Robonaut Mobile Autonomy: Initial Experiments

M. A. Diftler, R. O. Ambrose, S. M. Goza, K. S. Tyree
Automation, Robotics, and Simulation Division
NASA/Johnson Space Center
Houston, Texas, 77058
diftler@jsc.nasa.gov

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

Abstract—A mobile version of the NASA/DARPA Robonaut humanoid recently completed initial autonomy trials working directly with humans in cluttered environments. This compact robot combines the upper body of the Robonaut system with a Segway™ Robotic Mobility Platform yielding a dexterous, maneuverable humanoid ideal for interacting with human co-workers in a range of environments. This system uses stereovision to locate human teammates and tools and a navigation system that uses laser range and vision data to follow humans while avoiding obstacles. Tactile sensors provide information to grasping algorithms for efficient tool exchanges. The autonomous architecture utilizes these pre-programmed skills to form complex behaviors. The initial behavior demonstrates a robust capability to assist a human by acquiring a tool from a remotely located individual and then following the human in a cluttered environment with the tool for future use.

Keywords—dexterous robot, mobile, autonomy, humanoid, hazardous environment

I. Introduction

Humanoid robots offer great potential for working with humans on a variety of tasks. By definition, they are designed to perform an ever increasing set of tasks that are currently limited to people. Of specific interest here are tasks that currently require human level dexterity while working in dangerous arenas. To this end, NASA and DARPA (Defense Advanced Research Projects Agency) are jointly pursing the development of a mobile autonomous humanoid, Robonaut, for use in the hazardous environments of low earth orbit (LOE) and planetary exploration. Robonaut, which can also be teleoperated, is capable of interfacing with Extra-Vehicular Activity (EVA) systems that only have human interfaces and working with the same human rated tools designed for all NASA missions.

Humanoids are a new class of robots. Two of the most well known are the self-contained Honda Humanoid Robot [1] and its most recent descendent, ASIMO [2], which are able to walk and even climb stairs. Another member of the Japanese Humanoid family also inspired by the Honda robot is the HRP-2P [3] which is able to lie down and then stand back up and cooperatively move objects with humans. In the area of upper body capability, several prototypes have been built that are designed to work with humans. One of the first, Greenman [4], showed the benefits of a human teleoperating a humanoid robot. WENDY (Waseda Engineering Designed sYmbiont) [5] 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 [6] and Vanderbilt’s ISAC [7] 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. [8], provide an excellent survey of anthropomorphic robot evolution and suggest three characteristics that are most important towards making a better humanoid: human-like motion, human-like intelligence, and human-like communication.

Initially the NASA/DARPA Robonaut achieved these human-like characteristics solely through a human teleoperator directly controlling the system. Through an incremental process, more of the skills necessary to achieve the human-like capabilities necessary to perform EVA and planetary tasks are being reproduced within the Robonaut’s control system. These skills combine new software, sensors, and, most recently, mobility, to form the basis for an extremely flexible autonomous device.

II. Robonaut

With more than 40 degrees of freedom each, the
Robonaut systems (figure 1) are the first humanoids specifically designed for space [9]. They incorporate technology advances in dexterous hands, modular manipulators, and lightweight materials. Robonaut is human size, has an articulated waist, and two seven-DOF arms, giving it an impressive work space for interacting with its environment. It has a pan/tilt stereovision camera unit that provides ranging information for both teleoperators and machine vision algorithms. In addition to having the correct anatomy to work with EVA equipment, the Robonaut system is designed with on-orbit operations in mind. Robonaut’s single leg (figure 2) design includes a stinger to directly mate with the same Space Station worksite interface used by crew for stabilization.

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 [10]. 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. This API is a living interface and accommodates communication requirements for additional capability as the robot evolves. One of the most recent additions includes definitions for mobility commands and data.

III. Mobility

Robonaut can be configured for a variety of mobility options. In its zero-g configuration, it can be moved by a larger manipulator that connects with the grapple fixture on Robonaut’s back. The arms can be used to climb along truss handrails and the leg can be used for repositioning of the body once it is stabilized to a space structure (figure 2). On planetary missions, Robonaut can be integrated with a rover to make a modern day centaur (figure 3). For its first mobile autonomous assistant role, Robonaut is combined with a more readily available lower body.

A. Robotic Mobility Platform

The SegwayTM Robotic Mobility Platform (RMP), as shown in figure 4a, provides mobility for Robonaut. It is a derivative of the SegwayTM Human Transporter (HT). The HT is a two-wheeled motorized vehicle for transportation. It is capable of traversing a multitude of terrains. DARPA commissioned Segway™ to develop a computer-controlled version capable of balancing large payloads. This became the Segway™ RMP. The RMP is controlled via computer with velocity and turning rate as the primary controls. When these values are set to zero, the RMP will hold position even when external forces are applied. The RMP is capable of a zero turning radius and +/- 8 mph velocities.
The RMP was tested extensively to determine its capabilities. Control software was tested and the stock hardware was modified to suit the needs of Robonaut. A battery power distribution system was added along with training wheels for initial development and testing. One shortfall of a two-wheeled platform is its inability to stay upright if drive power is severed for any reason. The training wheels prevent this failure mode from causing damage to its robotic payload.

Robonaut’s leg, except for a single roll joint, was removed in preparation for mounting on the RMP. This joint provides a single waist DOF and allows Robonaut to pivot on top of its mobile platform providing more flexibility to the teleoperator. Figure 4b depicts this DOF. Robonaut on the RMP combines human like dexterity with low profile mobility making for an impressive and capable humanoid [11]. Sensing in the form of vision, a laser rangefinder, and tactile feedback provide the data needed for initial experiments in autonomous mobility and manipulation.

IV. Vision

The Robonaut stereo-based vision system capitalizes on object shape to track the pose (position and orientation) of well-defined objects, such as wrenches and tables, as well as variable-form structures, such as a human torso and limbs. To achieve tracking, the vision system operates over a series of stages, one stage cascading into the next [12]. Stages of processing are hierarchical, with each stage typically composed of a series of sub-processes. At the highest level, there are currently three stages of processing:

1) Image Server: This stage takes the stereo video streams from a pair of firewire cameras mounted in Robonaut’s head and generates sign-of-LoG (Laplacian of Gaussian) convolved (filtered) image pairs. LoG convolution accentuates any spatial variations (texture) in the grayscale image to promote matching between the stereo image streams.

2) Stereo Tower: This stage takes the binary images from stage (1) and performs patchwise (area) correlation to generate silhouette and range maps in real-time. A range map is a two-dimensional array of distance measurements corresponding to object surface points within a scene. Silhouette maps are binary derivatives of range maps. Each bit of the silhouette map, corresponding to a point in the scene, indicates whether surface material is measured within a specifically targeted distance range. The silhouette map provides a simple means of segmenting out objects of interest from the rest of the scene (See Figure 6: a-1, a-3).

3) Multi Tracker: This stage takes silhouette and range maps from stage (2) and matches them against large sets of 2D templates for acquisition, and 3D templates for pose estimation.

Pose estimation is the job of Multi Tracker, which is composed of a set of Object Trackers, each seeking a specific type of object. Each Object Tracker is composed of a series of cascading filters, designed to match sets of object-specific templates against incoming silhouette maps.

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 binary-packed templates can be accessed with a single instruction, and bitwise matching can be performed in parallel. The following section IV. A discusses how template matching filters are used to determine object pose in real-time.

![Examples of wrench silhouette templates: 2D binary templates of an adjustable wrench representing its silhouette as viewed from different distances and orientations.](image-url)
templates are true 2D silhouettes of the target object, appearance of the target object. In the final step, the fidelity templates, which more accurately reflect the undetermined DOFs, can afford to apply higher silhouettes. Later steps, which have fewer this flexibility, early steps must employ liberal to any remaining undetermined DOFs. To achieve (component of orientation). A step must be capable designed to "capture" a specific degree of freedom of capturing its target DOF while remaining tolerant to any remaining undetermined DOFs. To achieve this flexibility, early steps must employ liberal silhouettes. Later steps, 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 step, the templates are true 2D silhouettes of the target object, providing the greatest pose estimation precision in all degrees of freedom.

V. Navigation

For Robonaut, vision provides the goal and navigation is needed to reach the goal. To this end, obstacle avoidance, mapping and tracking functions were ported to Robonaut from another JSC robot designed for planetary surface exploration, the EVA Robotic Assistant [13]. The modular software from this project was easily transferred to the Robonaut platform, with only a few modifications required. Most of the ported navigation software is self-contained, but a few pieces have to communicate directly with Robonaut: sending driving and steering commands, receiving roll and pitch values, and receiving the human target position from the vision system. To enable the required communication, the navigation software was augmented to communicate with Robonaut's API.

In order to add obstacle avoidance to Robonaut, a sensor had to be added to detect obstacles. A SICK laser rangefinder was thus mounted on the front of the RMP slanting slightly downward such that the laser's rays hit the ground a few feet in front of the robot. By using the current roll and pitch values of the RMP, the intersections of the laser's rays and the ground (or objects) can be converted into the robot's coordinate system. An obstacle map is generated from the resulting sensor data, filling in "goodness" and "certainty" values for each cell of a discretized 2D map. These values are calculated based on height differences between adjacent laser rays and height values of individual laser rays relative to the robot. Adjustable parameters determine the threshold values which indicate obstacles.

For the EVA Robotic Assistant [13], the global position of the robot is also taken into account each time the laser data is obtained, based on GPS readings. Each generated laser map is merged into a global map for future use. For Robonaut, working primarily indoors, localization information is not available, so only the instantaneous laser maps are used. Obstacle avoidance in this case is more reactive to the immediate environment, with older data not available.

To navigate to a chosen target – a human in the case of this scenario – two steps are needed. First, the tracking software requires knowledge of the target location. For Robonaut's local-only system (no global information) the target's position is given relative to the robot itself. This information comes from the vision tracking software described above. The mobile base tracking software then takes this relative position and determines the steering and velocity commands needed to travel towards that target.

The second step in the navigation is to check for obstacles. The ground track arc that would result from the desired driving command is overlaid on the laser obstacle map, to determine if the arc intersects
any obstacles. If not, the driving command is allowed through and the robot proceeds on its course. If the arc does intersect an obstacle, then a modified driving command must be generated. A set of alternative arcs is defined for the robot ahead of time, and all of these arcs are overlaid on the laser obstacle map to check for obstacle intersections. Those arcs that are safe are also compared to the desired arc, and the closest safe match is selected. Point turns are also considered, if needed, as the RMP is quite adept at turning with a zero turning radius. The modified arc (or point turn) is then sent to the RMP, via the tracking software and the API, and the robot proceeds toward the target, avoiding any obstacles. The entire arc evaluation process occurs with a 5 Hz frequency, so once an obstacle is avoided, the robot can proceed back on course straight toward the target.

As mentioned earlier, parameters can be set to define what constitutes an obstacle for a given robot. The correlation between driving commands and ground track arcs is also parameterized and is set specifically for a given robot. These correlations had to be defined for Robonaut and the RMP at the start of the project. Additional parameters are also available to help adjust for the size of the robot, the desired buffer zone around the robot, the amount of space in which the robot can maneuver, and the existence of global localization. All these parameters are kept in a configuration file, to be read into the software at run time. Most of the parameters can be adjusted during run time as well.

VI. Grasping

The mobile Robonaut employs a tactile glove (see figure 7a), for autonomous grasping [14]. This glove is instrumented with 19 moderate resolution, resistance based, force sensors. Each finger joint is equipped with one sensor, and the thumb has multiple sensors to distinguish between the different thumb contacts that can be achieved. Three sensors are strategically located across the palm, and are very useful for determining contact when making tool and power grasps. In addition to providing good tactile data, the glove is rugged and designed to protect the sensors, provide excellent gripping surfaces and take the abuse associated with doing a wide range of EVA and planetary tasks.

![Figure 7. (a) Robonaut tactile glove          (b) Reflexive grasp](image)

A “grab reflex” commands the fingers to autonomously close when tactile stimulation is detected on the glove’s palm sensors (figure 7b). Upon contact with the object, the fingers continue to apply force. This is similar to the distal curl reflex observed in human infants. This simple primitive is one of many grasp primitives [15] based on tactile and force feedback available to build autonomous sequences.

The “grab reflex” along with the vision based primitives and the navigation routines form the basis for Robonaut’s initial mobile autonomous behavior. A sequencer combines these capabilities and coordinates Robonaut’s actions.

VII. Sequencer

The Sequencer monitors the status of the various processes and controls the activation and deactivation of various autonomy routines through Robonaut API calls. If communication is lost during part of the run or before a process is started, then the Sequencer flags the error and does not allow the routine to progress. In order for this to occur, all processes produce a heartbeat. This heartbeat is a simple data packet that emits a status at a constant deterministic rate.

The Sequencer also acts as part of the safety system. Processes can be aborted at any time with buttons located on the dialog. In addition, if the routine aborts mid stream, the Sequencer allows the operator to restart processes and continue from the abort point.

In normal operation, the operator presses one start button and the Sequencer coordinates the entire autonomous behavior. It starts and stops processes as required without operator intervention until the routine is complete. While it is running, the buttons on the dialog display status information describing the state of the various components. The Sequencer provides the operator with a simple means of monitoring system status and current state. Color coding gives a quick visual indication of state and where problems exist. If a problem occurs, the Sequencer allows quick recovery.
Using the sequencer described above, Robonaut is able to combine basic behaviors to form its more complex mobile autonomous behavior. In this case the task is to act as a tool-handling assistant. The sequence for this behavior is: scan the room searching for human heads and acquire one as a goal, proceed towards the human, avoid obstacles along the way using the SICK laser, stop in front of the human, scan for a tool in the human's hand, follow the human hand as the tool is positioned for hand off, reach out for the tool, close hand upon making palm contact with the tool and maintain a secure grasp, move the acquired tool to a stowed position, reacquire the human head, and follow the human with tool to worksite.

Figure 8 shows various sub-behaviors that make up this task. Robonaut has successfully performed this task more than 50 times in multiple environments. While mobility provides exceptional benefits: a large work area, an unlimited number of potential tasks, etc…, it also comes with a number of associated challenges: uncontrollable lighting, unpredictable obstacles, and radio frequency interference.

A. Lessons Learned

These challenges were apparent when Robonaut performed this task at the 2004 DARPATech conference in Anaheim, California. Lighting conditions varied significantly throughout a three day period during demonstrations. Parameter and camera tuning were required and in certain lower lighting situations, the human leader was required to stay closer to the robot so that it could maintain visual contact. This data is being fed back into future camera selection criteria and vision development priorities.

Visitors to the Robonaut demonstration area would occasionally be picked up by the vision system as the human teammate. The actual teammate could easily regain Robonaut’s attention by stepping between the robot and the visitor and becoming the most prominent human in the scene. A console operator monitoring the vision system could provide the same correction, by momentarily overriding the visually controlled head control system and redirecting Robonaut’s view. A similar vision challenge involved mirrors where Robonaut mistakenly would pick up the reflection of its teammate. Reflective surfaces can be found in a variety of space and planetary applications, and a mobile autonomous humanoid assistant must be able to deal with this challenge along with other confusing scenery. Additional sensing that will augment vision to help identify image features is currently being investigated.

Another issue not normally associated with well structured space and planetary activities is radio interference. Using a robotic assistant in environments where both digital and analog wireless communications can not be controlled obviously requires more care than is normally provided in a laboratory environment. When Robonaut’s wireless Ethernet system is overwhelmed by interference on its current operating channel, the robot will safe itself and wait for communications to clear up. This is a reasonable interim solution while more fault tolerant communications are investigated.

While the single SICK laser rangefinder works well for its intended task, the initial implementation has limitations. With no global localization, a persistent map is not available in case the robot needs to back up. The laser rangefinder also only has a single scan line. Objects below the beam of the laser cannot be seen. Even if the Robonaut originally sees a short object, if it gets too close, or turns suddenly to face the short object at close range, the laser cannot detect the obstacle. A related challenge dealt with point turns. When Robonaut found itself too close to an obstacle, the software determined that a point turn was necessary. However, without a persistent global map, the robot soon turned so far that the obstacle no longer appeared in the laser scan. The software would then decide to turn back the way it had come, generating an oscillating behavior. Modifications had to be made to correct this problem, insisting that a point turn be carried out long enough to clear the obstacle completely before allowing a different point turn or driving direction to be selected.

To further augment the SICK laser based obstacle avoidance system, infrared sensors are currently being tested on a second RMP that is slated as an upgrade to the current mobility system. Each Wany Robotics™ sensor provides a 15-degree cone for obtaining obstacle distance and has a range up to 3 meters. Figure 9 shows the custom configuration developed for Robonaut that will provide additional instantaneous distance information for navigation.
IX. Future Work

Robonaut has successfully demonstrated its initial abilities as a mobile autonomous humanoid assistant. Robonaut’s laboratory tactile and grasping capabilities easily transferred to a mobile base. Navigation software using a laser rangefinder developed for NASA’s EVA Robotic assistant also integrated well with Robonaut’s RMP base. The most challenging transfer from the laboratory to the field involved stereo vision, a demanding field in any arena. Much was learned from this integration experience and is being used to plan future work.

As noted above, work is underway to provide more sensing for Robonaut’s RMP base. Additionally, a rear facing laser rangefinder is being considered along with a pan-tilt unit that could produce a more complete environmental map.

New firewire cameras that provide a global shutter, which exposes all image lines at the same time, are being tested. By capturing the image in such a fashion, camera motion does not bias data within each image. This is expected to significantly improve stereo performance.

A next generation tactile glove with significantly more coverage will replace the existing one and provide data for more complex grasps. Additionally, force/torque sensors located within Robonaut’s forearms will provide data to better communicate intent between the human and the robot.

All of these additional capabilities will lead to a more robust version of Robonaut’s initial autonomous behavior and augment the basis for new behaviors.

Acknowledgment

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

References


