Challenges for Robot Manipulation in Human Environments

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 these carefully controlled robot realms, even the most sophisticated robot would be unable to pour you a drink. 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, yet 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.

I. To What End?

Commercially available robotic toys and vacuum cleaners inhabit our living spaces, and robotic vehicles traverse our roads. 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 the work of daily life. This vision has driven the recent growth of research into robot manipulation in human environments. Research in this area will someday lead to robots that can work alongside us in our homes and workplaces, extending the time an elderly person can live at home, providing physical assistance to a worker on an assembly line, or helping us with household chores.

II. 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.

A. Simulation

Within simulation, robots have performed sophisticated manipulation tasks such as grasping convoluted objects, tying knots, carrying objects around complex obstacles, and extracting objects from entangled circumstances. 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. Almost all 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 3D structure of the objects it is manipulating.

Fig. 1. Work with the AIST HRP-2 humanoid has combined high-level teleoperation with autonomous low-level perception and control. Here it is shown retrieving a drink from a refrigerator [1]. (Permission not yet acquired)

B. Controlled environments

In a carefully controlled environment, these assumptions can be met. 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 working. Mechanical feeders can enforce constraints on the pose of the objects to be manipulated. And 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 the sensor. Moreover, since the objects and tasks are known in advance, perception can be specialized and model-based. Whether by automated planning or direct programming, robots perform exceptionally well in factories around the world on a daily basis. Within research labs, successful demonstrations of 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.

C. 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, shown in Figure 2, to perform everyday tasks that would otherwise be beyond their abilities. Attendees of the workshop were in agreement that under human control today's robots can successfully perform sophisticated manipulation tasks in human environments, albeit slowly and with significant effort on the part of the human operator.

Fig. 2. The Manus ARM is designed to be attached to a wheelchair and has a joystick interface. It is controlled under close supervision by an operator. Researchers at UMass Lowell are expanding the autonomy of the arm in order to improve the assistance provided to the physically disabled. [1]. (Permission not yet acquired)

III. Human Environments

Human environments have a number of challenging characteristics that will usually be beyond the control of the robot's creator, such as the characteristics listed below:

- **people present**
  Users who are not roboticists or technicians will 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 present**
  For example, people, animals, and other robots may be in the presence of the robot.
- **dynamic variation**
  The world can change without the robot taking action.
- **real-time constraints**
  In order to interact with people and generally match the dynamics of the world, the robot must meet real-time constraints.
- **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.
- **need for specialized tools**
  Many tasks, such as cooking, assembly, and opening locks, require tools.
- **variation in object type and appearance**
  There can be one-of-a-kind objects and many instances of a particular type of object such as a screwdriver, and wear and tear on known objects.
- **non-rigid objects and substances**
  For example, deformable objects, cables, liquids, cloth, paper, gases, and air flow may need to be manipulated.
- **variation in the structure of the environment**
  For example, architecture, furniture, and building materials all vary from place to place.
- **sensory variation, noise and clutter**
  For example, lighting variations, occluding objects, background sounds, and dirty/sticky 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 fridge, 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.

Robots should not need to have this level of capability to be useful. However, 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 the environment for the benefit of the robot.

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

IV. 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.

A. 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. Arguably, this goal is unachievable within human environments, since there will almost always be hidden information such as occluded surfaces, the distribution of mass in objects, and detailed material properties. Nonetheless, the success of robots in simulation, in controlled environments, and under teleoperation suggests that perception is one of the most important challenges facing the field.

Within specialized perceptual research communities, relatively little emphasis is placed on the distinctive perceptual problems of robot manipulation. These problems differ in terms of the desired perceptual output, the data that are available as input, and the computational constraints. For

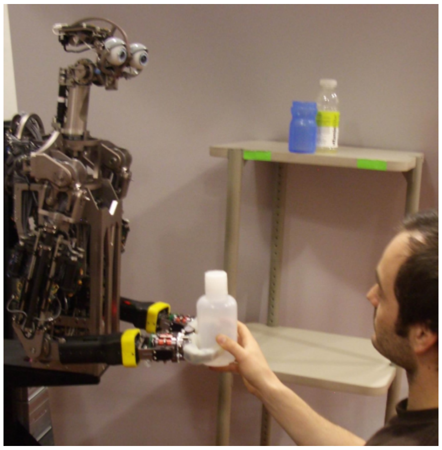


Fig. 3. The MIT CSAIL robot Domo works with a human to place cylindrical objects on a shelf [3].

Fig. 4. Using visual motion and shape cues, Domo can detect the tips of tool-like objects it is rigidly grasping. The tip of a tool is a feature relevant to control of the tool. The white cross shows the kinematic prediction for the tool tip, based on noisy 2D detections, and the white circle shows the mean pixel error of the prediction relative to hand labels that were solely used for evaluation. The black cross marks the hand-labeled tip [4]
.

example, visually recognizing an object does not necessarily map to a method for grasping the object, since an object's geometry can be more important than its identity. Although real-time constraints can be daunting, computation continues to become more affordable. Robot perceptual systems also have the opportunity to benefit from streaming sensory data, multiple sensing modalities, contact-based perception, and physical interaction with the environment.

1) 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. Similar principles also apply for other perceptual modalities. For example, when searching for a shelf, the MIT robot, Domo, can reach out into the world to physically find the shelf and in the process record an arm posture that makes contact with the shelf (see Figure 5). Likewise, the MIT robot Obrero, can reach out to the area near an object and tactilely find it and grasp it.

In our work at 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, Domo can use a single monocular camera to look for fast moving convex regions in order to robustly detect the tip of a tool and control it (see Figure 4). This method performs well in the presence of cluttered backgrounds and unrelated motion. For a wide

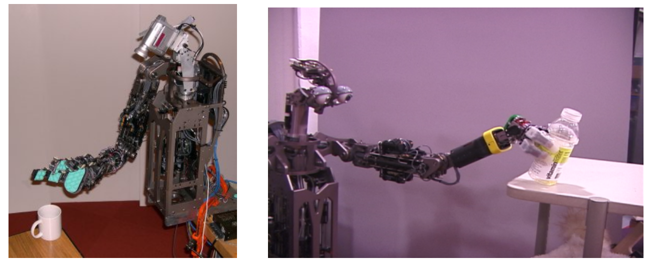


Fig. 5. The MIT robots Obrero (left) and Domo (right) use compliance and force control when reaching out into the world. Obrero (left) reaches in the direction of an object with its tactile sensors leading the way. Obrero is able to grasp the cup using tactile sensing [2]. Domo (right) reaches out to a shelf in order to place an object on it. Domo uses force control and compliance to let the object settle into place on the shelf [3].

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 position and force 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, 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. Detectors can be specialized to the task relevant features, and control can be specified in terms of these task relevant features. For our research, we have encoded several tasks, such as pouring, insertion, and brushing, in terms of the detection and visual servoing of task relevant features relative to one another (see Figure 3). As another example, when Domo transfers an object from one hand to the other, Domo visually servos the convex outline of its empty open palm towards the object. In this case, the contact surface of Domo's hand is a task relevant feature. During this process the visual motion of the hand helps Domo detect this surface and Domo maintains a posture that keeps the surface visible.

2) 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. Work from NASA/JSC on Robonaut (see Figure 8) and work from AIST on HRP-2 (see Figure 1), used model-based visual perception. Each robot had a small number of 3D models for known objects that could be matched and registered to objects viewed by the robot's stereo camera. As of yet, the ability of these vision systems to reliably scale to large numbers of everyday manipulable objects has not been demonstrated.

Ashutosh Saxena from Andrew Ng's group at Stanford presented very promising work on visually detecting "grasp points" on everyday objects using a single monocular camera (see Figure 6). The system was trained in simulation using a handful of rendered 3D models of everyday objects for which the "grasp points" had been hand labeled. 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. The scenes on which the algorithm was tested were fairly uncluttered and usually involved high-contrast objects placed against a low contrast, white background. The ability of this

Fig. 6. Work presented by Ashutosh Saxena from Andrew Ng's group at Stanford uses supervised learning to discover the visual grasp points (red points in top row) for objects, allowing the manipulator to grasp many everyday items (bottom row) [1]. (Permission not yet acquired)

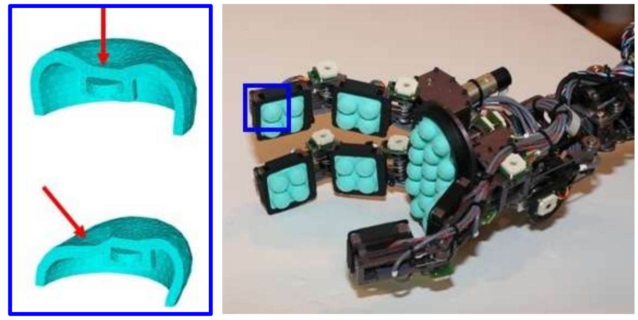


Fig. 7. A compliant hand and tactile sensor used on the Obrero platform at MIT. The tactile sensor is highly sensitive to normal and shear forces, providing rich sensory feedback as the robot grasps unmodelled objects [2].

particular solution to scale across large numbers of objects in realistic cluttered scenes is unclear. However, the approach demonstrates the powerful potential for learning task relevant features that directly 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 3D 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 offers the possibility of dramatically simplifying the acquisition of training data, training protocols, and preliminary evaluations of learned algorithms for robot manipulation in human environments.

3) 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. Researchers have had some success with gripper-mounted IR range sensors, and small force-torque load cells, but IR range sensors do not exploit contact and even the smallest load cells are insensitive and too large to cover a gripper.

Unfortunately, many traditional tactile sensing technologies do not fit the requirements of robot manipulation in human environments. For an example, consider a computer touch pad that uses force resistor sensors (FSR). These pads have high spatial resolution, low minimum detectable force (about 0.1N) and a good force range (7 bits). These features make the sensor work very well when a human finger, a plastic pen, or another object with a pointy shape comes in contact. However, if you place a larger object on the pad or the same pen at a small incident angle, the sensor is unlikely to detect contact unless the applied force is very large. This is a serious issue, since a robot would be unable to manually explore its surroundings without the risk of unduly altering the state of the world or causing damage.

The curvature of everyday objects varies considerably, and a robot will often not know the angle at which to expect contact. During exploration, a light touch is desirable, but when handling an object the robot's fingers must exert high forces. Research at MIT by Eduardo Torres-Jara, one of the authors, has demonstrated new tactile sensors that address these issues (see Figure 7). The sensor's protruding shape 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. By detecting the deformation of the compliant dome, the sensors can detect the magnitude and the direction of applied forces with great sensitivity. Conformation of the rubbery domes also increases friction when in contact with an object. Using these sensors and a behavior-based algorithm, the humanoid robot Obrero has been able to tactilely position its hand around low mass objects, grasp, lift and place them in a different location. In these tests, the force needed to avoid slippage and the conditions to release the object on a surface were also determined using tactile information [2]. No model or reconstruction of the object was used during grasping.

B. Learning

Today's top performing computer vision algorithms for detection and recognition rely on machine learning, so it seems almost inevitable that learning will play an important role in robot manipulation. However, the significance and nature of this role has yet to be fully determined. 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. However, as we have noted, 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 to optimize

Robots are unlikely to be able to directly sense all of the relevant properties of the world they are manipulating, or at least not in an efficient manner. By learning from the natural statistics of human environments, robots may be able to reliably infer some of these properties or select appropriate actions that implicitly rely on characteristics of the unobservable world. For example, if a robot were asked to fetch a drink for someone, it should be able to know that

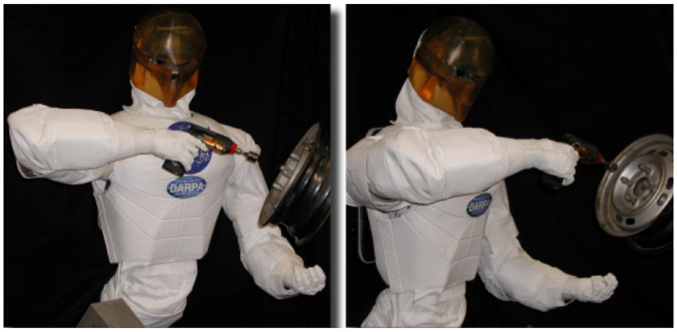


Fig. 8. Researchers at NASA/JSC have been moving from teleoperation to autonomy with their Robonaut platform. Here it autonomously learns when a teleoperator succeeds at manipulating a power drill [1]. (Permission not yet acquired)

the drink is more likely to be located in the kitchen than on the floor of the bedroom.

Also, learning helps address the 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 lab.

1) Structures for Learning: At the workshop, researchers presented robots that learned about grasping objects from autonomous exploration of the world, from teleoperation, and even 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. Unfortunately, developmental systems are still in their infancy and are difficult to design. Learning from teleoperation is advantageous since all of the relevant sensory input to the person, and output from the person, can be captured. Recent work presented by Chad Jenkins of Brown demonstrated autonomous discovery of task success and failure using data captured while Robonaut was teleoperated to grasp a tool or use a drill (see Figure 8). The work presented by Kaijen Hsiao from Tomas Lozano-Perez group at MIT, showed a method by which a simulated humanoid robot could learn whole-body grasps from human teleoperation of a simulated robot (see Figure 9). As previously discussed, in the research from Stanford, the a real robot learned to grasp objects from simulated data.

2) Commonsense for Manipulation: To what extent can the problems of manipulation in human environments be solved through knowledge or experience? This is an important unanswered question that relates to learning. Large databases containing examples of common objects, material properties, tasks, and other relevant information may allow much of the human world to be known to robots in a straightforward way. In a sense, this type of approach would be a direct extension of research in which a robot manipulates a few objects for which it has 3D models and associated task knowledge. If robots could reliably work with some parts of the world and avoid the parts of the world unknown to them, they may be able to perform useful tasks for us. Given the standardization that has occurred through mass production and the advent of

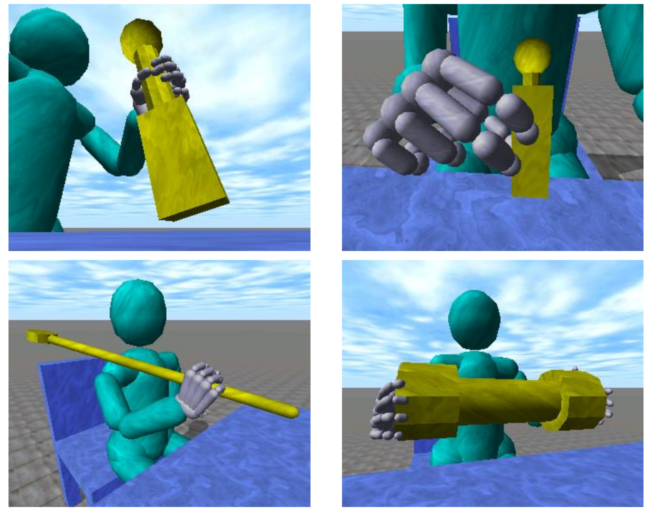


Fig. 9. In this work presented by Kaijen Hsiao from MIT, after a simulated robot learns from examples provided by teleoperation, it is able appropriately pick up objects it has never before encountered using whole-body grasps [1]. (Permission not yet acquired)

RFID tags, this approach seems plausible for some limited tasks. Moreover, if robots could easily be given additional knowledge and share it over the web, even some distinctive parts of the world might become accessible to them.

C. Working with People

For at least the near term, robots in human environments will be dependent on people. Fortunately, people tend to be present within human environments. As long as the robot's usefulness outweighs the efforts required to help it, full robot autonomy is unnecessary. Careful design can make robots intuitive to use, thereby reducing the required effort. 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, although the person's involvement is reduced to a few infrequent tasks that are beyond the capabilities of the robot.

By treating tasks that involve manipulation as a cooperative process, people and robots can perform tasks that neither one could perform as an individual.

1) 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. Of course, this does not make the problem any easier. Even under human control, most robots move slowly, require great effort by the human operator, and may not be dependable in everyday scenarios.

Besides giving a better idea of what current robots can do, teleoperation suggests a smooth path for progress. As shown with rehabilitation robots, teleoperated robots are already useful. By gradually incorporating more autonomy into the robots, researchers can increase their usability and expand areas to which they can be practically applied. Along these lines, Neo Ee Sian from AIST presented a teleoperated system that

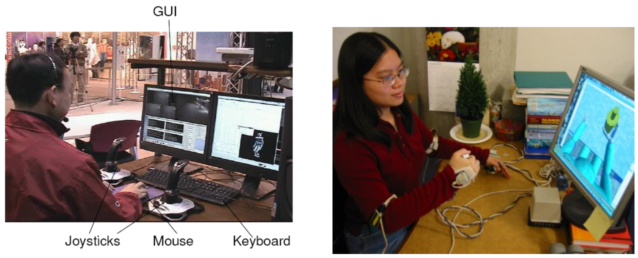


Fig. 10. Left: By integrating autonomous components with teleoperation, researchers at AIST have developed a system by which people can reliably control the highly complex HRP-2 humanoid robot to perform everyday manipulation tasks [1]. Right: Research by Kaijen Hsiao and Tomas Lozano-Perez at MIT, uses virtual teleoperation to provide examples of whole body grasps from which a simulated humanoid robot can learn [1]. (Permission not yet acquired)

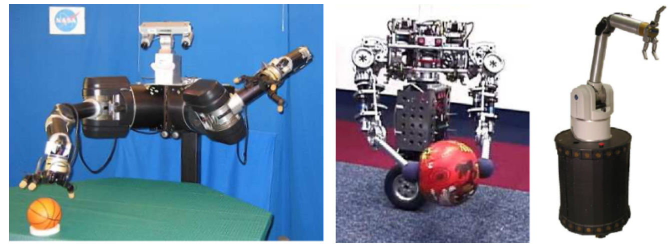


Fig. 11. Researchers at UMass Amherst have developed a variety of different platforms for manipulation research with distinct capabilities. These three platforms were used in work presented at the workshop from Rod Grupen's group (left & middle) and Oliver Brock's group (right) [1]. The left robot, Dexter, is a non-mobile humanoid robot. The middle robot, uBot-4 is a compact, dynamically stable, robot with the ability to bimanually grasp some objects off the floor when teleoperated. The right robot, UMan, is a single armed mobile manipulator with a dexterous robot arm (WAM arm by Barrett Technology) with kinematics similar to a human arm, and positioned so as to be able to access everyday objects. (permission not yet acquired)

allows a human operator to reliably command a very complex humanoid robot, HRP-2, to perform a variety of everyday tasks (see Figures 1 and 10). The system integrates various forms of low-level autonomous motor control and visual perception, as well as higher-level behaviors. The higher-level behaviors can be interrupted and corrected if the human operator notices a problem. Similarly, Holly Yanco's group at UMass Lowell are investigating improved interfaces to the Manus ARM that incorporate autonomous components, such as visual servoing, to help a disabled user grasp an object more easily (see Figure 2).

Prior to achieving full autonomy, one can imagine scenarios where the brains for semi-autonomous robots could be outsourced to people working at remote locations.

2) Human Interaction & Cooperation: Researchers have looked at techniques for cooperative manipulation that physically couple the robot and the person, such as carrying an object together, or guiding a person's actions with a Cobot manipulator. Robots with human like features (eg. humanoids) can also leverage a person's intuitive understanding of physical and social cues. 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 [3], we have shown that a person can intuitively work with a robot to place everyday cylindrical objects on a shelf. In this work, the humanoid robot, Domo, was able to cue a person to hand it an object by reaching towards the person with an open hand. In doing so, the person would solve the grasping problem for the robot. We believe that applications such as this, where people and robots work closely together to intuitively perform sophisticated manipulation tasks hold great promise for applications in areas such as manufacturing and healthcare.

3) Safety: Robots working with people must be safe for human interaction. Traditional industrial manipulators are dangerous and people are kept behind a fence, away from the robot. 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. The danger depends on the velocity, the mass and the compliance of the manipulator. Commercially available arms such as the Manus ARM, the Katana arm from Neuronics and the Kuka light-weight arm (based on the DLR arm) are beginning to address these issues. Also, research into manipulators that incorporate elastic elements in the robot's drive train has made progress, including work at MIT on Series Elastic Actuators and at Stanford on the DM2 manipulator.

D. Platform Design

Careful design and use of the robot's body can reduce the need for perception and control, compensate for uncertainty, and enhance sensing. For example, the body of the Roomba vacuum is a low-profile disc. This allows it to avoid getting snagged in tight spaces and trapped under beds, greatly limiting its need to sense the details of someone's home. In the following sections we discuss ways that researchers are addressing the challenges of human environments through the design and use of the robot's body.

1) On Human Form: Human environments tend to be well-matched to the human body and human behavior. 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 at table height. It is easier to perceive these objects if the robot's sensors are looking down at the surface, which requires that the robot's sensors be high off of the ground. Similarly, everyday handheld objects, such as tools, are designed to be grasped and manipulated using a human hand. A gripper that has a similar range of grasp sizes, will tend to be able to grasp everyday human objects. A direct approach to taking advantage of these structural properties of human environments is to create humanoid robots that emulate the human form, but mobile manipulation platforms can also emulate critical features such as a small footprint, sensors placed far above the ground, an approximately hand sized gripper, and robot arms with approximately human size and degrees of freedom (see Figure 11). [1]

[1]Attendees agreed that the lack of affordable off-the-shelf robotic platforms suitable for manipulation in human environments is a serious impediment to research.

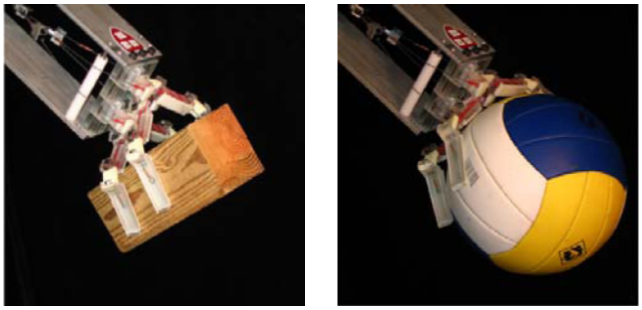


Fig. 12. A compliant grasper developed by Aaron Dollar and Robert Howe at Harvard leverages its physical design to grasp unknown objects [1]. (Permission not yet acquired)

2) 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 work with uncertainty.

Aaron Dollar and Robert Howe from Harvard optimized several parameters in the design of a robot hand so that it could better grasp objects with uncertain positions (see Figure 12). The hand is made entirely out of compliant, urethane materials of varying stiffness. It has embedded tactile and position sensors and is actuated by remote motors through tendons. The hand's compliance, combined with its optimized design, allows it to robustly form power grasps on a variety of objects. Remarkably, the hand is also robust to sustained impacts from a hammer.

Our humanoid robots developed at MIT use compliant, Series Elastic Actuators at all the joints of the arms and hands. They also have compliant, rubber skins on the fingers. This passive compliance allows them to safely explore unknown environments without risk of damaging the manipulator drive train. On our robot Domo, shown in 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 passive compliance allows the the object's flat base to stably align with the shelf surface [3]. On our robot Obrero [2], the compliance in the fingers (see Figure 7) allows the robot to gently come in contact with light objects without knocking them over and allows its hand to conform to unknown objects.

E. Control

Within perfectly modeled worlds, motion planning systems perform extremely well. Once the uncertainties of human environments are included, alternative methods for control become important. For example, control schemes must have real-time capabilities in order to reject disturbances from unexpected collisions and adapt to uncertain changes in the environment, such as might be caused by a human collaborator. As we've previously mentioned, in our work at MIT, we frequently use visual servoing, since we believe that tight, closed-loop, sensory-motor control is advantageous. Rob Platt and the Robonaut group at NASA/JSC, and Rod Grupen's group at UMass have explored ways to learn and compose real-time, closed-loop controllers in order to flexibly perform a variety of autonomous manipulation tasks. Oliver Brock's group at UMass Amherst is looking at ways to extend planning based approaches, so that they may rapidly adapt to changes in the world.

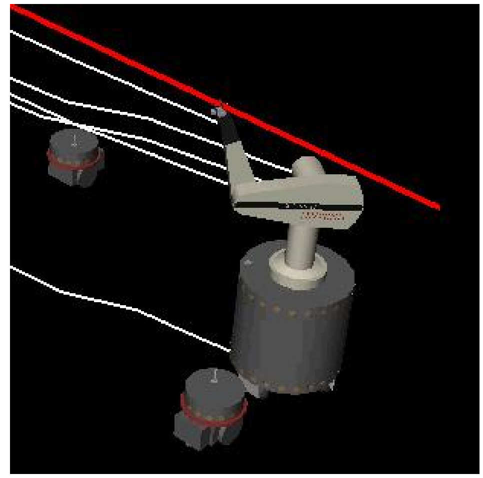


Fig. 13. The elastic roadmap planner, developed by researchers as UMass Amherst, allows for autonomous execution of mobile manipulation tasks in unstructured, dynamic environments. [5]. (Permission not yet acquired)

V. 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 grand challenges that encapsulate many of the important themes of this research domain. The agreed upon challenges were: cleaning and organizing a house, preparing and delivering an order at a burger joint, and working with a person to assemble a habitat (a tent).

Each challenge emphasizes different aspect of the field. A robot that can enter any home and clean up a messy room must adapt to the large variability of our domestic settings, understand the usual placement of everday objects, and be able to grasp-carry-and-place everyday objects including clothing. Preparing and delivering an order at a burger joint would require the robot to dexterously manipulate flexible materials and tools designed for humans, and perform a variety of small, but complex, assembly tasks. Assembling a habitat such as a tent requires that the human and robot cooperate in a coordinated fashion. It also requires that the assembly proceed with a partial ordering, and often requires coordination involving fixturing, insertion, and lifting.

A. Smooth Paths to Progress

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

1) By Approach: We expect progress to be made for each of the approaches we have discussed within this paper. Yet, the problem of robot manipulation for human environments necessitates integration of these approaches into solutions that are validated in the real-world on tasks with clear measures of success. These intelligent systems will combine hardware with software, and perception with action.

2) By Module & Algorithm: We would expect research to result in de facto standards for modules and algorithms that perform various important tasks, such as grasping. We already see this to some extent with face detectors, low-level vision algorithms, and machine learning algorithms. As researchers are able to make use of one another's components, progress will accelerate and they will be better able to verify one another's work through repetition.

3) From Semi-autonomy to Full Autonomy: As we have previously mentioned, a useful direction for progress is to focus on semi-autonomous, human-in-the-loop systems. This direction gives a clear path for incrementally increasing the autonomy of systems, while allowing humans to take-over when the system is having trouble.

4) From Simple to Complex Tasks: Technically, the Roomba could be considered the first successful autonomous mobile manipulator for the home. It manipulates dirt on the floor as it moves around the home. By narrowing the scope of a task, useful robots may be developed more quickly and serve as a base for further capabilities. Rather than push for highly complex tasks, many researchers are focusing on simpler, foundational 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 (eg. cylindrical) or the domains of the environment it interacts with (eg. accessible flat surfaces such as desks and tables).

B. Conclusion

Robot manipulation in human environments is a young research area, but one that is certain to expand rapidly in the coming years. Without advances in robot manipulation, many promising robotic applications will not be possible. We have presented our perspective on the challenges facing the field and proposed paths towards the long-term vision of robots that can work alongside us in our homes and workplaces as useful, capable collaborators.

Acknowledgment

The authors would like to thank all of the participants at the workshop, and give special thanks to those who presented a paper, poster, or demo. We'd also like to thank our coorganizers: Lijin Aryananda, Paul Fitzpatrick, and Lorenzo Natale, as well as our invited speaker Rod Brooks. We encourage the reader to go to the workshop's website for further information, including downloadable papers and posters: `http://manipulation.csail.mit.edu/rss06/` or to the official archived online proceedings: `http://www.archive.org/details/manipulation_for_human_environments_2006`

References

[1] C. Kemp, A. Edsinger, E. Torres-Jara, L. Natale, P. Fitzpatrick, and L. Aryananda *Proceedings of the Robotics Science and Systems Workshop on Manipulation for Human Environments*. http://manipulation.csail.mit.edu/rss06/ . Philadelphia, Pennsylvannia, August, 2006.

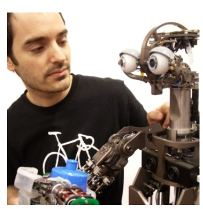
[2] Lorenzo Natale and Eduardo Torres-Jara, *A Sensitive Approach to Grasping*, Sixth international Conference on Epigenetic Robotics, Paris, France, 20-22 September, 2006.

[3] Aaron Edsinger and Charles Kemp, *Manipulation in Human Environments*, To appear in the Proceedings of the IEEE/RSJ International Conference on Humanoid Robotics, Genoa, Italy, 2006.

[4] Charles C. Kemp and Aaron Edsinger, *Visual Tool Tip Detection and Position Estimation for Robotic Manipulation of Unknown Human Tools*, Massachusetts Institute of Technology, CSAIL, AIM Tech. Report, AIM-2005-037, December, 2005.

[5] Yuandong Yang and Oliver Brock, Elastic Roadmaps: Globally Task-Consistent Motion for Autonomous Mobile Manipulation in Dynamic Environments. Proceedings of Robotics: Science and Systems, Philadelphia, USA, 2006.

Charles C. Kemp Charles C. Kemp is currently a Senior Research Scientist and director of the Center for Healthcare Robotics in the Health Systems Institute at Georgia Tech. He holds a Ph.D., M.Eng., and B.S. from MIT 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 the home and workplace.



Aaron Edsinger Aaron Edsinger is currently a doctoral candidate in computer science at the MIT Computer Science and Artificial Intelligence Laboratory. He holds a B.S. in Computer Science from Stanford University and a S.M. 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 manipulator and hand design, as well as hand-eye coordination.

Eduardo Torres-Jara Eduardo Torres-Jara received his Ingeniero degree in Electrical Engineering from Escuela Politécnica del Ejército, Ecuador, and his M.S. from MIT in Electrical Engineering and Computer Science. He is currently a Ph.D. candidate at MIT CSAIL. His current research interest is in *sensitive manipulation*. Manipulation that uses dense tactile sensing and force feedback as its main input. Manipulation that is about action as much as perception. His work includes tactile sensing and compliant actuator design, and behavior-based architectures.