Lindsey Dean
H. Shawn Xu
CIS II Project Plan

Voice Control of a Surgical Robot

Summary

Our project explores voice integration with the *da Vinci* Surgical System from Intuive Surgical®. We aim to provide surgeons with an alternative method of performing basic functions such as varying camera zoom or “clutch” by simply using voice commands. Similarly, we hope that the same can also be applied the menu and graphics overlay system that was previously developed for the *da Vinci* by the ERC CISST. Currently, many of these functions demand tedious and/or complicated gestures by surgeons' hands and/or feet to perform. Moreover, these gestures often require the surgical tools be temporarily locked in place and the surgery briefly paused. We hope that voice integration would allow surgeons to allocate more of their time during surgery to manipulating the surgical tools instead of to performing peripheral functions, and thus ultimately smooth the surgical process.

- **Students:** Lindsey Dean, H. Shawn Xu
- **Mentor:** Anton Deguet

Background, Specific Aims, and Significance

Currently surgeons operate the *Da Vinci*® machine using their hands and feet. The master console is equipped with foot pedals and the newer models even have kinesthetic combinations where the surgeon must tap their leg against the side to perform a specific task. While these gestures are meant to be natural for the surgeon it has been proposed that adding voice-to-command utility to the machine could further simplify the surgeon’s interactions with the machine.

The CIIST libraries have developed the 3D-UI which is capable of allowing the surgeon to interact with the images and find information such as distance between points. Currently, however the 3D-UI cannot be used at the same time as the surgeon manipulates the salve, but rather the surgeon’s hands control a mouse on the 3D-UI. Through voice command there is potential for the surgeon to have an additional way of interacting with the interface without losing the functionality of their hands to manipulate the machine.

Our specific aims are:

1. Increase ease of use of Da Vinci machine
2. Add voice-control capability to 3D-UI

Deliverables

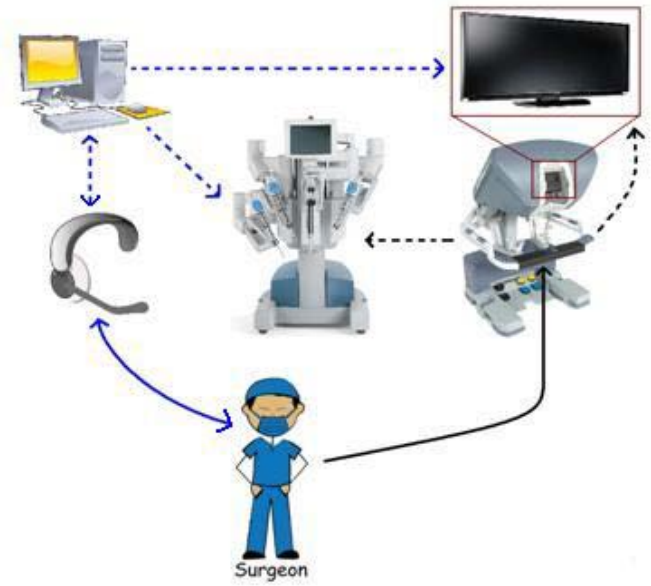
Give minimum, expected, and maximum specific, measurable deliverables for the project. If more than one, give a numbered list. Something along the following lines.

- **Minimum:**
  - Well-documented program that adds singular functionality
  - A video demonstration of voice control

- **Expected:**
  - Integration of voice and physical control (ie. Camera, clutch)
  - Add additional multi-state functionality
  - Additional demonstration(s) that show different functions voice can perform on Da Vinci

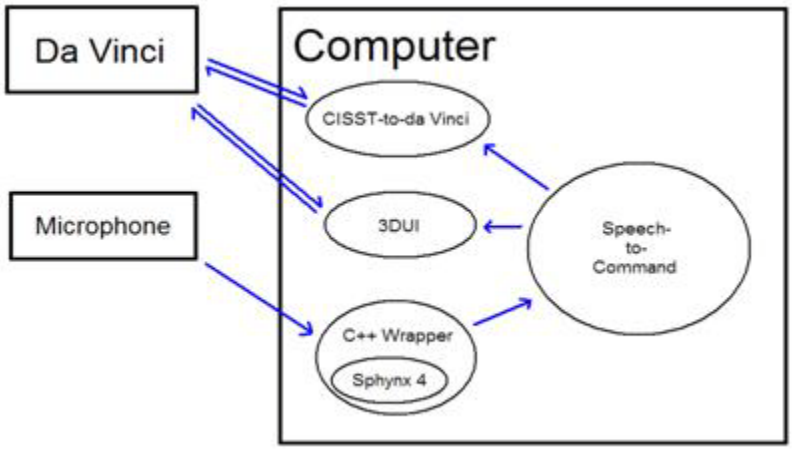
- **Maximum:**
  - Fully-functioning library of states and commands that can easily be expanded upon

Technical Approach



The surgeon will be able to speak commands into a microphone/headset to control what is displayed on the console display and the surgical robot. We also hope to implement speech feedback.

The general architecture of our software is as follows:



The surgeon's microphone communicates with Sphynx (which has a C++ wrapper). This in turn communicates with our speech-to-command system. This system sends commands to the already built CISTT-to-da Vinci and 3DUI systems. These can directly control the da Vinci robot and console display. We also hope that an additional layer of feedback and interfacing can be added through voice feedback.

The speech-to-command logic will be a state-based approach. The program will keep track of the current state of the system. When the surgeon wants to perform a certain voice controlled task, he first speaks the command to change the state of the system. Then he triggers the task with another command. Finally, he tells the system to revert back to the default state or to switch to another state. Each state will only have a few allowable commands, and this will reduce the risk of interference. We may also add universal commands (acceptable in all states), such as STOP.

Management Plan

Dependencies

	Being done to resolve...	Affects..	Resolve by...
Access to Mock OR	Allison Morrow	Exploratory Phase	2/21
NDA w. Intuitive Surgical	Alyssa	Exploratory Phase	2/21
JHED Access to Mock OR computer	Anton	Exploratory Phase	2/27
Time with Anton to put necessary software on computers	Anton	Design Phase	3/6
Sphynx 4 C++ Wrapper	CS undergrad should be done in 2 weeks	Design Phase	3/12
Video Camera	DMC has available to reserve	Implementation Phase	April

Milestones

March 13

- Finish design and list of states and commands we want to implement
- Fully flesh out software architecture

April 17

- Complete simple voice control program
- Finish first video demonstration

May 15

- Wrap up all coding

May 19

- Poster presentation

Timeline

(subject to change)

Bibliography

- Sevinc, Gorkem. *INTEGRATION AND EVALUATION OF INTERACTIVE SPEECH CONTROL IN ROBOTIC SURGERY*. Thesis. Johns Hopkins University, 2010. Print.
- A. Kapoor, A. Deguet, and P. Kazanzides, "Software components and frameworks for medical robot control," in Robotics and Automation, 2006. ICRA 2006. Proceedings 2006 IEEE International Conference on, 2006, pp. 3813-3818.
- "Sphinx-4." *CMU Sphinx - Speech Recognition Toolkit*. Web. <http://cmusphinx.sourceforge.net/sphinx4/javadoc/index.html>.
- Liu, Peter X., A.D. C. Chan, and R. Chen. "Voice Based Robot Control." *International Conference on Information Acquisition (ICIA)* (2005): 543. Web. <http://ieeexplore.ieee.org.proxy3.library.jhu.edu/stamp/stamp.jsp?tp=&arnumber=1635148>.