Marker Based Task-Level Teleoperated Manipulation

Rafael Cisneros[1]
rafael.cisneros@aist.go.jp

Takeshi Sakaguchi[1]
sakaguchi.t@aist.go.jp

Shuuji Kajita[1]
s.kajita@aist.go.jp

***Abstract*— This paper presents the method used by our team, AIST-NEDO, at the DARPA Robotics Challenge (DRC) to deal with the requested manipulation tasks by means of a task-level teleoperation, by considering a degraded communication between the user and the robot and that the environment was not known in advance. The method basically consists on the use of 3D models of objects (from now on referred as "markers") which, once aligned with the actual attitude of the real objects that they represent, provide a reference frame in which the motion can be described, in order to successfully realize a manipulation task in a non-structured environment. These markers can represent the object being manipulated, some reference object in the environment and the hands of the robot. This method is illustrated by means of describing three representative tasks (which where requested during the DRC) and presenting the corresponding results obtained during the competition.**

I. Introduction and Motivation

Disaster response is attracting attention from the robotics research community, and even more since the Fukushima Daiichi nuclear power plant accident that followed the 2011 Great East Japan earthquake and tsunami. As a concrete materialization of this increasing interest, a challenge is proposed by the American Defense Advanced Research Projects Agency (DARPA) to use robots in disaster-hit facilities that were made too hazardous for direct human operator intervention. It is worth noticing that the challenge does not impose any constraint on the design of the robot, but since the environment (industrial ladders, doors, valves, cars) as well as the tools (levers, drills, hammers) were meant to comply with the human morphology, it is a natural option to develop the necessary means to make the humanoid robots capable of performing inspection and disaster recovering actions inside a non-structured environment [1].

This environment can be considered to be "kind of" known in the sense that we know which actions are required in advance and that we have a rough idea of its spatial distribution, maybe altered due to the disaster itself. Then, only very limited assumptions about the structure of the environment can be made beforehand, in contrast to structured scenarios where semantic knowledge of their structure can be leveraged for highly autonomous robots operating in them [2].

It is also mandatory to consider that within a disaster-hit facility it is not possible to rely on a stable, wide bandwidth wireless communication system with the robot. The signal may be degraded and blackouts may occur frequently. Then, it is not feasible to consider a purely teleoperated robot. First, because of the high dimensionality of its control system, and second because the capabilities of the robot and the operator should include near real-time feedback without disruptions in the communications as well as transmission of large amounts of data to the operator. On the other hand, a fully autonomous robot navigating and interacting in an unconstrained environment should include extensive databases of information about possible objects of interest to be found, highly efficient grasping algorithms and the ability to react to unforeseen situations, which are still unsolved problems [3]. A feasible alternative is the development of supervised semi-autonomous high Degrees-Of-Freedom (DOF) robotic systems; that is, task-level teleoperated systems in which the operator cognitive burden is minimized by lowering the control space dimensionality [4], such that these operators function as supervisors setting high level goals, assisting the robot with complex perception tasks, directly changing robot parameters to improve its performance and making decisions when facing unexpected situations [2].

II. Related Work

III. Teleoperated Manipulation Method

Let us consider a humanoid robot equipped with a Laser Range Finder (LRF) placed at the head, as well as three cameras: at the head and at each hand. The LRF provides a point cloud of the environment, probably contaminated with noise due to the enviromental conditions of the disaster scenario; that is, even after a proper calibration it is not possible to consider that the 3D data precisely represents points belonging to objects in the environment, but within a certain amount of tolerance. On the other hand, the frame rate of the cameras, as well as the resolution, are intentionally set low foreseeing the effects of the degraded communication.

This information of the environment, together with the sensorial information providing the current state of the robot, are the only information available to the operator, which has to supervise the robot while performing the tasks by establishing high level goals, assisting with perception and changing parameters during the task. For this purpose, and under the circumstances stated above, we came up with a teleoperated manipulation method based on "markers" which assume minimum knowledge of the environment. This one is explained on the following.

[1] R. Cisneros, T. Sakaguchi and S. Kajita are members of the Humanoid Research Group of the National Institute of Advanced Industrial Science and Technology (AIST), 305-8568 Tsukuba, Japan

A. Approaching to the target

The robot must achieve a proper stance with respect to the object(s) representing the target of the task, such that they be inside of the dextrous workspace of the hands of the robot. To do that the robot must perform a first measurement of the environment in order to check, together with the head camera, if the manipulation target is represented by some set of points of the resulting point cloud. In such a case, a preliminar alignment of a 3D object representing this manipulation target (the *manipulation marker*) is first performed.

By using the Graphical User Interface (GUI) provided by Choreonoid [5], the Manipulation Marker is represented as the corresponding 3D model together with a set of arrows and rings which allow the operator to translate and rotate it with respect to its local reference frame, as depicted in Fig. 1.

Doing this alignment manually can be tedious, besides the fact that it may require a lot of time. To speed it up, it is possible to use a built in function included in the Point Cloud Library (PCL) that automatically alignes the marker with the best fit set of points, by providing maximum displacements, the number of iterations and some allowable error. However, ordinarily the robot will not be close enough to the object(s) for them to be measured with enough density of points, in such a way that the automatic alignment be prone to fail. One way to overcome this problem is to select one point of the point cloud belonging to the object and set this as the origin of the local reference frame of the Manipulation Marker, then the automatic alignment will lead to an alignment that may or may not require further small manual adjustments.

One way to improve the initial alignment of the Manipulation Marker is to use beforehand information of the possible attitude of the object with respect to the nearest wall. For example, if the manipulation target is a box attached to the wall, its front face will probably be parallel to it. Knowing this, it is just the matter to identify the plane of the wall (and maybe the floor), get its mathematical representation and use it to define an initial attitude of the Manipulation Marker, requiring little automatic or manual adjustments.

Once this is done, it is possible to define a proper stance of the robot (decided beforehand) with respect to the local reference frame of the Manipulation Marker. Then, by taking into account the height field of the floor (obtained from the point cloud) and avoiding the obstacles of the environment (walls and/or other objects), a proper footstep planning is performed in order for the humanoid robot to arrive to the desired stance [6] [7].

B. Grasping the target

Once humanoid the robot arrives to the desired stance, it has to perform another measurement. First, because of the positioning errors accumulated during its locomotion, and second, in order to obtain a more dense point cloud inteded for refining the alignment of the Manipulation Marker.

Having done this refinement, it is possible to describe the attitude of the hands of the robot with respect to the local reference frame of the Manipulation Marker in order to approach to the manipulation target and grasp it, push it or pull it as required. Once decided, the whole-body Inverse Kinematics (IK) solution must be found, taking into account the redundancy of the robot to avoid collisions with the environment (represented with the point cloud) [8].

This relative attitude of the hands needs to be previously decided, such that the task be effectively carried out by means of smooth motions; that is, avoiding singular configurations and critical postures that may compromise the stability of the robot. However, these ones can be modified during the execution of the task if necessary, by means of a Hand Marker, a 3D representation of the hand and a set of arrows and rings which allow the operator to translate and rotate it with respect to its local reference frame, calculating at the same time the resulting configuration of the humanoid robot. See Fig. 2.

C. Dealing with uncertainties

Grasping the target with the desired relative attitude can be critical for some manipulation tasks, specifically if the target size is small when compared to the hand. However, it is worth to consider that the point cloud has an intrinsec noise, mainly caused as a result of the sunlight which complicates

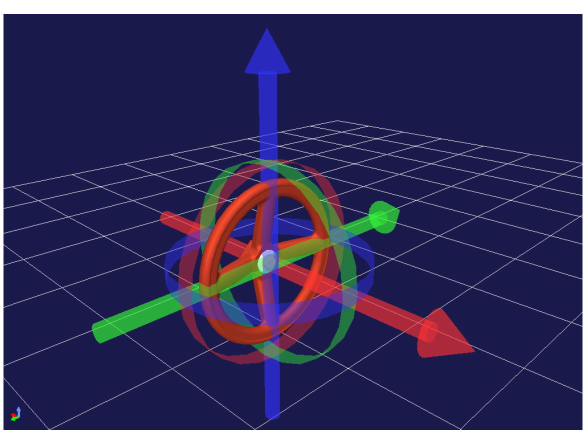

Fig. 1. Manipulation Marker of a valve.

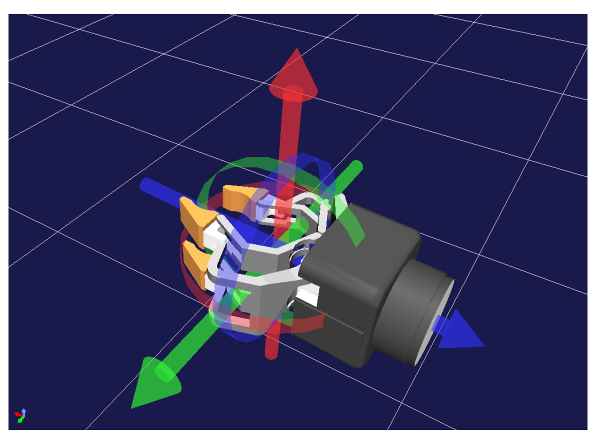

Fig. 2. Left Hand Marker.

measurements. Then, it is not advisable to rely completely on the point cloud to plan the grasping motion, as the shown points may not represent actual points on the objects.

One way to overcome this problem is to place the robot's hand within the view field of the LRF, and add some offset to the point cloud in order to match the corresponding points with the hand, whose attitude is known. This strategy practically improves the representation of the target objects by the point cloud but the precision may not be high enough. However, it can be used to approach to the manipulation targets (within some centimeters) before grasping them, with enough confidence that the hand is not going to collide unintentionally with the environment.

Having done this, it is possible to use an approaching strategy in which the hand intentionally collides with the target by means of a slow motion, in order to use the force sensor installed in the hand to stop the hand when it senses a force greater than some established offset. Then, the real position of some plane of the target can be known relative to the hand, whose attitude can be computed by using forward kinematics and the actual joint values. The grasping point on the target can then be reached by moving the hand a very small distance with respect to its local reference frame.

D. Manipulating the target

Once the target is grasped, the relative attitude between the Manipulation Marker and the Hand Marker(s) can be set to be constant. Then, it is possible to manipulate the target by translating / rotating the Manipulation Marker, such that the required configuration of the humanoid robot be automatically calculated by means of solving the corresponding whole-body IK problem. In this way, a broad range of manipulation tasks can be accomplished by following this scheme, as illustrated by some examples which are described as follows.

IV. Pulling and Inserting a Plug

One of the surprise tasks at the DARPA Robotics Challenge consisted of pulling out a plug from one socket and putting it back into another socket, in a set-up like the one shown in Fig. 3.

A. Detection of the socket and plug

In order to perform this task it is first necessary to identify the plug and the socket where it is originally inserted, by

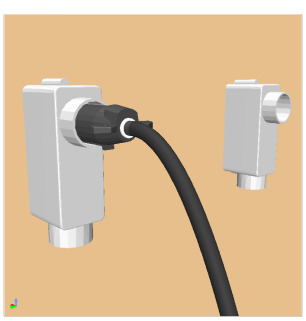

Fig. 3. Plug Task.

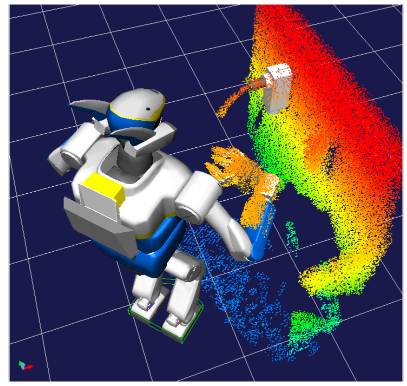

Fig. 4. Detection of the socket and the plug.

placing the corresponding Manipulation Marker within the poing cloud. The reason to identify both objects instead of just the plug is because almost half of it is not visible as it is inside of the socket, making it too difficult just to match the plug, given the low density of points belonging to it due to its small size.

By using a PCL's built in function it is possible to detect all the planes in the scene, one of them being the wall where the sockets are installed. Then, given that the front face of the sockets is parallel to this wall, it is possible to use the corresponding plane equation to calculate the socket's orientation with respect to the robot, in such a way that the plug is directed towards it. Having done this, one point belonging to the plug can be manually selected in order to compute an approximate initial position for the Manipulation Marker representing the socket and the plug. This one is later refined, after the robot has arrived to the desired stance and the point cloud has been adjusted by using the robot's hands as a reference, as seen in Fig. 4 (and explained in Subsection III-C).

B. Grasping the plug

Having detected an approximate attitude for the socket, the robot first approaches the plug with one hand while aligning the camera installed at the other hand with the axis of the plug. This is to be able to make slight adjustments of the grasping hand by using visual feedback.

The size of the visible part of the plug is not that big compared to the hand of the robot, and because of that the tolerance for grasping the plug is minimum. Then, it is required to grasp the plug at the point in which the medial side of the hand touches the cylindrical shaped part of the socket. This can be done by reducing the angle of the gripper and then, by moving the hand along the axis of the plug until sensing 30 N of force (enough for considering that it arrived to the socket). This configuration of the robot just before moving the hand towards to the socket is depicted in Fig. 5.

This strategy is very effective (when adjusted properly) as the grasp can be done every time at the desired point even in the presence of uncertainties, as shown in the dynamical simulation depicted in Fig. 6.

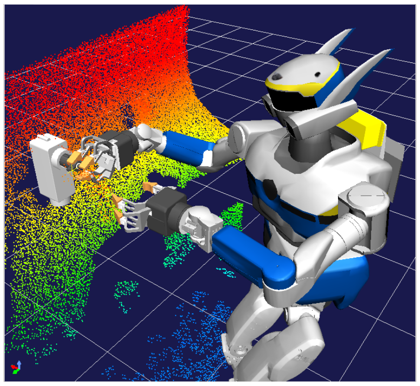

Fig. 5. Configuration of the robot before grasping the plug.

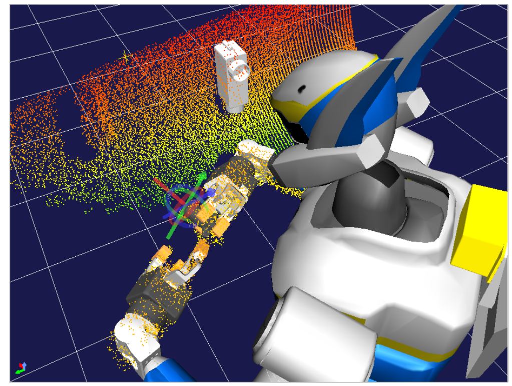

Fig. 7. After pulling the plug its attitude is adjusted.

C. Pulling and adjusting the plug

After grasping the plug, the robot pulls the plug. Due to the balancing process this motion may not be performed exactly along the axis of the plug, and it may hit the inner walls of the socket, modifying the planned relative attitude between the plug and the hand.

For this reason, before inserting the plug into the other socket, the robot first brings the plug in front of its chest, takes an updated point cloud, and uses the camera placed at the head and at the other hand to look at the plug from two perpendicular directions, as seen in Fig. 7. By using this information it is possible to fix the actual attitude of the plug with respect to the hand.

D. Inserting the plug

Once the Manipulation Marker representing the plug is fixed at the hand, it can further be properly aligned with the destination socket. This one can be represented with another marker (Object Marker), which can be adjusted within the point cloud before taking the plug to the pre-insertion position. However, because of inaccuracies of the point cloud, the camera at the other hand is used once again together with the camera at the head to look at the plug, and use this visual feedback to adjust its position to properly insert the plug within few movements. This is represented in Fig. 8.

Then, once the task is completed, the hand opens and an updated point cloud is taken, in order to correctly plan the returning motion without hitting the cable of the plug.

V. Opening a box and pressing a button

VI. Opening a door

VII. Results

VIII. Conclusions

References

[1] K. Bouyarmane, J. Vaillant, K. François, and A. Kheddar. Exploring humanoid robots locomotion capabilities in virtual disaster response scenarios. In *IEEE-RAS International Conference on Humanoid Robots*, 2012.

[2] Stefan Kohlbrecher, Alberto Romay, Alexander Stumpf, Anant Gupta, and Oskar von Stryk. Human-robot teaming for rescue missions: Team vigir's approach to the 2013 darpa robotics challenge trials. *Journal of Field Robotics*, 32(3), 2014.

[3] Alberto Romay, Stefan Kohlbrecher, David C. Conner, Alexander Stumpf, and Oskar von Stryk. Template-based manipulation in unstructured environments for supervised semi-autonomous humanoid robots. In *IEEE-RAS International Conference on Humanoid Robots*, 2014.

[4] Kapil D. Katyal, Christopher Y. Brown, Steven A. Hechtman, Matthew P. Para, Timotthy G. McGee, Kevin C. Wolfe, Ryan J. Murphy, Michael D.M. Kutzer, Edward W. Tunstel, Michael P. McLoughlin, and Matthew S. Johannes. Approaches to robotic teleoperation in a disaster scenario: From supervised autonomy to direct control. In *IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*, 2014.

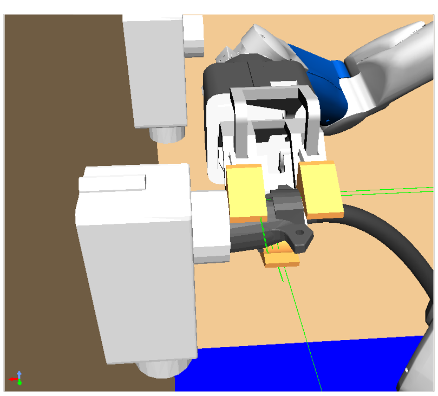

Fig. 6. The plug being effectively grasped at the desired point.

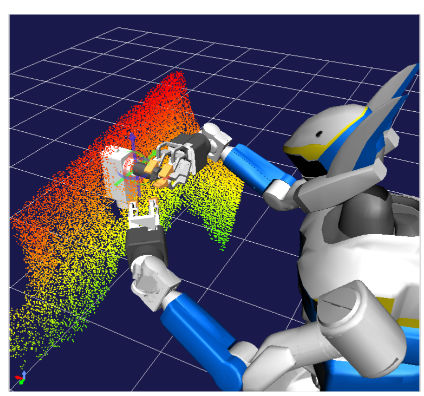

Fig. 8. The plug is inserted by using visual feedback.

Fig. 9. Door Task at the DARPA Robotics Challenge [9].

[5] S. Nakaoka, S. Kajita, and K. Yokoi. Intuitive and flexible user interface for creating whole body motions of biped humanoid robots. In *IEEE/RSJ Int. Conf. on Intelligent Robots and Systems (IROS)*, pages 1675–1681, Taipei, Taiwan, October 2010.

[6] Mitsuharu Morisawa, Shuuji Kajita, Fumio Kanehiro, Kenji Kaneko, Kanako Miura, and Kazuhiro Yokoi. Balance control based on capture point error compensation for biped walking on uneven terrain. In *IEEE-RAS International Conference on Humanoid Robots*, 2012.

[7] N. Perrin. *Footstep Planning for Humanoid Robots: Discrete and Continuous Approaches*. PhD thesis, University of Toulousse, 2011.

[8] Oussama Kanoun, Florent Lamiraux, and Pierre-Brice Wieber. Kinematic control of redundant manipulators: generalizing the task priority framework to inequality tasks. *IEEE Transactions on Robotics*, 27(4), 2011.

[9] DARPA. Darpa robotics challenge finals 2015. `http://www.theroboticschallenge.org/`, 2015.

[10] Evan Ackerman and Erico Guizzo. Darpa robotics challenge: Amazing moments, lessons learned, and what's next. `http://http://spectrum.ieee.org/automaton/robotics/humanoids/`, 2015.

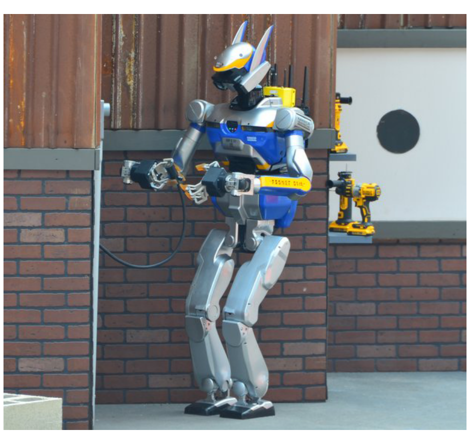

Fig. 10. Plug Task at the second day of the DARPA Robotics Challenge [10].

Fig. 11. Button Task at the rehersal of the DARPA Robotics Challenge [9].