

Surgical Skill Evaluation in Endoscopic Sinus Surgery

Vishwa Parekh under supervision of Daniel Mirota, Narges Ahmidi, Dr. Gregory Hager

Group 4, Computer Integrated Surgery II, Spring 2012

Abstract

The sinus anatomy is in vicinity of critical regions like carotid artery, which supplies blood to the brain and the optic nerve. Thus the surgical procedure involves a lot of critical movements to be performed by the surgeon. We want to develop a model that will evaluate how well these critical movements were carried out. This model will help train the surgeons better providing them quantitative feedback on their movements during surgery. This project involves development of a recording software to record the data from surgery, a registration software to register the motion of the tip of endoscope to the CT frame and an approach towards segmenting the motion into individual movements which can then be further analyzed.

Introduction

Endoscopic Sinus Surgery is a minimally invasive surgery. The surgical site is accessed using a laparoscopic device. The advantages a minimally invasive surgery are:

- Reduces operating time.
- Reduces the time required by the patient to recover.
- Prevents excessive loss of blood.

But a minimally invasive surgery requires a very high skill level on the surgeon's part. The major reasons for requirement of an increased skill level are:

- Less degrees of freedom to move the surgical tool.
- Monocular view of the surgical site.

And for our specific case of endoscopic sinus surgery, we have major critical regions like carotid artery, optic nerve, etc. Thus it is very important for the surgeon to have good surgical skills. It has been established in literature that good surgical skills come through practice with positive feedback. Thus, ultimate goal of this project is to develop a model that would analyze a series of surgical motions and determine the skill level of the surgeon and thus help the surgeon improve his skills. The complete technical approach can be summarized into following steps:

1. Data collection
2. Registration of motion of the tip to the CT scan
3. Segmentation of motion on the CT scan into different movements
4. Train a skilled surgeon model for each of these tasks using data recorded from surgeries conducted by experts.
5. Test the model using motion data from a surgeon with unknown skill level.

For the scope of this project, I have implemented steps 1 and 2. Since I have not been able to record any live surgery, I worked out Step 3 for a dataset recorded on a phantom and Step 4 and Step 5 would be done once data from live surgery is recorded. In the following sections, we will deal with each of these steps in the technical approach separately, followed by the experiments performed, results obtained, future work and the management summary.

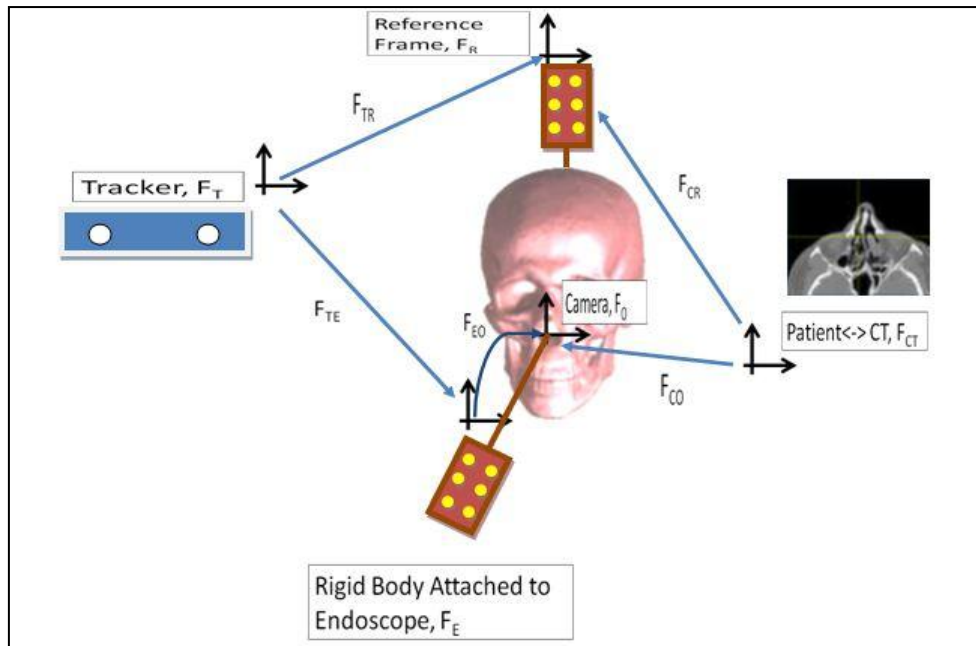


Figure 1. The complete setup in OR for data collection

Data collection

Figure 1 shows the complete setup used for recording the data from OR. For our project, we are interested in recording the endoscopic video and the marker coordinates recorded by the tracker

Technical Approach

In order to record the data from the surgery, a recording software was written. The video recording part was written using the CISST Stereovision library. The tracker used in the OR at Johns Hopkins Hospital

is Medtronic Stealthstation. Using the libraries provided by Medtronic and the 'saw' package on Stealthlink, the tracker recording part was written.

Finally, a recording software that records both the variables and can be start/paused by a pedal was written.

Software Testing:

The software was tested in the Mock OR, using the stealthstation in the Mock OR and a continuous DVI source from Sun Microsystems was used to test the software. In a successful run, the



Figure 2. Image from DVI source

software should be able to produce a video (.avi) file with the video from the source (Figure 2) and a .csv file which contains the location of all the markers with their timestamps.

The camera calibration routine for this software was written using the module for camera calibration from CISST Stereovision library.

Difficulties faced:

- The method in OpenCV to find sift features to track produced erroneous results for the version 2.2, but works perfectly well with the version 2.1. Because of this problem, it took a long time to figure out why the camera calibration routine was not working properly.
- There was a problem with CMake written for the stealthlink modules in saw. Because of this, there was a problem linking the libraries provided by medtronic to the code written using CISST libraries. This problem was resolved by sitting with Anton and changing the CMake files.
- The major problem that I faced and it took me almost 5 weeks to resolve it was the issue of conflicts between the libraries provided by medtronic and the MS Visual Studio standard libraries. There were certain symbols that were being defined by the Medtronic libraries and hence were creating conflicts.

But on turning the NODEFAULT lib option for the MS Visual Studio Standard libraries, certain symbols were not being recognized because they were defined only in these libraries. Finally, this problem was resolved by writing a .def file which contained the symbols defined in the MSVS standard libraries but not in the libraries provided by the medtronic.

- After everything was ready, a major dependency problem occurred due to relocation of the OR to the new building. Due to this there were no surgeries between April 23 and May 7 and hence I was not able to record a live surgery.

Registration Software

The aim of the registration procedure is to accurately determine the position of the tip of the endoscope in the CT coordinate frame. From figure 1, we are looking to calculate the transformation F_{CO} .

The complete outline for the Registration procedure is shown in Figure 3. The registration procedure can be broadly divided into three parts.

- Tracker Based Registration
- Video-CT Registration
- Registration Optimization

The complete registration procedure has been developed using the software TREK, which has been developed by I-STAR laboratory. It is developed over CISST and slicer and is useful in conducting point

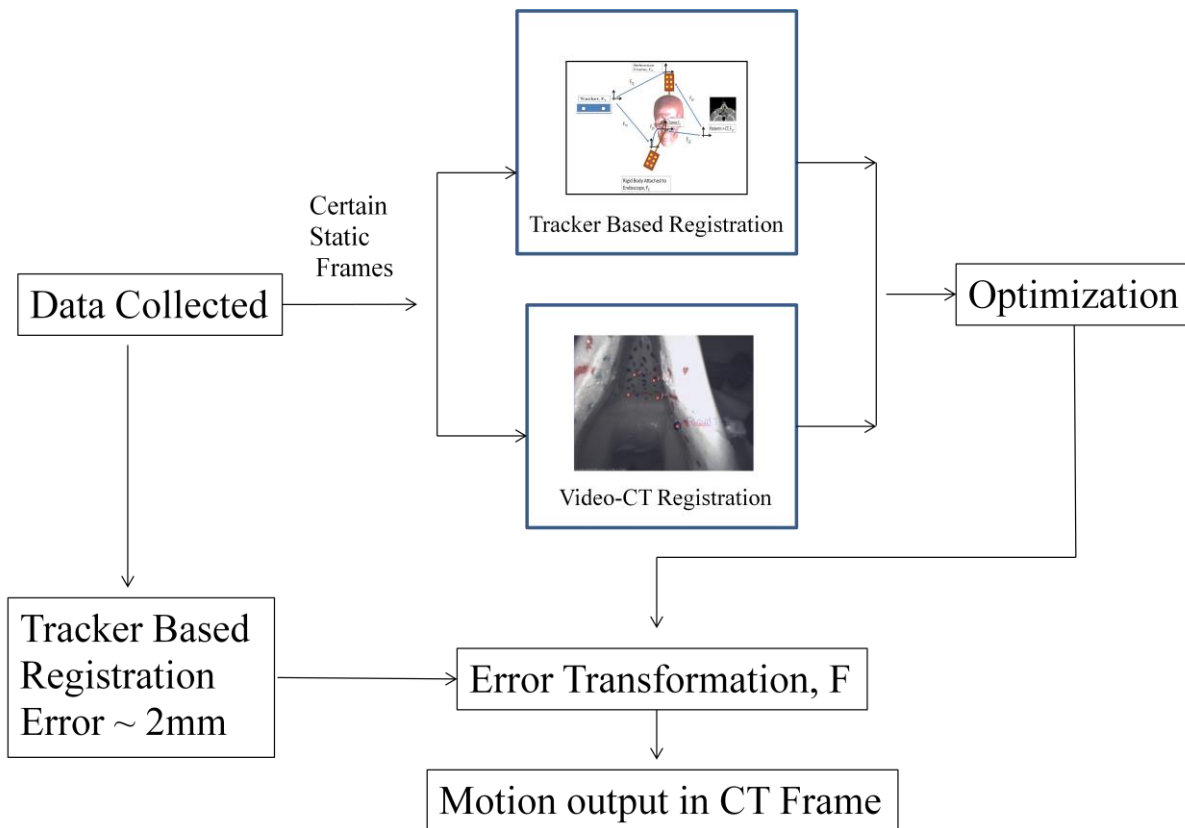


Figure 3. Complete registration procedure.

Tracker Based Registration:

The tracker based registration is obtained by solving the following equation:

$$\mathbf{F}_{CO} = \mathbf{F}_{CR} \mathbf{F}_{TR}^{-1} \mathbf{F}_{TE} \mathbf{F}_{EO}$$

The tracker based registration calculates the position of the tip of the endoscope for all the frames. The problem with this result is that, this registration can have an error of up to 2 mm. In order to refine the registration and reduce the error to as low as possible, Video-CT based 2D-3D landmark registration is carried out.

Video - CT registration.

The basic idea of Video-CT registration here is to segment certain landmarks in CT and register them to segmented landmarks in the endoscopic image. After registering these two, we will be able to determine the new position of the tip of the endoscope in CT coordinates for this particular frame. Now, repeat this procedure for a number of static frames and calculate the new pose for the tip of the endoscope for these frames.

Technical Approach

TREK software was used to mark the 2D landmarks on the image and to load the corresponding segmented landmarks in the CT coordinates. Then, these points were registered using the algorithm implemented in [2]. After running this

algorithm, the new pose for the tip of the endoscope is saved and this is repeated for a number of static frames.

Figure 4 shows an example of 2D-3D registration where the left image shows the camera pose before registration and the image on the right shows the camera pose after registration

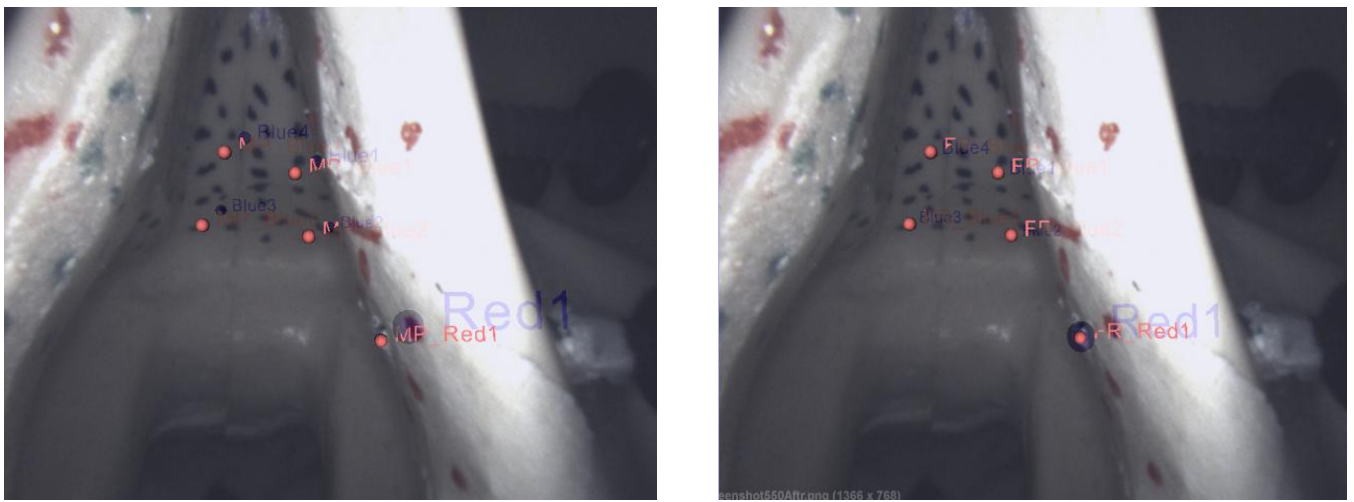


Figure 4. The image on the left shows the camera pose before registration and the image on the right shows the camera pose after registration. The blue points indicate the position of the landmarks as segmented from CT and the red points indicate the position of the landmarks in the image coordinates

Difficulty faced

The problem faced in this approach was the selection of the landmarks in the 2D image. The landmarks had to be selected very carefully to be able to account for the major degrees of freedom of the endoscopic tip.

If the landmarks are not selected properly, this transformation when used to refine the registration over all the frames may worsen the overall registration error instead of improving it. Hence it was a very difficult task to select the landmarks.

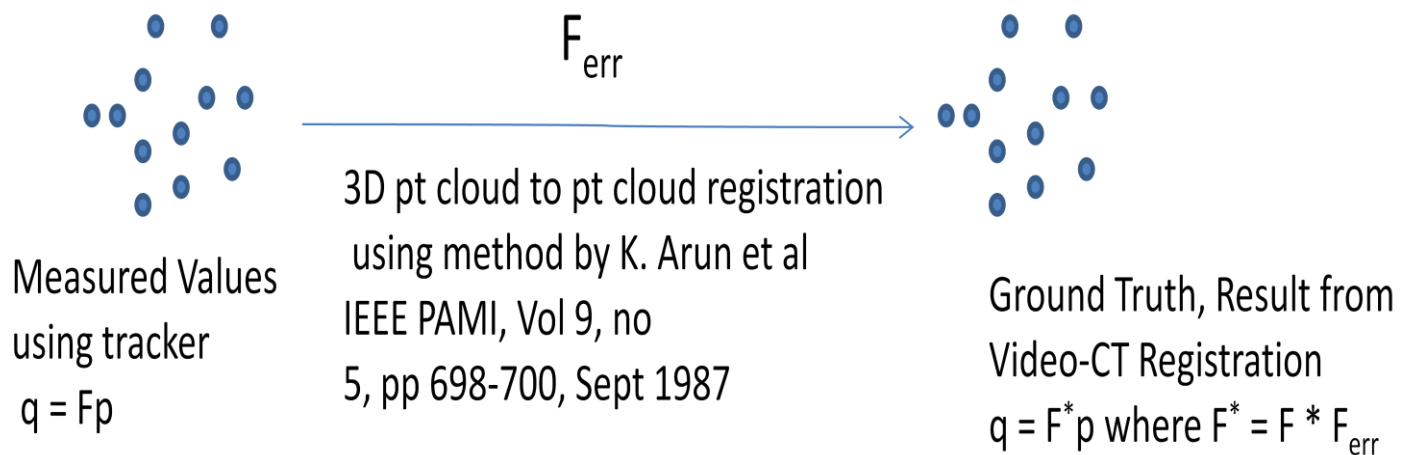


Figure 5: Illustration of optimization procedure

Registration Optimization

Figure 5 illustrates the basic approach to optimize the registration through all the frames. On the left, we have the values obtained for the position of the tip of the endoscope in CT coordinates from tracker based registration and on the right, we have the values obtained for the position of the tip of the endoscope in CT coordinates from Video-CT registration. These values are obtained for certain static frames. An error transformation is calculated by conducting point cloud registration between the two sets of data and is then applied to all the frames and the position of the tip of the endoscope in the CT coordinates is updated.

Difficulty faced

Again the problem here was to determine how good 2D-3D registration was for these frames and how many frames should be considered in order to get a good refinement over all the frames. Thus we are currently conducting experiments using different number of static frames to refine the registration and determining the improvement in the registration obtained.

Experiment conducted to test the registration procedure

The dataset for this experiment was obtained from a previous experiment of

moving the endoscope in a phantom. The tip of the endoscope for the above experiment was to be registered to the CT of the phantom.

For this purpose, 13 different frames where certain CT landmarks were clearly visible on the image were selected and 2D-3D registration was conducted for these frames. After registration, 10 frames were used to optimize the registration over all the frames and the other 3 were used to test the efficiency of registration. This was done for all the combinations and the result of this experiment was:

Error Before Refinement (in mm):
1.570135 2.213231 1.5551

Error After Refinement (in mm):
1.238383 0.949756 1.486395

Segmentation and Motion Analysis

After the tip of the endoscope has been accurately registered to the frame of the CT, the next part is segmenting the motion into different tasks. This is done manually by annotating the video from surgery.

After the complete surgery has been divided into different tasks, the feature vector that would characterize the task would be the location of the endoscope in the CT frame through the task. Since we will be dealing with different patients with CT scans taken at different times and at the different position and orientation of the patient, we would want to transform the coordinate frame from CT to something that would stay consistent among all the patients. Thus the coordinate frame is transformed from CT to a coordinate frame determined by the position of one or more anatomical landmarks of the patient in the CT frame.

Experiment conducted

Since the dataset collected was not from a surgery, the dataset did not have any motions that had tissue interaction. Thus for this particular dataset, the motions were manually segmented into approach into the sinus cavity, coming out from the cavity, etc.

This particular dataset had two instances of the approach into the sinus cavity.

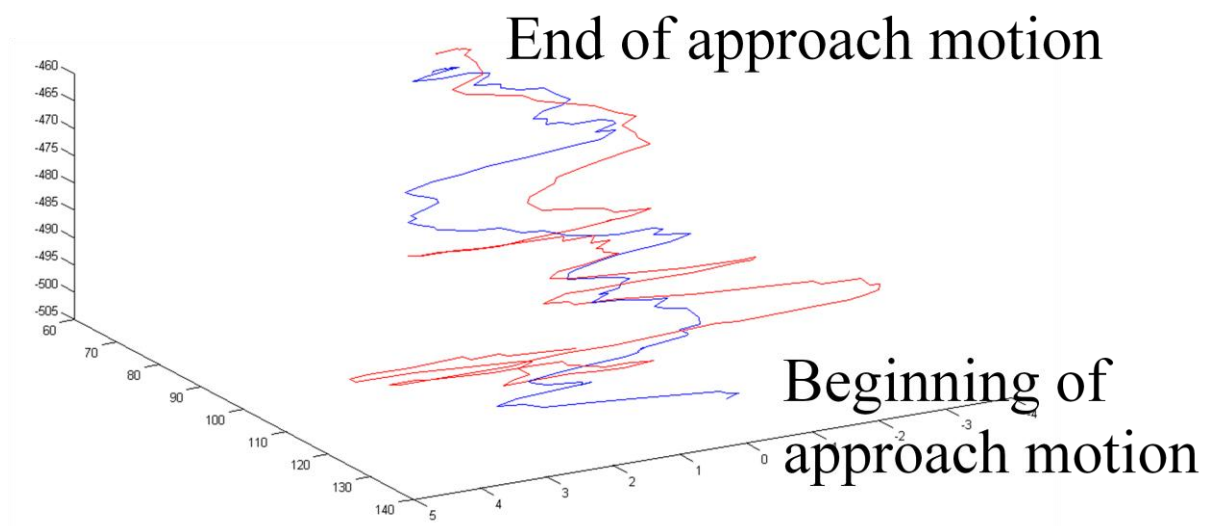


Figure 6 Illustration of the approach motion. Both red and blue plots represent the approach motions for different instances in time

Figure 6 shows the plot of the motion of the tip of the endoscope in the 3D Cartesian space. We can see that, for both the instances, the approach had a lot of side to side motion and jitter, clearly showing that the person handling the endoscope is a novice. Thus the basic idea here is to develop a model such that the differences between different skill levels are characterized.

Future Work

The next step in the future is to record data from expert conducted surgeries and train a Hidden Markov model for each of the different tasks, or some other model that would be able to characterize the differences between an expert and a novice conducted surgery. After the model is trained, it would be tested using unknown datasets from differently skilled surgeons.

Conclusion

The complete plan for deployment of the project is ready and there are a few things that need to be completed before the project deployment. This includes deciding on what kind of landmarks to select for 2D-3D registration and what would be the ideal number of frames for the optimization procedure.

Management Summary

Accomplished vs. Planned

I was able to achieve the expected deliverables in my plan. The reason I was not able to achieve more than the expected deliverables was due to issues regarding library conflicts between the Medtronic and the Visual Studio libraries.

Based on my initial plan, I had allotted a week of time for testing the recording software in the OR, but this actually took me 5 weeks to get the software ready. I recovered most of the time spent in the testing by completing the registration procedure in 1 week when the time allotted for it was 4 weeks. But the major setback I faced was unavailability of any surgical data due to relocation of the ORs. Because of this, I was not able to train a skill evaluation model and hence could not complete my maximum deliverables. Currently I am working towards completing my maximum deliverables.

Lessons Learned

- When developing a software on windows using libraries and codes from various sources, always check the compatibility between different versions and with libraries from VS.
- The quality of Landmark based 2D – 3D registration depends largely on the selection of landmarks for the registration. Hence very careful selection of anatomical landmarks is required.

Acknowledgements:

I am very thankful to

- My mentors, for constantly guiding me and being very patient with my doubts.
- Prof. Taylor, for giving me an opportunity to work on this project and for providing valuable feedback on my work and presentation skills.
- Haluk, to for being very cooperative and for providing constructive feedback on my work.

References

1. D. J. Mirota, H. Wang, R. H. Taylor, M. Ishii, G. L. Gallia, and G. D. Hager, "A system for video-based navigation for endoscopic endonasal skull base surgery," *medical imaging, IEEE transactions on*, ISS. 99, 2011.
2. Chien-ping Lu, Gregory D. Hager and Eric Mjolsness, 'fast and globally convergent pose estimation from video images', *IEEE transactions on pattern analysis and machine intelligence*, 2000
3. Carol E. Reiley and Gregory D. Hager, 'Decomposition of Robotic Surgical Tasks: An analysis of Subtasks and Their Correlation to Skill.
4. D. J. Mirota, A. Uneri, S. Schafer, S. Nithianathan, D. D. Reh, G. L. Gallia, R. H. Taylor, G. D. Hager, and J. H. Siewerdsen, "High-accuracy 3D image-based registration of endoscopic video to C-arm cone-beam CT for image-guided skull base surgery," in *Medical Imaging 2011: Visualization, Image-Guided Procedures, and Modeling*, 2011, p. 79640j-1.