Human-Robot Control Strategies for the NASA/DARPA Robonaut

M. A. Diftler
Automation and Robotics Dept.
Lockheed Martin
Houston, Texas 77058
diftler@jsc.nasa.gov

C. J. Culbert, R. O. Ambrose
Robotic Systems Technology Branch
NASA Johnson Space Center
Houston, Texas 77058
Chris.Culbert@jsc.nas.gov
Robert.Ambrose@jsc.nasa.gov

E. Huber
Metrica
Johnson Space Center
Houston, Texas 77058
huber@jsc.nasa.gov

W. J. Bluethmann
Hernandez Engineering, Inc.
Johnson Space Center
Houston, Texas 77058
bluethmann@jsc.nasa.gov

Abstract- The Robotic Systems Technology Branch at the NASA Johnson Space Center (JSC) is currently developing robot systems to reduce the Extra-Vehicular Activity (EVA) and planetary exploration burden on astronauts. One such system, Robonaut, is human scale, anthropomorphic, and designed to approach the dexterity of a space-suited astronaut. Robonaut can perform numerous human rated tasks, including actuating tether hooks, manipulating flexible materials, soldering wires, grasping handrails to move along space station mockups, and mating connectors. More recently, developments in autonomous control and perception for Robonaut have enabled dexterous, real-time man-machine interaction. Robonaut is now capable of acting as a practical autonomous assistant to the human, providing and accepting tools by reacting to body language. A versatile, vision-based algorithm for matching range silhouettes is used for monitoring human activity as well as estimating tool pose.

INTRODUCTION

The requirements for extravehicular activity (EVA) onboard the International Space Station (ISS) are considerable. These maintenance and construction activities are expensive and hazardous. Astronauts must prepare extensively before they may leave the relative safety of the space station, including pre-breathing at space suit air pressure for up to 4 hours. Once outside, the crew person must work very carefully to prevent damage to the suit.

Future human planetary exploration missions may involve habitat construction, systems maintenance, geological exploration, material’s processing, launch and landing preparations, scientific instrument manipulation, and other tasks that expose humans to dangerous or risky environments.

The Robotic Systems Technology Branch at the NASA Johnson Space Center (JSC) is currently developing robot systems will help reduce the amount of EVA and planetary exploration activities astronauts have to perform and also to serve in rapid response capacities. One such system, Robonaut, a humanoid robot, is capable of interfacing with external space station systems that currently have only human interfaces and working with the same human rated tools designed for all NASA missions. Robonaut development is also supported by the Defense Advanced Research Projects Agency (DARPA) Mobile Autonomous Robotic Software program.
Humanoids are a relatively new class of robots. One of the most well known is the self-contained Honda Humanoid Robot [1], which is able to walk and even climb stairs. In the area of upper body capability several prototypes have been built that are designed to work with humans. One of the first, Greenman [2], showed the benefits of a human teleoperating a humanoid robot. WENDY (Waseda Engineering Designed sYmbiont) [3] has a full upper torso on a wheeled base and is a prototype for a possible domestic humanoid. Several humanoids have been designed specifically to explore human-robot interaction. MIT’s Cog [4] and Vanderbilt’s ISAC [5] are both remarkable platforms for such work.

These are all impressive devices, but are still prototypes and of course evolving. Unlike natural evolution, researchers from around the world are experimenting with different techniques to improve their humanoids. Fukuda, et. al.[6], provide an excellent survey of anthropomorphic robot evolution and suggest three characteristics that are most important for making a better humanoid: human like motion, human like intelligence, and human like communication.

Through several stages of mechanical design and teleoperated tests, Robonaut has evolved into a highly dexterous mechanical device, capable of remote operation. Now that is has been proven mechanically, much of the development effort is shifting towards achieving greater autonomous control.

Robonaut is a complex device. Along with a large number of actuators (DOFs), Robonaut has many sensors to measure force/torque, tactile, joint position, and joint torque as well as a stereo camera pair and a microphone. This complexity represents a welcome challenge for autonomous research. Too often, the target device has little more than wheels for actuation – making it difficult to perform anything interesting, let alone practical. Robonaut poses the contrary challenge, through teleoperation demonstrations, Robonaut has demonstrated that it is physically capable of performing useful tasks – the difficulty lies in doing them autonomously.

Of particular interest to NASA, is an autonomous anthropomorphic robot that can work closely with humans, especially suited astronauts, providing assistance during assembly, maintenance, and exploration activities. Towards this goal, the development team is developing autonomous skills which enable Robonaut to track humans, accept and provide tools, prepare a work surface, etc. – providing similar functions as those of a surgical assistant in an operating room.

To monitor human activity and interact with objects in it’s environment Robonaut relies heavily on vision. Initially, work in machine vision from prior NASA/JSC robotics projects [7][8] was transitioned to Robonaut, enabling it to track spatially isolated objects including humans. Since then, the group has developed a more sophisticated, silhouette-matching-based vision algorithm, which is capable of tracking a wide variety of objects in full (6-DOF) pose.

Force/Torque and tactile sensing also play key roles in man-machine interaction. The Robonaut team has developed sophisticated tools for parsing temporal force/torque/tactile profiles enabling Robonaut to detect and react to a human’s touch during tool exchange. To orchestrate Robonaut’s complex suite of sensors and actuators, a systematic approach to control is necessary. The development team has developed a hierarchical state-based control environment that embodies many of the lower and mid-level traits common to the great number of robot control architectures found in the AI and robotics communities.

Through use of this architecture, a milestone in autonomous control of an anthropomorphic robot has been reached. Robonaut is now capable of providing practical assistance to a human during an assembly procedure.

Visual based autonomous capabilities have recently been added to Robonaut and provide an additional control mode for a human working with Robonaut. Robonaut can now differentiate between different tools, tracking multiple tools and humans in its workspace to better facilitate astronaut/Robonaut interaction.

**NASA/DARPA ROBONAUT SYSTEM**

The requirements for interacting with space station EVA crew interfaces and tools provided the starting point for the Robonaut design. The NASA/DARPA Robonaut shown in figure 1 is equipped with two seven degree of freedom arms, two dexterous five finger hands [9], a two degree freedom neck and a head with multiple stereo camera sets, all mounted on a three degree freedom waist to provide an impressive work space. The limbs can generate a maximum force of 20 lbs and torque of 30 in-lbs, the forces required to remove and install EVA orbital replaceable units (ORUs) [10].
Figure 2: Robonaut – Astronaut size comparison

Robonaut’s hands are very human like and are able to actuate most of the astronaut’s tools. Figure 1 shows the prototype Robonaut operating a tether hook which is used by astronauts to tether themselves and their tools. As shown in figure 2, this highly anthropomorphic robot is smaller than a suited astronaut and is able to fit within the same corridors designed for EVA crew.

TELEOPERATION

Robonaut’s initial and currently most dexterous control mode is teleoperation. Actually, an immersive version of teleoperation, telepresence is the chosen technique. Using a collection of virtual reality gear, the human operator immerses himself into the robot’s environment making control extremely intuitive. The operator wears a helmet with stereo screens, stereo headphones, and a microphone linked directly to the robot’s stereo cameras, stereo microphones, and speaker, respectively. From a sensory standpoint the human operator’s “presence” is shifted to the robot. (Figure 3)

Four Polhemus™ trackers provide data to control the arms, neck, and waist, providing very human like motion. Fully instrumented Cybergloves™ are worn on both hands to control the fingers. The mapping between human and robot is relative, permitting the operator to maintain a more comfortable pose while controlling the robot’s limbs.

Numerous human rated tasks have been performed under teleoperator control. Figure 4 shows Robonaut tying a knot, demonstrating the ease with which a human’s ability to work with soft flexible materials can be transferred through the telepresence control system. Similarly a human operating Robonaut can even thread a nut onto a bolt. These are difficult tasks for a robot and will likely stay within the class of teleoperator controlled functions for some time to come.

Other tasks that are relatively easy to perform under direct human control are good candidates for more shared control and automation. Figure 5 shows Robonaut moving along the outside of a simulated Space Station module by grasping hand rails in succession. Through a combination of computer vision and grasping algorithms this task will be performed autonomously in the near future. While more difficult to completely automate, the operator workload for the electrical connector installation can be reduced by using grasps and arm motion primitives, and force control.

Figure 4: Robonaut tying a knot (L) and threading a nut onto a bolt(R).

The telepresence control paradigm combines the best of two worlds: the durability of a robot designed to work in the extremes of space, and the flexibility of a human mind immersed in the robot’s environment. Most importantly, the human is able to quickly develop and test time saving control strategies that form the basis for shared control and automation.
**SHARED CONTROL**

While direct teleoperation is still the fastest way to perform high dexterity tasks, it is not the most efficient technique for all operations. By intelligently shifting portions of control to the robot in the form of low level skills and functions, operator workload can be significantly reduced for many tasks. The Robonaut control system responds to voice commands that activate and deactivate the following example skills that are a subset of what is currently available.

**Compliance Control** - At the Johnson Space Center teleoperators have experimented with a variety of force feedback devices with varying results. In general, it is beyond current force reflection technology for a teleoperator to "feel" all the components of force, torque, and tactile feedback during a multi-body assembly procedure. But even when a force feedback device is used, local compliance control at the robot is very useful.

By controlling the stiffness [11] of the Robonaut arms, assembly forces are substantially reduced and the teleoperator does not need to be as precise during constrained motion since the robot is moving to reduce forces that are a result of misalignment. Reductions in task time and operator workload have been achieved with the addition of compliance control for the tasks shown in figure 5.

**Hand Primitives** - Using techniques developed for the NASA DART robot [11] as a starting point, a set of hand primitives have been developed and are now available for Robonaut. These primitives simplify the operator's hand motions for specific grasps: pinch, tether, spherical, splint, and drill. The spatial configuration of the fingers is modulated by the human operator and mapped into one of these primitive grasp geometries. The teleoperator uses only a few human joints to control all 12 hand joints, resulting in a decreased workload. For example, the drill primitive freezes the command to all of Robonaut's fingers except the trigger finger. In this way, the teleoperator can relax his human fingers while Robonaut maintains a firm grasp on the drill. Similarly, in spherical grasp mode the robot's fingers are spread apart, but the human maintains a comfortable hand pose while manipulating an object.

**AUTONOMY**

In keeping with the biological theme that is at the basis for developing humanoids, automated functions developed for Robonaut are distributed into various control system nodes that are analogous to the human brain's anatomy. The lowest level functions include: actuator control, motion control, safety, compliance control, tactile sensing, etc. All of these functions are implemented as part of Robonaut's brainstem. Higher level functions such as vision, memory, and grasping are located in other parts of Robonaut's brain. All communication between the distributed control system nodes passes through a well-defined Application Programmer's Interface (API) that is analogous to the thalamus in a human brain.

Robonaut's higher-level "brain" function occurs with a hierarchical state-machine-base, control environment. Control skills are packed and interact as reusable control modules. Though intended for autonomous control, this environment was designed to allow a high degree of observability and controllability by a human operator to promote ease of control strategy development.

Being the primary component in Robonaut's autonomy skill set, the bulk of this section will discuss the vision algorithms used to track humans and tools in real-time. This section concludes with a discussion of recent experiments in using Robonaut as a human assistant.

**Visual Cortex** - In order to meet the goals for Robonaut autonomy, the vision system must be capable of estimating the pose of a variety of objects. Some objects, such as the tools (wrenches, screwdrivers, rails) that Robonaut handles, are well modeled; others, such as the human head and hands, vary considerably from one instance (person) to the next. Another goal is for the vision system to be real-time to support interactions with humans in a timely and practical fashion. Also, the vision system must be tolerant of clutter — the desired object must be disambiguated from many objects of similar shape and size. Finally, partially occluded objects should yield positive identification. The last requirement is important in cases where a tool is held within a person's grasp; a human needs to be tracked even if he or she is moving around; or a wrench is not quite fully imaged within the field-of-view.

To meet these objectives, the Robonaut team developed a template-based method that enables a computer to efficiently estimate the pose (position and orientation) of objects in cluttered environments. The first phase of the approach extracts the essence of an object's shape in the form of a binary range map. The second phase involves a multi-stage, template-based search within the range map to recognize target objects (of known appearance) and determine their pose. The following subsections provide background on the techniques employed by this method.

**Binary Range Maps** - A range map is a two-dimensional array of distance measurements corresponding to points within a scene. There are a number of devices and methods for obtaining range maps, for Robonaut, the range maps are generated by processing synchronized images from the stereo pair of cameras mounted in Robonaut's head. Greyscale images are Laplacian-of-Guassing convolved,
binarized, and area-correlated at high speeds through efficient use of the Pentium-MMX register set. However, this paper will focus on the use of range maps, not the means for generating the maps themselves.

First a depth map (See Fig. 6a) of suitable spatial and depth resolution must be obtained from a capable device. The pose estimation technique requires a binary depth map as input. Each bit of the binary map, corresponding to a point in the scene, indicates whether surface material was measured within a specifically targeted distance range. To produce a binary map, a conventional depth map is band-filtered. For example, if searching for an object between 3 and 5 meters away, individual depth measurements undergo band (high and low) thresholding to produce a binary depth map selective to that range.

The binary range map provides a simple means of segmenting-out objects of interest from the rest of the scene. Optimal segmentation is achieved when the target range corresponds to the range of depths presented by the target object's surface; thereby minimizing the inclusion of non-target objects (See Fig. 6).

**Figure 6: Band Filtered Depth Map**

- **a-1.** Color-coded depth map of human, wrench, and screwdriver.
- **a-2.** Binary depth map of (a-1).
- **b-1,2.** Human depth band-filtered maps.
- **b-3.** Matching human head silhouette template.
- **c-1,2.** Wrench depth band-filtered maps.
- **c-3.** Matching wrench silhouette template.

*Shape Templates* - Templates are used by the image processing community to select for specific object views or artifacts, such as shape, color, shading, or line intersections. Templates can be either synthetically generated or derived from imagery of real target objects at specific orientations and/or distances. Often the templates are matched against many different portions of an image, representing different points within the 3D world, in
attempt to find strong correlations (match values) that may reveal the target object’s location.

If unknown, an object’s orientation can be “captured” by applying the appropriate batch of templates. The method used with Robonaut uses 2D silhouette templates to search for objects (See Fig. 7a). By matching against the entire silhouette of an object, this method lacks much of the “brittleness” associated with the more common approach of edge-based matching.

![Figure 7a: Examples of Wrench Silhouette Templates](image)

2D binary templates of an adjustable wrench representing its silhouette as viewed from different distances and orientations.

Using templates to search a scene for complex objects presents the potential for a combinatorial explosion. This is especially true if the full 6-DOF pose of a complex object is required and the scene is cluttered with other objects and artifacts. If real-time performance is an issue, then it is important that template matches be made efficiently.

To locate objects within binary range maps, the Robonaut vision system uses binary templates. Match correlation values are simply computed by summing the XOR results between individual binary pixels. By keeping data compact and the operations simple, this approach to matching templates and depth maps is fast. Using the Multi-Media registers available on conventional desktop processors, entire rows of a binary-packed template can be accessed with a single instruction, and bitwise matching can be performed in parallel.

**Pose Estimation Method** - It is difficult to match to the orientation of an object if its position and scale within the scene aren’t known. Yet it is difficult to apply templates to finding an object’s position without knowing what it looks like – and its appearance is dependant on orientation (See Fig. 7a). In short, the problem of template-base, pose estimation is one of bootstrapping.

The Robonaut approach to this problem employs several successive stages of pruning. It starts by finding a small set of templates that will likely capture the target object in any pose within the given domain. This set of templates is generic (liberal) in form, and as a side effect, non-targeted objects may also match. Successive stages use templates that are increasingly specific to the target object. As templates become more specific, they increase in fidelity; shapes are sharper making matching requirements more precise. Upon each stage more specific, they increase in fidelity; shapes are sharper making matching requirements more precise. Upon each stage foreign (non-targeted) objects are “weed-out” and only target objects remain. (See Fig 8).

High fidelity matching occurs after significant pruning is performed by earlier stages. Many more templates are required to interrogate a candidate location, but only a small fraction of image pixels remain as candidates. The next few subsections explain the approach in detail.

By this method, templates are applied through successive stages to filter out target object candidates within the scene until only the “true” candidate(s) remain. Template fidelity is increased at each stage to gain an increasingly accurate estimate of object pose. Each stage re-assess match candidate locations within the scene and passes only the best remaining candidates to the next stage. Each stage narrows down the pose search by at least one degree of freedom. See Fig. 8 for a pose estimation sequence of an adjustable wrench.

Each stage of pose estimation employs templates designed to “capture” a specific degree of freedom (DOF, component of orientation). A stage must be capable of capturing its target DOF while remaining tolerant to any remaining undetermined DOFs. To achieve this flexibility, early stages must employ liberal silhouettes, which tend to be “fuzzy” depictions of the target object. Later stages, which have fewer undetermined DOFs, can afford to apply higher fidelity templates, which more accurately reflect the appearance of the target object. In the final stage, the templates are true 2D silhouettes of the target object, providing the greatest pose estimation precision in all degrees of freedom.

**Experimental Results** - Robonaut’s stereo vision system is a key component in most of Robonaut’s autonomous skill set. Hierarchical skill sets are currently being built which combine to create complex, continuous, interactive scenarios demonstrating Robonaut as a practical human-assistant.
Figure 8: Multi-Stage Pose Matching of Adjustable Wrench

a-1. Color-coded confidence map for scale (distance (S)) match.

a-2. Template for matching scale at any orientation about Z-Y-X.

a-3. Best match patch from binary depth map.

a-4. Correlation between template (a-2) and patch (a-3).

a-5. Anti-correlation between (a-2) and (a-3).

b-1. Confidence map for Z-rotation (in-plane).

b-2. Template matching S, Z-rotation and any orientation about Y-X.

c-1. Confidence map for Z-Y rotation (in-plane).

c-2. Template matching S, Z-Y rotation and any orientation about X.


d-2. Template fully matching the pose of the imaged wrench.
One of the more interesting scenarios demonstrated recently was Robonaut as a tool handling assistant. Robonaut scans the room searching for human heads. Once found Robonaut locks-on to the human, panning and tilting his head as necessary to keep the human centered in his field of view (FOV). If the human comes close and stays there Robonaut looks down. Several possible threads of interaction occur from this point as follows:

1. If Robonaut finds a tool (in the human's hand) he takes it from the human.
2. If Robonaut finds an empty hand and Robonaut "knows" that he already possesses a tool, then he hands the tool to the human.
3. If Robonaut sees neither hand nor tool after several seconds of both electronic and mechanical (neck pan and tilt) searching, then he returns to human scan mode.

In cases (1) and (2) Robonaut's interaction with the human is sophisticated. When reaching out the robot constantly monitors the human's hand location, attempting to match it with his own. Only when the human's hand has stabilized in its position does Robonaut's hand perform a final engagement move during which force/torque sensors in his wrist are monitored for contact.

At the conclusion of a tool exchange, a confirmation test is performed. If the human refuses to give-up the tool, Robonaut recognizes a resistive force signature as he gently attempts to pull the tool away — resulting in an immediate release. Conversely, if the human fails to maintain a firm grasp of an object being handed to him, this too is recognized by a lack of force — resulting in the retention of the object within Robonaut's grasp.

For all possible outcomes of interaction Robonaut's control system is designed to "unwind" gracefully. If a tool of interest momentarily "disappears", Robonaut's hands start moving back to a neutral (home) position. When a tool interaction completes, with either success or failure, Robonaut looks up to reacquire the human. If the human suddenly leaves, the robot returns to human search mode. Using this approach to robot control, the system can be operated constantly — always ready to assist humans.

**CONCLUSION**

This report presents an overview of the visual methods used to enable Robonaut to interact with humans in an autonomous manner. This method employs several key innovative features that make it robust and fast.

The three different control strategies presented above: teleoperation, shared control, and automation, are designed to provide flexibility. These strategies combine together to form a general distributed control model shown in figure 9. Within this framework an autonomous hierarchical control system has demonstrated the ability to orchestrate sophisticated man-machine interaction strategies.

A key component of Robonaut's autonomous skill set is a flexible vision system; capable of locating both well modeled objects (such as a wrench) and loosely modeled objects, such as a "generic" human head. Through a combined strategy for perception and control, Robonaut now demonstrates a key milestone: practical assistance to a human during an assembly operation.

![Figure 9: Distributed control](image)

Robonaut's control system is continuing to evolve. Additional and improved sensors and algorithms will lead to new skills that will give both Robonaut teleoperators and humans working directly with Robonaut more capability and options in performing space based and planetary activities.

**ACKNOWLEDGEMENT**

This work is sponsored by NASA and the Mobile Autonomous Robot Software (MARS) program in the DARPA Information Processing Technology Office (IPTO)

**REFERENCES**


13. **AUTHOR BIOGRAPHIES**

Robert O. Ambrose received his B.S. and M.S. in Mechanical Engineering from Washington University in St. Louis, and his Ph.D. from the University of Texas at Austin. His publications include topics in robot design, space environmental modeling, actuator development, kinematics, kinetics, bio-mechanics, interactive design software and non-linear optimism. He has built robotic systems for space, microelectronics, nuclear and agricultural applications. Systems that he has designed include multiple manipulators, force feedback joysticks, gantries, walking machines and wheeled robots. He is currently employed at NASA’s Johnson Space Center, where he serves as the Robonaut Project Leader.

William Bluethmann is a senior engineer work for Hernandez Engineering, Inc. at NASA Johnson Space Center and currently is the software lead for the Robonaut project. His research interests include manipulator control, force control, distributed autonomy, and task level automation. He earned a PhD. in Mechanical Engineering from the University of Kansas. He may be reached at bluethmann@jsc.nasa.gov.

Chris Culbert is Chief of the Robotic Systems Technology Branch in the Automation, Robotics, and Simulation Division at NASA's Johnson Space Center (JSC). He received his B.S. in Aerospace Engineering from the University of Arizona and his M.S. in Space Sciences from the University of Houston – Clear Lake. He has published on expert system technology, verification and validation methodologies, advanced automation technologies, and the use advanced robotics for space. Contact him at Chris.Culbert@jsc.nasa.gov.

Myron A. Diftler received his B.S.E in Mechanical Engineering from Princeton University, his M.S. in Electrical Engineering from Yale University and his Ph.D. in Mechanical Engineering from Rice University. His research interests include robotic hands, grasping, and automated assembly. Dr. Diftler has published robotics papers in the fields of mechanism design, dynamics and control, and teleoperation, and shares a patent for a Compact Dexterous Robotic Hand. He is currently employed with Lockheed Martin at the NASA Johnson Space Center as a senior robotics engineer. Contact him at difftler@jsc.nasa.gov.

Eric Huber received his BS in Mechanical Engineering from Rutgers University in 1985 and his MS in Mechanical Engineering (Robotics) from Stevens Institute of Technology in 1991. Since 1986 he has developed vision systems for mobile robots and manipulators at Picatinny Arsenal, General Electric Astro-Space, Goddard Space Flight Center, and Johnson Space Center. His publications include topics in machine vision, stereo vision, real-time tracking, pose estimation, human gesture recognition, and intelligent control of robots. Contact: eric@roboteyes.com.