Voice Control of da Vinci® Checkpoint Presentation

Lindsey Dean and H. Shawn Xu
April 26, 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)
## Dependencies

<table>
<thead>
<tr>
<th>Task</th>
<th>Status</th>
<th>Planned Date</th>
<th>Date Accomplished</th>
</tr>
</thead>
<tbody>
<tr>
<td>Access to Mock OR</td>
<td>x</td>
<td>2/21</td>
<td>2/28</td>
</tr>
<tr>
<td>Signed NDA with Intuitive</td>
<td>x</td>
<td>2/21</td>
<td>2/28</td>
</tr>
<tr>
<td>Account on DaVinci PC</td>
<td>x</td>
<td>2/27</td>
<td>3/3</td>
</tr>
<tr>
<td>Software Installation: Sphinx4 (w/ C++ wrappers), ITK, CISST libraries, 3DUI, DaVinci Wrappers, Robotorium</td>
<td>x</td>
<td>3/12</td>
<td>3/28</td>
</tr>
<tr>
<td>Anton’s Time</td>
<td>–Has been very dependable and believe will continue to be so</td>
<td>Cont. through 4/26</td>
<td>4/26</td>
</tr>
<tr>
<td>Voice Package Functioning on PC</td>
<td></td>
<td>Mid–April (Next Week)</td>
<td>4/21</td>
</tr>
</tbody>
</table>
Completed Tasks

- We have met our minimum deliverable.

- We have successfully built our first demo (to be shown tomorrow at 9:30AM)
  - Allows surgeon to control 3DUI with voice
Logic of Demo

- **AreYouTalkingToMe**
  - Voice Control

- **SelectMode**
  - Distance
  - Map
  - Stop Listening

- **Measurement**
  - Quit
  - Restart
  - Freeze
  - Start

- **Map**
  - Add Marker
  - Show
  - Hide
  - Quit
## Milestones & Progress

<table>
<thead>
<tr>
<th>Milestone</th>
<th>Status</th>
<th>Planned</th>
<th>New Expected</th>
<th>Accomplished</th>
</tr>
</thead>
<tbody>
<tr>
<td>Overcome logistical dependencies</td>
<td></td>
<td>2/27</td>
<td>--</td>
<td>2/28</td>
</tr>
<tr>
<td>Ready for Software Architecting</td>
<td></td>
<td>3/17</td>
<td>--</td>
<td>3/28</td>
</tr>
<tr>
<td>Approved Documentation of Software Framework</td>
<td></td>
<td>3/23</td>
<td>--</td>
<td>4/10</td>
</tr>
<tr>
<td>Working Demo of Voice Control of DaVinci</td>
<td>Live demo on 9/27 at 9:00AM</td>
<td>4/17</td>
<td>--</td>
<td>4/25</td>
</tr>
<tr>
<td>Incremental improvement of first voice demo</td>
<td></td>
<td>--</td>
<td>--</td>
<td>--</td>
</tr>
<tr>
<td>Wrap up all coding – produce final report</td>
<td></td>
<td>5/15</td>
<td>--</td>
<td>--</td>
</tr>
</tbody>
</table>
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
Current Status

- Solved previous difficulties

- Sphinx 4 on Windows
  - Tinkered with configuration
  - Still not perfect, but working well enough for proof-of-concept demonstration

- Shared vs. static library conflict on Windows
  - Problem solved (THANKS ANTON)
Moving Forward

- **Clean up logic**
  - Currently a lot of the behavior logic is very confusing (was meant for 3DUI)
  - We hope to refactor the code so that adding additional functionality or connecting other components is more intuitive

- **Implement additional functionality**
  - Connect with *da Vinci®* console/arms directly
    - Control camera?
    - Select which arms the console controls?
Deliverables Unchanged

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

- **Expected (75%)**
  - 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

<table>
<thead>
<tr>
<th>Exploration</th>
<th>Feb 20</th>
<th>Feb 27</th>
<th>Mar 6</th>
<th>Mar 13</th>
<th>Mar 27</th>
<th>Apr 3</th>
<th>Apr 10</th>
<th>Apr 17</th>
<th>Apr 24</th>
</tr>
</thead>
<tbody>
<tr>
<td>Initial Design</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Implementation</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Get separate pieces working</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Combine pieces</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Build 1st demonstration</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Analyze/modify</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Add additional functionality</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Document</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Final Paper &amp; Presentation</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

- **Spring Break**:
  - Mar 27
  - Apr 3
  - Apr 10
  - Apr 17
  - Apr 24

- **Today**:
  - May 1
  - May 8
  - May 15

---

**ERC | CISST**

```
LABORATORY FOR Computational Sensing + Robotics
THE JOHNS HOPKINS UNIVERSITY
```
Summary

- We are excited to have met our minimum deliverable, a working demo of voice control on the *da Vinci®*

- After tomorrow’s live demo, we will continue to add functionality
Thank You!