repeated with random ordering of the samples. After completion of the three trials with the given instrument, the subject moved on to the next instrument (placed on a second simulator) and repeated the ranking procedure described above. Finally, the same ranking was done using the third instrument.

The second experiment used the same method as the first, but here the task was to rank 5 silicone balls from smallest to largest. To restrict the subjects from discriminating by using only their eyes, we filled the intestines with water so that the balls were completely hidden (see Figure 5).

Although we conducted two separate experiments, one with focus on hardness and one with focus on size, the subjects performed both at the same time, meaning that they performed all trials for either hardness or size first and without a longer break proceeded directly to the next (either hardness or size depending on which experiment they started with). The order of the two experiments was randomized between subjects.

The subjects were encouraged to talk during all trials and tell what they felt and saw, and they were video recorded. This was done to ensure that no information was lost and to double check the results. All subjects were given a short presentation of all instruments prior to the experiments. They were also informed about the scope of the experiments, how the tests would be conducted and asked to complete two questionnaires. The subjects were not allowed to see the silicone balls before or during the experiments.

2.4 Participants

Nine surgeons with varying surgical experience and 6 medical students/medical doctors participated in the experiments. The test persons were from 23 to 53 years old, with an average of 35.2 years. Of the 15 subjects, 9 were male and 6 female. Five of the participants had performed more than 200 surgical interventions and were classified as experienced surgeons, while 4 of the participants (with 15-200 interventions) were classified as surgeons with some experience. The remaining 6 participants were described as unexperienced. Surgeons familiar with laparoscopic surgery were preferred since using laparoscopic equipment is rather complicated, and special training is needed to feel comfortable using it. Surgeons are also the target group for the device, and feedback from them is of utmost importance. We included the 6 completely unexperienced participants too see if the differences were noticeable.

2.5 Data collection

Altogether each subject had 18 different tasks to complete, 3 trials for F, LI and LIS with respect to both hardness and size. In each trial the subject ranked the silicone rubber balls from 1 to 5, with 1 being the softest or smallest and 5 being the hardest or biggest. The data was not characterized as "true" or "false". Instead, we recorded the ranking the subject gave and compared this value with the right value. The reason we did this was that we wanted to characterize it as a bigger mistake if the subject exchanged a 1 and a 5 than if he exchanged a 3 and a 4. If a subject could not discriminate between two balls, for instance with hardness 3 and 4, he was told to rate them with the same number (the choice of which rating to give (3 or 4) was done by comparing with a softer and a harder ball and determining which one they were closer to, e.g. a 2 or a 5).

We used Friedman's test, which is a two-way analysis of variance by ranks and a nonparametric analog to one-way repeated ANOVA (Analysis of Variance). The test is an extension of the binomial sign test for two dependent samples to a design involving more than two dependent samples (k>2). The reason why we use a nonparametric test is that we cannot assume that such a small data set has a normal (Gaussian) distribution. We also have repeated measures since we test the same subject several times. Additionally, we have more than two data sets (instruments), and hence Friedman's test is suitable (Sheskin 2000). If the result of the Friedman two-way analysis by ranks is significant, it indicates that there is a significant difference between at least two of the sample medians in the set of k medians. Although nonparametric tests are not based on assumptions of normality, randomization is still required, and the tests are less powerful than parametric tests.

With Friedman's test we compare all 3 instruments simultaneously, but when the value of χ_{α}^2 is significant it does not indicate whether just two or, in fact, more than two conditions differ significantly from each other. In this case, tests designed to compare only two samples provide a more effective alternative (Lehman, 1975). In our experiments we did not use paired observations but divided our subjects into blocks. A block is a portion (e.g. two seeds in the same pot) of the experimental material that is expected to be more homogeneous than the aggregate (e.g. all seeds not in the same pot) (Box, Hunter and Hunter 2005). By confining comparisons to those within blocks, greater precision is usually obtained because the differences between associated blocks are eliminated. We consider the results for each test subject (for hardness or size) as one block. Since we compare two instruments we get 6 observations in each block. Based on this we can perform a Wilcoxon two-sample test and compute the rank sum for each of the instruments. The Wilcoxon two-sample test is an appropriate nonparametric alternative to the parametric two-sample t-test, which is extensively used in problems that deal with inference about the population mean or in problems that involve comparative samples (Walpole, Myers and Myers, 1998). Although we do not assume normality, randomization is still required in the same way as for Friedman's test. The calculations were done using MATLAB'S Statistics Toolbox (The MathWorks Inc.).

3 Results

Figure 9 shows the average error and the standard deviation of the error for each instrument for both hardness and size. Here the magnitude of the error in each case is taken into account, meaning that if a subject ranked a value 4 silicone rubber ball as a number 2, the error has the absolute value 2. As the figure shows, the fingers are superior to laparoscopic instruments for discrimination both in the hardness and the size case. For hardness, the laparoscopic instrument with sensor resulted in average error per trial of 1.38, while the conventional laparoscopic instrument led to an average error of 1.98. The difference is not as pronounced in the size case, where LIS has an average error of 1.69 versus 1.96 for LI.

The subjects were given a post test questionnaire, where they were asked to rate their own performance and to answer some questions about the test and the equipment. We also wanted their subjective opinion about the necessity of including tactile feedback on laparoscopic instruments. Therefore, we asked the surgeons whether the missed tactile feedback when performing laparoscopic interventions. They had spilt opinions on this question, but 53 % meant that it tactile feedback could be somewhat useful, while 33.3 % did not know. The remaining felt that it would not be helpful.

We also asked them what they thought of the tasks they had to perform, since it is always important to know if the difficulty of the task has the right level. None of the subjects felt it was too easy. In fact, 11 of the subjects thought the tasks had the right level of difficulty, while the remaining 4 meant it could have been easier.

Most of the subjects felt that the experiments were somewhat comparable with real surgery and only 1 of the 15 subjects felt that the extra information shown on the laptop (as in Figure 3) was confusing.

3.1 Experiment 1 - Hardness

Hypothesis H1 : When discriminating between balls with varying hardness, a conventional laparoscopic instrument and a laparoscopic instrument with visually presented tactile information perform equally well as gloved fingers, i.e. F=LI=LIS. The alternate hypothesis is H1: not H0, i.e. $F \neq LI \neq LIS$.

A Friedman test, with a critical p-value=0.05, was conducted to test the above hypothesis against the alternate hypothesis H1. Since we compare k=3 different instruments we have 2 degrees of freedom, and this results in $\chi_\alpha^2 = 5.99$. The Friedman test gave $\chi_r^2 = 17.43 > 5.99$ ($p=7.42 \times 10^{-5}$, meaning that H0 is rejected.

As mentioned before a significant value of χ_r^2 does not say anything about the pair wise comparison of the instruments. Therefore we performed an analysis based on the Wilcoxon two-sample test. We have 3 paired comparisons (F versus LI), (F versus LIS) and (LI versus LIS). After employment of Wilcoxon two-sample test the total rank sums, WHF, WHLI and WHLIS, are assumed to be approximately normal (here W indicates the Wilcoxon two-sample test, H stands for Hardness, and F/LI/LIS indicates which instrument is considered). Hence, the following cases were tested:

H0: WHF = WHLI versus H1: WHF < WHLI

H0: WHF = WHLIS versus H1: WHF < WHLIS

H0: WHLI = WHLIS versus H1: WHLIS < WHLI

When comparing F and LI, we found that $z_{WHLI}=1.86>1.65=z_{0.05}$ (with a 0.05 level of significance), meaning that we can reject H0 and conclude that F is significantly better than LI for hardness. From the table of the area under the normal curve we find $P_{WHLI}=P(z>1.86) = 0.0307$. In other words, we can reject H0 at a level of significance of approximately 0.03.

Performing the same analysis for F versus LIS resulted in $z_{WHLIS}=1.15<1.65=z_{0.05}$. Hence we cannot reject H0, and we conclude that F is not significantly better than LIS for hardness.

When comparing LI and LIS, we found $z_{WHLI}=0.46<1.65=z_{0.05}$ and we cannot reject H0. In other words, we can conclude that LIS is not significantly better than LI for hardness.

Figure 10 shows the errors done by each individual subject in the hardness case. As can be seen from the plot, subject 2 did a lot of errors with all instruments. Except for subject 1, 4, 12, 13 and 15, all performed better with LIS than LI. Gloved fingers were better or as good as both LI and LIS for all subjects.

Figure 11 shows how the error percentage varies with the amount of experience with laparoscopic surgery. This plot does not take the magnitude of the error into consideration. As can be seen from the figure, the overall error percentage was 5.8% for F, 32% for LI and 21.3% for LIS. The error rate when using gloved fingers is low for all groups. The experienced surgeons have better performance using conventional laparoscopic instruments than the other two groups. For the laparoscopic instrument with sensor it is interesting to notice that the unexperienced subjects have better performance than both the other groups.

The balls used in the experiment ranged in hardness from glass to foam rubber. Between these extreme points, three silicone rubber balls with different hardness were used. As figure 12 shows, the softest ball (foam rubber) stands out as the easiest to distinguish for all instruments. The glass ball was easy to identify with the fingers, but harder with the other two instruments (LI and LIS). Compared to laparoscopic instruments, the instrument with sensor seemed to improve the subjects' performance for identification of hardness 2, 3 and 5 (where 5 is the glass ball).

As previously mentioned, we did not characterize errors as "false" or "true", but by the magnitude. When counting the number of times the subjects gave a wrong ranking (could be several errors per trial), we found that of the total 133 errors (out of 675 possible) made for hardness, 108 of them had magnitude 1, meaning that they had characterized for instance a hardness 1 as a hardness 2. 18 errors had a magnitude of 2 and the remaining 7 were of magnitude 3.

The time elapsed for each trial was recorded (the time consumed to rank 5 objects). Table 1 shows the average time for ranking with each instrument in the hardness case. The table also shows how the time varies with experience. Notice that the experienced subjects are somewhat faster with LIS than LI, as opposed to the other two groups which are faster with LI than LIS.

When rating their own performance, only 3 of the subjects felt they performed better with LI than LIS. Four of the subjects thought they performed equally well with LI and LIS and the remaining 8 felt they performed better with LIS than LI. All meant that they had performed best with the fingers.

3.2 Experiment 2 - Size

Hypothesis H0: When discriminating between balls with different sizes, a conventional laparoscopic instrument and a laparoscopic instrument with visually presented tactile information perform equally well as gloved fingers, i.e. F=LI=LIS. The alternate hypothesis is H1: not H0, i.e. $F \neq LI \neq LIS$.

The Friedman test for size gives $\chi_r^2 = 19.23 > 5.99$ ($p=4.78 \times 10^{-5}$), and H0 is rejected. We performed the same paired comparisons for size as for hardness, and the following cases were tested:

H0: WSF = WSLI versus H1: WSF < WSLI

H0: WSF = WSLIS versus H1: WSF < WSLIS

H0: WSLI = WHLIS versus H1: WSLIS < WSLI

When comparing F and LI, we found that $z_{WSLI}=2.01>1.65=z_{0.05}$. We can reject H0 and conclude that F is significantly better than LI for size. From the table showing the area under the normal curve we find P_{WSLI}=P(z>2.01) = 0.0222. In other words, we can reject H0 at a level of significance of approximately 0.02.

For the comparison of F and LIS, the calculations resulted in z_{WSLI}=1.86>1.65= z0.05, hence we can reject H0 and conclude that F is significantly better than LIS for size.

When comparing LI and LIS we found z_{WSLI}=0.39<1.65= z0.05, meaning that we cannot reject H0. Therefore we conclude that LIS is not significantly better than LI for size.

The errors done by the individual subjects in the size case are shown in Figure 13. As we can see, the fingers are superior for all subjects, except for subject 11 and 15, where LIS is slightly better or equally good. The difference between LI and LIS is not as noticeable for the ranking size.

Figure 14 shows how many times the subjects ranked an object incorrectly in the size case. For all subjects the error percentage was 5.3% for the fingers, 29.8% for conventional laparoscopic instruments and 28.9% for the laparoscopic instrument with sensor. As opposed to the hardness case, the experienced surgeons performed better than the other two groups with LIS.

In the same way as with ranking of hardness, the subjects made more errors with some of the objects (see Figure 15). The smallest ball (0.5 mm in diameter) was responsible for 45% of the errors. Note that out of the 65 times the subjects made an error with the smallest ball, 46 of them were due to the fact that he could not identify a ball at all (although there was always an object present). The subjects also had some trouble finding the next smallest ball using LI and LIS, and in some cases they exchanged sizes 3 and 4. The easiest ball to discriminate was the largest.

Of the 144 errors done in the size case, 116 of them had magnitude 1, 24 had magnitude 2 and 4 had magnitude 3. The average time consumed per trial in the size experiment is shown in table 2. As we can see, the experienced surgeons were slightly faster than the other groups for both LI and LIS.

In the rating of themselves, 6 of the 15 subject thought that they performed better with LI than LIS when ranking the sizes. Three subjects thought they performed equally well and the remaining 6 felt they performed better with LIS than LI. All rated the fingers higher than the other two instruments.

4 Discussion

Our experiments show that the fingers are superior for palpation, both compared to conventional laparoscopic instruments and our laparoscopic instrument with visual feedback of the tactile image. This is not surprising, and it indicates that there is a definite potential for improving today's laparoscopic instruments. Our laparoscopic instrument with sensor did not prove to be significantly better than conventional laparoscopic instruments (LI), neither for hardness nor size, but a positive trend could be noted in the results, especially for hardness, where F (fingers) did not prove to be significantly better (with a 95 % confidence interval) than LIS (laparoscopic instrument with sensor). It is quite surprising that the sensor worked well for hardness, since it is presented as a shape and not as a compliant object. This indicates that humans are able to translate shape information into information about hardness in a useful way. It should be kept in mind, though, that in all the discrimination tasks for hardness, the objects had the same size. In a real laparoscopic surgery, this will usually not be the case. It is not clear how crucial this information is, since the surgeon does receive some tactile and haptic clues related to the size of the object through the handle (given that the laparoscopic instrument does not have a relieve spring). Using active palpation (for instance by sliding back and forth over the object) will also provide the surgeon with information about the size. Frequently, the surgeons use the length of the end effector to measure the size of objects inside the body. This measure gives an estimate of the size, especially relative to other adjacent tissues or objects. If more accurate information about size is needed, measurements of the angle between the jaws of the end effector can be implemented by using a position sensor either on the handle or closer to the end effector (depending on the requirements for accuracy). Another possibility is to measure the total pressure applied to the sensor array and present this data to the surgeon, either visually or by coding of vibratory signals.

For hardness, the unexperienced subjects seemed to find the sensor most helpful. There can be many reasons for this, one of them being that they are not familiar with the conventional laparoscopic instruments and do not have the necessary techniques to utilize them as well as an experienced surgeon. Therefore, any extra information to help them discriminate between the objects could be useful. The experienced surgeons, on the other hand, have the advantage that they are used to both the instruments and the 2D view of the objects, and thus it is easier for them to analyze the compliance of the objects without help from the sensor. The experienced surgeon also has the advantage of knowing that intestines are quite strong and endurable, and therefore use more force when investigating the balls. In a couple of cases, the experienced surgeons used too much force, resulting in damage to the intestine.

In contrast to the hardness case, the experienced surgeons actually performed better with the sensor in the size experiment. It is likely that the reason for this is, again, technique. While the experienced surgeons slid the grasper along the intestine or used a poking technique, the less experienced and unexperienced subjects used grasping. In this way, the experienced surgeon was able to ensure that the object under investigation did not slip, and he could corner it and use visual feedback to tell the size (often by comparing with the length of the grasper). In some cases when the objects were too small and the subjects were not able to corner them, sliding the grasper with tactile feedback over the intestine helped them identify the object. When using LIS, the poking technique was not as useful as with LI, due to the bigger size of the grasper. Some subjects reported that the size of the grasper in LIS compared with LI made it harder to distinguish between the objects, while some reported that it was easier. In any case, the difference in size of the grasper jaws of LI and LIS or the improved quality of the reusable grasper used in LIS compared to LI may have affected the results. Therefore, it should be kept in mind that any improved results of LIS compared to LI could be due to both the visual presentation of the tactile image and the size or quality of the grasper. As opposed to conventional graspers, the grasping surface of the laparoscopic instrument with sensor was smooth. This seemed to be a bigger problem for the unexperienced subjects than the experienced ones.

In the experiments the subjects made more errors with some of the balls. In Figure 12 we see that the instrument with sensor seemed to improve the subjects' performance for identification of hardness 3 and 5 but not for hardness 4. This is a strange result since the improved performance when identifying hardness 3 and 5 should have affected hardness 4 as well. It is probable that this is a coincidence and a mere result of the underlying stochastic process, rather than an indication of that hardness 4 is more difficult to identify than hardness 3 and 5.

As mentioned before, we recorded the time spent on each trial. The experienced surgeons were, in general, faster than the other two groups, and this is not surprising owing to the experience with laparoscopic surgery. All groups performed faster with gloved fingers than both LI and LIS, and ranking with LIS was most time consuming in both experiments. The probable reason for this is that it takes time to get used to the extra source of information. The experienced surgeons have many years of training using the conventional instruments, and very little training with the new equipment. A follow-up study where some subjects go through a training program over several weeks will therefore be interesting. However, the study performed in this paper is still valuable since it is always a goal to make the equipment as intuitive and simple to use to reduce the required training.

Most of the subjects did not find the extra source of information confusing, although it was suggested that the information should be placed on the same screen as the endoscopic video and that information about the pressure (numbers) should be more visible. It should be noted that the visually presented tactile information competes with the endoscopic video for the surgeon's attention. In addition, it is also debatable if the brain is able to translate visual images into meaningful tactile impressions fast enough for it to contribute to the surgeon's tactile understanding of the object under investigation. In accordance with the findings of Simpson's E.P.P. concept we found that sensory substitution is difficult and expect that feeding back the tactile sense rather than providing sensory substitution will be more efficient.

Our future goal is to incorporate a tactile display on the same grasper as the tactile senor array (Ottermo et al. 2004), and in this case graphically presented data will be redundant. Conducting the same experiments with a tactile display could give completely different results. In conclusion, gloved fingers proved to be better than both LI and LIS for palpation. Despite this, a visual feedback of the tactile image seemed to be useful for the subjects who fully understood how to use it, especially for hardness discrimination. This however, is just a small step in the right direction, as the overall goal should be to make an instrument as good as the human hand.

References

1. Bholat OS, Haluck RS, Murray WB, Gorman PJ, Krummel TM. Tactile feedback is present during minimally invasive surgery. J Am Coll Surg 1999; 189(4): 349-55.
2. Bicchi A, Canepa G, Rossi DD, Iacconi P, Scilingo EP. A sensorized minimally invasive surgery tool for detecting tissutal elastic properties. Proc IEEE Int Conf on Robotics and Automation, Minneapolis, MN, 1996; 884–8.
3. den Boer KT, Herder JL, Sjoerdsma W, Meijer DW, Gouma DJ, Stassen HG. Sensitivity of laparoscopic dissectors. Surg Endosc 1999; 13: 869–73.
4. Box GEP, Hunter JS, Hunter WG. Statistics for experimenters. John Wiley & Sons, 2005.
5. Jackman SV, Jarzemski PA, Listopadzki SM, Lee BR, Stoianovici D, Demaree R, Jarret TW, Kavoussi LR. The EndoHand: comparison with standard laparoscopic instrumentation. J Laparoendosc Adv Surg Tech A 1999; 9(3), 253–8.
6. Lehman EL. Nonparametrics. Statistical methods based on ranks. McGraw-Hill, 1975.
7. MacFarlane M, Rosen J, Hannaford B, Pellegrini C, Sinanan M. Force-feedback grasper helps restore sense of touch in minimally invasive surgery. J Gastrointest Surg 1999; 3(3), 278–85.
8. McLaughlin, ML, Hespanha JP, Sukhatme GS. Touch in Virtual Environments. Prentice Hall, 2002.
9. Ottermo MV, Stavdahl Ø, Johansen TA. Palpation instrument for augmented minimally invasive surgery. Proc IEEE/RSJ Int Conf on Intelligent Robots and Systems, Sendai, Japan, 2004; 3960–4.
10. Sedaghati R, Dargahi J, Singh H. Design and modeling of an endoscopic piezoelectric tactile sensor. Int J Solids and Structures 2004; 42, 5872–5886.
11. Sheskin DJ. Handbook of Parametric and Nonparametric Statistical Procedures. Chapman & Hall/CRC, 2000.
12. Simpson DC. The control and supply of a multimovement externally powered upper limb prosthesis, Proceedings of the 4th International Symposium on External Control of Human Extremities, 1973, pp. 247–254.
13. Simpson DC. The choice of control system for the multimovement prosthesis: extended physiological proprioception (E.P.P.). The control of Upper-Extremity Prostheses and Orthoses, Thomas, 1974, chapter 15, pp. 146–150.

14. Tholey G, Desai JP, Castellanos AE. Force feedback plays a significant role in minimally invasive surgery: Results and analysis. Ann Surg 2005 241(1), 102–9.
15. Wagner CR, Stylopoulos N, Howe RD. The role of force feedback in surgery: Analysis of blunt dissection. Proc 10th Int Symp on Haptic Interfaces for Virtual Environment and Teleoperator Systems, Orlando, Fl, 2002.
16. Walpole RE, Myers RH, Myers SL. Probability and statistics for engineers and scientists. Prentice Hall, sixth edn, 1998.

Figure legends

Figure 1: The picture shows the full tactile sensor array system (TactArray), which measures the tactile pressure distribution between objects in direct physical contact.

Figure 2: A close-up view of the end effector of the laparoscopic grasper with the TactArray attached.

Figure 3: Visually presented tactile information (pressure distribution).

Figure 4: Simulator for laparoscopic training.

Figure 5: The picture shows the intestine divided into pockets, where the intestine in the back is filled with water to avoid visual exposure of the sizes.

Figure 6: The objects that were put in the intestines. Each pocket contained one of these balls.

Figure 7: Experiment setup showing all three laparoscopic simulators, one for each instrument, gloved fingers (F), laparoscopic instrument (LI) and laparoscopic instrument with sensor (LIS).

Figure 8: Experiment setup when the laparoscopic instrument with sensor (LIS) was used. The computer behind the surgical simulation trainer presents the pressure distribution visually.

Figure 9: Plot showing average magnitude of the error and standard deviation for each instrument (gloved fingers (F), laparoscopic instrument (LI) and laparoscopic instrument with sensor (LIS)) for both hardness and size, respectively.

Figure 10: Plot showing the total magnitude of error per instrument (gloved fingers (F), laparoscopic instrument (LI) and laparoscopic instrument with sensor (LIS)) for each subject in the hardness case.

Figure 11: Plot showing how often a wrong ranking was given for each instrument (gloved fingers (F), laparoscopic instrument (LI) and laparoscopic instrument with sensor (LIS)) in the hardness case. The plot also shows the difference between surgeons with experience, some experience and no experience.

Figure 12: Plot showing the percentage of total error caused by each hardness for gloved fingers (F), laparoscopic instrument (LI) and laparoscopic instrument with sensor (LIS).

Figure 13: Plot showing the total magnitude of error per instrument (gloved fingers (F), laparoscopic instrument (LI) and laparoscopic instrument with sensor (LIS)) for each subject in the size case.

Figure 14: Plot showing how often a wrong ranking was given for each instrument (gloved fingers (F), laparoscopic instrument (LI) and laparoscopic instrument with sensor (LIS)) in the size case. The plot also shows the difference between surgeons with experience, some experience and no experience.

Figure 15: Plot showing the percentage of total error caused by each size for gloved fingers (F), laparoscopic instrument (LI) and laparoscopic instrument with sensor (LIS).