Voice Control of *da Vinci*® Checkpoint Presentation

Lindsey Dean and H. Shawn Xu
April 7, 2011

Mentor: Anton Deguet

Background

- The *da Vinci*® is a robotic teleoperated surgical system
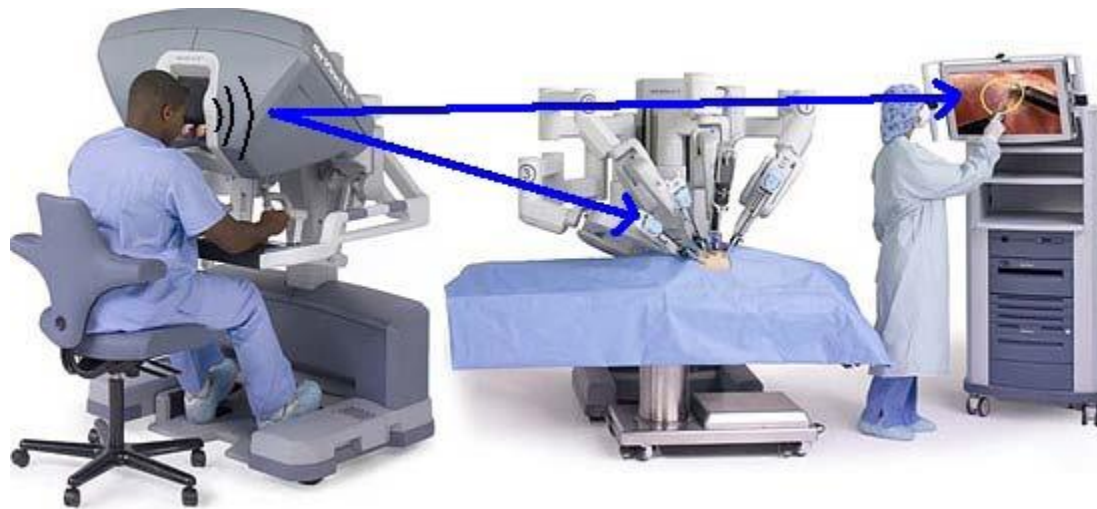- Controlled by surgeon at HD workstation with hands and feet

PROBLEM:
Complex gestures, stop-start procedures

SOLUTION:
Allow surgeon to control certain parts of the system via voice

Background (continued)

Courtesy of www.IntuitiveSurgical.com

Milestones & Progress

	Milestone	Status	Planned	New Expected	Accomplished
✔	Overcome logistical dependencies		2/27	––	2/28
✔	Ready for Software Architecting	All necessary files on computer	3/17	––	3/28
✔	Finished Design of Software		3/23	––	4/7
	Working Demo of Voice Control of DaVinci	Need voice program working on PC	4/17	4/21	––
	Incremental improvement of first voice demo		4/20	4/28	––
	Wrap up all coding – produce final report		5/15		––

Dependencies

|	Status	Planned Date	Date Accomplished
Access to Mock OR	x	2/21	2/28
Signed NDA with Intuitive	x	2/21	2/28
Account on DaVinci PC	x	2/27	3/3
Software Installation: Sphinx4 (w/ C++ wrappers), ITK, CISST libraries, 3DUI, DaVinci Wrappers	x	3/12	3/28
Anton's Time	-Has been very dependable and believe will continue to be so	Cont. through 4/26	--
Voice Package Functioning on PC	-Currently having difficulties with optimal configuration	Mid-April (Next Week)	--

What We've Been Up to...

▸ **Exploration:** We and Anton have brainstormed several examples that we think would best demonstrate practical uses of voice control.

▸ **Initial Design:** We have a general design for building our demos (see next slide).

▸ **Implementation:**
  ◦ We have gathered, installed/built all the necessary separate pieces
  ◦ We are currently still debugging Sphinx 4 – this will be an ongoing problem (see Difficulties)

ERC | CISST

LABORATORY FOR
Computational
Sensing + Robotics
THE JOHNS HOPKINS UNIVERSITY

Design

- Three separate components
  - *da Vinci*®
  - 3D User Interface
  - Voice Recognition

- Our demo will be a 4th internal component (voice user interface)
- It will
  - Define contexts
  - Define grammars for each context
  - Listen for events from voice recognition
  - If input is defined in current context, either change context or communicate with *da Vinci*® or 3DUI

- Easy to add/modify behaviors: simply change context/grammar definition and change links between provided and required interfaces
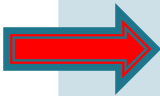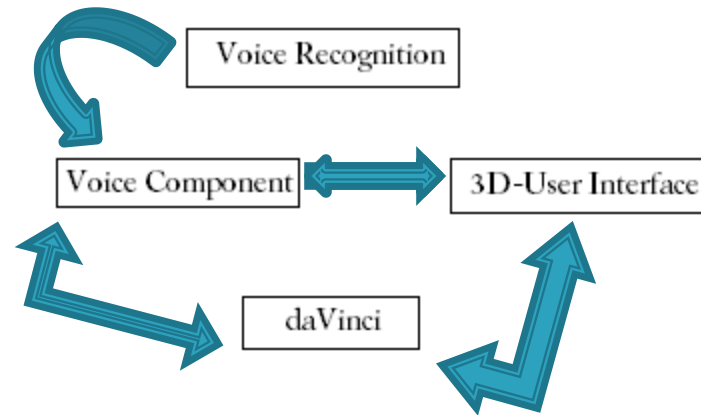
What Has Been Completed

- Created a Cmake file containing all the necessary libraries (CISST, CISSTdaVinci, Robotorium, VoiceDemo) – VoiceDemo is our program which will use keywords delivered to it from Sphinx to trigger events – have been going through tutorials to identify which blocks of code can be isolated to trigger events

- Currently voice will not override will simply be another form of interface – surgeon can both use "select" gesture for 3D-UI on master console or can say which menu option they would like

- Preliminary program which is activated by keyword "voicecontrol" which then allows the system to listen for certain set of words depending on which word is selected the next set of keywords capable of triggering an event may be a simple yes/no, or another list of keywords comprising a menu

Current Status

▸ We are able to modify and have tested PC's communication with voice recognition and *da Vinci*® /3DUI separately

▸ Compiling source files for project to ensure proper compilation among all the elements – Sphynx, CISST3DUI, CISST

▸ This has been more troublesome than anticipated

Milestones & Progress

	Milestone	Status	Planned	New Expected	Accomplished
	Overcome logistical dependencies		2/27	––	2/28
	Ready for Software Architecting	All necessary files on computer	3/17	––	3/28
	Finished Design of Software		3/23	––	4/7
→	Working Demo of Voice Control of DaVinci	Need voice program working on PC	4/17	4/21	––
→	Incremental improvement of first voice demo		4/20	4/28	––
→	Wrap up all coding – produce final report		5/15	5/15	––

Current Difficulties

- Sphinx 4 on Windows
  - Out-of package, it just doesn't work well
  - We must tinker with the configuration file
  - Two options
    1. "always" match with grammar – this makes the computer recognize random sounds... (CURRENT BEST OPTION)
    2. allow out-of-grammar match – this makes the computer not recognize anything...

- Shared vs. static library conflict on Windows
  - Sphinx 4 is a JAVA program with a C++ wrapper and since it runs in a virtual environment, MUST be compiled as a dynamic library
  - CISST-to-daVinci and other components currently compile as static library
  - When PC is pulling from both static and shared libraries a run-time error is thrown due to the possibility of competing versions of code
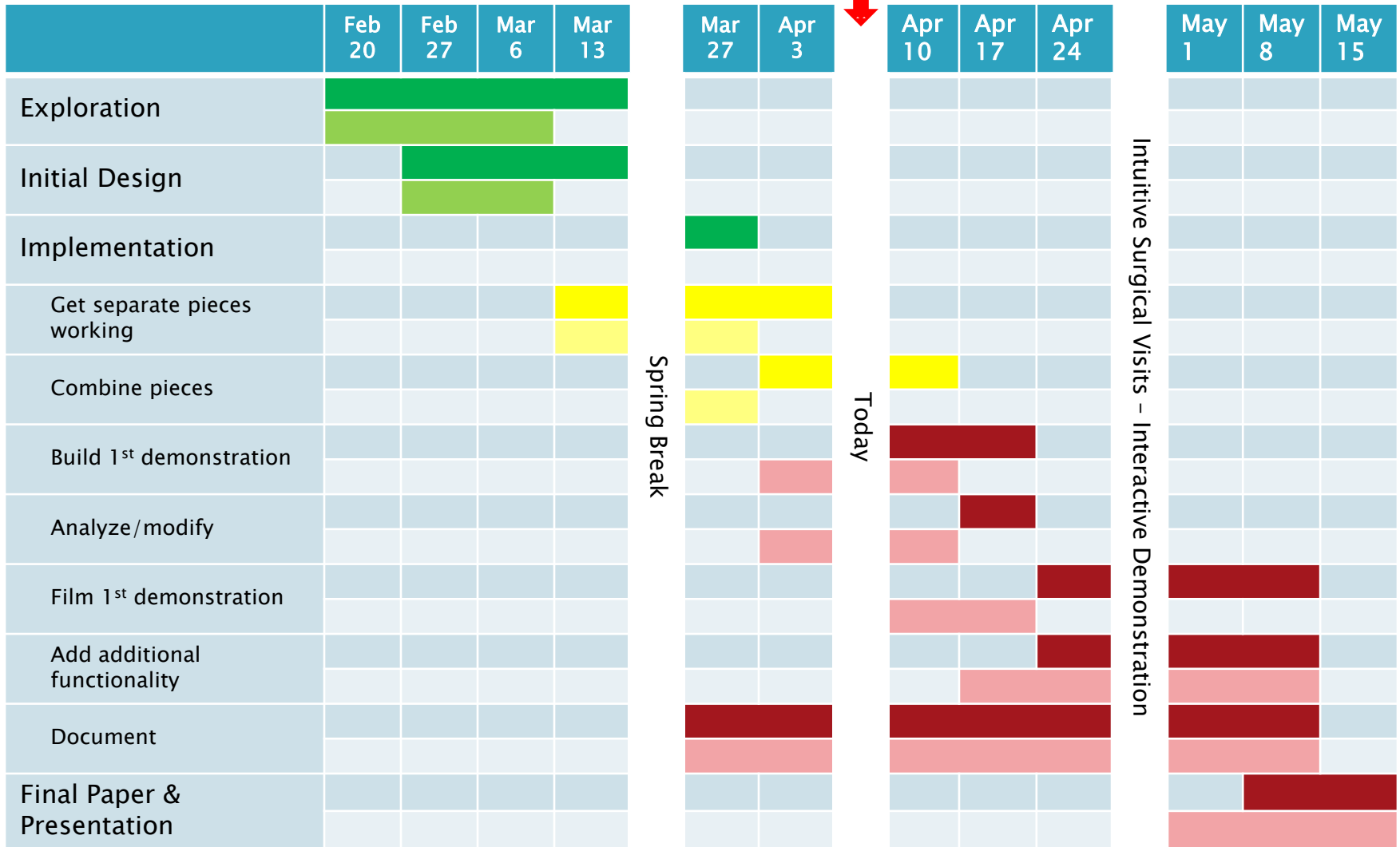
Moving Forward

▸ **Solve difficulties** – Anton is one of our greatest dependencies but is also very dependable in working very closely with us to ensure we are progressing and able to reach our deliverables

▸ Build 1st demo
▸ Test/analyze/modify
▸ Film video demonstration (more concerned with Intuitive Demonstration now)

▸ Add functionality

Deliverables Unchanged

▸ Minimum (100%)
  ◦ Well-documented demo program that adds singular functionality
  ◦ A video demonstration of voice control

▸ Expected (75%)
  ◦ Add multi-state functionality
  ◦ Additional features in demo program that show different functions voice can perform on *da Vinci*®

▸ Maximum (→0%)
  ◦ Fully-functioning library of states and commands that can be easily expanded upon

Timeline

	Feb 20	Feb 27	Mar 6	Mar 13		Mar 27	Apr 3		Apr 10	Apr 17	Apr 24		May 1	May 8	May 15
Exploration															
Initial Design															
Implementation															
Get separate pieces working															
Combine pieces															
Build 1st demonstration															
Analyze/modify															
Film 1st demonstration															
Add additional functionality															
Document															
Final Paper & Presentation															

Spring Break

Today

Intuitive Surgical Visits – Interactive Demonstration

Summary

▸ We are slightly behind our original schedule.

▸ We still hope to meet our original expected deliverable: a demo that shows several features of the *da Vinci*® system and 3DUI being controlled by voice.

- We'd like to acknowledge our mentor Anton Deguet for all of his support through this project

- Thank you. Questions?