Human-Robot Interaction for Cooperative Manipulation: Handing Objects to One Another

Aaron Edsinger
Computer Science and Artificial Intelligence Laboratory
Massachusetts Institute of Technology
edsinger@csail.mit.edu

Charles C. Kemp
Health Systems Institute
Georgia Institute of Technology
charlie.kemp@hsi.gatech.edu

***Abstract*— For manipulation tasks, the transfer of objects between humans and robots is a fundamental way to coordinate activity and cooperatively perform useful work. Within this paper we demonstrate that robots and people can effectively and intuitively work together by directly handing objects to one another.**

First, we present experimental results that demonstrate that subjects without explicit instructions or robotics expertise can successfully hand objects to a robot and take objects from a robot in response to reaching gestures. Moreover, when handing an object to the robot, subjects control the object's position and orientation to match the configuration of the robot's hand, thereby simplifying robotic grasping and offering opportunities to simplify the manipulation task.

Second, we present a robotic application that relies on this form of human-robot interaction. This application enables a humanoid robot to help a user place objects on a shelf, perform bimanual insertion tasks, and hold a box within which the user can place objects. By handing appropriate objects to the robot, the human directly and intuitively controls the robot. Through this interaction, the human and robot complement one another's abilities and work together to achieve results. [1]

I. Introduction

Robots that work alongside people in their homes and workplaces could extend the time an elderly person can live at home, provide physical assistance to a worker on an assembly line, or help with household chores. Human environments present special challenges for robot manipulation since they are complex, dynamic, uncontrolled, and difficult to perceive reliably. By working with people, robots can more easily overcome these challenges and provide worthwhile services.

The transfer of objects between humans and robots is a fundamental way to coordinate activity and cooperatively perform useful work. A robot can leverage a person's familiarity with cues such as physical contact, reach direction, and grasp shape in order to facilitate object transfer. Within this paper, we look at cooperative manipulation tasks involving a single human, a single robot, and objects that can be considered to be exclusively controlled by the robot, exclusively controlled by the human, or briefly in transition between these two states. A wide variety of tasks fit this structure.

Many possibilities exist for transfering control of an object between a robot and a human. For potential users of manipulation assistance technology, the tasks cited as being most useful include preparing food, picking items up from the floor, and placing items on a shelf [12]. These tasks could potentially include some form of cooperative object hand-off. Recent results with the NASA Robonaut [5], the AIST HRP-2 [11], and the HERMES robot [2] have included the handing of objects between a humanoid and person. However, these projects have yet to consider object transfer in detail. Within [3], Breazeal et. al. maintain that during a collaborative task, a robot requires understanding of a person's intentions and desires in order to behave as a partner rather than just a tool. However, within this paper we show that a robot and a human can effectively and intuitively work together using only simple social and physical cues.

In this paper, we first present experimental results that demonstrate that subjects without explicit instructions or robotics expertise can successfully hand objects to a robot and take objects from a robot in response to reaching gestures. Moreover, when handing an object to the robot, subjects control the object's position and orientation to match the configuration of the robot's hand, thereby simplifying robotic grasping and offering opportunities to simplify the manipulation task.

We then present a robotic application that relies on this form of human-robot interaction. This application enables a humanoid robot to help a user place objects on a shelf, perform bimanual insertion tasks, and hold a box within which the user can place objects. By handing appropriate objects to the robot, the human directly and intuitively controls the robot. Through this interaction, the human and robot complement one another's abilities and work together to achieve results.

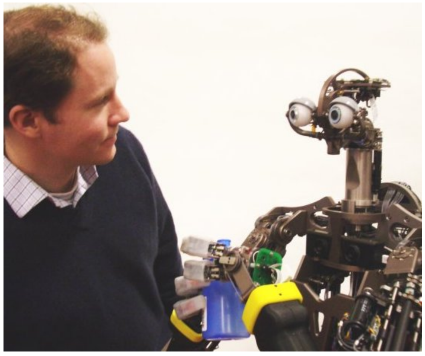

Fig. 1. The humanoid robot Domo is designed to assist people with everyday manipulation tasks.

[1]This work was sponsored by Toyota Motor Corporation: Autonomous Manipulation Capabilities for Partner Robots in the Home.

II. IMPLEMENTATION

Our work is implemented on the 29 degree-of-freedom humanoid robot, Domo, pictured in Fig. 1. In this section we present key components of Domo's design that allow it to safely interact with a person, hand objects to a person, and receive objects from person.

A. Safe, Physical Interaction

In order to directly hand objects to one another, the human and robot must work in close proximity. This requires great attention to safety. Researchers have developed a variety of approaches for safe robots [8], [1]. Our robot, Domo, uses passive compliance and force sensing actuators throughout its body [7]. These Series Elastic Actuators (SEA) lower the mechanical impedance of its arms, allowing for intrinsically safe physical interaction with a person given moderate end-effector velocities [10], [14]. The Head Injury Criterion (HIC) is a commonly used index to evaluate robot manipulator safety [13]. An HIC value near 100 can be safe for human contact while an HIC of 1000 is potentially fatal. As we have discussed in [6], given a hand velocity $1.0m/s$ Domo's manipulator has an HIC of approximately 167, while the Puma 560 has a substantially higher estimated HIC of 550.

B. Detecting Object Transfer

Domo detects when an object has been placed in its hand, attempts to grasp the object, and then detects whether or not the grasp has been successful. Domo also detects when the user attempts to acquire an object from Domo's grip. This section describes the implementation of this functionality and indicates the names of the relevant perception and control modules in *italics*.

1) Detecting Hand Velocity: Domo must decide when to close its gripper in order to grasp an object being handed to it by a human. Many options exist for detecting when an object has been placed in a robot's hand, including tactile sensing and range sensors in the gripper. Moreover, autonomous grasping often requires that a similar decision be made. For the work described here, Domo lowers the stiffness of the virtual springs used to control its arm (*StiffnessAdapt*) and monitors the velocity of its hand in order to detect when an object has been placed in its hand, or is being pulled out of its hand (*ContactDetect*).

We used support vector regression (SVR) with a Gaussian RBF kernel, as implemented in the LIBSVM package [4], to model the maximum expected hand velocity in the absence of external disturbances as a function of arm stiffness [6]. When Domo's hand moves with a velocity greater than this maximum expected velocity (*ContactDetect*), Domo compliantly closes its hand using force control (*PowerGrasp*).

2) Detecting a Successful Grasp: Upon closing its hand, Domo estimates the resulting grasp aperture, which gives a strong indication of whether or not Domo has successfully grasped an object. The grasp aperture, typically defined as the distance between the thumb and forefinger, is a common measure used when studying human manipulation [9]. On a robot, the grasp aperture can be used to estimate the size of an unknown, grasped object. For example, the grasp aperture created by a power grasp on a cylinder is proportional to the cylinder's diameter.

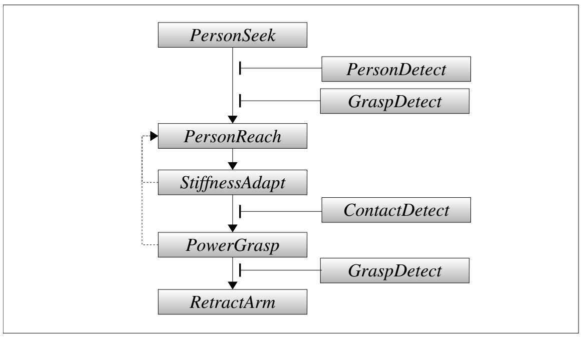

Fig. 2. The *AssistedGrasp* module gains assistance from a person in order to grasp an object. Control state transitions (arrows) occur contingent on perceptual feedback (bars).

On Domo, there is substantial compliance in the finger. The compliant fingertip and skin also allow the finger surface to deform during grasping. These factors would significantly complicate model-based estimation of the grasp aperture. Instead, we once again used support vector regression (SVR) with a Gaussian RBF kernel to perform supervised, offline learning. The resulting function takes the hand's four joint angles as input and outputs an estimate of the object's diameter. Training data was gathered for 50 power grasps formed on five cylindrical objects of known diameters between 25mm and 75mm [6].

GraspDetect signals that a stable grasp has been made on an object. It relies on three conditions to detect a grasp. First, it monitors the net torque applied by the fingers. If it is positive and above a threshold, then the hand is assumed to be closed or closing with significant force. Second, if the net angular velocity of the fingers is close to zero, it is assumed that the fingers are in a stable state. Third, if the estimated grasp aperture is greater than $20mm$, then the fingers are assumed to be wrapped around an object and not resting on the palm of the hand. If all three conditions are true, then *GraspDetect* signals a stable grasp.

C. Transfer of Objects between a Human and Robot

The *AssistedGrasp* algorithm puts all of the previous modules together to acquire an object from the human. The algorithm is depicted in Fig. 2. If at least one hand is empty, *AssistedGrasp* can be activated. First, it employs *PersonSeek* to find a person in the room. If *PersonDetect* finds a face at a location, then *PersonReach* reaches below this location, approximately toward the person's midriff.

Once *PersonReach* has achieved its target and the arm is nearly stationary, *StiffnessAdapt* lowers the stiffness of the arm. This increases the likelihood of *ContactDetect* given small contact forces. As the person gives Domo the object, small displacements of the hand are sensed by *ContactDetect*. *PowerGrasp* then closes the fingers around the object. If *GraspDetect* signals success, *RetractArm* brings the grasped object to the robot's side. However, if *ContactDetect* or *GraspDetect* fail, then *PersonReach* is reactivated and the robot cues the person again.

While *AssistedGrasp* takes an object from a person, the *AssistedGive* module hands a grasped object back. Its implementation is nearly identical to *AssistedGrasp,* except that *GraspRelease* is used instead of *PowerGrasp*.

III. Testing Cooperative Manipulation

When designing a robotic application that involves cooperative manipulation, one must consider the roles of both the human and the robot. By placing the human "in the loop", a robot can be useful without achieving full autonomy. However, in order for applications to be successful, the cooperation must result in a net benefit for the human. Through human-like form and behavior, robots may be able to reduce the burden of cooperation. Humans could potentially use robots intuitively without specialized instruction.

Socially understood gestures and cues that facilitate collaboration may be relatively easy for a robot to generate. The human can then use the task context to interpret these signals appropriately. For example, *AssistedGrasp* implicitly assumes that a collaborator will understand the reaching cue as a request for an object, and consequently hand an appropriate object to the robot. People exhibit similar behavior when interacting with one another. For example, one would expect a person to pass a coffee cup, instead of a dinner plate, to a waitress holding a coffee pot.

Over time, as a person becomes familiar with a robot, we would expect the person to adapt his or her behavior to make better use of the robot and increase the chances of task success. As we will show with the following experiments and a demonstration application, during *AssistedGrasp,* subjects intuitively hand an object to the robot in a pose that anticipates both the robot's use of the object and the limitations of the robot's grasping.

In this section we present experimental results that demonstrate that subjects without explicit instructions or robotics expertise can successfully hand objects to a robot and take objects from a robot in response to reaching gestures. Moreover, when handing an object to the robot, we show that subjects control the object's position and orientation to match the configuration of the robot's hand, thereby simplifying robotic grasping.

A. The Give and Take Experiment

1) Experimental Setup: As shown in Fig. 3, the subject sits in front of the robot. The robot is at a table and an oblong box (60mm $\times$ 85mm $\times$ 200mm) sits on the table. The box is instrumented with a an inertial measurment unit (IMU) to measure its orientation. The subject is told only that the robot is performing an unspecified visual hand-eye calibration task, and that whenever the robot reaches to them, he or she is to place the box in the robot's hand. This explanation is to deter the subject from explicitly considering the way in which he or she hands the box to the robot. In a single trial, the robot reaches to the subject with its hand open in a power-grasp configuration (preshaped). The orientation of the open hand is varied per trial. The subject places the box in the robot's hand and the robot grasps the box, brings the box up to its cameras, appears to inspect it, and then lowers its arm in one of two ways.

In the first case, the robot reaches towards the person, bringing the box just in front of and above the table edge nearest the subject. It says the word "Done" and pauses for one second. It then releases its grasp, dropping the box onto the table, and retracts its arm. In the second case, the robot does an identical action as in the first case, but this time it reaches just past the table edge. Unless the subject takes the box from the robot, it falls to the floor. This marks the end of a trial. The robot pauses for 5 seconds and then initiates the next trial. Six trials are performed with each subject.

At the start of each experiment, the box is aligned to the robot's body and the IMU is calibrated with respect to the world frame $\{W\}$. We define the vector $\mathbf{b}^W$ as the longest edge of the box. We define the power-grasp axis as $\mathbf{z}^H$ in the hand's coordinate frame $\{H\}$. This axis corresponds to the long axis of a cylinder held in a power grasp. In frame $\{W\}$ this axis is $\mathbf{z}^W$. The angle between $\mathbf{z}^W$ and $\mathbf{b}^W$ is defined as the grasp alignment error. During each trial, we measure the average grasp alignment error during the 500ms just prior to the grasp being formed. We also vary the wrist rotation for each of the six trials such that the angle formed between $\mathbf{z}^W$ and gravity is $[0°, -45°, 90°, 0°, -45°, 90°]$.

2) Experimental Hypothesis: This experiment considers the following three questions:

1) When a subject hands the robot the box, do they adjust its orientation to match the pose of the robot's hand?
2) Will the subject correctly interpret the robot's reaching gesture, vocalization, and pause as a social cue to take the object?
3) Can a small incentive such as not having to pick up the object increase the subject's willingness to respond to the social cue?

We use the measured grasp alignment error to answer the first question. We would expect to see $\mathbf{b}^W$ track $\mathbf{z}^W$ as it varies between the three wrist orientations. The second question is more difficult to confirm. For each experiment, we measure the take-back rate as the number of trials a subject reached to take the box back from the robot. We expect that the subject will take back the box when the robot performs its reaching gesture, vocalization, and pause. The subjects are never instructed to take the object back. In order to address the third question, the trials are varied so that for half of the subjects the robot drops the cylinder on the table, and for the other half the robot drops the cylinder on the floor. If dropping the cylinder on the floor serves as an incentive for the subject to take the box back prior to the drop, we should see an increase in the take-back rate. Importantly, the arm postures achieved by the incentive-reach and the no-incentive-reach are nearly identical, in order to minimize the differences between the reaching cues.

3) Experimental Results: Prior to the experiments, we first measured the average grasp alignment error when we deliberately oriented the box to match the robot's grasp. From repeated trials the mean error was 8.9°. Next, we measured the range of grasp alignment errors that are possible when the object is placed in the robot's open hand, by freely rotating the box within the open hand (preshaped). In this case the distribution of grasp errors was fairly uniform between $0°$ and $60°$. We tested the experiment using 10 subjects (6 female, 4 male in the ages of 18-55). All subjects were naive to the experimental objectives and had little if any prior experience working with robots. In Fig. 4, we see the grasp alignment errors for each of the six trials of a typical subject. All subjects matched the orientation of the offered box to the orientation of the robot's hand

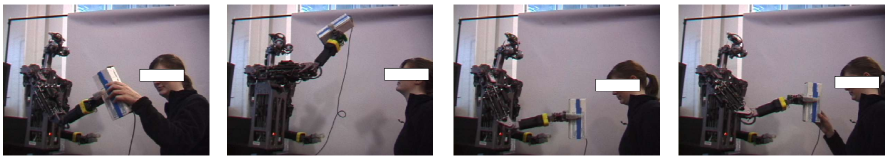

Fig. 3. One trial of the Give and Take experiment.

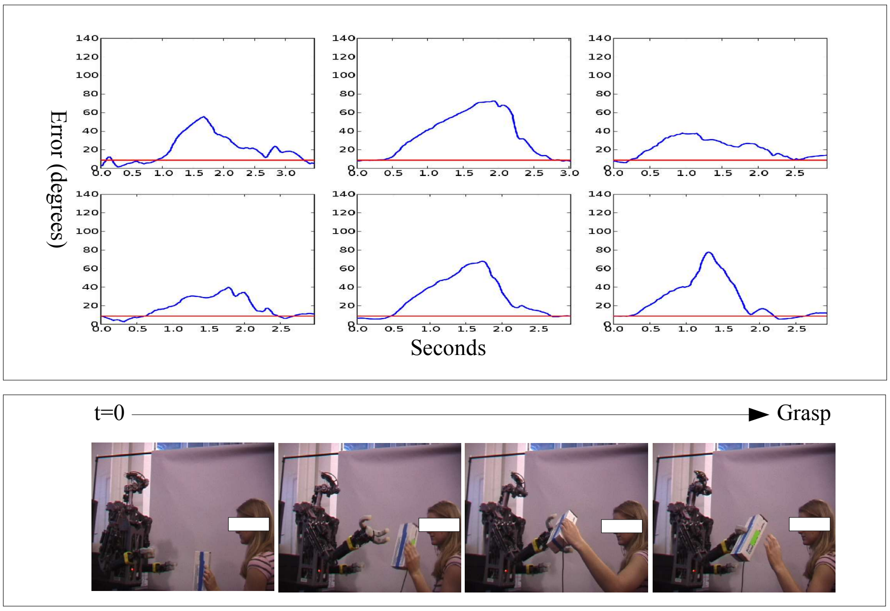

Fig. 4. The grasp alignment errors from six trials with a typical subject during the Give and Take experiment. We see that for nearly all trials, the subject aligns the box within a few degrees of the best expected performance. (Top) Blue shows the grasp alignment error (degrees) of the box with respect to the open grasp configuration (preshaped). Red (horizontal) shows the mean error achieved when we deliberately aligned the box with the grasp. The X axis shows the trial time, in seconds, starting when the reach commenced and ending when the grasp was initiated. (Bottom) The execution sequence from the initiation of the reach until the grasp is formed.

with surprising accuracy. In Fig. 5, we see that the average error for each subject is near the average error achieved by one of the authors. We also see that, when offered the box by the robot, $90theboxwasheldoverthefloor,100$one or more times, and the subjects let the box drop to the ground in only 1 out of 30 trials. The possibility of the box falling to the ground appears to influence the subject's behavior since subjects allowed the robot to drop the box onto the table in 11 out of 30 trials. This experiment only scratches the surface of the potentially rich interactions that may occur during cooperative manipulation. However, it shows quantitatively that people will intuitively adapt to and assist the robot without instruction. We would expect that more substantiative assistance could be given if the person possessed greater contextual knowledge about the task and the robot could generate more nuanced cues.

IV. Application

In this section we show that a robot and a person can work together by directly handing objets to one another, and that this form of interaction can support a variety of everyday manipulation tasks. These results are presented in more detail elsewhere [6].

As shown in Fig. 6, a collaborator can ask the robot to take an object (*AssistedGrasp*), give back an object (*AssistedGive*), insert one object into another (*BimanualInsert*), place an object on a shelf (*ShelfPlace*), or hold a box while objects are placed in it (*BimanualFixture*). These manual skill behaviors run concurrently, allowing a person to vocally request them at any time. If the collaborator notices that Domo is failing at a task, he or she can provide vocal (*VocalRequest*) or contact (*ContactDetect*) feedback to alert the robot. If Domo accidentally drops an object (*GraspDetect*), the person can pick it up and ask the robot to grasp it again (*AssistedGrasp*). Alternatively, at anytime the person can ask Domo to hand over a grasped object (*AssistedGive*). Through these behaviors, the robot and the person work as a team. The person intuitively provides task-level planning and guides the robot's action selection. By handing objects to the robot and taking objects from the robot, the person directly and unambiguously specifies the objects that the robot should use, while also helping the robot grasp them in a manner appropriate for the task.

The following example scenario for cooperative manipulation illustrates the utility of these behaviors:

1) Domo is positioned at a table cluttered with objects and near a shelf. Domo first physically verifies the location of the shelf.
2) A person asks for help in preparing a drink. He hands Domo a cup and bottle of juice. Domo "pours" the juice into the cup.

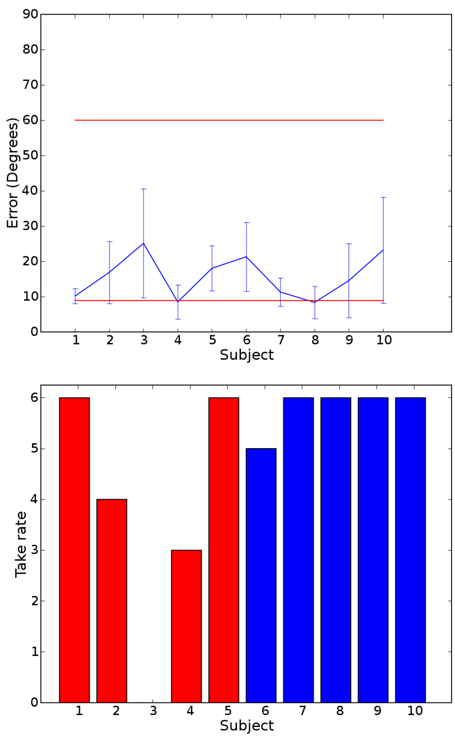

Fig. 5. (Top) Results showing that the 10 subjects intuitively aligned the cylinder's axis with the grasp axis of the robot's hand when handing it to the robot. For each of 6 trials, we measure the average alignment error, in degrees, during the 500ms prior to grasping. The error bars show the mean error and standard deviation for each subject. The red lines indicate the best and worst expected performance. (Bottom) Results showing the number of trials each subject took the box back from the robot when cued. For subjects 1-5 (red), the robot dropped the box on the table, while for subjects 6-10, the robot provided incentive by dropping the box on the floor if the subject did not take it back.

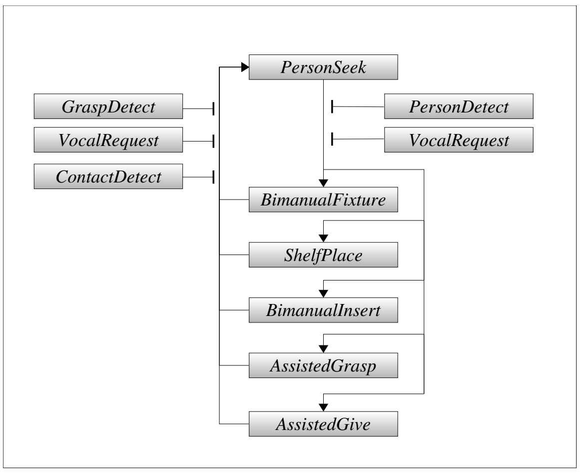

Fig. 6. A collaborator can compose a task by coordinating its manipulation skills using voice cues (*VocalRequest)* while the robot tracks the person in the scene (*PersonSeek*, *PersonDetect*). The person can ask the robot to take an object (*AssistedGrasp*), give back an object (*AssistedGive*), insert one object into another (*BimanualInsert*), place an object on a shelf (*ShelfPlace*), or hold a box while objects are placed in it (*BimanualFixture*). The robot can reattempt a manual skill if failure is signaled (*GraspDetect, VocalRequest, ContactDetect*).

3) Domo hands the bottle of juice back to the person.
4) The person now hands Domo a spoon. Domo inserts the spoon into the cup and "stirs" the drink.
5) Domo hands the spoon back to the person and then places the prepared drink on the shelf.
6) Next, the person asks for help in putting away groceries. He hands Domo a box of crackers. Domo passes the box to the other hand and puts them upright on the shelf.
7) The person hands Domo a paper bag of coffee and Domo places it on the shelf as well.
8) Now, the person asks for help in clearing off the table. He hands Domo a box and Domo grasps it with both hands.
9) Domo keeps the box near the person as he goes about clearing the table into it.
10) Finally, the task is done and Domo lowers the box onto the table.

As shown in Fig. 7, a very similar scenario was realized by Domo and the author as one consecutive task, punctuated by vocal requests for the robot, over the course of 5 minutes. Of course, other scenarios are possible using this approach. For example, Domo could assist a person working on an assembly line by holding a tool tray for the person, putting tools away, holding a tool and then handing it back when the person is ready, and performing the insertion of two parts during assembly.

V. Discussion

Within this paper, we demonstrated that a robot and a human can effectively and intuitively work together by directly handing objects to one another. This mode of interaction takes advantage of the complementary skills of the human and the robot, and offers several distinct benefits. Objects can be efficiently transferred between the human and the robot with minimal delay between the human's hand and the robot's end effector. The human solves a potentially difficult grasping problem for the robot by directly placing the object within the robot's hand in a favorable configuration. As we have shown empirically, the human does not require explicit instruction or expertise to work with robots in this way, possibly because of the similarity with the way people work with one another. The robot reaches toward the person, which can simplify the transfer for the human. This could be especially important for assistive applications for people with motor impairments. As we have demonstrated, handing objects to one another can also serve as a useful interface, since by handing an object to the robot, the human implicitly commands the robot to manipulate that particular object in the near future.

An interesting avenue for future work would be the refinement of the robot's reaching gestures. Reaching into the human's interpersonal space has social, ergonomic, and biomechanical implications that are worthy of study. Continued research in this area may lead to human-robot interaction that is more intuitive, more comfortable, and more efficient.

References

[1] A. Bicchi, G. Tonietti, M. Bavaro, and M. Piccigallo. Variable stiffness actuators for fast and safe motion control. In B. Siciliano,

Fig. 7. In this sequence, Domo assists in a variety of manual tasks. (A) Domo begins at a cluttered table. (B) A shelf appears and Domo verifies its location. (C-D) A juice bottle and cup are handed to Domo. (E) Domo visually guides the bottle into the cup. (F-G) Now, Domo is handed a spoon and it "stirs" the drink. (H) Finally, Domo puts the finished drink on the shelf. (I-L) A box of crackers is handed to Domo's right hand. It transfers them to the left hand and places them upright on the shelf. (N-0) A bag of coffee beans is handed to Domo. This time, it uses its left hand because of the person's proximity. It then puts the bag on the shelf. (P) Domo forms a bimanual grasp on a box. (Q-R) Domo keeps the box near the person as they clean up the table and put items in the box. (S-T) Now that the task is done, Domo lowers the box onto the table

O. Khatib, and F.C.A. Groen, editors, *Proceedings of ISRR 2003*, Springer Tracts in Advanced Robotics (STAR). Springer Verlag, 2003.

[2] R. Bischoff and V. Graefe. Design principles for dependable robotics assistants. *International Journal of Humanoid Robotics*, 1(1):95–125, 2005.

[3] C. Breazeal, A. Brooks, D. Chilongo, J. Gray, G. Hoffman, C. Kidd, H. Lee, J. Lieberman, and A. Lockerd. Working collaboratively with humanoid robots. In *Proceedings of IEEE-RAS/RSJ International Conference on Humanoid Robots*, Santa Monica, CA, 2004.

[4] Chih-Chung Chang and Chih-Jen Lin. LIBSVM: a library for support vector machines, 2001.

[5] M. Diftler, R. Ambrose, K. Tyree, S. Goza, and E. Huber. A mobile autonomous humanoid assistant. In *Proceedings of the 4th IEEE/RAS International Conference on Humanoid Robots*, 2004.

[6] Aaron Edsinger. *Robot Manipulation in Human Environments*. PhD thesis, Massachusetts Institute of Technology, February 2007.

[7] Aaron Edsinger-Gonzales and Jeff Weber. Domo: A Force Sensing Humanoid Robot for Manipulation Research. In *Proceedings of the 2004 IEEE International Conference on Humanoid Robots*, Santa Monica, Los Angeles, CA, USA., 2004. IEEE Press.

[8] O. Khatib, K. Yokoi, O. Brock, K. Chang, and A. Casal. Robots in human environments: Basic autonomous capabilities. *International Journal of Robotics Research*, 18(684), 1999.

[9] M. Mon-Williams and J.R. Tresilian. A simple rule of thumb for elegant prehension. *Current Biology*, 11(13):1058–1061, July 2001.

[10] G. Pratt and M. Williamson. Series Elastic Actuators. In *Proceedings of the IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS-95)*, volume 1, pages 399–406, Pittsburg, PA, July 1995.

[11] N. Sian, K. Yoki, Y. Kawai, and K. Muruyama. Operating humanoid robots in human environments. In *Proceedings of the Robotics, Science & Systems Workshop on Manipulation for Human Environments*, Philedelphia, Pennsylvania, August 2006.

[12] C. Stanger, C. Anglin, W. Harwin, and D. Romilly. Devices for assisting manipulation: A summary of user task priorities. *IEEE Transactions on Rehabilitation Engineering*, 2(4):256–265, December 1994.

[13] J. Versace. A review of the severity index. In *Proceedings of the 15th Stapp Car Crash Conference*, pages 771–796, 1971.

[14] M. Zinn, O. Khatib, B. Roth, and J.K. Salisbury. Playing it safe: human-friendly robots. *IEEE Robotics & Automation Magazine*, 11(2):12–21, June 2004.