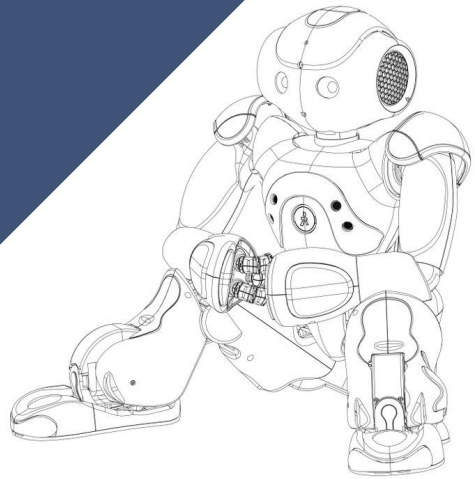
Visual and Aural Telepresence via NAO Robot



Chelsea Cantone (CoE), Theresa Pham (CoE), Daniel Ponsini (EE)

Advisor: Dr. Seung-yun Kim

May 2nd, 2018

TCNJ THE COLLEGE OF NEW JERSEY

Overview

Introduction

- Problem Definition & Need Identification
- Humanoid Robot Background
- Project Goals
- Team Breakdown

Specifications

- High-Level Look at System
- System Block Diagram
- Quantitative Specifications

Subsystem Overview

- Web Application Interface
- User and Robot Systems

Project Management

- Budget
- Hours

Conclusion

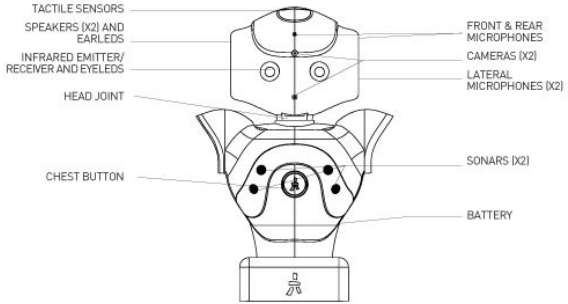
Questions

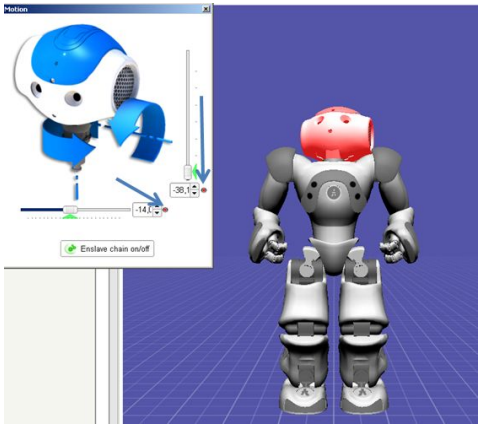
Problem Definition & Need Identification

- Applications for **telepresence/teleoperation devices**
  - Remote in to lectures, classes, conferences
  - Navigate dangerous terrain
  - Work remotely as tour guides, security, consultants, etc.
- Will allow people in far away places or people with disabilities to **immerse** themselves and physically **interact** in a **remote environment**

Humanoid Teleoperated Robot

- SoftBank Robotics NAO Robot
- Ideal for HRI (human robot interaction)
  - 23 inches (58 centimeters)
  - 25 Degrees of freedom to mimic human motion
  - Microphones (x 4) and speakers (x 2)
  - Cameras (x 2)





Project Goals

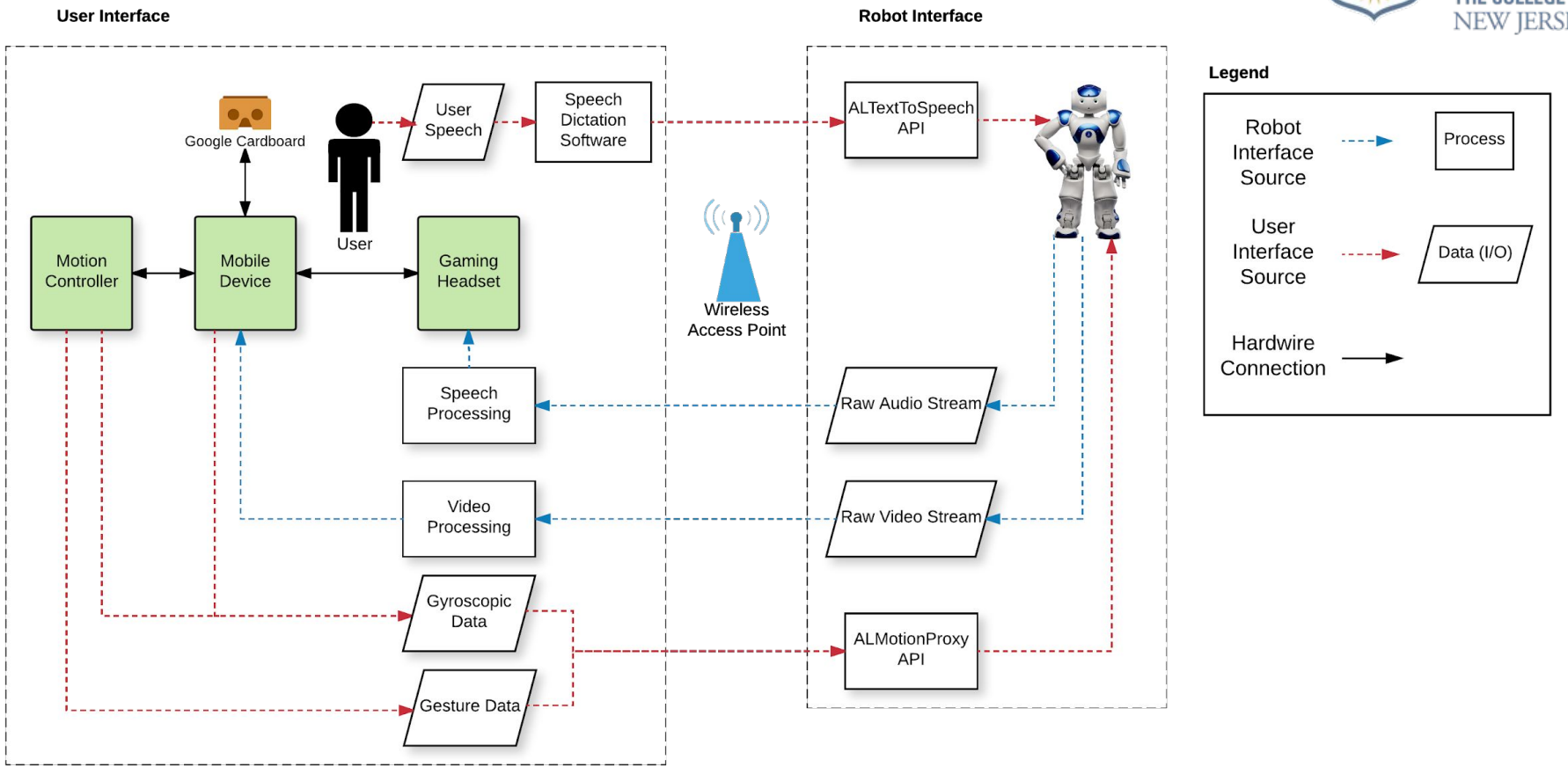
- To **aurally** and **visually** engage the user and audience using the NAO robot
- To develop **wireless communication** between the NAO and a user wearable headset
  - connecting movement using **gyroscope** data
- To develop a **web/mobile application** which receives a **live video and audio stream** from NAO
- To incorporate *Motion-Based Humanoid Robot Controller* project

Team Breakdown

Team Member	Work
Chelsea Cantone	To implement audio streaming and speech processing between the NAO and its operator
Theresa Pham	To work on video processing and develop the web/ mobile application for video streaming to the operator
Daniel Ponsini	To work on controlling NAO movement based on sensor information and establish wireless communication between the NAO and headset

High-Level Look at the System

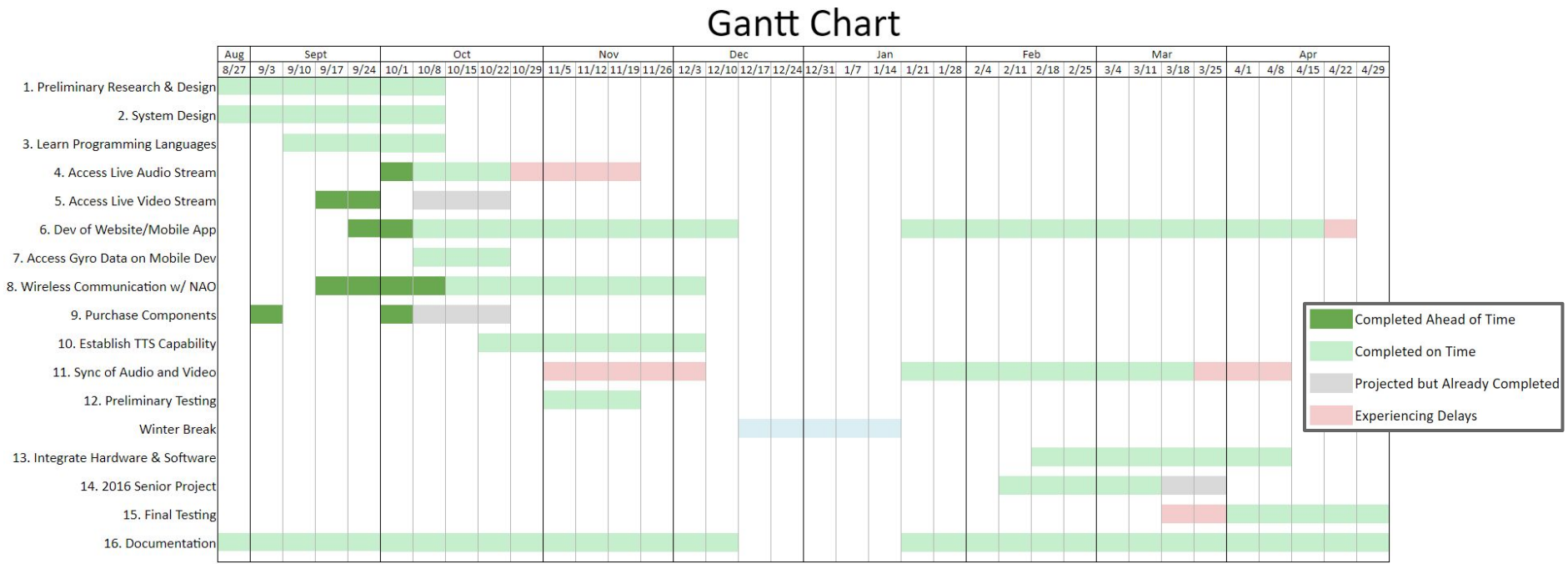
Detailed Block Diagram



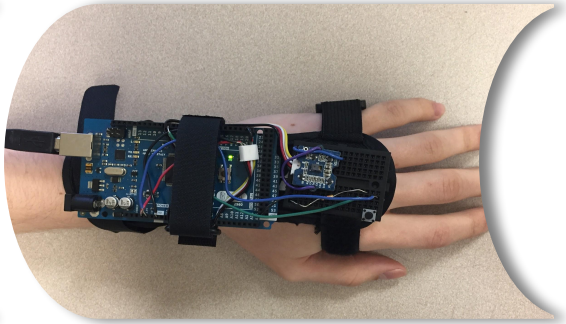
Quantitative Specifications

Specification	Quantitative Value
Horizontal Field of View:	60 degrees
Vertical Field of View:	50 degrees
Headset Weight:	< 2.5 lbs
Arm Controller Weight:	< 1 lb
Battery Life:	1 - 2 hrs
Video Resolution:	320×240 pixels
Tolerated Latency:	< 1 second
Framerate:	2 fps
Movement sensitivity:	+/- 10 degrees
Degrees of Freedom in Teleoperated Robot:	6
Wireless Range:	32 meters
Wireless Transmission Reliability:	< 3% frame loss rate

Schedule


Gantt Chart
Aug
Sept
Oct
Nov
Dec
Jan
Feb
Mar
Apr
8/27 9/3 9/10 9/17 9/24 10/1 10/8 10/15 10/22 10/29 11/5 11/12 11/19 11/26 12/3 12/10 12/17 12/24 12/31 1/7 1/14 1/21 1/28 2/4 2/11 2/18 2/25 3/4 3/11 3/18 3/25 4/1 4/8 4/15 4/22 4/29
1. Preliminary Research & Design
2. System Design
3. Learn Programming Languages
4. Access Live Audio Stream
5. Access Live Video Stream
6. Dev of Website/Mobile App
7. Access Gyro Data on Mobile Dev
8. Wireless Communication w/ NAO
9. Purchase Components
10. Establish TTS Capability
11. Sync of Audio and Video
12. Preliminary Testing
Winter Break
13. Integrate Hardware & Software
14. 2016 Senior Project
15. Final Testing
16. Documentation
Completed Ahead of Time
Completed on Time
Projected but Already Completed
Experiencing Delays

System Components



Web Application	Headset	Arm Controller
▪ Where the user interacts with the system ▪ Accessible from any smart device	▪ Live video stream ▪ Phone gyroscope data ▪ Live audio stream ▪ User speech input	▪ Leg motion control using gesture sensor ▪ Speech control using button

Web Application Interface

Pages:

- Home Page
- Start Page
- Main Page
- End Screen/Redirection Page

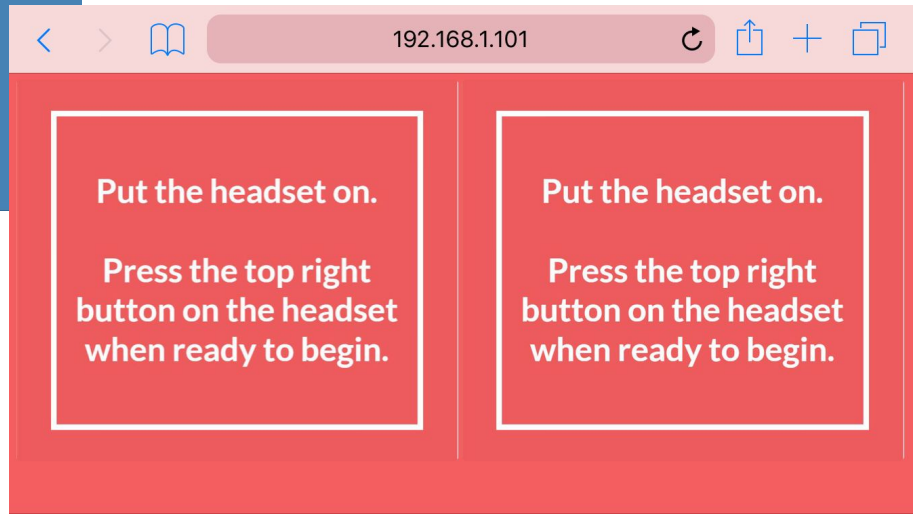
Utilized:

- Python: Flask (web framework)
- Front-End Languages: HTML, JavaScript

To Start Application

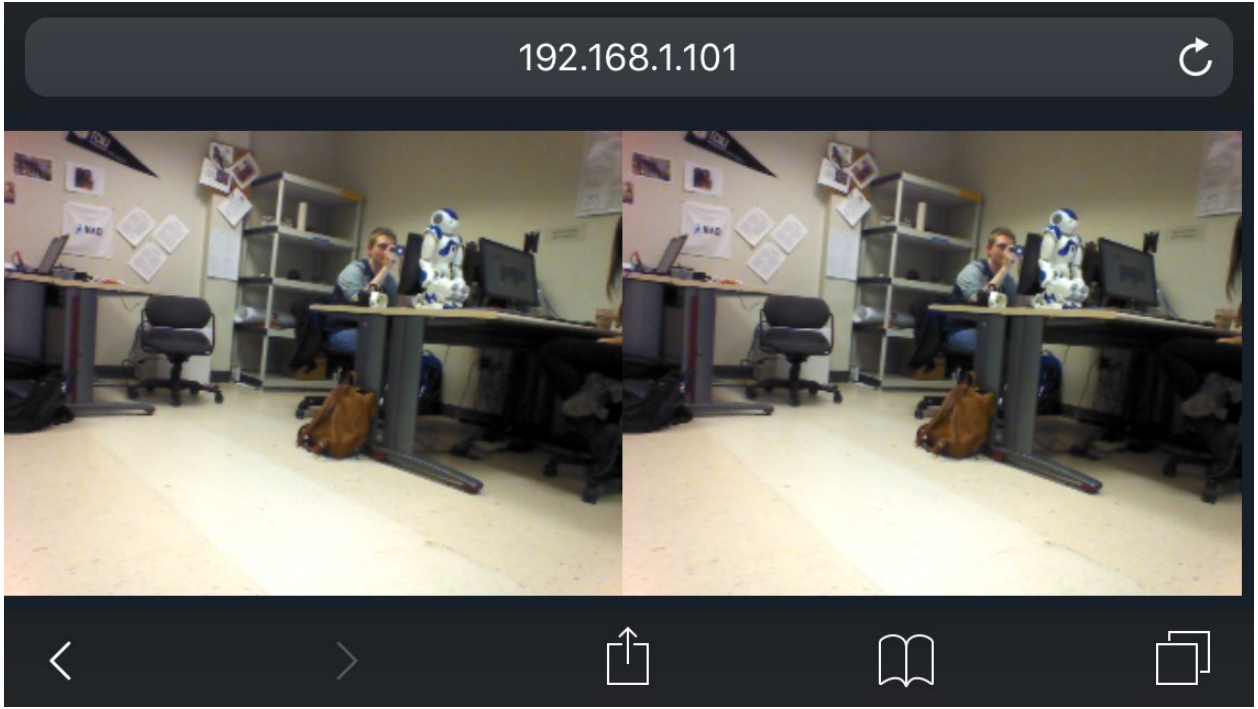
TCNJ THE COLLEGE OF NEW JERSEY

Home Page



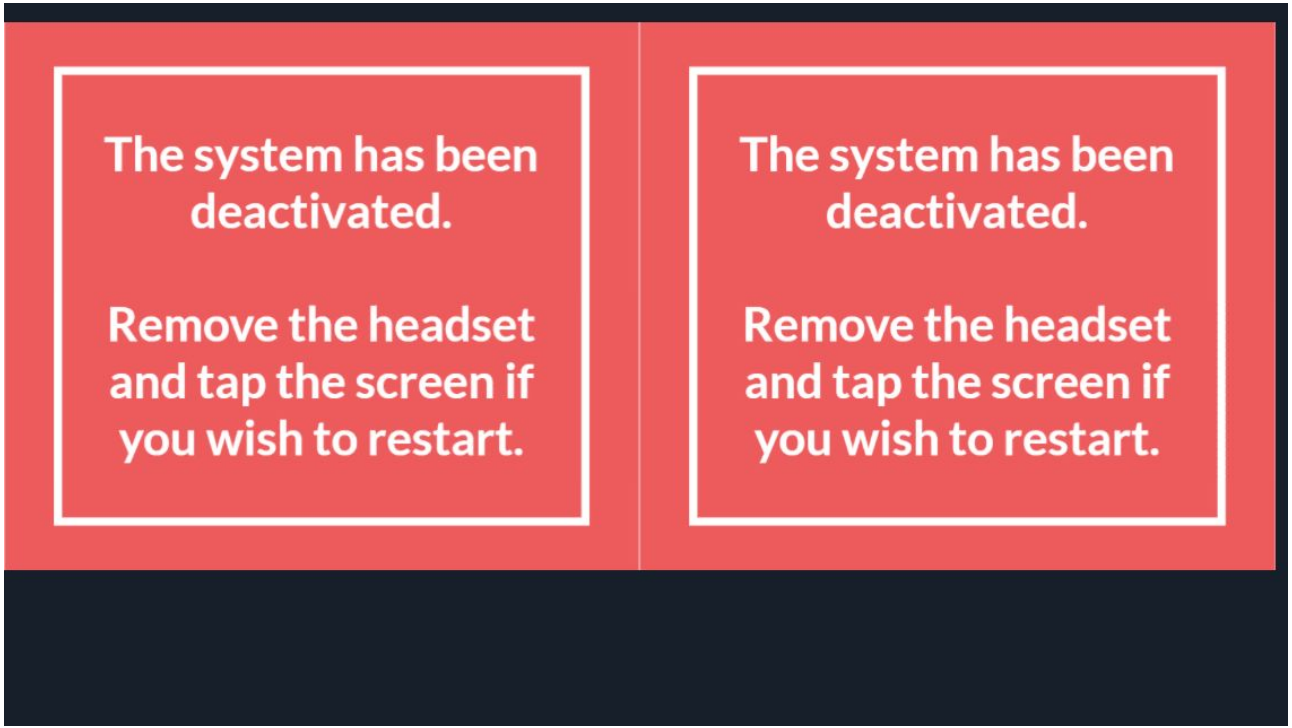
Start Page

Main Page



Main Page
Resolution: 320x240
Latency: ~0.6 seconds

End of System



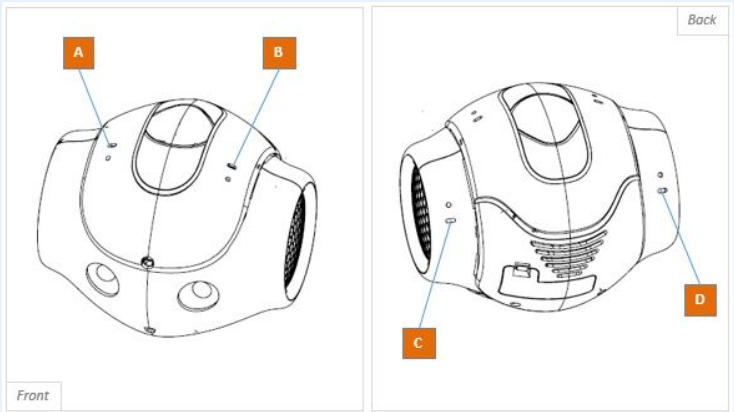
End Screen/Redirection Page

Headset Controller

- Google Cardboard holds phone for video streaming and head motion control
- Headphones project audio stream and allow for speech from user through robot
- Headset is hands-free by adding velcro strap that connects over the head as well as on the sides

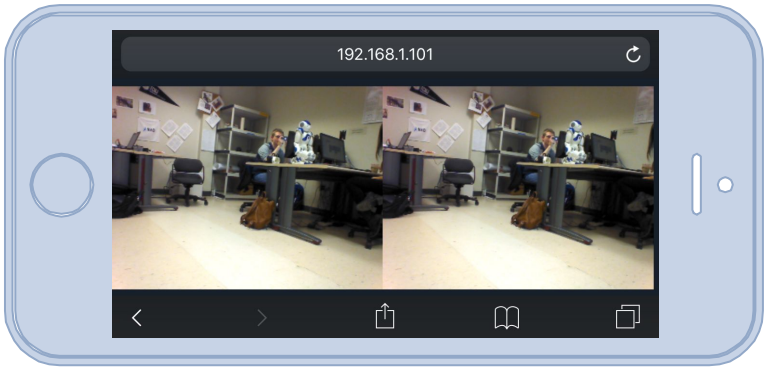
Live Audio Streaming

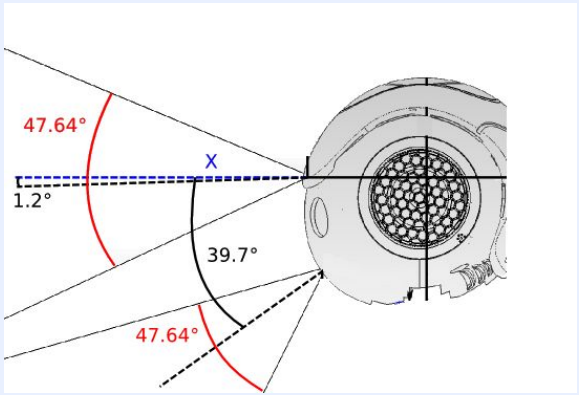
- Subscribe to NAO microphones using NAOqi SDK
- Filter and reshape received data to reduce noise
- Play sound in chunks using a deque
- Used a bandpass filter



Live Video Streaming

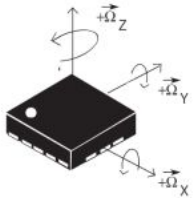
- Capture NAO camera frames and project stereoscopic image to headset with JavaScript
- Resolution 320 × 240
- Framerate 2 fps



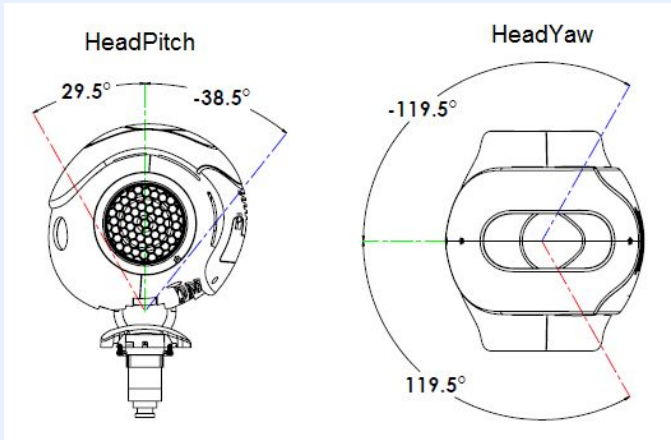


Gyroscope Data on Mobile Device

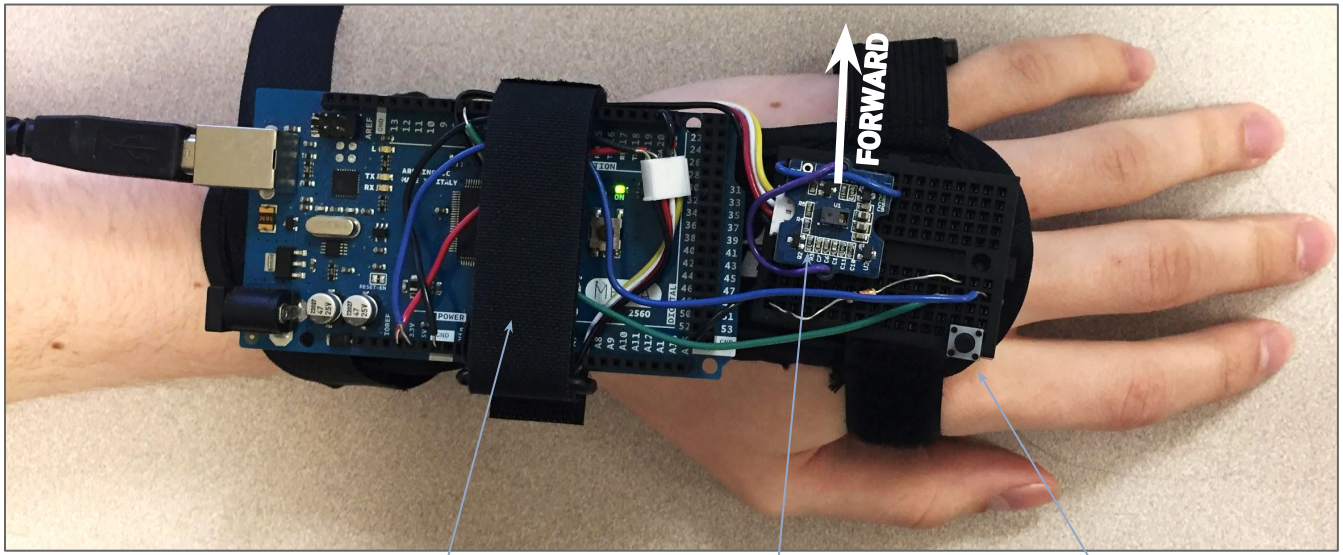
- Receive and process data from phone's internal gyroscope using JavaScript
- Initializes an origin at startup
- Constrain gyroscopic data to NAO's allowable range of movement



Gyroscope



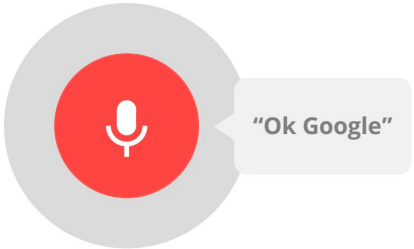
Arm Controller



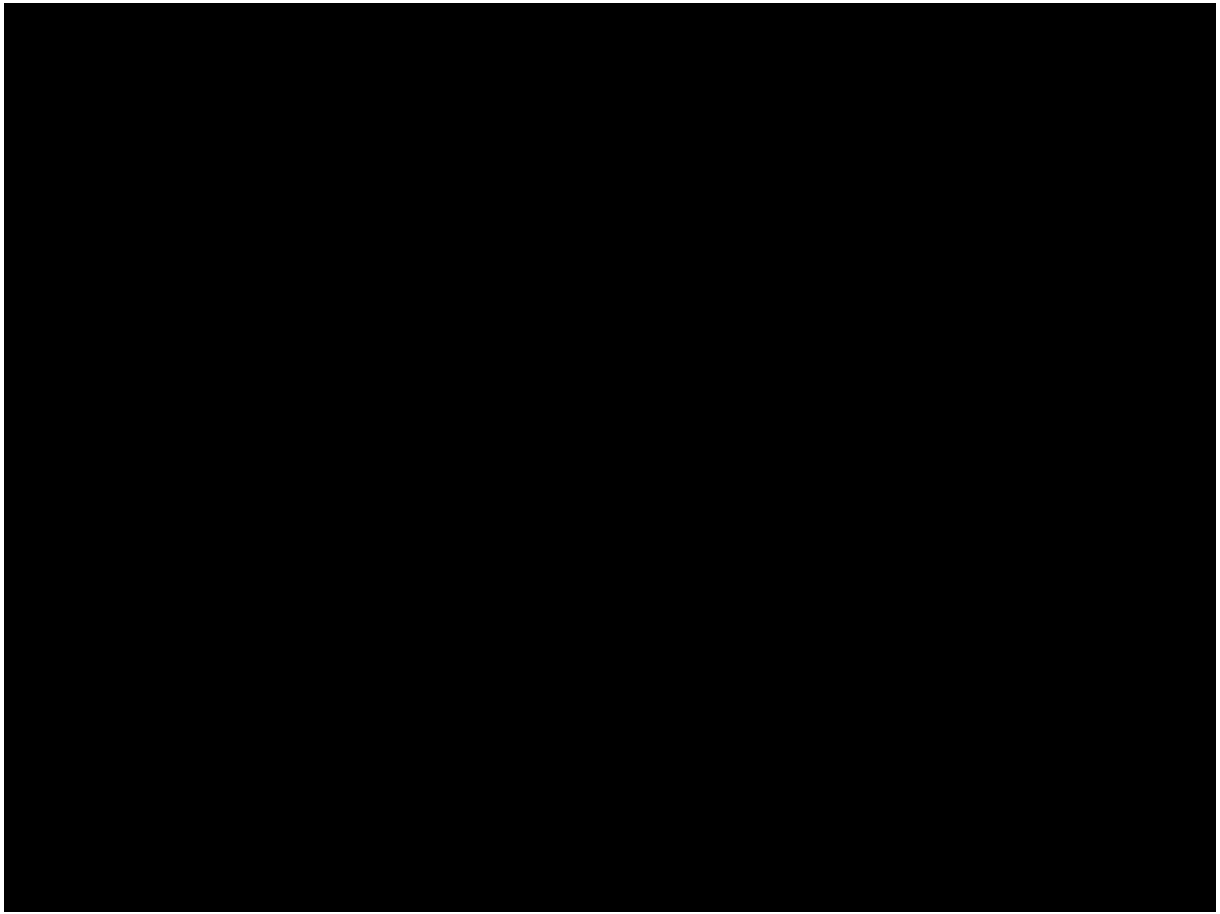
- Serial communication between Arm Controller and Web Application
- Gesture sensor provides commands for walking forward, backwards, left, right, turning, and stopping
- Press for robot text-to-speech
  - Command for “System Deactivate”

Text-to-Speech with Google Speech Recognition API

- Text-to-Speech prompted by button press
- Python script converts speech-to-text utilizing Google Speech Recognition API
- Processing speed varies depending on length of speech
- Error message notifies robot's audience if Google Speech Recognition cannot accurately distinguish human speech
  - "I am trying to say something, let me try again"
- Command for "Deactivate System"



Demonstration



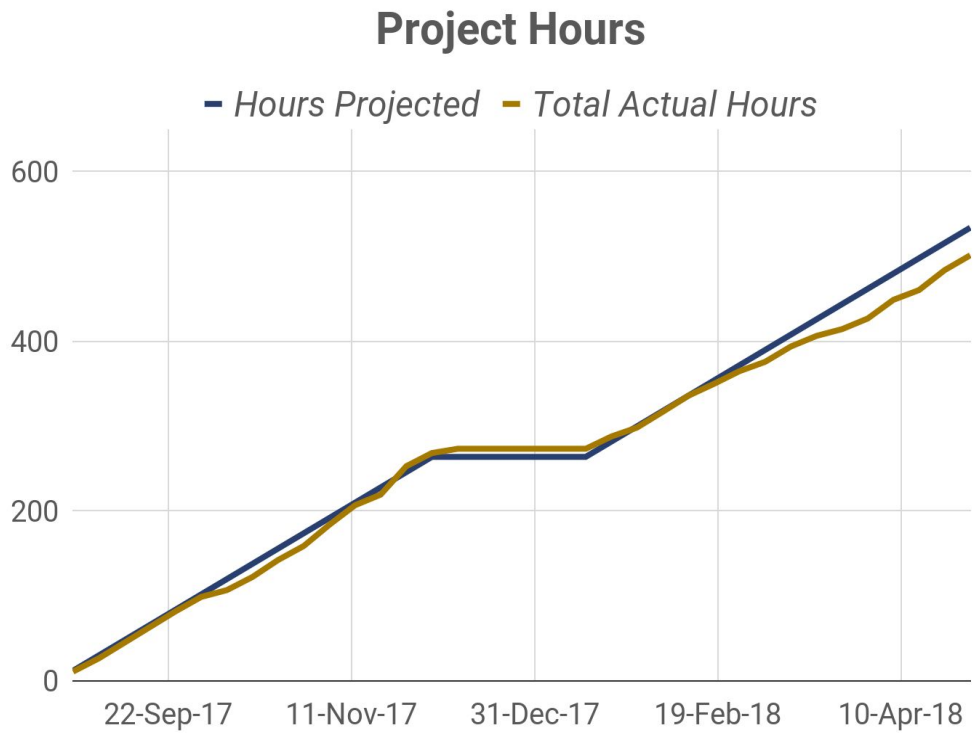
Budget

Total Budget: $300

Total Spent: $280.92

Item	Quantity	Cost per Quantity ($)	Total Cost ($)
Gaming Headset	1	$ 128.46	$ 128.46
Google Cardboard	2	$ 19.99	$ 39.98
Mic #1 (USB)	1	$ 21.99	$ 21.99
Mic # 2 (KIMAFUN)	1	$ 49.99	$ 49.99
Headstraps (2 pack)	1	$ 5.99	$ 5.99
Headstrap w/ Overhead Piece	1	$ 5.99	$ 5.99
Female to Female Connectors (6 pack)	1	$ 6.99	$ 6.99
NAO Robot	1	$9,500.00	N/A
Miscellaneous (extra parts, shipping, etc.)	N/A	$8.63	$8.63

Total Project Hours



Conclusion

- The team successfully developed a telepresence system
  - Incorporated and optimized arm controller project to include in system
- Specifications were met overall
- Tested with several users
- Operator is able to project their voice and physical presence to a remote location, and receive sensory feedback on their environment to capture an **immersive experience**
- Audience can be **more receptive** to a speaker with a physical presence

Questions?