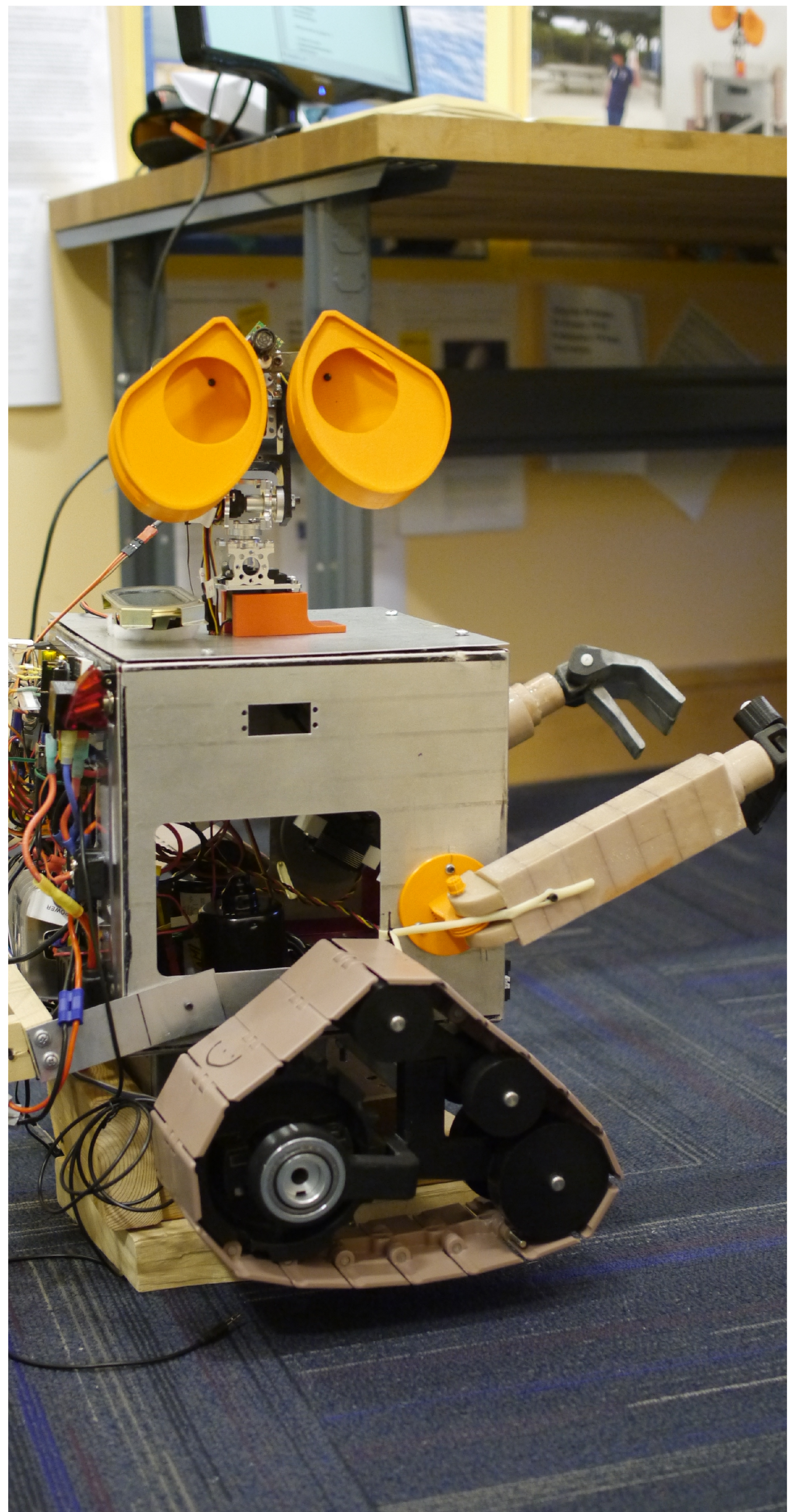


# Interactive Gesture Control of a Robot: The WALL-E Project

## Integrating mobile robotics and computer vision to build a robot capable of interactive gesture control

Computational Robotics, ENGR 3599

Presented by Adela Wee ('14) and Sophie Li ('17)

## ABSTRACT

As the price of integrated sensor and computing packages drops, the availability of robotic technologies is quickly increasing. The challenges of developing robotic technologies today is heavily focused on developing behaviors and sensors to make robots more compatible with humans. Our expressive mobile robot based off the robot from the Disney-Pixar movie of the same name, *WALL-E*, was designed with the goal of demonstrating more natural forms of communication through the form of gesture control. This makes communication intuitive without putting the burden on the human to learn specific commands through traditional implements like a keyboard or a mouse.

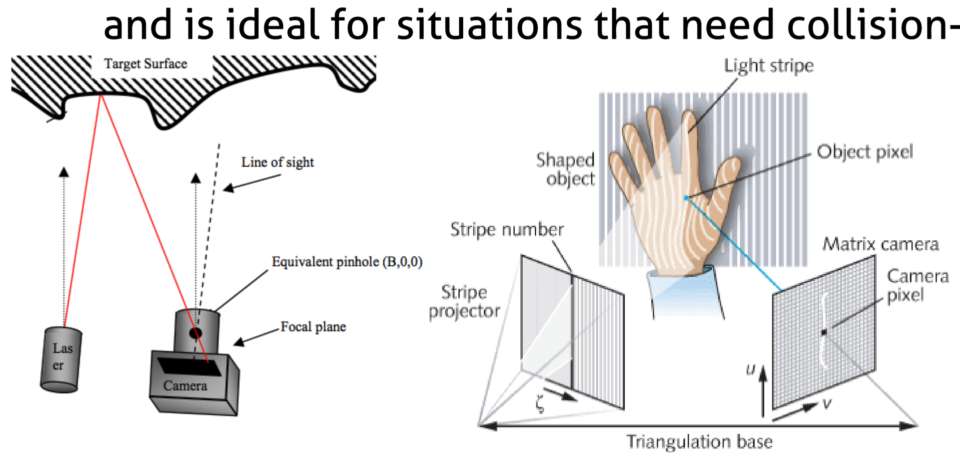
## INTRODUCTION

In order for robots to become more integrated into our daily lives, we must be able to have more natural interactions with them. More natural forms of communication involve both explicit and implicit communication. Explicit communication is limited to specific environments, and implicit communication can add clairty, contextual input, and efficiency. If robots were able to partner with humans as effectively as dogs do with their handlers, the social and technical roadblocks to having human-robot teaming and collaboration would be hugely reduced.

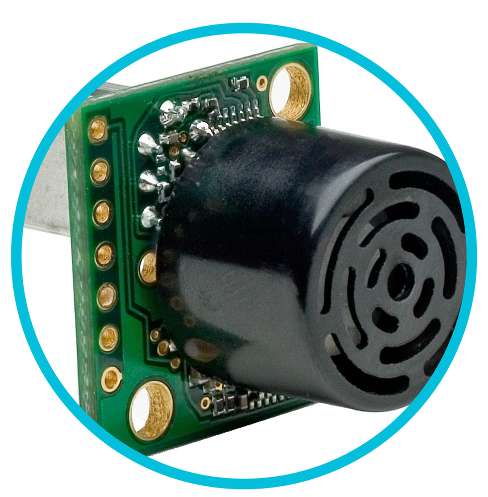
We all have a predisposition towards nonverbal cues, as we can easily intuit others emotional states through discussions and personal interactions. Gestures also replace speech when we cannot or do not want to verbalize our thoughts and help regulate the back-and-forth flow of speech in a conversation. Robots will need to be able to interact with a lay audience which will not necessarily be used to interacting with a robot through a computer and a screen, but are generally more familiar with cues such as pointing in a specific direction or waving.

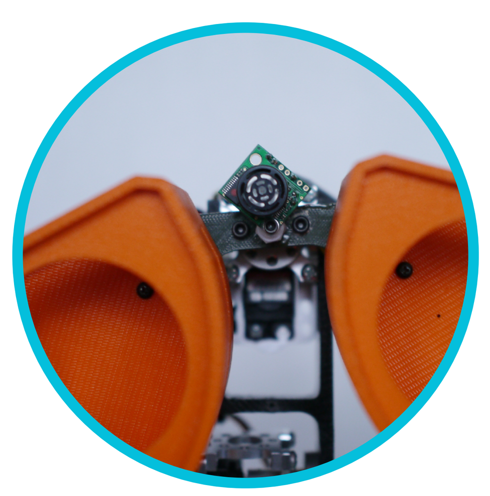
## SENSING CAPABILITIES

- Microsoft *Kinect* for *XBox360* : RGB-D camera that uses structured light with an infrared laser and cameras to determine how far away objects are
- Ultrasonic Sonar: This sonar can accurately sense obstacles in front of it and is ideal for situations that need collision-avoidance behaviors.




The *Kinect* uses structured light-- a camera and laser diffraction to measure the distances of objects it sees in the field of view. It correlates this to the RGB image to build 3D models. The *Kinect's* range is roughly 4-20 ft.





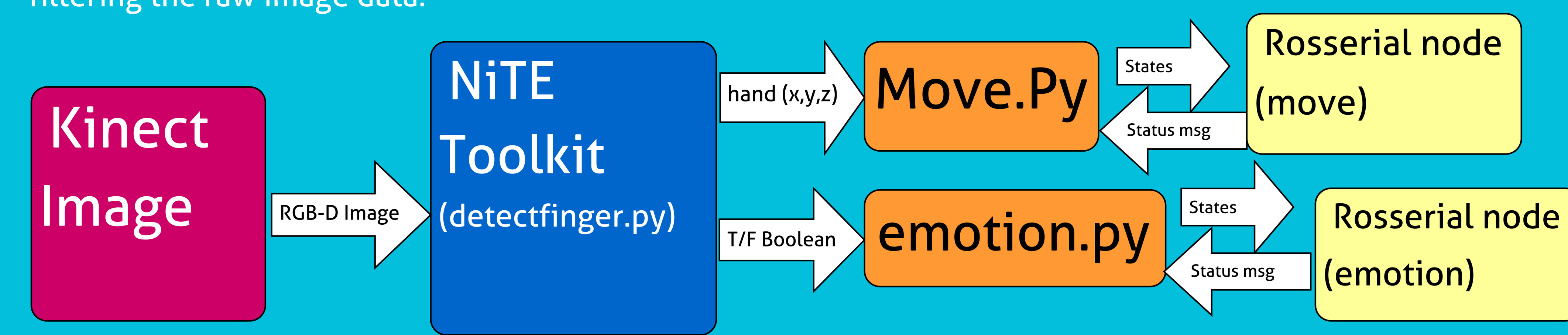
An ultrasonic sensor helps build WALL-E's awareness of the environment immediately in front of the robot to ensure that the robot doesn't collide with anyone.

For future iterations, we would have the head sweep from side to side to develop better obstacle avoidance behaviors.

## SOFTWARE

The software utilized in detecting human-like figures has become even more advanced with better machine learning algorithms and filters, resulting in facial feature tracking, gesture tracking, and 3D mapping of indoor spaces. It is now possible to quantitatively measure human body language, as other researchers have done to construct things like interpersonal trust in social interactions.

Our software is based off the Open-source OpenNI NiTE toolkit, which allowed us to focus on higher-level problems such as imbuing personality and figuring out what to do with the processed data as opposed to filtering the raw image data.




## CONCLUSION

We have demonstrated that it is feasible to interface with the Microsoft *Kinect* through OpenNI and NiTE to effectively control a mobile robot. In accomplishing this project, we wrote a custom makefile that successfully compiles all OpenNI and NiTE dependencies. Our work can be used as a foundation in future projects that want to take advantage of the completed gesture recognition capabilities in the Kinect. WALL-E, as a mobile platform built completely in ROS, successfully demonstrates integration with high level vision algorithms, ROS, and hardware serial devices. Our work provides a foundation for future development of interactive robots using a toolset at Olin students are familiar with (ROS, C++ and Python).

## FUTURE WORK

In the span of 4 weeks, we weren't able to fully exploit NiTE's capabilities for gesture recognition and feature detection. Several modules like scene analysis, push/pull, and circle drawing detection were left untouched. If combinations of these modules are included in future iterations, we could greatly increase the range of gestures WALL-E could respond to. Additionally, all processing is currently done on a laptop that WALL-E must carry around. A future version could take advantage of the size and power of modern embedded computers. Though this would require additional research into interfacing with an OS that's not as transparent as Ubuntu 14.04 , the use of an embedded computer would greatly increase WALL-E's mobility.

## ACKNOWLEDGEMENTS

Many thanks to Paul Ruvolo, for his magical C++ and ROS skills, Victoria Coleman and Victoria Preston for their debugging help, Heather Boortz for ROS support, and Drew Bennett for giving us lab space and equipment. Thanks for helping make this vision a reality!

OLIN INTELLIGENT VEHICLES LAB

Olin College of Engineering