Project 16: Da Vinci Intelligent Surgical Assistance

Chris Paxton (cpaxton3@jhu.edu)

February 20, 2014

1 Summary

My goal is to develop tools to automate parts of a surgical procedure, so that the machine can make procedures faster and more efficient. Surgeries can be hours long, and contain many repetitive motions; automating parts of a task makes a surgeon's job easier and lets them focus on the patient. The Da Vinci robot has a third arm which is often used for supplementary tasks like cutting threads. A surgeon cannot control all three arms at once, so they need to clutch to switch which arm is being controlled. Automating small parts of the procedure that need the third arm would help reduce task complexity and decrease load on the user.

I want to focus on applying this to a simple collaborative block-passing procedure first, then to a suturing example based on surgical training procedures, leveraging the large amount of information collected by the Language of Surgery project.

Mentors I plan on working under the guidance of members of the CIRL lab. Kelleher Guerin and Jonathan Bohren are senior PhD students working on human/machine collaboration; they have I will also be working with Professor Greg Hager. Some recent work has been done on automating parts of the suturing procedure by Pieter Abbeel's group at UC Berkeley [8].

1.1 Background and Significance

Human-robot collaboration is increasingly important as robots become more capable of contributing to skilled tasks in the workplace. Robotic Minimally Invasive Surgery (RMIS) is a part of this trend. Partial automation would decrease the load on surgeons during long procedures by automating repetitive sub-tasks, and it would improve surgeon performance if procedures are being performed over long distances in conditions of high latency. The Language of Surgery project at JHU has collected a large amount of surgical data used for skill classification and for providing feedback to surgeons in training [9, 1]. Recent work has also looked into automatic segmentation of video and kinematic data from these surgical procedures.

I am building off previous work done by Sebastian Bodenstedt and Nicolas Padoy under Professor Hager that used a Gaussian Mixture Regression model to learn when to rotate a ring during a wire walking task [2]. The task was to manipulate a small ring down a length of wire bent into a variety of shapes.

1.2 Specific Aims

The goal of the project is to be able to recognize when specific components of procedures have been completed and enable the system to react appropriately to assist the user. To accomplish this, I need to combine knowledge of the environment attained through machine vision and knowledge of the expected kinematics and effects of a user's motion at different steps in a procedure. Prof. Hager's lab has a large database of surgical procedures, together with manually assigned labels and robot kinematics and camera positions, collected for the Language of Surgery project [9]. I can use this data to learn about the task structure of real-life suturing procedures. Unfortunately, the camera calibration is not very useful right now because the dataset was collected from different robots and the robots' camera focal lengths can change between trials. I need to come up with a robust stereo registration approach using the robot tooltips and kinematics.

1. Stereo registration for the Da Vinci to perform 3D reconstruction of scenes and extract objects of interest.

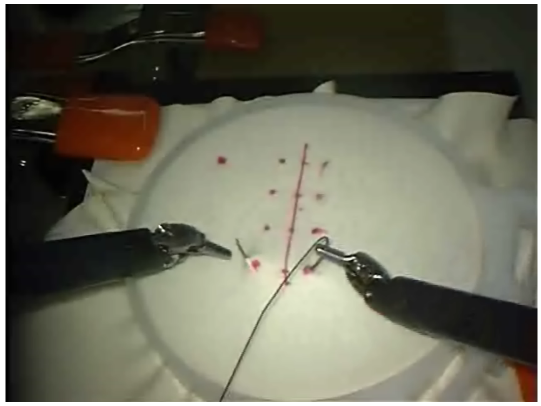

Figure 1: Needle drive from the publicly released Language of Surgery data set.

2. Complete a simple peg passing task with cooperation between user and computer controlled arms of the Da Vinci. In this task, the user takes a peg with one manipulator, hands it to a second manipulator, and then clutches to switch to a third manipulator arm and takes the peg. The approach described by this project should learn how to take control of the third arm and grab the peg, without the need for the user to take control.

3. Assist suturing task in a test Da Vinci surgical procedure through control of the third arm. The robot should be able to grab a suture needle after a needle drive, as per Fig. 1.

2 Deliverables

2.1 Minimum Deliverables

Simple Stereo Registration and Reconstruction calibrate the robot using chessboard images to perform 3D reconstruction of scenes, without finding tooltips. This is already possible, and will be useful if the proposed approach for stereo reconstruction fails.

Adapt Formal Algorithmic Approach The approach for motion and task modeling described below is well established, but needs to be formalized and adapted into an algorithm suited for the specific task. This is very important, and so it needs to be done as soon as possible. Expected by 2/28/2014.

Model Task Components Using this algorithmic approach and a stero registration approach, I should be able to learn an IOC model for individual components of a task. Even if other components of the project fail, this model can still be useful for assessing surgical skill or providing real-time feedback if a user starts to deviate from the expected trajectory. Expected by 3/14/2014.

2.2 Expected Deliverables

Tooltip-based Stereo Registration and Reconstruction Develop a procedure for stereo calibration and reconstruction from collected surgical videos with different camera intrinsics. Expected by 3/7/2014.

Peg Passing Task Finish models to the point where a simple block-passing task can be accomplished through partial task automation. Expected by 4/11/2014.

2.3 Maximum Deliverables

Suturing Task Adapt approach to assist a user in a simple suturing task as described above. Expected by 4/25/2014.

Semi-Automation Toolkit We are also interested in intention recognition and semi-automation for industrial robotics and for high-latency telemanipulation projects. If all other work is successful, I will adapt this project into a toolkit that can be applied to new tasks and robots such as the WAM arms and the Raven surgical robot. Expected by 5/7/2014.

3 Technical Approach

Stereo Registration and Reconstruction Using the available video data for stereo reconstruction is difficult because the camera intrinsic parameters of the robots change from trial to trial, because the robots' camera focal distances can change. In addition, trials are recorded on different robots.

Recent work has been able to identify the Da Vinci tooltips in video [7]. I want to use this to find locations of the tooltips in all collected video data, then use this together with the available camera position and tooltip.

Motion Model Prior work from Prof. Hager's group used Gaussian Mixture Models to determine when rotations needed to occur [2]. I am interested in using a different and hopefully more robust approach to model how interactions should occur. Some recent work in Inverse Optimal Control (IOC) has dealt with learning in continuous environments from locally optimal examples [5], and recent work submitted to IROS by Amir Masoud and myself under Prof. Hager has looked at learning how to incorporate new environmental information into demonstrated trajectories [3].

The approach I plan on using for modeling agents' motion is based on maximum-entropy IOC [10]. In this case, we maximize the probability of actions a from state x based on observed expert trajectories:

$$P(a_0|s_0) = \frac{1}{Z}\sum_t r(s_t, a_t) \tag{1}$$

In this case, however, we also need to take into account noisy environmental features. Motions need to be in relation to observed features of interest (needle, suture points, peg being passed) and the tissue. Previous work has looked at this problem before through the use of hidden variable Markov Decision Processes for activity forecasting [4]. This has also been used to predict the intention of an actor, which is useful for predicting when intervention should take place.

Task Model While the IOC component is capable of modeling individual segments of a complex task, we also need some idea of how different task components fit together. Luckily, our surgical data has already been manually labeled and segmented with a set of rigorously tested and well-defined definitions available on the Language of Surgery wiki. We also know that, when performing third-arm tasks, the user will clutch to switch arm control, providing an easy segmentation of which parts of the task the software will be responsible for handling. Previous work in temporal planning and hierarchical control has elaborated on how to combine multiple sub-tasks(for example, a PR2 plugging itself into a wall in [6]).

4 Requirements

To develop intelligent assistance for surgical procedures, I need access to a robot, task models the robot can perform, and a set of training data. As of Feb. 2014, I already have access to the BB API necessary for read/write instructions to the Da Vinci robot, and I can use the robot in Hackerman for research and development. I also have access to collected surgical data already, and can collect more using the Da Vinci for specific tasks.

I will also be using the open-source CISST and OpenCV libraries. CISST has a number of useful tools for robotics, but it also has a video codec necessary for recent data collected at MISTIC. OpenCV has tools to perform stereo calibration and 3D reconstruction. Both of these systems are already set up on my laptop and workstation. I will use NLOPT, an efficient cross-platform nonlinear optimization library written in C++, to solve necessary parts of the IOC problem for modeling motions.

It may be possible to speed up development with the use of a simulator instead of performing all experiments on the actual Da Vinci; however, at present it is unclear when or if this will happen. The Mimic simulator in question (at Johns Hopkins Bayview) is capable of simulating deformable materials and threads, but we need to wait on the company itself to see whether we can access the position of the needle and thread during the task and to be able to read out the kinematics. Current plans assume I will not be able to use the simulator this semester.

Summary of Primary Requirements

- CISST
- OpenCV
- NLOPT
- Da Vinci
- Intuitive Surgical BB read/write API
- Materials for peg-passing and suturing tasks
- Access to collected surgical video and kinematics data

Summary of Optional Requirements

- Access to Mimic surgical simulator

5 Project Schedule

I will have biweekly meetings with Prof. Hager on Fridays at 9:30 pm and weekly meetings with the group on Mondays at 11:00 am. Additional meetings will be scheduled as necessary. I have outlined the project deadlines below in Table 1.

Deliverable	Deadline
Project Proposal	2/20/2014
Study Related Papers	2/25/2014
Adapt Formal Algorithmic Approach	2/28/2014
Project Presentation	3/04/2014
Tooltip-based Stereo Registration and Reconstruction	3/7/2014
Model Task Components	3/14/2014
Peg Passing Task	4/11/2014
Suturing Task	4/25/2014
Semi-Automation Toolkit	5/07/2014
Poster Session	5/09/2014

Table 1: Schedule for Intelligent Surgical Assistance project.

References

[1] Narges Ahmidi, Yixin Gao, Benjamín Béjar, S. Swaroop Vedula, Sanjeev Khudanpur, René Vidal, and Gregory D. Hager. String motif-based description of tool motion for detecting skill and gestures in robotic surgery. In *Medical Image Computing and Computer-Assisted Intervention–MICCAI 2013*, Nogoya, Japan, 2013. Springer, Springer.

[2] Sebastian Bodenstedt, Nicolas Padoy, and Gregory Hager. Learned partial automation for shared control in tele-robotic manipulation. In *2012 AAAI Fall Symposium Series*, 2012.

[3] Amir M. Ghalamzan E., Chris Paxton, Gregory Hager, and Luca Bascetta. Robot learning from demonstration: from imitation to emulation. 2014.

[4] Kris M Kitani, Brian D Ziebart, James Andrew Bagnell, and Martial Hebert. Activity forecasting. In *Computer Vision–ECCV 2012*, pages 201–214. Springer, 2012.

[5] S. Levine and V. Koltun. Continuous inverse optimal control with locally optimal examples. In *Proceedings of the 29th International Conference on Machine Learning, ICML 2012*, volume 1, pages 41 – 48, 2012.

[6] Conor McGann, Eric Berger, Jonathan Bohren, Sachin Chitta, Brian Gerkey, Stuart Glaser, Bhaskara Marthi, Wim Meeussen, Tony Pratkanis, Eitan Marder-Eppstein, et al. Model-based, hierarchical control of a mobile manipulation platform. In *4th workshop on planning and plan execution for real world systems, ICAPS*, 2009.

[7] Austin Reiter, Peter K Allen, and Tao Zhao. Appearance learning for 3d tracking of robotic surgical tools. *The International Journal of Robotics Research*, page 0278364913507796, 2013.

[8] John Schulman, Ankush Gupta, Sibi Venkatesan, Mallory Tayson-Frederick, and Pieter Abbeel. A case study of trajectory transfer through non-rigid registration for a simplified suturing scenario. In *Intelligent Robots and Systems (IROS), 2013 IEEE/RSJ International Conference on*, pages 4111–4117. IEEE, 2013.

[9] Lingling Tao, Gregory D. Hager, and René Vidal. Segmentation and recognition of surgical gestures from kinematic and video data. In *Medical Image Computing and Computer-Assisted Intervention–MICCAI 2013*, Nogoya, Japan, 2013. Springer, Springer.

[10] Brian D Ziebart, Andrew L Maas, J Andrew Bagnell, and Anind K Dey. Maximum entropy inverse reinforcement learning. In *AAAI*, pages 1433 –1438, 2008.