Critical Paper Review
Computer Integrated Surgery II
Vishwa Parekh
April 19, 2012

Paper:
Title: Review of methods for objective surgical skill evaluation
Authors: Carol E. Reiley, Henry C. Lin, David D Yuh, Gregory D. Hager

Project Background:
Title: Surgical Skill Evaluation in Endoscopic Sinus Surgery.
Goal: Mathematically model various surgical movements in an endoscopic sinus surgery from the motion of the endoscopic tool tip through the surgical procedure.

Relevance of the paper to this project:
This paper provides a great background and a very good starting point for anyone interested in evaluating surgical skills. This paper shows how motion analysis is a useful measure of surgical skill evaluation, talks about different ways to capture motion in a surgical procedure and discusses a wide variety of methods that have been used to build an evaluation model and how they compare with each other.
My project is surgical skill evaluation for endoscopic sinus surgery, which is a minimally invasive kind of surgery. This paper gives an idea of what would be a good measure for surgical skill evaluation for this surgery and what kind of tracking method would allow me to capture useful information. Finally, it gives me an idea of what kind of data modeling method should be tried in my specific case and provides a brief introduction to the method and a reference to look for in order to implement it.

Summary of the paper:
In this paper, authors discuss the history of surgical skill evaluation, how it has evolved over the years, what is the current state of art and compare the current methods in surgical skill evaluation.
In figure 1, the performance of an expert surgeon in a 4 throw suturing task is compared to the performance of an intermediate surgeon. This figure clearly shows that the plot on the left has more clearly defined actions, an efficient use of workspace and throws are very distinct.

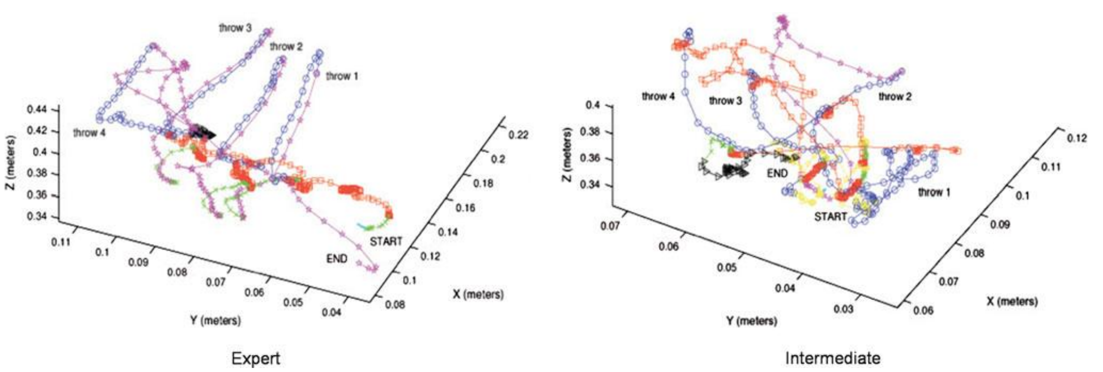


Fig. 1 Comparison - Expert Surgeon vs. Intermediate Surgeon

[C. Reiley, H. Lin, D. Yuh, and G. Hager, "Review of methods for objective surgical skill evaluation," Surgical Endoscopy, pp. 1-11, 2010.]

This also gives an idea that computationally modeling of motion analysis looks a very promising measure for surgical skill evaluation. The authors also discuss an intuition that motor skills are learned over time and improves with practice.

Data collection Methods

In order to evaluate motion, we need good data collection methods. The skills required for an open surgery are different that those required for minimally invasive surgery. Brief overview of the different kinds of surgical procedures and the corresponding data collection methods discussed in the paper are as follows:

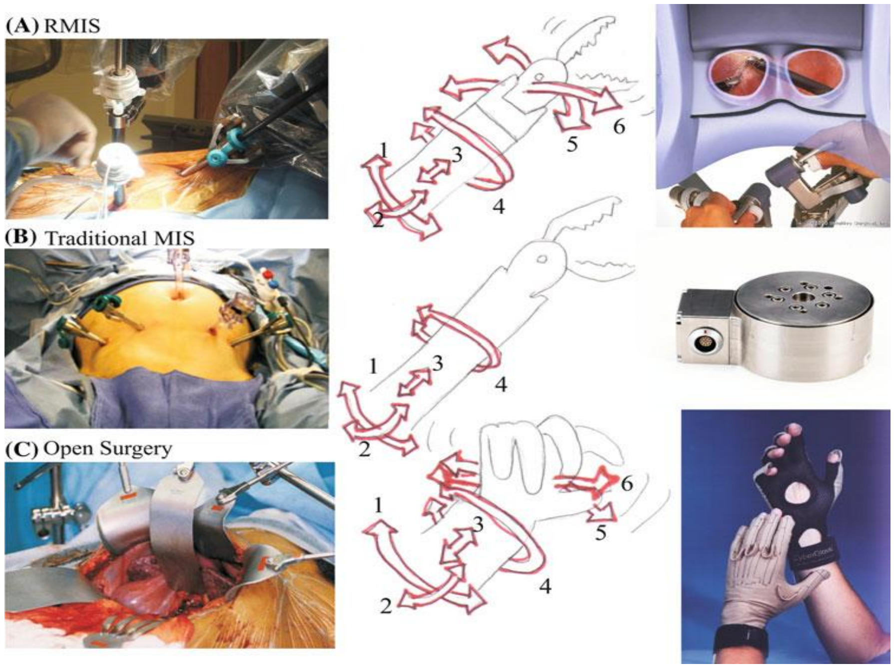


Figure 2. Different surgical procedures and the various motion information available in each of them

[C. Reiley, H. Lin, D. Yuh, and G. Hager, "Review of methods for objective surgical skill evaluation," Surgical Endoscopy, pp. 1-11, 2010.]

Open Surgery:

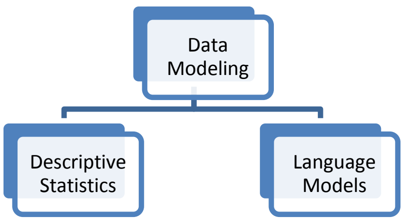
In this procedure, the surgeon can freely work with many different tools. Thus the hand motion of the surgeon is tracked in this procedure using a wireless data glove, to capture all the 23 degrees of freedom of motion.

Minimally Invasive Surgery:

In this procedure, the surgeon works with the target region using specialized laproscopic instruments. This limits the freedom of motion of the tool being operated by the surgeon and the view of surgeon is restricted to monocular view. Different kinds of Trackers are used to track the motion of the tool by attaching a rigid body to the endoscope or by using other methods.

Robotic Minimally Invasive Surgery:
In this procedure, surgical robots like Intuitive Surgical's da Vinci surgical system are used.
Using these systems allows the surgeon to regain more degrees of freedom than Minimally invasive surgery. These systems are capable of providing velocity and position of information for all the joints, stereoscopic video, etc. The motion is captured using the information from the system.

Data Modeling:



Descriptive Statistics: In this method of data modeling, descriptive statistics on velocity, position and in some cases, torque and force is provided in real time to the user. The authors discuss that studies have shown that this method is good for an individual to improve his surgical skills but not good enough to differentiate an expert from a novice.

Language Models: This method is analogous to forming a dictionary of different words in particular language and the efficiency of a speaker is analyzed in how well his pronunciation of these words are. In this case, the different words are the different surgical movements and the pronunciation corresponds to the skill or the efficiency of the surgeon is conducting these movements.
The Language models can be formed at three different levels:

1. Procedure Level: This includes the complete surgical intervention. This model would be able to provide effectiveness of the surgical procedure but will not be able to comment on the surgeon's skill level.
2. Task Level: This includes different tasks in a surgical procedure like Suturing or Dissection.
3. Surgeme Level: This includes subtasks like needle insertion, position needle, reach for the needle, etc.

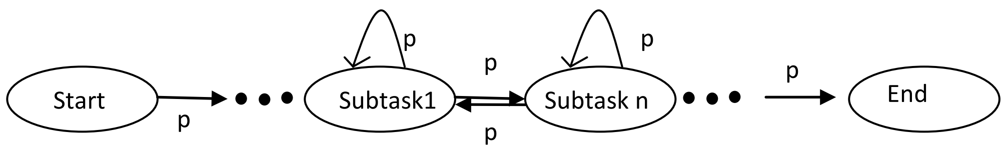
In the next section the authors discuss the different methods that have been implemented to model the surgical skills at both the task and subtask (surgeme) levels. Different segmentation techniques have been used by different groups.

Segmentation methods:
Manual Segmentation: In this type of segmentation, experts annotate the data from the surgical procedure manually into different tasks. Each of these annotations now form a feature vector of the complete position/velocity, etc. during the procedure. In order to reduce the dimension of this feature vector, most of the groups used vector quantization.

Automatic Segmentation: Here the tasks are automatically segmented based on supervised learning from available training data, and then using LDA to perform discriminative dimensional reduction and then Bayes classifier is used for classification.

Different type of modeling techniques used by various groups:

Hidden Markov models: In this technique, each of the surgeme/subtask is regarded as a hidden state in the model, the emission would be the feature vector emitted for that particular subtask and the hidden markov models are trained for each of the expertise in the surgeons.
The basic structure of the hidden markov model would look something like:



The p_s here represent the different probability values for the transitions.
The assumption here is that there are certain subtasks in the procedure that the expert surgeons use more often than the novice surgeons and in the same way there are certain subtasks that expert surgeon never use and are used only by the novice surgeons.
Thus the probability values for transitions and emission inside these states would be trained differently for surgeons with different expertise.

Now, when we have motion data for a surgeon with unknown skill level, we can run this data through all the 3 models and see which model gives us the best value for log likelihood.

Accuracy Results reported in [2]

- 84 % in a simulated needle task.
- 90% in a simulated retinal peeling task
- 97% for classification accuracy

Critique on this model:

1. This model once trained, would be able to predict the expertise level of a set of surgical motions performed by a surgeon in a very accurate manner as demonstrated by the results.
2. The model will have to be trained under the assumption of known skill level of the performing surgeon and this is a fuzzy value to consider.
3. The size of the training set that would be able to distinguish the skill levels would be a good number and it might be difficult to get these recordings for inexperienced surgeons.

Sequence matching: In this technique, the complete motion for a skilled surgeon is learnt and modeled as a sequence of characters, similar to a genetic sequence. After this, when a new series of sequence comes in, skill level of the surgeon is determined by scoring against the sequence for an expert surgeon.

No experimental results have been provided by the author.

Critique on this model:

1. This model would require training data from surgeries recorded only for an expert surgeon which is easy to get.
2. The model should be able to capture the transition between subtasks better than hidden markov models.
3. The scoring system would have to be very efficient to quantitatively determine the skill level of the surgeon and designing such a scoring system looks a bottleneck for this method.

Critique on the paper:

As a review, this paper does a very good job of summarizing the methods in surgical skill evaluation starting from the earliest in the history to the state of the art methods, analyzing each of these methods carefully. I will now go through the major topics and experiments discussed in the paper analyzing each of them.

- Experiment: Compare the performance of an expert surgeon to a novice in a 4 throw suturing task using Da Vinci by plotting their motion in the cartesian coordinate system. (Fig. 1)

  Pros:
  - The figure very clearly showed the difference between the two skill levels and showed that using tool motion as a feature vector to build the model can give great results.
  - The figure did a great job of making clear what the methods for surgical skill analysis are looking to model.

  Cons:
  - It would have been better to have listed precisely what was the criteria of surgeon being called experienced and the criteria for the surgeon being called an intermediate.

- Description of the different methods of surgical procedure (Fig. 2):

  Pros:
  - This section clearly describes why each of the surgical procedure would require a different kind of model to evaluate the skill level.
  - The clear detail on what should be tracked and modeled and how it should be done gives a very good idea on where to begin and how to go about when trying to build a model for a particular surgical procedure.

  Cons:
  - I found this section very interesting and the only part in the section which I felt could have better was the tracking methods described to capture the motion for minimally invasive surgeries.

- Modeling the surgical skill level

  Pros:
  - Figurative description of each level of surgical subtask at which motion can be modeled was helpful to understand the methods of modeling.

- The paper gives a brief overview on the segmentation methods to segment the overall motion into different tasks and subtasks. This part was important because otherwise it would have been very difficult to understand what each of the methods discussed in this section were trying to model and how did they get that data.
- The discussion of different techniques to model skill level and task and subtask level was highly detailed. This made it easy to follow the course of evolution of these techniques and how the efficiency of modeling has varied through the time. Most importantly, this section was able to provide a good basic intuition about each of the techniques, how the model is trained in each of them and how an unknown sample is classified.

Cons:

- The description of modeling surgical skill at procedure level was inadequate and was unable to provide a good idea of how it is being done.
- It was good that the paper touched the segmentation methods briefly while explaining the methods of modeling, but I feel there should have been a separate section before the description of modeling techniques to emphasize the importance of this step and discuss the state of the art segmentation methods that have been used by various research groups.

Relating back to my project:

The goal of my project is to model the surgical skill level of an endoscopic sinus surgery. This is a minimally invasive procedure and I am modeling the motion of the endoscopic camera during the procedure. This paper gave me an idea as to how should I segment my motion data as I now have an idea what my subtask motions would look like. Further, the paper introduced me to many different modeling methods which I went on to read in the respective papers and now have an idea of what different kinds of methods I might want to try to model the surgical data I would get.

References

1. *C. Reiley, H. Lin, D. Yuh, and G. Hager, "Review of methods for objective surgical skill evaluation," Surgical Endoscopy, pp. 1-11, 2010.*
2. *C. S. Hundtofte, G. D. Hager, and A. M. Okamura,"Building a Task Language for Segmentation and Recognition of User Input to Cooperative Manipulation Systems," 10th International Symposium on Haptic Interfaces for Virtual Environment and Teleoperator Systems, 2002, pp. 225-230.*
3. *Nagy I, Mayer H, Knoll A, "The Endopar system for minimally invasive robotic surgery" Technical report, TUM (2003)*