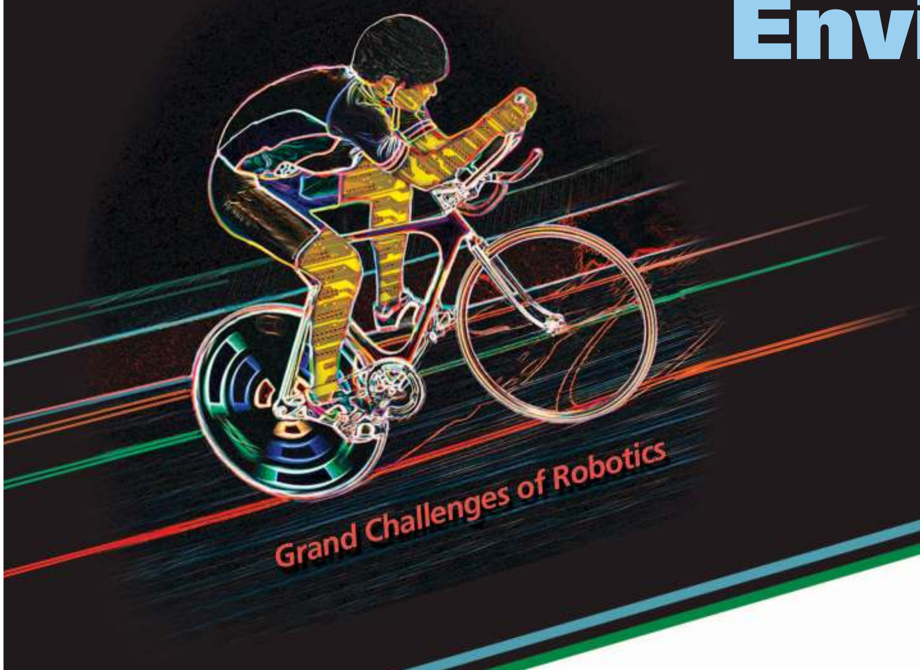
Challenges for Robot Manipulation in Human Environments



Developing Robots that Perform Useful Work in Everyday Settings

BY CHARLES C. KEMP, AARON EDSINGER, AND EDUARDO TORRES-JARA

Within factories around the world, robots perform heroic feats of manipulation on a daily basis. They lift massive objects, move with blurring speed, and repeat complex performances with unerring precision. Yet outside of carefully controlled settings, even the most sophisticated robot would be unable to get you a glass of water. The everyday manipulation tasks we take for granted would stump the greatest robot bodies and brains in existence today.

Why are robots so glorious in the factory, but so incompetent in the home? At the *Robotics Science and Systems Workshop: Manipulation for Human Environments* [1], we met with researchers from around the world to discuss the state of the art and look toward the future. Within this article, we present our perspective on this exciting area of robotics, as informed by the workshop and our own research.

To What End?

Commercially available robotic toys and vacuum cleaners inhabit our living spaces, and robotic vehicles have raced across the desert. These successes appear to foreshadow an explosion of robotic applications in our daily lives, but without advances in robot manipulation, many promising robotic applications will not be possible. Whether in a domestic setting or the workplace, we would like robots to physically alter the world through contact.

Robots have long been imagined as mechanical workers, helping us in our daily life. Research on manipulation in human environments may someday lead to robots that work alongside us, extending the time an elderly person can live at home, providing physical assistance to a worker on an assembly line, or helping with household chores.

Today's Robots

To date, robots have been very successful at manipulation in simulation and controlled environments such as a factory. Outside of controlled environments, robots have only performed sophisticated manipulation tasks when operated by a human.

 1070-9932/07/$25.00©2007 IEEE

Simulation

Within simulation, robots have performed sophisticated manipulation tasks such as grasping convoluted objects, tying knots, and carrying objects around complex obstacles. The control algorithms for these demonstrations often employ search algorithms to find satisfactory solutions, such as a path to a goal state, or a set of contact points that maximize a measure of grasp quality. For example, many virtual robots use algorithms for motion planning that rapidly search for paths through a state space that models the kinematics and dynamics of the world [2]. Most of these simulations ignore the robot's sensory systems and assume that the state of the world is known with certainty. For example, they often assume that the robot knows the three-dimensional (3-D) structure of the objects it is manipulating.

Controlled Environments

Within controlled environments, the world can be adapted to match the capabilities of the robot. For example, within a traditional factory setting engineers can ensure that a robot knows the relevant state of the world with near certainty. The robot typically needs to perform a few tasks using a few known objects, and people are usually banned from the area while the robot is in motion. Mechanical feeders can enforce constraints on the pose of the objects to be manipulated. In the event that a robot needs to sense the world, engineers can make the environment favorable to sensing by controlling factors such as the lighting and the placement of objects relative to a sensor. Moreover, since the objects and tasks are known in advance, perception can be specialized and model-based.

Factories are not the only controlled environments in which robots perform impressive feats of manipulation. Researchers often simplify the environments in which they test their robots in order to focus on problems of interest. So far, successful demonstrations of research robots autonomously performing complicated manipulation tasks have relied on some combination of known objects, simplified objects, uncluttered environments, fiducial markers, or narrowly defined, task-specific controllers.

Operated by a Human

Outside of controlled settings, robots have only performed sophisticated manipulation tasks when operated by a human. Through teleoperation, even highly complex humanoid robots have performed a variety of challenging everyday manipulation tasks, such as grasping everyday objects, using a power drill, throwing away trash, and retrieving a drink from a refrigerator (Figure 1). Similarly, disabled people have used wheelchair mounted robot arms, such as the commercially available Manus ARM (Figure 2), to perform everyday tasks that would otherwise be beyond their abilities. Attendees of the workshop were in agreement that today's robots can successfully perform sophisticated manipulation tasks in human environments when under human control, albeit slowly and with significant effort on the part of the human operator.

Human Environments

Human environments have a number of challenging characteristics that will usually be beyond the control of the robot's creator. The following list briefly describes some of these characteristics.

- **People are present**
  Users who are not roboticists may be in the same environment and possibly close to the robot.
- **Built-for-human environments**
  Environments and objects will usually be well-matched to human bodies and capabilities.
- **Other autonomous actors are present**
  For example, pets and other robots may be nearby.
- **Dynamic variation**
  The world can change without the robot taking action.
- **Real-time constraints**
  In order to interact with people and match the dynamics of the world, the robot must meet real-time constraints.

Figure 1. *Work at AIST with the HRP-2 humanoid has combined high-level teleoperation with autonomous perception and control. This work has enabled a user to reliably control a complex humanoid robot to perform sophisticated manipulation tasks, such as retrieving a drink from a refrigerator [3].*

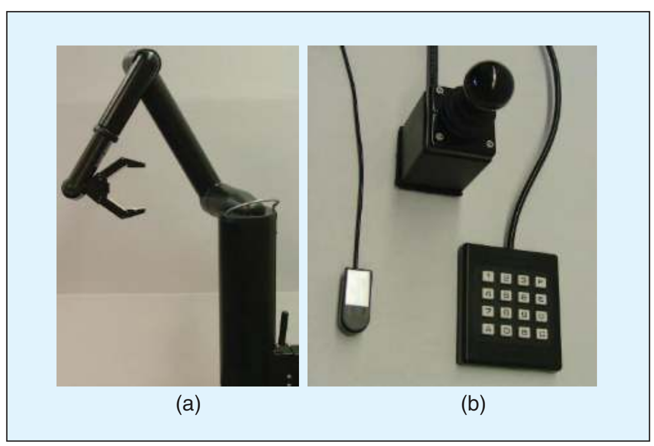
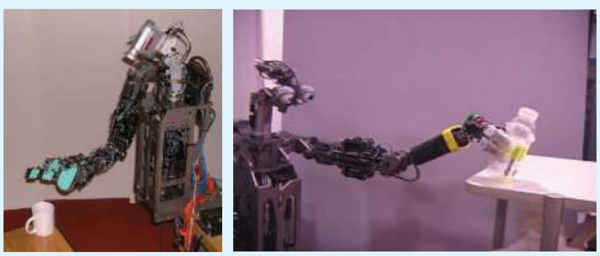


Figure 2. *(a) The Manus ARM is designed to be attached to a wheelchair and controlled by the wheelchair's occupant using (b) interfaces. Researchers at UMass Lowell are working to add sensing and semi-autonomous behaviors to this system in order to make it easier to use [4].*

- **Variation in object placement and pose**
  For example, an object may be placed in a cabinet, on a table, in a sink, in another room, or upside down.
- **Long distances between relevant locations**
  Tasks will often require a mobile manipulator, such as when moving objects from one room to another.
- **Need for specialized tools**
  Many tasks, such as cooking, assembly, and opening locks, require tools.
- **Variation in object type and appearance**
  For example, there can be one-of-a-kind objects and objects that have changed due to wear and tear.
- **Nonrigid objects and substances**
  For example, deformable objects, cables, liquids, cloth, paper, and air flow may need to be manipulated.
- **Variation in the structure of the environment**
  For example, architecture, furniture, and building materials vary from place to place.
- **Architectural obstacles**
  For example, robots can encounter cabinet doors, drawers, doors, and stairs.



(a) (b)

Figure 3. *(a) MIT robots Obrero and (b) Domo use compliance and force control to safely reach out into the world. Obrero reaches in the general direction of an object and then finds and grasps it haptically [8]. Domo initially reaches out toward a shelf in order to confirm its location and find a posture for placing objects. Once Domo has an object in hand, it reaches for the shelf with this posture and uses force control and compliance to let the object settle into place. [5].*

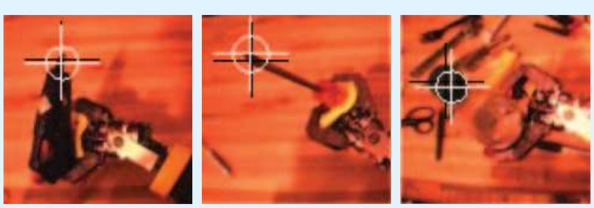


Figure 4. *Using visual motion and shape cues, the robot Domo can detect the tip of a tool-like object it is rigidly grasping. The white cross shows the kinematic prediction for the tool tip, and the white circle shows the mean pixel error of the prediction relative to hand-labeled tips (black cross). [5].*

- **Sensory variation, noise and clutter**
  For example, lighting variation, occluding objects, background sounds, and unclean surfaces are not uncommon.

People handle these issues daily. If you were at a friend's house for the first time and you were told to get a drink out of the refrigerator, you would most likely have no difficulty performing the task even though at some level everything would be different from your previous experiences. In fact, most cooks could walk into a well-stocked kitchen that they've never seen before and cook a meal without assistance.

Although robots should not need to have this level of capability to be useful, a human's great facility with such dramatic variation has a very real impact on the types of environments people inhabit. Even especially well-organized people live within highly variable environments, and engineers will rarely have the opportunity to tightly control these environments for the benefit of the robot.

How can roboticists develop robots that robustly perform useful tasks given these issues?

Approaches

Researchers are pursuing a variety of approaches to overcome the current limitations of autonomous robot manipulation in human environments. In this section, we divide these approaches into five categories (perception, learning, working with people, platform design, and control), which we discuss using examples drawn from the research presented at the workshop.

Perception

Robot manipulation in simulation and in controlled environments indicates that robots can perform well if they know the state of the world with near certainty. Although robots in human environments will almost always be working with uncertainty due to their limited view of a changing world, perceptual systems have the potential to reduce this uncertainty and enable robust autonomous operation. As such, perception is one of the most important challenges facing the field. Within this section, we discuss distinctive aspects of robot perception for manipulation with an emphasis on visual and tactile sensing.

Active Perception and Task Relevant Features

Through action, robots can simplify perception. For example, a robot can select postures in order to more easily view visual features that are relevant to the current task. Similarly, a robot can reach out into the world to physically sense its surroundings (see Figure 3).

In our work at the Massachusetts Institute of Technology (MIT), our robots often induce visual motion to better perceive the world. For instance, by rotating a rigidly grasped tool, such as a screwdriver or pen, the robot Domo can use monocular vision to look for fast moving convex regions in order to robustly detect the tip of a tool and control it (see Figure 4) [5]. This method performs well in the presence of cluttered backgrounds and unrelated motion. For a wide variety of human tools, control of the tool's tip is sufficient for its use. For example, the use of a screwdriver requires precise control of the tool blade relative to a screw head, but depends little on the details of the tool handle and shaft.

Encoding tasks in terms of task relevant features, such as the tip of a tool or the contact surface of a hand, offers several advantages. Tasks can be more easily generalized, since only the task relevant features need to be mapped from one object to another object, and irrelevant features can be ignored. Similarly, behaviors can be designed to enhance the detection of these features through postures or active perception. For our research, we have encoded tasks such as pouring, insertion, and brushing in terms of task relevant features that Domo detects and then visually servos with respect to one another (see Figure 5). Further research will be required to determine how well these methods extend to other tasks and objects.

Vision

Vision is probably the most studied modality for machine perception. Much of the research presented at the workshop involved some form of machine vision. For example, research from NASA/JSC with Robonaut (see Figure 6) and research from AIST with HRP-2 (see Figure 7) use model-based approaches to visual perception. Each robot has a small number of 3-D models for known objects that can be matched and registered to objects viewed by the robot's stereo camera in order to enable the robots to perform tasks such as opening a refrigerator or picking up a geological sample box with two hands [3], [6]. So far, the ability of these vision systems to reliably scale to large numbers of everyday manipulable objects has not been demonstrated.

A. Saxena from Stanford presented very promising work on visually detecting locations at which to grasp everyday objects using a single monocular camera [7]. The researchers trained the detector in simulation using rendered 3-D models of five object types (book, cup, pencil, block and cocktail glass) on which the researchers had marked locations called grasp points. Using the resulting grasp point detector, a robot arm was able to grasp and lift a variety of everyday objects outside of the training set (see Figure 8). The algorithm was tested on scenes that were fairly uncluttered and usually involved high-contrast objects placed against a low contrast, white background. It is unclear if this method will scale to large numbers of objects in diverse, realistically cluttered scenes, but the results generated significant interest among the workshop attendees.

This approach demonstrates the powerful potential for learning task relevant features that map to actions, instead of attempting to reconstruct a detailed model of the world with which to plan actions. In particular, it shows that at least some forms of grasping may be defined with respect to localized features such as grasp points instead of complicated configurations of 3-D contact points. This work also indicates that learning that has taken place in simulation can sometimes be transferred to robots operating in the real-world. If this holds true for other domains, it could dramatically simplify the development of autonomous manipulation capabilities for robots in human environments.

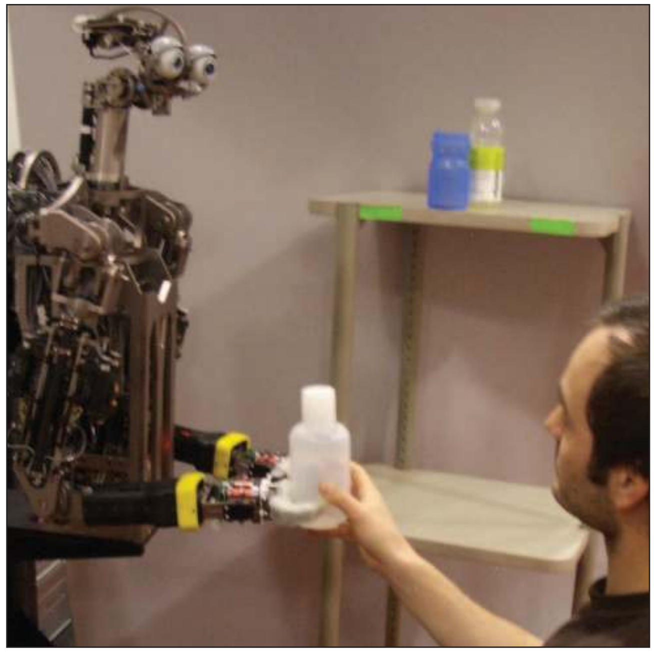


Figure 5. *MIT CSAIL robot Domo works with a human to place objects on a shelf [5].*

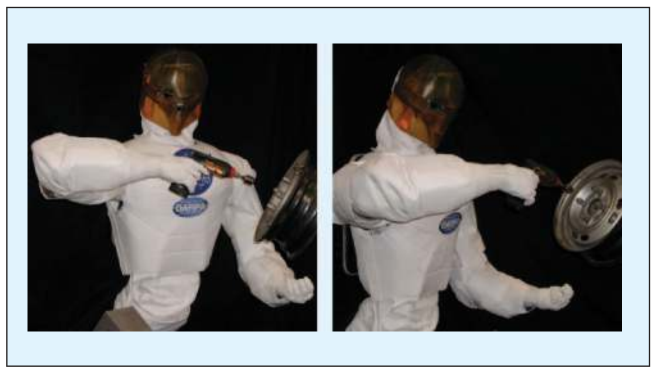


Figure 6. *Researchers from Brown and Vanderbilt have been developing methods that enable Robonaut to learn to behave autonomously from teleoperated examples. At the workshop, they presented a method that discovers instances of success and failure from teleoperated examples of Robonaut using a power drill [11].*

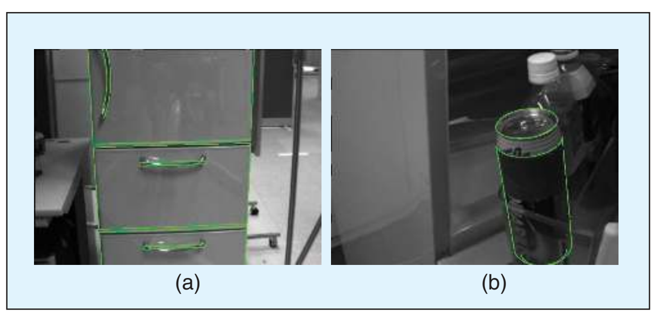


Figure 7. *Work at AIST with the HRP-2 humanoid robot uses 3D models of objects (overlayed on these images) to perceive the world. In the left image, the robot finds the refrigerator so that it can open it to retrieve the can. In the right image, the robot finds the can sitting inside the refrigerator [3].*

Whether in a domestic setting or the workplace, we would like robots to physically alter the world through contact.

Tactile Sensing

Since robot manipulation fundamentally relies on contact between the robot and the world, tactile sensing is an especially appropriate modality that has too often been neglected in favor of vision based approaches. As blind people convincingly demonstrate, tactile sensing alone can support extremely sophisticated manipulation.

Unfortunately, many traditional tactile sensing technologies, such as force sensing resistors (FSRs), do not fit the requirements of robot manipulation in human environments due to a lack of sensitivity and dynamic range. Researchers are seeking to develop new tactile sensors that take advantage of advances in materials, microelectromechanical systems (MEMS), and semiconductor technology [9]. Current sensors rarely provide directional information and tend to perform poorly when the incident angle of contact deviates significantly from the direction that is normal to the sensing surface. These are serious issues in human environments, since a robot must use low force interactions to manually explore its surroundings without unduly altering the state of the world or causing damage, and since the robot will rarely be able to control the exact angle at which its tactile sensors make contact with the world.

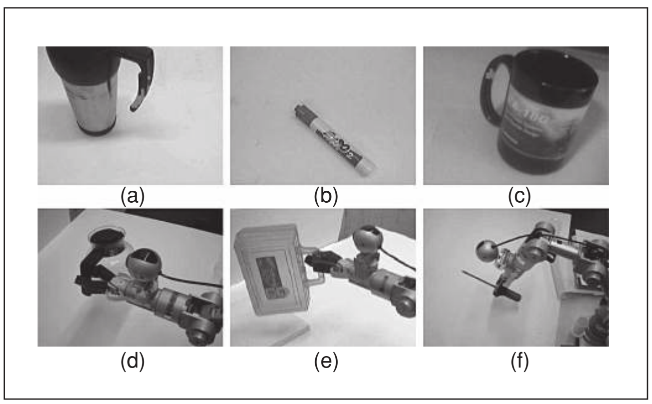


Figure 8. *Work presented by A. Saxena from Stanford uses supervised learning to train a system that visually detects grasp points on objects. (a)–(c) Resulting detections as red points. (d)–(f) These grasp points enabled a robot arm to grasp many everyday objects, including objects outside of the training set [7].*

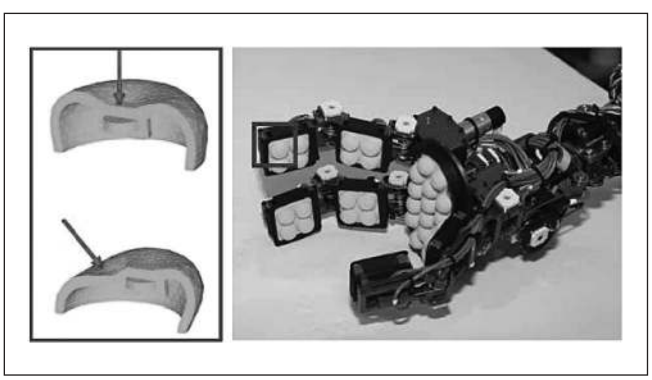


Figure 9. *A compliant hand and tactile sensor used on the Obrero platform at MIT. The tactile sensor is highly sensitive to normal and shear forces, providing rich sensory feedback as the robot grasps unmodelled objects [8].*

Researchers are addressing these challenges through novel sensor designs. Recent work at MIT by E. Torres-Jara has developed sensors with a protruding shape that allows them to easily make contact with the world from many directions in a similar way to the ridges of a human fingerprint or the hairs on human skin (see Figure 9). By measuring the deformation of the compliant dome, the sensors can estimate the magnitude and the direction of applied forces with great sensitivity. Conformation of the rubbery domes also distributes the force applied to the surface of an object, which reduces the stress. Using these sensors and a behavior-based algorithm, the humanoid robot Obrero has been able to tactilely position its hand around low mass objects and then grasp, lift and place them in different locations without using an explicit object model [8].

Learning

Today's top performing computer vision algorithms for object detection and recognition rely on machine learning, so it seems almost inevitable that learning will play an important role in robot manipulation. Explicit model-based control is still the dominant approach to manipulation, and when the world's state is known and consists of rigid body motion, it's hard to imagine something better. Yet robots cannot expect to estimate the state of human environments in such certain terms, and even motion planners need to have goal states and measures of success, which could potentially be learned.

Even with dramatic advances in sensing technologies, some relevant properties of the world are likely to remain hidden, such as occluded surfaces or the distribution of mass within an object. By learning from the natural statistics of human environments, robots may be able to infer unobservable properties of the world or select appropriate actions that implicitly rely on these unobserved properties. For example, if a robot were asked to fetch a drink for someone, it should be able to know that a drink is more likely to be located in the kitchen than on the floor of the bedroom.

Learning can also help address problems of knowledge acquisition. Directly programming robots by writing code can be tedious, error prone, and inaccessible to non-experts. Through learning, robots may be able to reduce this burden and continue to adapt once they've left the factory.

Opportunities for Learning

At the workshop, researchers presented robots that learned about grasping objects from autonomous exploration of the world, from teleoperation, and from simulation. If robots could learn to manipulate by autonomously exploring the world, they could potentially be easier to use and more adaptable to new circumstances. S. Hart from University of Massachusetts, Amherst, presented work on a developmental method that enables a humanoid robot to autonomously learn to reach for an object and grasp it [10]. Although the deep level of autonomy that developmental systems seek to achieve would be highly desirable, these types of learning systems are still in their infancy. Learning from teleoperation is advantageous since all of the relevant sensory input to the person, and output from the person, can be captured. O. Jenkins from Brown presented manifold learning methods for the autonomous discovery of task success and failure from unlabeled examples. This work used data captured while Robonaut was teleoperated to grasp a tool or use a drill (see Figure 6) [11]. K. Hsiao from MIT showed a method by which a simulated humanoid robot could learn whole-body grasps from examples that had been generated by the teleoperation of a simulated robot (see Figure 10) [12]. As previously discussed, research from Stanford showed that a real robot could learn to grasp objects from simulated data [7].

Common Sense for Manipulation

To what extent can the problems of manipulation in human environments be solved through knowledge? Large databases containing examples of objects, material properties, tasks, and other relevant information may allow much of the human world to be known to robots in a straightforward way. This type of approach could be a direct extension of research in which a robot manipulates a few objects for which it has 3-D models and associated task knowledge, or it could be coupled with offline machine learning methods that have been trained on the database. If robots could reliably work with some parts of the world and avoid the parts of the world unknown to them, they might be able to perform useful tasks for us. Given the standardization that has occurred through mass production and the advent of radio frequency identification (RFID) tags, this approach seems plausible for some limited tasks. If robots could easily be given additional knowledge and share it over the web, even uncommon parts of the world might become accessible to them.

Working with People

By treating tasks that involve manipulation as a cooperative process, people and robots can perform tasks that neither one could perform independently. For at least the near term, robots in human environments will be dependent on people. As long as a robot's usefulness outweighs the efforts required to help it, full robot autonomy is unnecessary.

Semi-Autonomous Teleoperation

From results in teleoperation, we can infer that computers with human-level intelligence could perform many useful and impressive tasks with today's robots. Unfortunately, computers are unlikely to have this level of ability any time soon, and even under human control most robots move slowly, lack dependability in everyday scenarios, and require great effort by the operator. By gradually incorporating autonomy into teleoperated robots, researchers can increase their usability and expand the areas to which they can be applied. One can even imagine scenarios in which the brains for semi-autonomous robots could be outsourced to people in remote locations.

At the workshop, N. Sian from AIST in Japan presented a teleoperated system that enables a human operator to reliably command a very complex humanoid robot to perform a variety of challenging everyday tasks (see Figures 1 and 11) [3]. The system integrates various forms of low-level autonomous motor control and visual perception, as

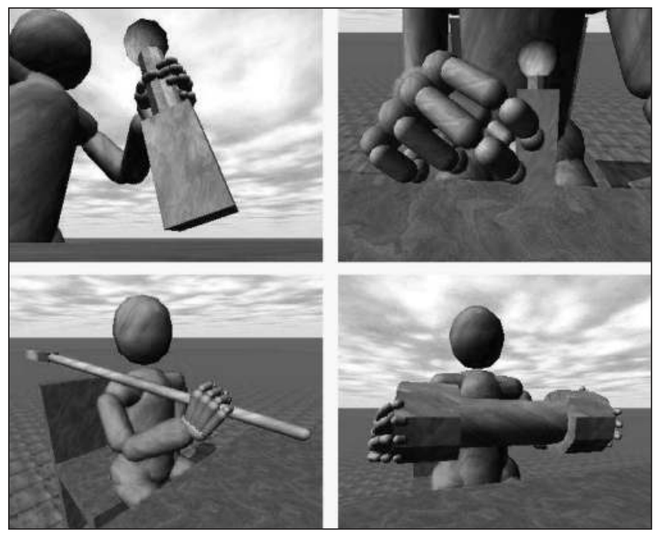


Figure 10. *K. Hsiao from MIT showed research on whole-body grasping that enables a simulated robot to learn from teleoperated examples. After training, the simulated robot is able to appropriately pick up objects it has not previously encountered using whole-body grasps [12].*

(a) (b)

Figure 11. *(a) By integrating autonomous components with teleoperation, researchers at AIST have developed a system by which people can reliably control the highly complex HRP-2 humanoid robot to perform sophisticated manipulation tasks in human environments [3]. (b) Research by K. Hsiao and T. Lozano-Perez at MIT uses virtual teleoperation to provide examples of whole body grasps from which a simulated humanoid robot can learn [12].*

Human environments have a number of challenging characteristics that will usually be beyond the control of the robot's creator.

well as higher-level behaviors. The higher-level behaviors can be interrupted and corrected if the human operator notices a problem. Similarly, H. Yanco's group at UMass Lowell is investigating improved interfaces to the Manus ARM that incorporate autonomous components in order to help a disabled user grasp an object more easily (see Figure 2) [4].

Human Interaction and Cooperation

Humans and robots can also work together while in the same physical space. Human environments tend to be occupied by humans, so robots have the opportunity to benefit from human assistance. For example, the initial version of the commercially successful Roomba relies on a person to occasionally prepare the environment, rescue it when it is stuck, and direct it to spots for cleaning and power. The robot and the person effectively vacuum the floor as a team, with the person's involvement reduced to a few infrequent tasks that are beyond the capabilities of the robot.

Researchers have looked at techniques for cooperative manipulation that physically couple a robot and a human. For example, humans and robots have carried objects together, and robot arms have helped guide human actions by resisting undesirable motions. Robots can also use social cues and physical cues to make cooperative manipulation more intuitive. Through eye contact, a vocal utterance, or a simple gesture of the hand, a robot may indicate that it needs help with some part of a task. In our work at MIT [5], we have shown that a person can intuitively work with a robot to place everyday objects on a shelf. In this work, the humanoid robot, Domo, was able to cue a person to hand it an object in a favorable way by reaching towards the person with an open hand. In doing so, the person solved the grasping problem for the robot.

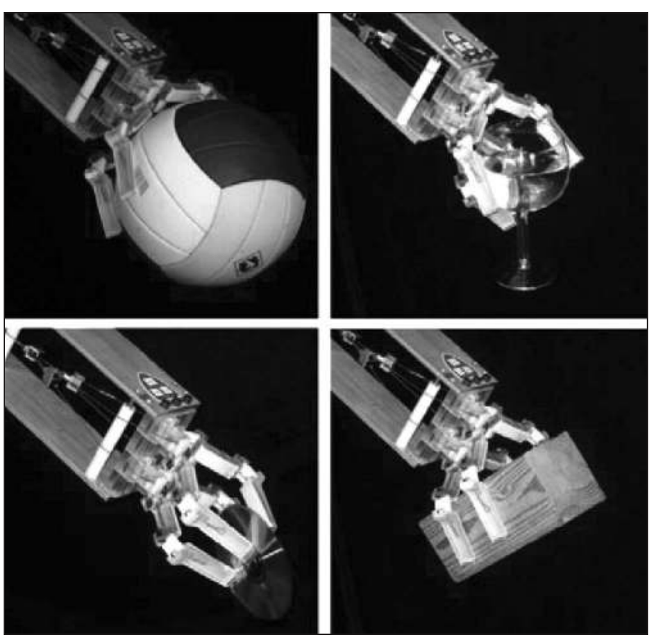


Figure 12. *A compliant grasper developed by A. Dollar and R. Howe at Harvard University leverages its adaptive physical design to robustly grasp unknown objects [16].*

Platform Design

Careful design of the robot's body can reduce the need for perception and control, compensate for uncertainty, and enhance sensing. Human environments have very different requirements from industrial settings, and attendees of the workshop agreed that the lack of suitable off-the-shelf robotic platforms is a serious impediment to research.

Safety

Robots that work with people must be safe. Traditional industrial manipulators are dangerous, so people are usually prohibited from being in a robot's workspace when it is in motion. Injury commonly occurs through unexpected physical contact, where forces are exerted through impact, pinching, and crushing. Of these, impact forces are typically the most dangerous, depending on the velocity, the mass and the compliance of the manipulator [13].

Commercially available arms such as the Manus ARM, the Katana arm from Neuronics, and the KUKA lightweight arm (based on the DLR arm) are beginning to address these issues. The Manus ARM incorporates several safety mechanisms, including current limits for the motors and slip-couplings that limit impact forces. The Katana arm is lightweight with low speed and low power. The KUKA arm is lightweight with force controlled joints that allow it to actively adjust its compliance through closed-loop control using feedback from torque sensors at the joints.

Researchers have also developed manipulators with passively compliant joints that incorporate elastic elements. For example, Stanford has developed Distributed Macro-Mini Actuation (DM^2) specifically for human-friendly robots, and many robots at MIT, including Domo and Obrero, have used Series Elastic Actuators (SEAs) [13]. These actuation methods have additional advantages, since their compliance is not wholly dependent on closed-loop force control. These manipulators still have compliance in the event of an unexpected impact beyond the bandwidth of their closed-loop force control.

Designing for Uncertainty

Traditionally, industrial robots have eschewed passive physical compliance at the joints in favor of stiff, precise, and fast operation. This is a reasonable design tradeoff when the state of the world is known with near certainty. Within human environments, compliance and force control are more advantageous since they help the robot safely interact with people, explore the environment, and cope with uncertainty.

A. Dollar and R. Howe from Harvard optimized several parameters in the design of a robot hand so that it could better grasp objects with uncertain physical properties (see Figure 12) [16]. The hand, driven by a single actuator, is made entirely out of compliant urethane materials of varying stiffness. It has embedded tactile and position sensors and is actuated by a remote motor through tendons. The hand's compliance, combined with its optimized adaptability, allows it to robustly form power grasps on a variety of objects in the presence of large sensing uncertainties. Remarkably, the hand is also robust to sustained impacts from a hammer.

Our humanoid robots developed at MIT use series elastic actuators in all the joints of the arms and hands. They also have compliant rubber skin on their fingers. This passive compliance allows them to more safely explore unknown environments and adapt to geometric uncertainty. On our robot Domo (see Figure 3), this compliance helps it to transfer unknown objects between its hands and place them on a shelf. When transferring an object between its hands, the grasped object passively adjusts to the bimanual grasp. When placing an object on a shelf, the compliance helps the object's flat base to stably align with the shelf surface [5]. On our robot Obrero, compliance in the fingers (see Figure 8) enables the robot to gently come into contact with objects without knocking them over, and helps its hand to conform to unknown objects [8].

On Human Form

Human environments tend to be well-matched to the human body. Robots can sometimes simplify manipulation tasks by taking advantage of these same characteristics. For example, most everyday objects in human environments sit on top of flat surfaces that can be comfortably viewed and reached by a human. A robot can more easily perceive and manipulate these objects if its sensors look down on the surfaces and its manipulators easily reach the surfaces. Similarly, everyday hand-held objects, such as tools, are designed to be grasped and manipulated using a human hand. A gripper that has a similar range of grasps will tend to be able to grasp everyday human objects. A direct approach to taking advantage of these properties of human environments is to create humanoid robots that emulates the human form, but mobile manipulation platforms can selectively emulate critical features such as a small footprint, sensors placed high above the ground, an approximately hand-sized gripper, and robot arms with approximately human size and degrees of freedom (see Figure 13).

Control

Within perfectly modeled worlds, motion planning systems perform extremely well. Once the uncertainties of dynamic human environments are included, alternative methods for control become important. For example, control schemes

Perception is one of the most important challenges facing the field.

must have real-time capabilities in order to reject disturbances from unexpected collisions and adapt to changes in the environment, such as might be caused by a human collaborator. Many researchers are looking at ways to extend planning methods so that they will perform well under these circumstances, including O. Brock's group at UMass Amherst and researchers at the University of North Carolina, Chapel Hill, Carnegie Mellon University, the University of Illinois at Urbana-Champaign, and Stanford [1], [2], [15]. Other researchers are addressing these issues with robust closed-loop controllers that make use of rich sensory feedback. For example, R. Platt and the Robonaut group at NASA/JSC and R. Grupen's group at UMass Amherst have explored ways to learn and compose real-time, closed-loop controllers in order to flexibly perform a variety of autonomous manipulation tasks in a robust manner [6], [10]. At MIT we often use hand-coded behavior-based controllers that specify tasks in terms of visual servoing and other forms of feedback driven control [5].

Grand Challenges

At the end of the workshop, we held a discussion on the topic of grand challenges for robot manipulation in human environments. As a group, we arrived at three challenges that encapsulate many of the important themes of this research domain. The agreed upon challenges were: cleaning and organizing a house, preparing and

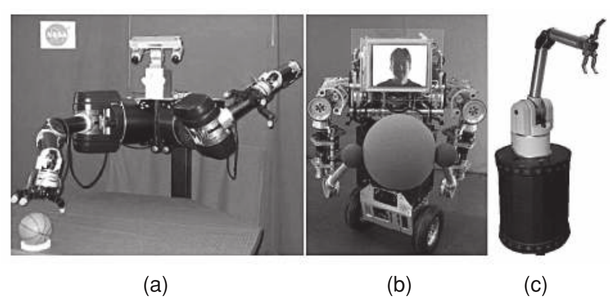


Figure 13. *Researchers at UMass Amherst have developed a variety of platforms for manipulation research with distinct capabilities. These three platforms were used in work presented at (a) and (b) the workshop from R. Grupen's group and (c) O. Brock's group [10], [14], [15]. (a) Dexter is a nonmobile humanoid robot. (b) uBot-4 is a compact, dynamically stable robot with the ability to bimanually grasp some objects off of the floor when teleoperated. (c) UMan is a mobile manipulator with a single dexterous robot arm (WAM arm by Barrett Technology) positioned to access everyday objects in human environments.*

Careful design of the robot's body can reduce the need for perception and control, compensate for uncertainty, and enhance sensing.

delivering an order at a burger joint, and working with a person to cooperatively assemble a large structure [1].

Disordered House to an Ordered House (Reversing Entropy to Create Beauty)

A robot that can enter a home and clean up a messy room must adapt to the large variability of our domestic settings, understand the usual placement of everyday objects, and be able to grasp, carry, and place everyday objects, including clothing.

Preparing and Delivering an Order at a Burger Joint

Preparing and delivering an order at an unmodified burger joint would require a robot to dexterously manipulate flexible materials, work with tools designed for humans, and perform a variety of small, but complex, assembly tasks. A mechanized burger joint with specialized machinery might be more practical, but this challenge emphasizes manipulation capabilities that would be of use to a broad set of applications, including small scale manufacturing tasks and meal preparation (e.g., a short-order cook or a domestic assistant).

Outdoor Party Preparation

Cooperatively assembling a large structure would require that a human and robot cooperate in a very direct way with whole body manipulations, fixturing, insertions, and lifting. One specific example, would be preparing a backyard or park for an outdoor event that involves setting up a tent, chairs and tables. A group of people and robots might work together to achieve this goal, and people could instruct the robots about their specific desires (much like a furniture moving situation).

Smooth Paths to Progress

Even though some aspects of these challenges appear within reach, nearly all of the participants agreed that it would be premature for researchers to directly pursue them. In this spirit, we conclude with several plausible paths for incremental progress towards these goals.

By Approach

We expect progress to be made along each of the approaches we have discussed within this article. However, the problem of robot manipulation in human environments necessitates the integration of these approaches into functional systems that can be validated in the real-world on tasks with clear measures of success.

By Module, Platform, and Algorithm

We would expect research to result in de facto standards for software modules, hardware platforms, and algorithms. We already see this to some extent with face detectors, low-level vision algorithms, and machine learning algorithms. Since manipulation in human environments is a systems level problem, sharing components will be especially important for progress so that researchers can build on one another's contributions, compare approaches, and generate repeatable results.

From Semi-Autonomy to Full Autonomy

Another smooth path for progress is the gradual automation of manipulation tasks that currently require a human. As illustrated within this article, robots that are teleoperated are excellent candidates for partial automation, and many researchers are already following this path. In general, semi-autonomous, human-in-the-loop systems offer the opportunity for robots to perform useful tasks in the near term with a human present to take-over when the robot gets into trouble.

From Simple to Complex Tasks

The Roomba could be considered the first successful autonomous mobile manipulator for the home, since it manipulates dirt on the floor. The Roomba partially automates a common household task. By narrowing the scope of a task, useful robots could be developed more quickly in order to drive progress and serve as a foundation for further capabilities. Rather than push for highly complex tasks, many researchers are focusing on simpler, core capabilities such as grasping everyday objects, fetching and carrying objects, placing objects, being handed objects by a person, and handing objects to a person. These tasks can be further constrained by limiting the types of objects the system works with (e.g., hand sized cylindrical objects) and the types of places in which it operates (e.g., accessible flat surfaces such as desks and tables). Over time, these constraints could be progressively loosened (e.g., objects that work with objects that require two hands, and robots that open cabinets to access more flat surfaces).

Conclusion

Within this article, we have presented our perspective on the challenges facing the field as informed by the workshop and our own research. We have discussed potential paths to the long-term vision of robots that work alongside us in our homes and workplaces as useful, capable collaborators. Robot manipulation in human environments is a young research area, but one that we expect to grow rapidly in the coming years as more researchers seek to create robots that actively help us in our daily lives.

Acknowledgment

The authors would like to thank the anonymous reviewers for their insightful comments and suggestions. They also thank the participants at the workshop with special thanks to those who presented their work and contributed to the discussions. They would especially like to thank O. Brock for his support and enthusiastic participation. Finally, they wish to thank keynote speaker R. Brooks and their fellow workshop organizers: L. Aryananda, P. Fitzpatrick, and L. Natale.

The authors encourage readers to go to the workshop's website for further information: http://manipulation.csail.mit.edu/rss06/, or to the official online proceedings [1].

Keywords
Assistive robots, grand challenges, human environments, robot manipulation, robotic grasping

References

[1] C. Kemp, L. Aryananda, A. Edsinger, P. Fitzpatrick, L. Natale, and E. Torres-Jara, eds. *Proc. RSS Workshop: Manipulation for Human Environments*, http://www.archive.org/details/ manipulation_for_human_environments_2006, Philadelphia, PA, Aug. 2006.

[2] S.M. LaValle, *Planning Algorithms.* Cambridge, U.K.: Cambridge Univ. Press, 2006.

[3] N. Sian, T. Sakaguchi, K. Yokoi, Y. Kawai, and K. Maruyama, "Operating humanoid robots in human environments," *Proc. RSS Workshop: Manipulation for Human Environments*, Philadelphia, PA, Aug. 2006.

[4] K. Tsui and H. Yanco, "Human-in-the-loop control of an assistive robot arm," *Proc. RSS Workshop: Manipulation for Human Environments*, Philadelphia, PA, Aug. 2006.

[5] A. Edsinger and C. Kemp, "Manipulation in human environments," *Proc. IEEE/RSJ Int.Conf. Humanoid, Robotics*, Genoa, Italy, 2006.

[6] R. Platt, R. Burridge, M. Diftler, J. Graf, M. Goza, and E. Huber, "Humanoid mobile manipulation using controller refinement," *Proc. RSS Workshop: Manipulation for Human Environments*, Philadelphia, PA, Aug. 2006.

[7] A. Saxena, J. Driemeyer, J. Kearns, and A. Ng, "Learning to grasp novel objects using vision," *Proc. RSS Workshop: Manipulation for Human Environments*, Philadelphia, PA, Aug. 2006.

[8] L. Natale and E. Torres-Jara, "A Sensitive approach to grasping," *Proc. 6th Int. Conf. Epigenetic Robotics*, Paris, France, Sep. 2006.

[9] V. Maheshwari and R.F. Saraf, "High-resolution thin-film device to sense texture by touch," *Science* 312 (5779), 1501, Jun., 2006.

[10] S. Hart, S. Ou, J. Sweeney, R. Grupen, "A Framework for learning declarative structure," *Proc. RSS Workshop: Manipulation for Human Environments*, Philadelphia, PA, Aug. 2006.

[11] C. Jenkins, A. Peters, and R. Bodenheimer "Uncovering success in manipulation," *Proc. RSS Workshop: Manipulation for Human Environments*, Philadelphia, PA, Aug. 2006.

[12] K. Hsiao and T. Lozano-Perez, "Imitation learning of whole-body grasps," *Proc. RSS Workshop: Manipulation for Human Environments*, Philadelphia, PA, Aug. 2006.

[13] M. Zinn, O. Khatib, B. Roth, and J.K. Salisbury, "Playing It safe: A new actuation concept for human-friendly robot design," *IEEE Robot. Automat. Mag.,* vol. 11, no. 2, pp.12–21, Jun., 2004.

[14] P. Deegan, B. Thibodeau, and R. Grupen, "Designing a self-stabilizing robot For dynamic mobile manipulation," *Proc. RSS Workshop: Manipulation for Human Environments*, Philadelphia, PA, Aug. 2006.

[15] D. Katz, E. Horrell, Y. Yang, B. Burns, T. Buckley, A. Grishkan, V. Zhylkovskyy, O. Brock, and E. Learned-Miller, "The UMass mobile manipulator uMan: An experimental platform for autonomous mobile manipulation," *Proc. RSS Workshop: Manipulation for Human Environments*, Philadelphia, PA, Aug. 2006.

[16] A. Dollar and R. Howe, "A Robust compliant grasper for unstructured environments," *Proc. RSS Workshop: Manipulation for Human Environments*, Philadelphia, PA, Aug. 2006.

Charles C. Kemp is currently a Senior Research Scientist in the Health Systems Institute at the Georgia Institute of Technology, Atlanta. He holds the B.S., M.Eng., and Ph.D. degrees from MIT, Cambridge, MA, in electrical engineering and computer science. His research focuses on the development of intelligent robots with autonomous capabilities for healthcare applications such as home health, rehabilitation, telemedicine, and sustainable aging. He is especially interested in robots that autonomously perform useful manipulation tasks within human environments.

Aaron Edsinger is currently a postdoctoral researcher at the MIT Computer Science and Artificial Intelligence Laboratory, Cambridge, MA. He holds the B.S. degree in computer science from Stanford University, Stanford, CA, and the S.M. and Ph.D. degrees from MIT. His doctoral work is on developing robots that can work with people in everyday settings. This includes research on behavior-based architectures, bimanual manipulation, compliant manipulators, hand design, and hand-eye coordination.

Eduardo Torres-Jara holds the Ingeniero degree in electrical engineering from Escuela Politécnica del Ejército, Ecuador, and the M.S. and Ph.D. degrees from MIT, Cambridge, MA, in electrical engineering and computer science. He is currently a postdoctoral researcher at MIT CSAIL. His current research interest is in *sensitive manipulation*: manipulation that uses dense tactile sensing and force feedback as its main input, and is about action as much as perception. His work includes tactile sensing, compliant actuator design, and behavior-based architectures.

Address for Correspondence: Charles C. Kemp, Health Systems Institute, Georgia Tech and Emory University, 901 Atlantic Drive, Suite 4100, Atlanta, Georgia 30332-0477. E-mail: charlie.kemp@hsi.gatech.edu.