Introduction

Surgical procedures, while they have been essential in the field of medicine, are often very invasive and can have long lasting side-effects. Image Guided Therapy (IGT) seeks to reduce the tissue trauma by allowing greater control of the surgical procedure and having real-time feedback on the effect of the intervention. In IGT, the physician uses surgical instruments that allow for precise tracking in conjunction with images such as ultrasound, computed tomography (CT), or magnetic resonance imaging (MRI) in order to directly or indirectly guide the procedure. Surgical navigation software is often used in the operating room for the physicians to review images, plan the procedure, and monitor the progress. However, because the physicians cannot touch anything outside the sterile field during the procedure, they require a technician on hand just to navigate the software, whether to manipulate the viewer or to scroll through the images. The goal of this study was to create a voice controlled interface for a surgical navigation software in order to alleviate the requirement for a technician to operate it. We implemented the interface as a plug-in for an open-source navigation software, 3D Slicer, to receive voice commands and manipulate Slicer itself using the SpeechRecognition Python library and the Qt Toolkit. The module recognizes commands such as rotating the 3D viewer, saving the scene, and scrolling through the images. With further improvements, such an interface can potentially make the operating process more streamlined and efficient for end-users.

Materials and Methods

Medical Image Software Platform: 3D Slicer
• Free and open source platform for image analysis and scientific visualization allowing rapid prototyping, customization, and testing of new algorithms
• Capabilities include handling Digital Imaging and communications in Medicine (DICOM) images, automatic image segmentation, and tracking of devices for image-guided procedures.
• Built with C++, Qt, and Python

Programming Language: Python
• Common scripted language widely accepted in the scientific community
• Allows quick prototyping and testing.
• Works on 3D Slicer

Voice Recognition Software: Speech Recognition
• Speech and language processing library for Python
• The library gives the option to use multiple different API’s, including the Google Web Speech API and CMU’s PocketSphinx API

Workflows:
• The user commands 3D Slicer with his/her voice after pressing the start button (Fig. 1) to call a registered task, such as changing view orientation.
• Once the voice command is converted to a string, the module looks up a dictionary of available voice commands and calls a corresponding callback function (Fig. 2). This architecture makes the code readable and modular, reduces parsing time.
• The callback function calls Slicer’s Application Program Interface (API) to call its functions.

Discussion

Findings
• For more complicated tasks such as the slice location operations, we found the average time with the voice to be shorter and more consistent, with a smaller standard deviation.
• All other actions such as the 3D Viewer operations had a faster time with the computer, and this is likely because those operations only require 1-2 mouse clicks or button presses.
• The fewer mouse clicks and button presses also explain the smaller standard deviations.

Challenges
• A period of around 4-5 seconds of combined initialization and processing time was required when using the voice, significantly adding to the total time compared to using the mouse.
• The initialization and processing time was subtracted from the voice data in order to make a fair comparison between the two methods.
• Most of the obstacles encountered during the usage of the module was simply in processing the speech, which is a limitation on the API used to record and parse the voice.
• More samples will be needed to obtain conclusive results.

Conclusions
• Our interface provides the foundation for a voice controlled tool that can in the future help surgeons and physicians, with the potential to make the operating process easier and more streamlined by removing the requirement for a technician to manage 3D Slicer.
• In the future, we hope to reduce initialization and processing time to make the module faster and more responsive by using more advanced API’s or 3rd-Party hardware such as the Amazon Echo.
• With further improvements, we plan to have the module be an official extension added to 3D Slicer that can be downloaded and used by anyone.

Results

Hypothesis:
• The Voice command interface improves the usability of the medical image software compared to regular mouse-based operation.

Experiment Method
• Five Volunteers were enrolled.
• The volunteers were asked to complete 26 tasks in four categories with the voice and with the mouse (Table 1).
• A built-in timer was used in the module to record the time for the voice, and a human-operated timer was used for the non-voice tests.
• For operations with the voice, estimated durations for initialization (3 seconds) and processing (1 second) were excluded.
• The times to complete the given tasks accomplished with the voice were compared against the times accomplished with the mouse using a paired Student’s-t-test. A p-value less than 5% is considered statistically significant.

Table 1: Total list of all the operations the voice controlled module can perform: 3D Viewer Operations, which are operations that manipulate the 3D Viewer, Axes Orientation Operations, which deal with the orientation of the 3D Viewer Camera, Slice Location Operations, which handle image scrolling, and Data Operations, which deal with patient data.

Table 2: The times it took to complete operations using the voice and using the mouse. The average time and standard deviation were calculated overall and for each of the four subgroups. P-values based on a T-Test are also shown.

References


Acknowledgements
Special thanks to Ahmed Mahrin, Andrew Zheng, Dian Ren, and Pedro Moreira for acting as test subjects for this project. We would also like to thank the RISE program and Brigham and Women’s Hospital for this wonderful opportunity.