Remotely Teleoperating a Humanoid Robot to Perform Fine Motor Tasks with Virtual Reality – 18446

Jordan Allspaw*, Jonathan Roche*, Nicholas Lemiesz**, Michael Yannuzzi*, and Holly A. Yanco*
* University of Massachusetts Lowell
** Wentworth Institute of Technology

ABSTRACT
In this paper, we describe our ongoing work to develop cooperative control of NASA's R5 Valkyrie humanoid robot for performing dexterous manipulation tasks inside gloveboxes commonly found in many nuclear facilities. These tasks can be physically demanding and provide some element of risk to the operator when done by a person in situ. For example, if a glove is ruptured, the operator could be exposed to radioactive material. In many cases, the operator has low visibility and is unable to reach the entire task space, requiring the use of additional tools located inside the glovebox. Such tasks include cleaning particulate from inside the glovebox via sweeping or vacuuming, separating a specific amount of a compound to be weighed on a scale, or grasping and manipulating objects inside the glovebox.

There is potential to move the operator to a nearby, safe location and instead place a humanoid robot in the potentially hazardous environment. However, teleoperating a humanoid robot to perform dexterous tasks at a comparable level to a human hand remains a difficult problem. Previous work for controlling humanoid robots often involves one or more operators using a standard 2D display with a mouse and keyboard as controllers. Successful interfaces use sensor fusion to provide information to the operator for increased situation awareness, but these designs have limitations. Gaining proper situation awareness by visualizing 3D information on a 2D screen requires time and increases the cognitive load on the user. Instead, if the operator is able to visualize and control the robot properly in three dimensions, it can increase situation and task awareness, reduce task time, reduce the chance of mistakes, and increase the likelihood of overall task success.

We propose a two-part system that combines an HTC Vive virtual reality headset with either the Vive handheld controllers, or the Manus VR wearable gloves as the primary control. The operator wears the headset in a remote location and can visualize a reconstruction of the glovebox, created live by sensor scans from the robot and with sensors located inside the glovebox for a perspective traditionally unavailable to operators. By using the controllers or gloves to control the humanoid robots hands directly, they can plan actions in the virtual reconstruction. When the operator is satisfied with the plan, the actions are sent to the real robot. To test this system we have created a mockup of a glovebox that is accessible by Valkyrie, as well as several tasks that are a subsample of the tasks that might be required when working in a real glovebox.

BACKGROUND
Gloveboxes are often used to handle nuclear material or perform experiments in a nuclear environment. These gloveboxes are enclosed structures designed to safely house radioactive material, while providing access to allowing safe access to professionals via ports on the side with built in gloves. The tasks that are performed inside these gloveboxes vary widely, ranging from measuring compounds, using electrical equipment and numerous other tasks often involving fine manipulation of objects. In addition, these gloveboxes require significant maintenance with tasks such as cleaning and removing excess refuse. There are many safety features built into their design and protocol for use, but accidents can still occur

[1]. When an accident occurs, the operators in the immediate vicinity are at the greatest risk, so there is a desire to perform necessary experiments or maintenance tasks with the operators in a remote location. One solution is to deploy a robotic agent to operate the glovebox, with the supervisors and operators in a nearby, remote location. If the robot is able to perform the necessary tasks in a safe and reliable manner, this would increase the safety of the operators in the event that something does go wrong.

A typical glovebox would be a rectangular compartment with two glove slots located such that a person standing would be easily able to put both arms inside, however in reality gloveboxes vary widely. They can range in size from small workspaces to as large as a room. In addition, while two glove ports located at a specific height approximately shoulder-width apart is a common layout, some only have a single porthole, and some have many portholes located at different heights and positions. When our team toured the Savannah River Site (SRS) in South Carolina, our hosts detailed the wide range of tasks and environments where a glovebox-capable robot could be useful, ranging from measuring compounds, using electrical equipment, and cleaning or maintenance tasks. With these tasks in mind, a robotic solution needs to be able to handle a wide range of situations that could arise.

Figure 1. *Example glovebox setups in use (Left from [2], Right from [3]).*

To meet the task requirements, the robot would need to have manipulators capable of grasping and using tools commonly used in a constrained glovebox environment. The robot would also need to be able to position itself and possibly move between different glove portholes to perform the tasks as required. One proposed robotic platform that could easily change its position would be a humanoid robot. To test this case, the team is using the R5 Valkyrie created by NASA [4]. The R5 Valkyrie stands at 6 feet (1.83 meters) tall, with two 7 degree of freedom (DOF) arms and 6 DOF hands. Her hands are shaped very similar to a person's, with 3 fingers and an opposable thumb. This means that she is able to grasp and operate similar tools to that of a human, as well as operate in similar environments with minimal redesign.

INTRODUCTION

While significant research has been conducted with robots in domains such as telepresence, homecare, and warehouse delivery systems, by comparison, controlling humanoid robots is far less explored. The largest exploration of the use of humanoid robots was conducted during the DARPA Robotics Challenge (DRC) where teams competed to perform tasks, such as opening a door, turning a valve, and walking up stairs [5]. One lesson established by research is that full autonomy can be very time consuming to implement and adapt to new situations [6]. However, by using a shared control strategy, where some components are handled autonomously and some are handled by the human operator, the benefits of each can be maximized while reducing development time [7]. Automated perception is an example of a task that is very difficult to work with in changing environments, yet tends to be trivial and quick for human operators with the right information. Even if the final goal is an autonomous solution, it can be desirable to start with a skilled operator first. With this in mind, we are first pursuing a shared control solution, where most of the decision making is performed by a skilled knowledgeable operator. The interface therefore needs to present the information and controls to allow the operator to perform their duties to a similar level to as if they were actually there.

INTERFACE DESIGN

Our proposed solution is to allow a skilled operator in a remote location to control the robot by a virtual reality (VR) headset. Using a commercially available headset, called the HTC Vive, and combined optionally with the Manus VR gloves, the operator can visualize and control Valkyrie. The HTC Vive is a VR headset that has built-in head tracking for both position and orientation [8], which allows the operator to navigate around a virtual reconstruction of the world by physically moving around in their remote open space. Doing so can allow for quick and accurate mental reconstruction of the remote world where the robot is located. The HTC Vive comes with two controllers, one for each hand, that each have a joystick, several buttons, and the same built-in tracking as the headset. As an alternative to the handheld controllers, the operator can wear the Manus VR gloves which allow the system to accurately track the operator's fingers. By combining this with tracking sensors attached to the wrist, the team can track the position of the operator's hand and fingers. With this entire setup, the operator can visualize and interact with a virtual reconstruction of what the robot sees.

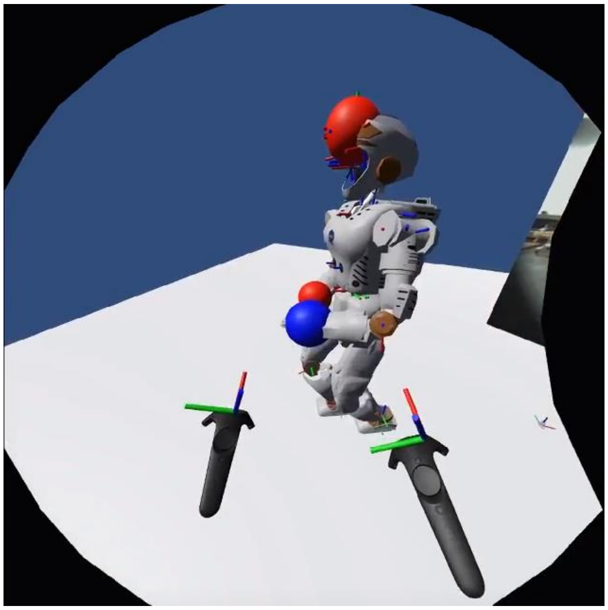

Figure 2. *Using the Vive controllers while viewing a 3D model of the robot.*

Egocentric and Exocentric Design

VR has enormous potential for a variety of ways for interacting with a robot, some of which are simply recreating concepts from traditional interfaces and some of which are possible only in a system like VR. To help categorize the different types of controls and visualizations, we break them down to either egocentric or exocentric. Egocentric interfaces, or interfaces where one sees the world from the position of the robot, tend to be better for navigation type tasks. Exocentric interfaces, or interfaces where one sees the world from an external point of view, tend to be better for understanding the environment's structure [7]. Many interfaces will combine elements of both, or otherwise allow the operator to switch between them, depending on which works better for the task. The team has incorporated this design by allowing the operator to switch between an egocentric or exocentric viewpoint. One example of the system working is where the operator starts out as a disembodied avatar, able to navigate around the virtual world at will. They can see the robot's position, as well as the information displayed by sensors such as the point-cloud generated by Valkyrie's lidar sensor. Using this information, the operator can build an accurate mental model of the area and plan out their tasks. Then the operator can switch to an egocentric view, where they are seeing the world directly from the perspective of the robot. Here, they can control the robot to perform the task while maximizing their ability to directly control the robot.

Virtual Reality Controls

Our team has allowed for several different ways for the operator to send commands to the robot. Starting in an exocentric view, the operator can "grab" one of Valkyrie's hands in the virtual world, by pressing a button with the Vive wands or making a fist while wearing the Manus glove, and move their own hands to the desired position. The robot will then plan the path to reach each desired position and the operator can watch the actions performed. The operator can choose to either have the robot follow in real-time, or to queue up many commands at once and send them all together. Alternatively, the operator can take an egocentric robot view, and move their own hands with the controller or gloves, with the robot mimicking immediately. Similarly, for controlling the head, the operator can either grab the head and pull it towards where the robot should look, or switch to an egocentric view and simply look where they want the robot to look using the Vive headset.

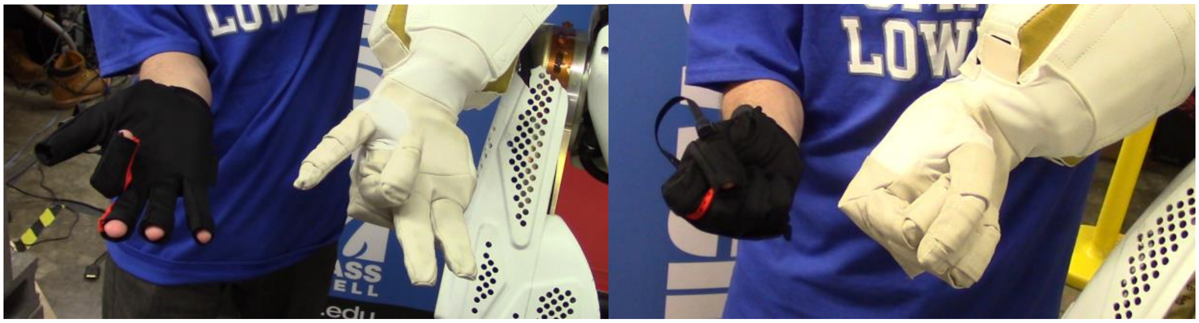

Figure 3. *Using the Manus VR glove to control Valkyrie's hand.*

VISUALIZING ROBOT INFORMATION

One of the traditional difficulties of complex robotic systems lies in visualizing information for operators.

Often times, operators will have to spend time deciding what is shown (or not shown) on an interface or navigating around an application to be able to get access to a specific piece of information. These processes can increase task time and cognitive load. Some interfaces attempt to solve this problem by autonomously selecting what information is applicable at any point in time to display, while other interfaces will choose to simply display everything, which can lead to instances of operator overload. All traditional interfaces still have the issue where there is only so much data that can be visualized by the user and easily reacted to on a 2D screen. A virtual reality device allows the operator to see an entire artificially-constructed world and allows data to be visualized in different places that are specifically relevant to the information being shown.

Specific benefits of this approach are evident in certain pieces of data, such as foot force sensors in a humanoid robot. Rather than numerical readings on a screen that take up precious screen space and time to read and interpret, the operator can visual force vectors located directly on the robot's feet. Reinterpreting information and displaying it in a virtual 3D world enables data to be laid out in a manner that can be significantly easier to visualize and understand. This process is called sensor fusion and has been found to be successful in 2D interfaces. For example, in the DARPA Robotics Challenge (DRC) Finals, a meta-analysis showed that "balancing capabilities of the operator with those of the robot and multiple sensor fusion instances with variable reference frames, positively impacted task performance" [6]. Specifically, it was found that "increased sensor fusion with common reference frames from an adjustable perspective is beneficial for remote teleoperation, and even more so by displaying two varying perspectives of the same data streams to increase the operator's situation awareness" [6]. Traditional interfaces leverage sensor fusion to great effect, but there is still a limit to how much information you can overlay on a 2D screen. The nature of VR could allow for more information to be presented in a context-sensitive nature, so that the operator only needs to be aware of it if they wish to be.

Robot State Information

The first and most important thing an operator can view is the current state of the robot. Many robots are capable of tracking their own joint movements, and so we display this information by simply updating the robot model in the virtual world with the correct values; see Figures 2 and 4 for examples. The operator is able to see the robot position as it moves around its environment, and most importantly, easily see other sensor data discussed below, relative to the robot.

Traditional Camera Streams

One of the most common sensors available on a robot is a standard camera image. Simple robots may have a single camera located on the robot, while more complex robotic systems can have many cameras spread out to give different vantage viewpoints, located both on the robot and in the environment.

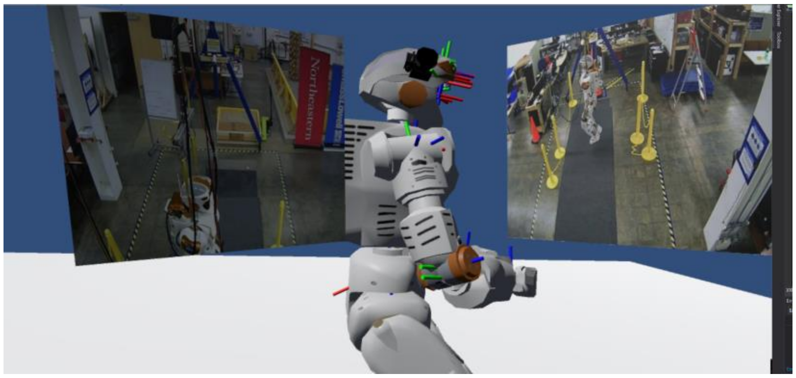

Figure 4. *Viewing multiple camera viewpoints. Both windows are from cameras placed external to the robot.*

In order to allow an operator to view these camera streams inside the virtual world, our team created virtual interactive windows. The operator can interact with these windows by "grabbing" them, either with the controller or glove. Figure 4 shows an example with two virtual windows located near the robot. In this example, both camera views are showing video feeds from external cameras located near the actual robot. The operator can choose to reposition the camera views, resize, change which camera is streaming, or remove them entirely with ease.

Visualizing Point Clouds

In this interface, significant time has been spent visualizing the robot's 3D environment scans. Robotic sensors, as well as special depth cameras, are able to take 3D scans, also known as point clouds, that can be processed or displayed to an operator. These point clouds are very similar to high-definition photographs in which each pixel has not only an x- and y-coordinate, but also a z-coordinate, representing the pixel's distance away from the camera. A large part of robotic interfaces are often spent visualizing point clouds and other pieces of depth data that make up the robot's environment. Virtual reality is uniquely capable of representing this data on more than just a two-dimensional screen. When it is shown on a screen, the only understanding of depth comes from the operator moving the camera around in a scene and relative depths are understood from the parallax effect. Because VR headsets render a separate image to each eye, VR has the distinct advantage of giving the user stereo-depth perception, the ability to see the depth of an image without the user having to move it around or make assumptions about the size of objects.

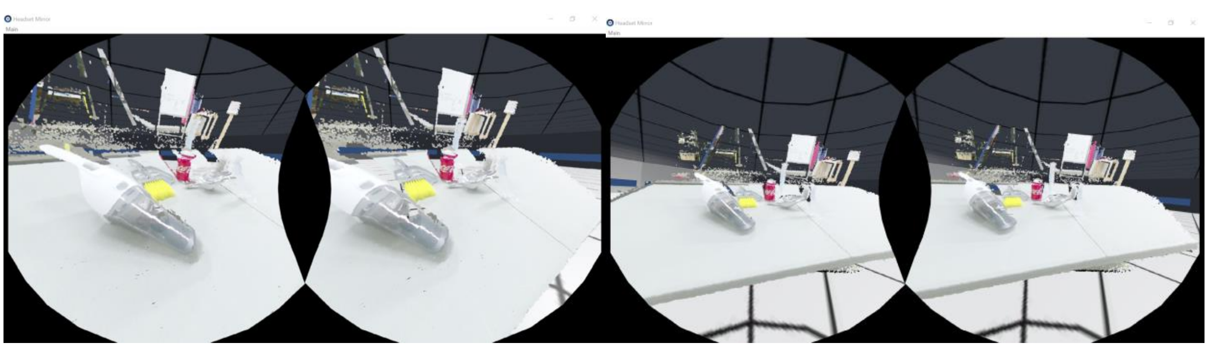

Figure 5. *Two different vantage points when visualizing a point cloud.*

When point cloud data is displayed, it is traditionally done in an application similar to CAD modeling applications. The operator has the ability to move the camera around in a scene, zoom in or out on a specific component, and to change the orientation of the environment. This works well on a 2D display, but in 3D, a different approach has been taken. In the VR world, the operator can take advantage of a feature known as room scale that allows for a one-to-one correlation between the user's motion in the real world and the camera's motion in the virtual world. Essentially, if there is a large or interesting obstacle in the environment, such as a table with objects for the robot to grab, the operator can simply stand up and walk around in the virtual reality setup to get a different view of the obstacle. For simplicity's sake, there is also the ability for the user to teleport around with their controller, but regardless of where they are, the one to one correlation of their motion and their perspective in the virtual world offers unique and potentially better methods of understanding and responding to the robot's environment.

One consistent problem with depth sensors, just like cameras, is that they can not detect things outside of their line of sight. This means that an operator can't see anything outside the robot's immediate field of view, or behind an obstacle. In a glovebox environment, this could be potentially problematic, especially if the robot's hands were to get in the way of its sensors and the objects that it is manipulating. The glovebox environment also offers unique solutions to this problem. In a traditional, dynamic environment, it is generally impossible to pre-install sensors and equipment to help a robot operate. However, in a controlled glovebox, you could easily pre-install additional sensors inside, allowing you information from several vantage points. By taking the data from multiple sensors, the interface can automatically combine and register them into a single depth render, significantly reducing the shadow effect of obstacles that would be presented if the object was seen from only one perspective. This further increases the benefit provided by the user's ability to walk around and naturally see from different vantage points.

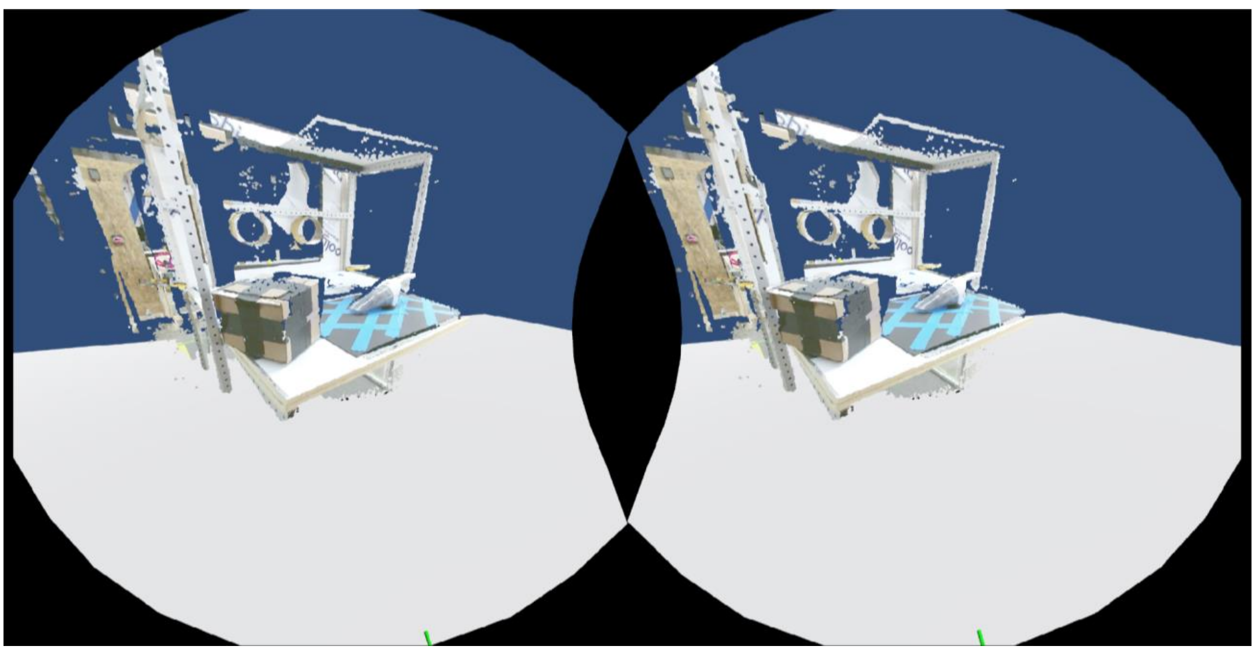

Figure 6. *Point cloud of a sample workspace, visualized in VR.*

Conclusions

We have outlined our current work and methodology for constructing a VR interface to control a humanoid robot. Our team's goal is to provide a human operator the ability to control a robot in a remote environment in a safe and accurate manner, using the specific case of an operator attempting to perform fine manipulation in a constrained environment, such as a glovebox. We believe that in some situations VR could provide benefits such as increased task awareness and situation awareness when working in remote environments. Going forward, we are looking to determine in what ways and what situations VR can provide a benefit, and how to maximize those. In addition, we are looking at determining how information can be adapted to display in 3D, in order to provide a clear benefit over that same information displayed on a 2D screen.

Acknowledgements
The work described in this paper was supported in part by the Department of Energy under award DE-EM0004482, by NASA under award NNX16AC48A, and by the National Science Foundation under award 1451427.

REFERENCES

1. Type B Accident Investigation Board Report, Employee Puncture Wound at the F-Tru Waste Remediation Facility at the Savannah River Site on June 14, 2010, available online http://energy.gov/ehss/downloads/type-b-accidentinvestigation-board-report-employee-puncture-wound-f-tru-waste
2. LLC, Cleatech. "Portable Glove Box System: S-2200." *Portable Glove-Box System S-2200 from Cleatech*, 1 Jan. 2016, www.laboratory-supply.net/gloveboxes/portable_glove_box.html.
3. "SEA-III Gloveboxes." Germfree, www.germfree.com/product-lines/laboratory-equipment/class-iii-biosafety-cabinets/sea-gloveboxes/sea-gloveboxes/.
4. Fallon, Maurice. "Perception and estimation challenges for humanoid robotics: DARPA Robotics Challenge and NASA Valkyrie." *Unmanned/Unattended Sensors and Sensor Networks XII*. Vol. 9986. International Society for Optics and Photonics, 2016.
5. Polido, Henrique. *DARPA Robotics Challenge*. Diss. Worcester Polytechnic Institute, 2014.
6. Johnson, Matthew, et al. "Team IHMC's lessons learned from the DARPA robotics challenge trials." *Journal of Field Robotics*32.2 (2015): 192-208.
7. Ferland, François, et al. "Egocentric and exocentric teleoperation interface using real-time, 3D video projection." *Human-Robot Interaction (HRI), 2009 4th ACM/IEEE International Conference on*. IEEE, 2009.
8. "VIVE VR SYSTEM." *VIVE™ | VIVE Virtual Reality System*, www.vive.com/us/product/vive-virtual-reality-system/.
9. Norton, Adam, et al. "Analysis of human–robot interaction at the DARPA Robotics Challenge Finals." *The International Journal of Robotics Research* (2017): 0278364916688254.