BUILDING AN INTERACTIVE ROBOT: CHANGING THE WAY ROBOTS INTERPRET AND REACT TO OUR ACTIONS

Adela H. Wee[*] and Andrew A. Bennett[†]
Olin College, 781-292-2522, adela.wee@students.olin.edu, www.olin.edu

In order for mobile robots to become more integrated into the daily life of average citizens, non-specialists must be able to communicate with them via a more natural interface. Through technologies like the Microsoft *Kinect*, tracking body movements and facial expressions is now possible with inexpensive and much smaller hardware. This project combines an expressive mobile robot with *Kinect* perception technology to facilitate interaction with a variety of people who are not trained roboticists. The behaviors developed could then form the foundation for many practical applications ranging from household use to military missions.

INTRODUCTION

The envisioned widespread adoption of mobile robots by the public requires a fundamental change in human-robot interaction. Currently, programming, teaching, or even operating a robot is difficult and requires extensive user training. An operator commands a teleoperated robot based on video feedback or via line of sight using a keyboard and mouse or a set of buttons and joysticks. Not only is this approach fairly cumbersome, but is not always practical even for trained professionals, especially in military applications or other high-stress environments. Further, acceptance by the public requires trust between human and robot,[‡] which can be significantly enhanced if the former can intuit the latter's behavior and is confident that his or her instructions have been clearly understood.

Although many robots today are controlled through some sort of wireless controller like that used on Microsoft's *Xbox 360* or a custom-designed joystick pad, imagine a world where robots were operating all around us. We have all found universal TV remotes are not necessarily easy to use or even compatible with the hardware we have. A similar problem exists in the world of mobile robots, where different payloads and manipulators make it difficult for a variety of dissimilar vendor products to be compatible with the same operator control unit. In fact, the Space and Naval Warfare Systems Center Pacific (SSC Pacific) developed the multi-robot operator control unit (MOCU) for this very purpose.[1] If we could control robots using more intuitive communication methods, then working with a robot would be as natural as talking to another human being or a

---

[*] Student Research Assistant, Mechanical Engineering, Olin Intelligent Vehicles Lab, 1000 Olin Way, MB 616, Needham, MA 02492-1200.

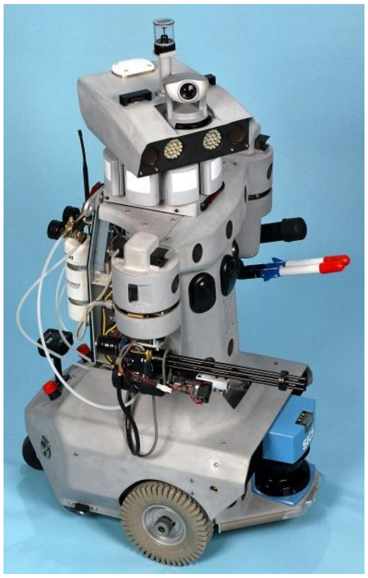
[†] Associate Professor of Mechanical Engineering, Mechanical Engineering, Olin Intelligent Vehicles Lab, 1000 Olin Way, AC 229, Needham, MA 02492-1200.

[‡] One could argue that teleoperated robots inherently trust humans.

pet dog. This would make it easier to incorporate robots into other aspects of our lives and alleviate the need to specifically train users.

A current movement in social robotics research does in fact seek to improve a robot's ability to engage in a more intuitive relationship that includes detection and reaction to human feelings and actions.[2] Feasibility prototypes developed for military applications at SSC Pacific collect state-variable information about their human partners through sensors embedded in the user's clothing, shoes, and equipment.[3] However, this approach does not yet take advantage of nonverbal cues observed through facial expressions and verbal communication. Designed and built under the *Warfighter's Associate* project at SSC Pacific, such robots include *ROBART III*,[3] shown in Figure 1a, which served as a proof-of-concept vehicle for the development of a voice-recognition command-and-control interface. Non-military robots such as Cynthia Breazeal's *Kismet*, shown in 1b, along with others from the MIT humanoid group (and later Media Lab) have been developed to interact with humans with the functional equivalence of a two-year old.[2] *Kismet*, albeit non-mobile, could detect people in close proximity and interact through facial expressions and sounds, and required verbal or visual stimulation to stay engaged. The expressivity of *Kismet*, however, has not yet been combined with as perceptive a ground robot as *ROBART III*.

1a.



1b.

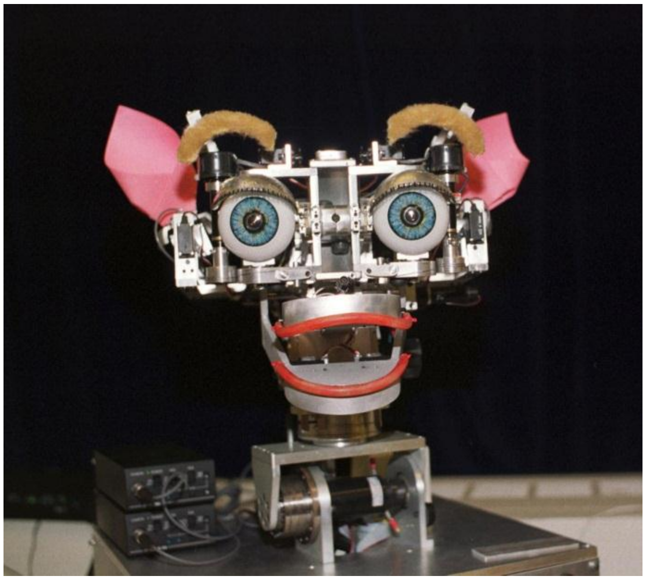


Figure 1a. Among other things, *ROBART III* was used as a proof-of-concept vehicle for a voice-recognition interface under the *Warfighter's Associate* project at SSC Pacific. Figure 1b. *Kismet* was a socially interactive robot built at MIT that could give emotional feedback through facial expressions and sounds, but could not move around the room. Images courtesy of SSC-Pacific Unmanned Systems Group and Plastic Pals.

The number of practical robotic applications increases significantly if we provide a more intuitive means of communication between a human and a robot. It is readily accepted that humans give off subtle nonverbal cues that convey information about various physiological states.[4] This is an indicator that nonverbal communication plays a key role in our day-to-day interaction, and the incorporation of such a capability would have a large impact on practical robotic applications for everyday use. If a robot can consistently detect and interpret human emotion in a reliable way, react to its surrounding environment, and intuitively communicate with humans, many new applications would open up which are currently impractical for teleoperation. As demonstrated under the *Warfighter's Associate* project, an intelligent robot should infer what it should do based on the information it perceives about the environment and the observed actions of any nearby humans.

With regard to perceiving its environment, robots obviously do not have the same rich awareness that we do of the world. Most use a rotating single-beam laser to acquire an accurate but planar (2D) view of their surroundings. An interactive robot immersed in our three-dimensional world, however, requires more robust 3D perception. In the not too distant past, 3D perception was prohibitively expensive in terms of both acquisition cost and power consumption, necessitating numerous tradeoffs with sensor resolution and computational resources. Tracking people and objects is now possible with relatively inexpensive and smaller 3D devices like the Microsoft *Kinect*, albeit with certain performance limitations. With the appropriate algorithms, the *Kinect's* near-infrared ranging and RGB video capabilities can facilitate an improved approach to human-robot interaction through detection of subconscious human feelings in the form of body posture and facial expressions. Unlike other *Kinect*-related research that involves computer-simulated virtual robots responding to body movements and/or facial expressions, this project applies *Kinect* technology to an actual mobile robot operating in a real environment, facilitating a more dynamic relationship as the robot interacts with people who are not trained roboticists.

This robotic platform is also different from most in that it focuses on the coupling of a perception sensor like the *Kinect* with a robot that has expressivity; in other words, a robot that can communicate through facial and body expressions as well as voice feedback, using a set of actuators, computers, and sensors that are much cheaper than those of other research-only robots. This combination allows for a more natural intuitive interface that reduces the control burden imposed upon the human. *Kismet* was primarily a research tool that was computationally expensive, requiring 15 behind-the-scene computers, and was not designed to move around a room to engage with different people.[*] In order to make a sufficiently expressive robot that would be readily accepted without creating negative reactions due to too-humanlike forms (known as the uncanny valley),[5] we decided to base our platform on the character *WALL-E* from the Disney-Pixar movie of the same name. This inspirational robot was specifically chosen because it communicates almost exclusively in nonverbal fashion throughout the movie (as opposed to Star Wars' *C3PO*), and is highly expressive as can be seen in figure 2. Additionally, *WALL-E* is an adorable non-threatening robot in a compact form, and therefore subconsciously appealing to humans.

---

[*] http://web.mit.edu/newsoffice/2001/kismet.html

Figure 2. Disney Pixar's *WALL-E* was the inspiration for our robot, as it communicates almost exclusively through nonverbal cues and expressions. Image courtesy of Pixar Times.[*]

BACKGROUND

Thoughts on robot autonomy

The concept of a teleoperated robot, operated from a distance by a human using some sort of physical controller, is not something new. In fact, teleoperated ground robots have been around since World War I, and during WWII Germany deployed thousands of the *Goliath* explosive-charge carrier robots shown in figure 3.[6] For the most part, however, it is impractical to teleoperate a robot whenever it is much easier to simply do the task yourself, although such a control paradigm often serves a valid purpose by increasing operator standoff from danger in certain military missions like explosive ordnance disposal (EOD). Teleoperation is largely impractical for most other scenarios, however, because of reduced situational awareness and excessive control burden for the operator.

---

[*] http://pixartimes.com/2012/08/03/watch-real-life-wall-e-comes-alive/

Figure 3. The WWII German *Goliath* explosive charge carriers are an early example of teleoperated robots. Image courtesy of Military History Monthly magazine.[*]

Teleoperation is just the low end of a wide spectrum of robotic command-and-control paradigms. As we move toward more human-independent control strategies, robotic behaviors change significantly. The ultimate goal from a military perspective would be a robot that acts in similar and predictable fashion to deployed troops in the field, intelligently carrying out high-level orders. Today's teleoperated robots are typically sent into dangerous scenarios while the humans stay a safe distance away. However, extended missions where a robot works in close proximity to its human partner inherently require a robot to be intelligent, hence the term proximal autonomy.[7] Such robots have the cognitive awareness to react and respond to a given situation,[3] with the ability to carry out an appropriate supporting task. Under this control paradigm, robots continuously evaluate the state of their surrounding environment as well as that of their human counterpart, and then respond accordingly, creating a robust framework for intuitive human-robot teaming.

The success of proximal autonomy involving a more team-oriented relationship between human and robot is heavily dependent on the development of an intuitive interface. With a more natural means of interaction, immersed robots are much better equipped to effectively address real-world problems. One such development attempt has been Honda's *Advanced Step in Innovative Mobility (ASIMO)* robot, which was recently tasked with serving as an interactive tour guide at the Miraikan Science Museum in Tokyo.[†] In this particular application, *ASIMO* doesn't have voice recognition, and can only detect tour groups via six sensors in the ceiling. When people gathered around to take pictures with their smartphones, *ASIMO* was unable to detect the difference between someone raising their hand to ask a question and someone holding a camera to take

---

[*] http://www.military-history.org/articles/back-to-the-drawing-board.htm/attachment/british-soldiers-with-captured-german-goliath-tracked-mines

[†] http://japandailypress.com/hondas-asimo-disappoints-as-robot-museum-guide-fails-to-wow-audiences-0331671

a picture.[*] With fairly robust voice recognition now rather commonplace in the cellphone and automotive industries, it's somewhat surprising such capability has not been integrated into this particular platform. There are only about 100 pre-programmed questions you can ask it, and those must be rather awkwardly submitted through a touchscreen. While *ASIMO* is an impressive example of "innovative mobility" involving a legged bipedal robot that has been under development since 2000,[†] Honda appears somewhat challenged in finding a practical application for the current configuration, unless you count a robot that walks around and helps travelers purchase train tickets a practical application.[‡] In essence, *ASIMO* only solves one part of the problem—practical legged-locomotion for a human-like robot operating in benign environments, and has not yet effectively addressed the bigger goal of being able to naturally interact with humans.

Figure 4. Honda's *Asimo*, Tokyo's Mirakan Science Museum's new tour guide, as it was unveiled to the public in July. *Asimo* cannot differentiate between someone asking a question or raising their smartphone in the air to take a picture—and can only be asked predetermined questions through a touchscreen interface. Image courtesy of Japan Daily Press.[§]

There have been several relatively intelligent robots created for specific applications ranging from security to healthcare. For example, it is clearly technically feasible to build a mobile security robot, and many such platforms have been developed over the past 30 or so years, both prototype and production. The first of these was *ROBART I*, as shown in figure 5a, which in 1981 had onboard sensors to detect intruders, fire, smoke, flooding, etc., could autonomously patrol a household environment, and was able to automatically recharge itself as needed.[8] Soon after, *ROBART II* (figure 5b) was created to provide better assessment capability for a mobile security

---

[*] http://www.usatoday.com/story/driveon/2013/07/06/honda-robot-asimo/2494143/

[†] http://asimo.honda.com/downloads/pdf/honda-asimo-robot-fact-sheet.pdf

[‡] http://www.utsandiego.com/news/2013/jul/04/tp-robot-having-trouble-in-new-job/

[§] http://japandailypress.com/hondas-asimo-disappoints-as-robot-museum-guide-fails-to-wow-audiences-0331671 robot, adding on higher-level algorithms to navigate and create models of its environment, mostly using off-the-shelf electronic parts from RadioShack, with only three component failures in its 14 years of operation.[9][*] The component technologies developed on *ROBART II* were later applied to warehouse security, where Cybermotion *SR2* robots, shown in figure 5c, served as intermediaries between security personnel at a remote guard shack and fixed-place motion sensors mounted to structural elements of the building. Such lower-cost fixed sensors are very prone to nuisance alarms, so the robots would investigate a suspected disturbance and then advise the human guard, who otherwise would have to go check each incident in person. As roving patrols, these fully autonomous vehicles did not require much human interaction or direction in order to carry out their duties.[10]

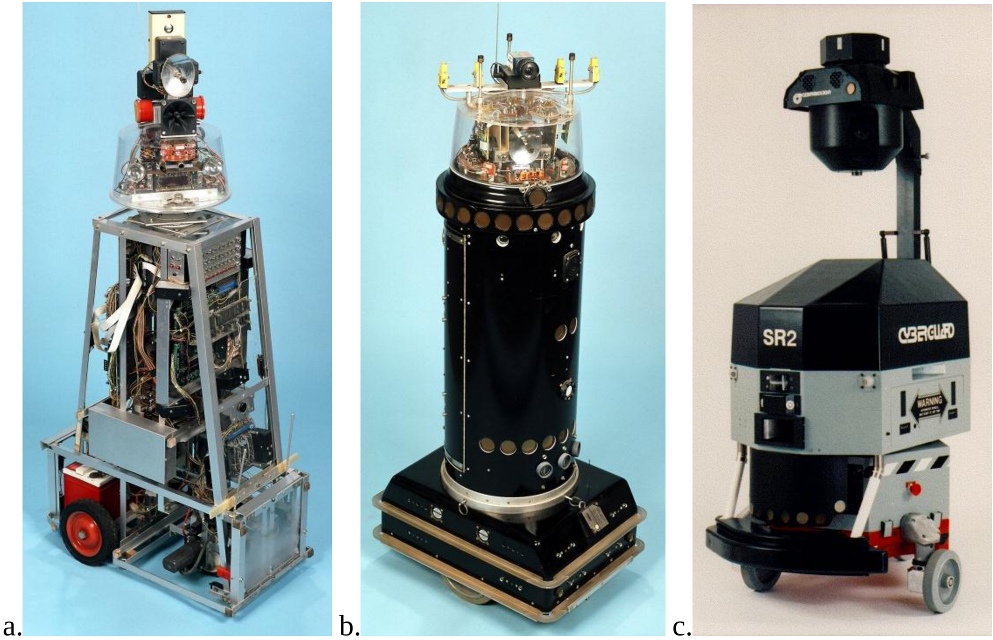


Figure 5a. *ROBART I* was an early prototype security robot that had onboard sensors to detect intruders, fire, smoke, and flooding. Fully autonomous, it could patrol a household environment and recharge itself without human assistance. Figure 5b. *ROBART II* was the second-generation security robot that had an improved sensor suite with more elaborate intelligence, planning, and navigational capabilities. Figure 5c. The Cybermotion *SR2* was a production security robot deployed for warehouse patrol. Images courtesy of SSC Pacific Unmanned Systems Group.

Healthcare applications, on the other hand, require a far more sophisticated platform that is aware of the many factors associated with the elderly, and also readily accepted by the people with whom it interacts. This incredibly difficult challenge is being addressed by research groups all over the world, particularly the Japanese.[11] An example of such a platform used for stroke-rehabilitation, designed by the USC Robotics Research Lab under the direction of Maja Mataric,

---

[*] http://www.public.navy.mil/spawar/Pacific/Robotics/pages/ROBART_2.aspx

would ask patients to engage in physical-therapy exercises and track their progress, all seemingly well-received by the patients and physical therapists.[12] As the aging population grows, intelligent interactive robots can help address the need for better healthcare.

Such robots in the future will require increasing levels of human-interaction skills. If we were to give intelligent mobile robots the ability to perceive body language, such that they could detect confusion, curiosity, or even suspicious behaviors while patrolling a public area, it would be much easier to focus effort towards scenarios where we could increase efficiency and safety using more automated systems.

Requirements and challenges presented

Within the realm of autonomous robots there are currently three major challenges: perception, intelligence, and natural interaction. In a recent *MIT Technology Review*,[13] John Leonard of MIT states, “I see how hard it is to do anything with robots…People and robots working together can happen much more quickly than robots simply replacing humans.” Even after many robots were introduced into factory automation in the 1980s, most systems are still not intelligent or advanced enough to consistently recognize everyday objects. Obtaining the necessary high-resolution data traditionally has required expensive instrumentation such as scanning laser rangefinders, or significantly increased processing power if cheaper cameras were employed for stereo ranging.

In order to track moving humans, either a 360-degree sonar array like that on *ROBART II* or RGB cameras with tracking algorithms like those on *ROBART III* were typically implemented on early systems. Neither option provided any detailed information regarding facial expression or body language, but instead gave only give a rough position of perceived human presence relative to the robot.[3] Newer developments like the *Kinect* and other combined depth (i.e., near-infrared ranging) and RGB camera systems allow for the application of human-tracking algorithms developed for the gaming industry, which provide relatively low-cost high-resolution data (both images stream at 640 x 480 at anywhere from 12-30 fps, depending on processor speed).

Onboard intelligence has also come a long way since the old EPROM chips used to store *ROBART II’s* code. Not only have we developed sophisticated software architectures that allow robots to think reflexively as well as actively, but we have also shown that these robots are capable of practical applications in the field, as evidenced by the numerous projects and achievements of the SSC Pacific Unmanned Systems Group over the past several years.[*] A form of more natural and intuitive interaction, however, is generally missing from the field of mobile robots. Although many universities have departments focused on artificial intelligence, very few have made robust and reliable robotic systems that are able to move around the lab—in fact, many of these robots are designed to be mounted to a fixed platform and be controlled by multiple processors behind a curtain, just as *Kismet* was. Compared to moving around a lab, moving around the real world is even more challenging.

In order to integrate robots into our lives, we need to better detect subconscious human feelings given off through body posture and facial expressions. The *Kinect*, being low cost and able to track motion without any markers or worn devices, has spurred research in the tracking of nonverbal behaviors and actions. Researchers at USC are focusing on using the *Kinect* depth sensor to unobtrusively measure breathing rates and “leg jiggling” or fidgeting.[4] Additionally, the *Kinect* is being used for modeling nonverbal trust behaviors at MIT, where researchers have found that a

---

[*] http://www.public.navy.mil/spawar/Pacific/Robotics/Pages/default.aspx

set of four cues (leaning-back, face-touching, crossing-arms, and hand-touching) were usually predictive of a distrusting behavior, while another set of three cues (leaning-forward, arms-in-lap, and having an open-arms pose) were usually indicative of higher levels of trust. The MIT effort attempted to automate the process of detecting these cues through 3D motion-capture and gesture recognition algorithms with moderate success.[14] With the development the *Kinect* and other RGB-D cameras, methods of natural and intuitive interaction through recognition of subconscious behaviors are on their way to becoming an eventual reality.

Related military research developments

In the mid-2000s, a few feasibility prototypes were built to specifically demonstrate innovative control strategies for significantly improved human-robot teaming. Under the *Birddog Warfighter Sensor System*, Science Applications International Corporation (SAIC) equipped a surrogate air-soft gun with an inertial measurement unit (IMU), and added a sensor on the weapon's safety to collect information about the state of the human, which along with the voice commands, was then passed to a *Segway Robotic Mobility Platform (RMP)*.[15] The objective was to aid the robot in "reasoning about the actions and intentions of humans based on their location and weapon status."[15] Location information for both human and robot was obtained through conventional GPS.

A similar concept had been explored by SSC Pacific that initially focused on voice commands spoken through a noise-cancelling microphone.[3] Through the Center for Commercialization of Advanced Technology (CCAT), SSC Pacific consequently funded further development of a second-generation of the instrumented SAIC weapon. Also under CCAT, SSC Pacific funded the University of Michigan to deliver an improved personal-odometry system consisting of an inertial measurement unit (IMU) embedded in the heel of a standard military boot that allowed the robot to track the human's location if it lost line of sight.[15] The state-variable information collected by both the boot and the weapon were passed to an iRobot *All Terrain Robotic Vehicle (ATRV)*, significantly reducing the dependence on voice control. For example, the robot could execute a following behavior while its pan-tilt-zoom camera automatically tracked the movements of the human's weapon. If the human switched off the weapon safety, the robot would then place itself as a defensive screen between the Warfighter and the perceived threat.[15] These pioneering efforts were aimed at furthering a better understanding of intuitive human-robot interaction, with the ultimate goal of achieving some degree of "artificial empathy," where the robot continuously gains insight into the human's current state and intentions.[*]

Olin College also had a similar research endeavor in 2011 with Scientific Systems Company, Inc (SSCI) and Spatio Systems LLC under a Small Business Technology Transfer (STTR) sponsored by the Office of Naval Research (ONR). The project used an autonomous John Deer *Gator XUV* and focused on implementing easy-to-use control, based on voice and gesture input. The vehicle demonstrated a real-world application of Direction Understanding in Naturalistic Environments (DUNE), a natural language understanding system, which allowed it to correctly recognize and understand specific routes and locations.[†]

---

[*] Everett, H.R., personal conversation January 2013. Technical Director of Unmanned Systems Group, SPAWAR-SSC Pacific, San Diego, CA

[†] http://www.youtube.com/watch?v=JCF7bRnT5vc

Related human-robot interaction research

The field of social robotics, which emerged in the early 2000s, is primarily concerned with human perception of robots within society, and finding ways to better integrate them into our lives. One of the first examples built was Cynthia Breazeal's *Kismet*, primarily designed to engage people in natural and expressive face-to-face interaction,[2] which has been on display in the MIT Museum since 2002.[*] The *Kismet* experience showed that robots could be expressive, that we could impose personalities and a character onto them, and that children even perceived *Kismet* like a baby and were subconsciously driven to "parent" it, in a way.[16]

Another notable lab is the Robotics, Health, and Communication Lab run by Laurel Riek at Notre Dame. One of her research pursuits is in social-context learning, where machine learning is applied to social scenes to predict the actions within them, with the eventual goal of using the results to increase the intelligence of robots in social contexts.[†] We can learn a lot from the work in this field, particularly when it concerns emotional response and feedback, and how robots are perceived in social contexts.

The *WALL-E* contribution

Our *WALL-E* is an attempt at combining and augmenting all the relevant component technologies to achieve a practical solution for proximal human-robot interaction. If we're going to have realistic acceptance and usage of robots working directly with humans, we have to build more intelligent systems that not only understand what we're subconsciously signaling, but also comprehend our vocal commands and physical movements. In other words, augmenting localization and collision-avoidance intelligence developed for the mobile-robot world with the social-interaction capabilities recently introduced in the last 5-10 years. Mataric and Breazeal have developed platforms that interact and respond to cues in the lab. Everett and others at SSC Pacific have made mobile robots that are beginning to fill the role of assisting the Warfighter in the field. *WALL-E* is going to make teamwork and non-verbal communication an effective solution to controlling an expressive robot that will be able to perceive social cues and understand commands while autonomously operating within a building, and eventually outdoors.

---

[*] http://www.plasticpals.com/?p=30191

[†] http://www3.nd.edu/~rhclab/projects.html

TECHNICAL DETAILS

This section pertains to the technical aspects of *Wall-E's* construction to date (figure 6). Low cost and ease of repair and maintenance were significant factors in the current design.

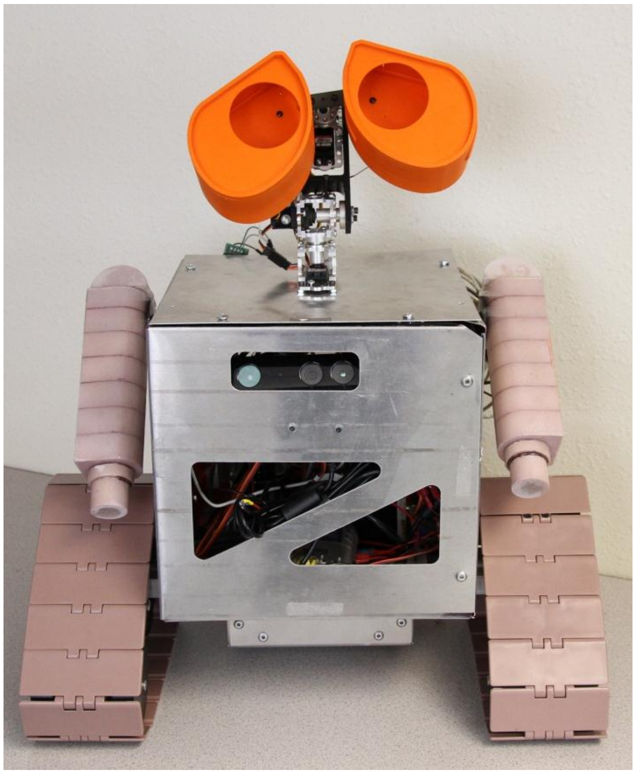


Figure 6. *WALL-E*, as of 8 July 2013, assembled with arms, tracks, head, and *Kinect*, along with working electronics.

Overview of *WALL-E's* hardware

Microsoft *Kinect*—The *Kinect* provides reasonably advanced machine-vision capabilities in a small, portable, easily interfaced, and relatively inexpensive package. It has two cameras, one is a 3D near-infrared range sensor that uses reflected beams to detect where objects are in the field of view, with a resolution of 640 x 480, and the other camera is RGB video with a resolution of 640 x 480. Through these two simplistic sensors it is possible to build a visual 3D map of an interior space. Four built-in microphones calculate the angle from which sound is coming within the forward hemisphere. This system was primarily chosen for use on *WALL-E* because there have been several algorithms written that in theory should make interaction with humans easier than previously possible.

Pololu *Maestro 12 Servo Controller*—Small 12-channel servo controller that allows for ease of animation using inexpensive hobby servos controlled via USB.

Dimension Engineering *Sabertooth 25A Dual Motor Controller*—Compact and powerful motor driver typically used for ground vehicles. It drives up to two motors at 12V 25A each.

Dimension Engineering *Kangeroo x2*—Controller that interfaces to the *Sabertooth* and allows for motion feedback control. It's very versatile and can accept input from limit switches, phase-quadrature shaft encoders, etc.

AM Equipment (AME) *218-series 12V 212-in-lb left-hand and right-hand gear motors*—used typically in cars as wiper motors, these are incredibly high torque and fairly compact—which is great if you do not want to construct your own gearbox and/or do not have space to put in a conventional spur gearbox.

Tank-tread material—typically known as table-top conveyor chain, it's used for transporting and putting on caps on bottles and in carton transport areas of a factory. Fairly easy to take apart and reassemble—just need to push out the pin that connects the links together.

Body construction—most other mechanical parts consist of 3D-printed ABS plastic and 5052 0.005" aluminum sheet. The metal, although thin, is extremely rigid when folded into a box-like shape, and the ABS is strong enough to take the load of being the frame for the drivetrain. All parts can be quickly disassembled for quick and easy transportation.

Arm construction—the arms are currently actuated with only one rotational degree of freedom at the shoulder. Designed to be powered by hobby servos (maximum 300 in-oz of torque), the lightweight composite arms were constructed by laying fiberglass and epoxy over rigid insulation foam machined using a waterjet cutter.

Discussion of software framework used

The current software architecture is based on the *Olin Robotics Software Architecture,*[*] which uses a *Sense-Think-Act* paradigm. Modeled after the human brain, there is a forebrain, midbrain, and hindbrain, in which high-, mid-, and low-level algorithms are processed. It answers "Where am I?," "What is around me?," and "How am I?" Since the robot is in a fairly early stage of development where intuitive interaction is valued over obstacle avoidance, it has not been possible to implement the usual software architecture, so we have a midbrain that processes sensor data from the *Kinect* and filters it to determine what commands it should send.

In the future this may change, as we are considering conversion to a combination of Willow Garage's *Robot Operating System (ROS)* and National Instrument's *LabVIEW*, dedicating one processor to run *ROS* and open-source *Kinect* drivers and another to run all the motor drivers and other sensor inputs (such as actuate other cameras or collect additional sensor data). As more sensor inputs and more arbitration is built in, we may also adopt a reflexive architecture.[9]

Expressivity of the robot itself

All of these components mean nothing if the robot cannot carry out its fundamental purpose in this proposed research—to interact with people. The head has servos to pan and tilt the head proper, actuate both eyebrows, pivot each eye individually, as well as tilt both eyes from side-to-side. Additionally, each arm can articulate at the shoulder. This allows *WALL-E* to express emotions such as sad, confused, surprised, and concerned as shown in figures 2a, b, c, and d, respectively. The precise timing of the animation from key frame to key frame changes the expression

---

[*] Barrett, David, personal conversation August 2011. Professor of Mechanical Engineering and Design at Olin College, Needham, MA.

perceived by the user, a phenomenon well-documented by animators.[17] This is a quality that sets *WALL-E* apart from most application-based robots.

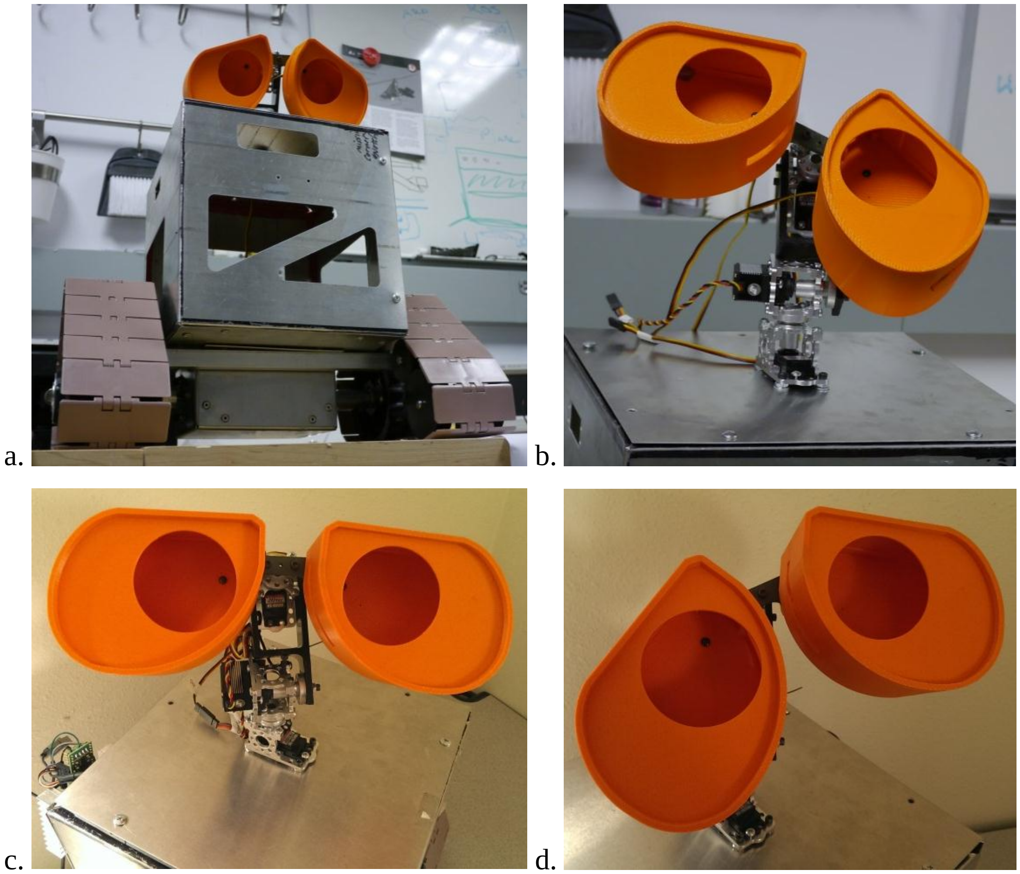


Figure 7. The expressivity of our *WALL-E* clone, achieved through the seven degree-of-freedom head, two arms, and actuated tracks. The perceived expression by the human changes based on the acceleration of the servos to and from various poses. Figures 7a, b, c, and d correspond to sad, confused, surprised, and concerned, respectively.

CONCLUSION

The three most significant technology areas involved are perception, intelligence, and natural interaction. Most organizations focus on some subset, but typically not all. By way of example, Honda picked “Innovative Mobility” and the challenges associated with bipedal movement, Breazeal devoted her efforts to social interaction and understanding the barriers to communicating with robots, and SPAWAR focused on collecting useful state variables describing not only the environment but also the human to make a more intelligent vehicle that could achieve some degree of “artificial empathy,” making it less dependent upon voice commands. In order to achieve a working prototype that we can further develop, we need to address all of these areas to achieve success. The CCAT program demonstrated that much of the supporting technology pertaining to each of those specialty areas is already out there, and that it is possible to harvest the developments of others in the community as a whole in order to make more rapid progress. Over the past 6 months we have developed an 11-degree-of-freedom research platform that can communicate through facial expression and interpret human movements based on image data from the Microsoft *Kinect.* We hope that our efforts at Olin College with *WALL-E* and our other autonomous vehicles will become known for integrating a large subset of the technologies and improvements made by other research organizations to conduct research and experiments on subconscious behaviors and responses from users.

As robots begin to play a larger role in our lives, it will be necessary to better understand the implications and needs of an interactive mobile robot that communicates naturally. It seems that a future involving robots working with humans in teams is inevitable; the applications are everywhere and will continue to increase, especially in our fast-paced digital world where technology and social media are already intertwined in our lives. The field of robotics is a disruptive technology—as computers and other hardware get smaller and more easily affordable, the challenges of building a mobile robot focus more upon systems integration and social interactions. Robots have already changed our lives through the cars we buy and the products we eat—who says they can't change them even more through intelligent and expressive interaction?

FUTURE WORK

Looking ahead, we plan to enhance *WALL-E's* capabilities by integrating natural language interaction using onboard *Kinect* microphones, adding on additional sensors for better real-world feedback and obstacle avoidance, as well as making it *ROS* compatible.

ACKNOWLEDGEMENTS

Olin College Intelligent Vehicles Lab, especially Dave Barrett, Garland O'Connell, Lesley Yu

SSC Pacific Unmanned Systems Group, especially Bart Everett

REFERENCES

[1] D.N. Powell, G. Gilbreath, M.H. Bruch, "Multi-robot operator control unit," SPIE Proc. 6230: *Unmanned Systems Technology VIII*, Defense and Security Syposium, Orlando, FL, 2006.

[2] C.L. Breazeal, "Designing Sociable Robots," Massachusetts Institute of Technology, Boston, Massachusetts, pp 15-25, 2002.

[3] H.R. Everett, et all, "Towards a Warfighter's Associate: Eliminating the Operator Control Unit," SPIE Proc. 5609:*Mobile Robots XVII*, Philadelphia, PA, 2004.

[4] N. Burba, et all. "Unobtrusive Measurement of Subtle Nonverbal Behaviors with the Microsoft Kinect," IEEE VR Workshop on Ambient Information Technologies, proceedings, pp 10-13, 2012.

[5] M. Saerbeck, A.J.N. van Breemen, "Design Guidelines and tools for creating believable motion for personal robots," 16th IEEE International Symposium on Robot & Human Interactive Communication, Korea, Proceedings, pp 386-391, 2007.

[6] H.R. Everett. *Unmanned Systems of World Wars I & II.* Draft awaiting publishing. July 8, 2013.

[7] H.R. Everett. "Evolving Paradigms of Human-Robot Interaction," poster, SD 932, SSC Pacific, San Diego, CA. October 2010.

[8] H.R. Everett. "ROBART I: In Retrospect." Technical Document 3199, Space and Naval Warfare Systems Center, San Diego, CA, December 2005.

[9] H.R. Everett, G.A. Gilbreath, and T.T. Tran, “Modeling the Environment of a Mobile Security Robot,” NOSC Technical Document 1835, Naval Oceans Systems Center, 1990.

[10] R.T. Laird, H.R. Everett, et all. “Early User Appraisal of the MDARS Interior Robot.” American Nuclear Society, 8th Topical International Meeting on Robotics & Remote Systems, Pittsburgh, PA, 1999.

[11] E. Broadbent, R. Stafford, and B. MacDonald. “Acceptance of Healthcare Robots for the Older Population: Review and Future Directions.” *International Journal of Social Robotics.* Vol. 1, Issue 4, 2009, pp. 319-330.

[12] J. Eriksson, M. Mataric, and C.J. Winstein. “Hands-off Assistive Robotics for Post-Stroke Arm Rehabilitation.” *Journal of Neuroengineering and Rehabilitation.* Vol 4, No. 5, 2007.

[13] D. Rotman. “How Technology is Destroying Jobs.”*MIT Technology Review.* Vol. 116, No. 4, 2013

[14] J.J. Lee, B. Knox, C. Breazeal. “Modeling the Dynamics of Nonverbal Behavior on Interpersonal Trust for Human-Robot Interactions.” AAAI Spring Symposia Series Symposium on Trust and Autonomous Systems, pp 46-47, 2013.

[15] D. Fellars, et all. “Enhancing Man-Portable Robot Functionality through Integration of New Sensor Payloads.” AUVSI Unmanned Systems North America, San Diego, CA, 2008.

[16] S. Turkle, C. Breazeal, O. Daste, B. Scassellati. “Encounters with Kismet and Cog: Children Respond to Relational Artifacts.” *Digital Media: Transformations in Human Communication.* New York, New York: Peter Lang Publishing, 2006, pp 313-330.

[17] R. Williams. *The Animator’s Survival Kit, Expanded Edition: A Manual of Methods, Principles, and Formulas for Classical, Computer, Games, Stop Motion, and Internet Animators*. London, United Kingdom: Faber & Faber, 2nd ed., 2012.