International Conference on Virtual, Augmented and Mixed Reality

Augmented, Mixed, and Virtual Reality Enabling of Robot Deixis

Tom Williams*, Nhan Tran, Josh Rands, and Neil T. Dantam

Colorado School of Mines, Golden CO, USA

Abstract. When humans interact with each other, they often make use of *deictic gestures* such as pointing to help pick out targets of interest to their conversation. In the field of Human-Robot Interaction, research has repeatedly demonstrated the utility of enabling robots to use such gestures as well. Recent work in augmented, mixed, and virtual reality stands to enable enormous advances in robot deixis, both by allowing robots to gesture in ways that were not previously feasible, and by enabling gesture on robotic platforms and environmental contexts in which gesture was not previously feasible. In this paper, we summarize our own recent work on using augmented, mixed, and virtual-reality techniques to advance the state-of-the-art of robot-generated deixis.

Keywords: Human-Robot Interaction, Deixis, Nonverbal Interaction, Teleoperation

1 Introduction

When humans interact with each other, they often make use of *deictic gestures* [1] such as pointing to help pick out targets of interest to their conversation [2]. In the field of Human-Robot Interaction, many researchers have explored how we might enable *robots* to generate the arm motions necessary to effect these same types of deictic gestures [3–8]. However, a number of challenges remain to be solved if effective robot-generated deictic gestures are to be possible *regardless of morphology and context*. Consider, for example, the following scenario:

A mission commander in an alpine search and rescue scenario instructs an unmanned aerial vehicle (UAV) "Search for survivors behind that fallen tree." The UAV can see three fallen trees and wishes to know which its user means.

This scenario presents at least two challenges. First, there is a problem of morphology. The UAV's lack of arms means that generating deictic gestures may not be physically possible. Second, there is a problem of context. Even if the UAV had an arm with which to gesture, doing so might not be effective; picking out far-off fallen trees within a forest may be extremely difficult using traditional gestures.

* The first author can be reached at twilliams@mines.edu. The first three authors are with the MIRRORLab (mirrorlab.mines.edu). The last author is with the DYALab (dyalab.mines.edu).

doi: 10.1007/978-3-319-91581-4_19

Recent advances in augmented and mixed reality technologies present the opportunity to address these challenges. Specifically, such technologies enable new forms of deictic gesture for robots with previously problematic morphologies and in previously problematic contexts. For example, in the previous example, if the mission commander were wearing an augmented reality head-mounted display, the UAV may have been able to pick out the fallen trees it wished to disambiguate between by circling them in the mission commander's display while saying "Do you mean *this tree*, *this tree*, or *this tree*?"

While there has been little previous work on using augmented, mixed, or virtual reality techniques for human-robot interaction, this is beginning to change. In March 2018, the first international workshop on Virtual, Augmented, and Mixed-Reality for Human-Robot Interaction (VAM-HRI) was held at the 2018 international conference on Human-Robot Interaction (HRI) [9]. The papers and discussion at that workshop make it evident that we should begin to see more and more research emerging at this intersection of fields.

In this paper, we summarize our own recent work on using augmented, mixed and virtual reality techniques to advance the state-of-the-art of robot-generated deixis, some of which was presented at the 2018 VAM-HRI workshop. In Section 2, we begin by providing a framework for categorizing robot-generated deixis in augmented and mixed-reality environments. In Section 3, we then discuss a novel method for enabling mixed reality deixis for armless robots. Finally, in Section 4 we present a novel method for robot teleoperation in Virtual Reality, and discuss how it could be used to trigger mixed-reality deictic gestures.

2 A Framework for Mixed-Reality Deictic Gesture

Augmented and mixed-reality technologies offer new opportunities for robots to communicate about the environments they share with human teammates. In previous work, we have presented a variety of work seeking to enable fluid natural language generation for robots operating in realistic human-robot interaction scenarios [10,11] (including work on referring expression generation [12,13], clarification request generation [14], and indirect speech act generation [15–17]). By augmenting their natural language references with visualizations that pick out their objects, locations, and people of interest within teammates' head-mounted displays, robots operating in such scenarios may facilitate conversational grounding [18,19] and shared mental modeling [20] with those human teammates in ways that were not previously possible.

While there has been some previous work on using visualizations as "gestures" within virtual or augmented environments [21] and video streams [22], as well as previous work on generating visualizations to accompany generated text [23–26], this metaphor of visualization-as-gesture has not yet been fully explored. This is doubly true for human-robot interaction scenarios, in which the use of augmented reality for human-robot communication is surprisingly underexplored. In fact, in their recent survey of augmented reality, Billinghurst et

doi: 10.1007/978-3-319-91581-4_19

al. [27] cite intelligent systems, hybrid user interfaces, and collaborative systems as areas that have been under-attended-to in the AR community.

Most relevant to the current paper, Sibertsiva et al. [28] use augmented reality annotations to indicate different candidates referential hypotheses after receiving ambiguous natural language commands, and Green et al. [29] present a system that uses augmented reality to facilitate human-robot discussion of a plan prior to execution. There have also been several recent approaches to using augmented reality to non-verbally communicate robots' intentions [30–36] These approaches, however, have looked at visualization alone, outside the context of traditional robot gesture. We believe that, just as augmented and mixed reality open up new avenues for communication in human-robot interaction, human-robot interaction opens up new avenues for communication in augmented and mixed reality. Only in *mixed-reality human-robot interaction* may physical and virtual gestures be generated together or chosen between as part of a single process. In order to understand the different types of gestures that can be used in mixed-reality human-robot interaction, we have been developing a framework for analyzing such gestures along dimensions such as embodiment, cost, privacy, and legibility [37]. In this paper, we extend that framework to encompass new gesture categories and dimensions of analysis.

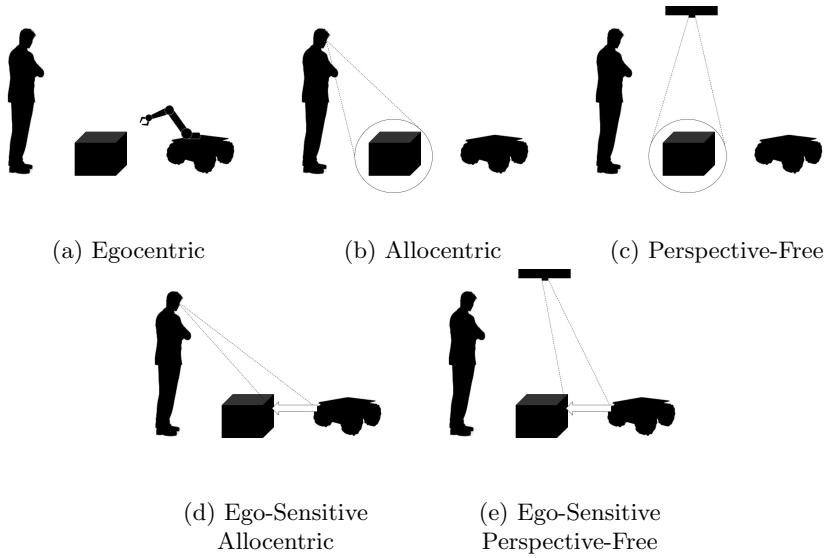
2.1 Conceptual Framework

In this section, we present a conceptual framework for describing mixed-reality deictic gestures. A robot operating within a pure-reality environment has access to but a single interface for generating gestures (its own body) and accordingly but a single perspective within which to generate them (its own)[1]. A robot operating within a mixed-reality environment, however, may leverage the hardware that enables such an environment, and the additional perspectives that come with those hardware elements. For robots operating within mixed-reality environments, we identify three unique hardware elements that can be used for deixis, each of which comes with its own perspective, and accordingly, their own class of deictic gestures.

First, robots may use their own bodies to perform the typical deictic gestures (such as pointing) available in pure reality. We categorize such gestures as *egocentric* (as shown in Fig. 1a), because they are generated from their own perspective. Second, robots operating in mixed-reality environments may be able to use of head-mounted displays worn by human teammates. We categorize such gestures as *allocentric* (as shown in Fig. 1b) because they are generated using only the perspective of the display's wearer. A robot, may, for example, "gesture" to an object by circling it within its teammate's display. Third, robots operating in mixed-reality environments may be able to use projectors to change how the world is perceived for all observers. We categorize such gestures as *perspective-free* (as shown in Fig. 1c) because they are not generated from the perspective of any one agent.

[1] Excepting, for the purposes of this paper, robots who are distributed across multiple sub-bodies in the environment [38].

doi: 10.1007/978-3-319-91581-4_19



(a) Egocentric (b) Allocentric (c) Perspective-Free

(d) Ego-Sensitive Allocentric (e) Ego-Sensitive Perspective-Free

Fig. 1: Categories of Mixed-Reality Deictic Gestures

In addition, robots operating in mixed-reality environments may be able to perform multi-perspective gestures that use the aforementioned mixed-reality hardware in a way that connects back to the robot's perspectives. A robot may, for example, gesture to an object in its teammate's display, or using a projector, by drawing an arrow from itself to its target object, or by gesturing towards its target using a virtual appendage that only exists in virtuality. We call the former class *ego-sensitive allocentric gestures* and the latter class *ego-sensitive perspective-free gestures*.

2.2 Analysis of Mixed-Reality Deictic Gestures

Each of these gestural categories comes with its own unique properties. Here, we specifically examine six: perspective, embodiment, capability, privacy, cost, and legibility. These dimensions are summarized in Table 1.

The most salient dimensions that differentiate these categories of mixed-reality deictic gestures are the perspectives, embodiment, and capabilities they require. The perspectives required are clearly defined: egocentric gestures require access to the robot's perspective, allocentric gestures require access to the human interlocutor's perspective, and perspective-free gestures require access only to the greater environment's perspective. The ego-sensitive gestures connect their initial perspective with that of the robot. Those categories generated from or connected to the perspective of the robot notably require the robot to be embodied and co-present with their interlocutor; but only the egocentric category requires the robot's embodied form to be capable of movement.

doi: 10.1007/978-3-319-91581-4_19

Table 1: Analysis of Mixed-Reality Deictic Gestures

Category	HW	CG	Per	Emb	Cap	Pri (L)	Pri (G)	Cost (G)	Cost (M)	Leg (D)	Leg (S)
Ego	Rob	Yes	Rob	Yes	Yes	Low	High	High	Low	Low	Low
Allo	HMD	No	Hum	No	No	High	Low	Low	High	High	High
P-F	Pro	No	Env	No	No	Low	Low	Low	Low	High	High
ES Allo	HMD	Yes	Rob+Hum	Yes	No	High	Low	Low	High	TED	ED
ES P-F	Pro	Yes	Rob+Env	Yes	No	Low	Low	Low	Low	TED	ED

Dimensions: HW = Hardware; CG = Connection to Generator; Per = Perspective; Emb = Embodiment; Cap = Capability; Pri = Privacy (Local/Global); Cost; Leg = Legibility (Dynamic/Static)
Perspectives: Ego = Egocentric; Allo = Allocentric; P-F = Perspective-Free; ES = Ego-Sensitive
Features: Rob=Robot; HMD = Head-Mounted Display; Pro = Projector; Hum = Human; Env = Environment; TED = Time and extent dependent; ED = Extent dependent

The different hardware needs of these categories result in different levels of privacy. Here, we distinguish between *local* privacy and *global* privacy. We describe those categories that use a head-mounted display as affording high local privacy, as gestures are only visible to the human teammate with whom the robot is communicating. This dimension is particularly important for human-robot interaction scenarios involving both sensitive user populations (e.g., elder care or education) or in adversarial scenarios (e.g., competitive [39], police [40], campus safety [41], or military domains (as in DARPA's "Silent Talk" program) [42]). On the other hand, we describe egocentric gestures as having high *global* privacy, as, unlike with the other categories, information about gestural data need not be sent over a network, and thus may not be as vulnerable to hackers.

These categories of mixed-reality deictic gestures also come with different technical challenges, resulting in different computational costs. From the perspective of energy usage, egocentric gestures are expensive due to their physical component (a high *generation cost*). On the other hand, gestures that make use of a head-mounted display may be expensive to maintain due to registration challenges (a high *maintenance cost*).

Finally, these gestures differ with respect to legibility. In previous work, Dragan et al. [43] defined the notion of the legibility of an action, which describes the ease with which a human observer is able to determine the goal or purpose of an action as it is being carried out. In later work with Holladay et al. [5], Dragan then applies this notion to deictic gestures as well, analyzing the ability of the final gestural position to enable humans to pick out the target object. We believe, however, that this is really a distinct sense of legibility from Dragan's original formulation, and as such, we first refine this notion of legibility as applied to deictic gestures into two categories: we use *dynamic legibility* to refer to the degree to which a deictic gesture enables a human teammate to pick

doi: 10.1007/978-3-319-91581-4_19

out the target object *as the action is unfolding* (in line with Dragan's original formulation), and *static legibility* to refer to the degree to which the final pose of a deictic gesture enables a human teammate to pick out the target object after the action is completed (in line with Holladay's formulation).

The gestural categories we describe differ with respect to both dynamic and static legibility. Allocentric and perspective-free gestures have high dynamic legibility (given that there is no dynamic dimension) and high static legibility (given that the target is uniquely picked out). Egocentric gestures have low dynamic legibility (relative to allocentric gestures) given that their target may not be clear at all as the action unfolds, and low static legibility, as the target may not be clear after the action is performed either, depending on distance to the target and density of distractors. The legibility of multi-perspective gestures depends on how exactly they are displayed. If they extend all the way to a target object, they may have high static legibility, whereas if they only point toward the target they will have low static legibility. Dynamic legibility depends both on this factor, as well as temporal extent. If a multi-perspective gesture unfolds over time, this may decrease the legibility (although it may better capture the user's attention).

2.3 Combination of Mixed-Reality Deictic Gestures

Finally, given these classes of mixed-reality deictic gestures, we can also reason about combinations of these gestures. Rather than explicitly discuss all 31 non-empty combinations of these five categories, we will briefly describe *how* the gestural categories combine. Simultaneous generation of gestures requiring different perspectives results in both perspectives being needed. The embodiment and capability requirements of simultaneous gestures combine disjunctively. The legibilities and costs of simultaneous gestures combine using a *max* operator, as the legibility of one gesture will excuse the illegibility of another, but the low cost of one gesture will not excuse the high cost of another. And the privacies of simultaneous gestures combine using a *min* operator, as the high privacy of one gesture does not excuse the low privacy of another.

3 Enabling Deictic Capabilities for Armless Robots using Mixed-Reality Robotic Arms

In the previous section, we presented a framework for analyzing mixed-reality deictic gestures. Within this framework, the gestural categories that have received the least amount of previous attention are the ego-sensitive categories which connect the gesture-generating robot with the perspective of the human viewer or with the perspective of their environment. In this section, we present a novel approach to *ego-sensitive allocentric gesture*. Specifically, we propose to superimpose mixed-reality visualizations of robot arms onto otherwise armless robots, to allow them to gesture within their environment. This will allow an armless robot like a wheelchair or drone to gesture just as if it had a physical

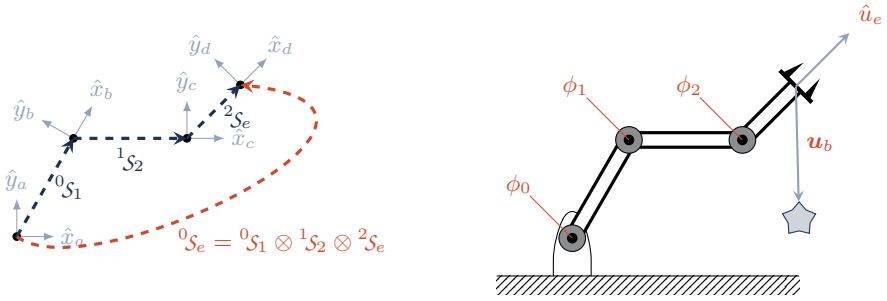
doi: 10.1007/978-3-319-91581-4_19

arm, even if mounting such an arm would not be mechanically possible or cost effective. Unlike purely allocentric gestures (e.g., circling an object in ones' field of view), this approach emphasizes the generator's embodiment, and as such, we would expect it to lead to increased perception of the robot's agency, increased likability of the robot, and promote positive team dynamics.

In this section we present the preliminary technical work necessary to enable such an approach. Specifically, we present a kinematic approach to perform this kind of mixed-reality deictic gesture. Compared to motion planning, a purely kinematic approach is more computationally efficient, a potential advantage for low-power embedded systems that we may wish to use for AR displays. The trade-off is that the kinematic approach is *incomplete*, so it may fail to find collision-free motions for some cluttered environments. However, collisions are not an impediment for virtual arms, thus mitigating the potential downside of purely kinematic motions.

Our approach applies dual-quaternion forward kinematics and Jacobian damped-least-squares inverse kinematics.

3.1 Kinematics



(a) (b)

Fig. 2: Kinematic diagrams for diectic gestures. **(a)** the local coordinate frames ("frames") of a serial manipulator. **(b)** a schematic of a serial manipulator with vectors for pointing direction and the vector to a target object.

Forward Kinematics We adopt the conventional model for serial robot manipulators of kinematic chains and trees [44–49]. Each local coordinate frame ("frame") of the robot has an associated label, and the frames are connected by Euclidean transformations (see Fig. 2a).

We represent Euclidean transformations with dual quaternions. Compared to matrix representations, dual quaternions offer computational advantages in efficiency, compactness, and numerical stability. A dual quaternion is a pair of

doi: 10.1007/978-3-319-91581-4_19

quaternions: an *ordinary* part for rotation and *dual* part for translation. Notationally, we use a leading superscript to denote the parent's local coordinate frame p and trailing subscript to denote the child frame c. Given rotation unit quaternion $\hbar$ and translation vector $\mathbf{v}$ from p to c, the transformation dual quaternion ${}^{p}\mathcal{S}_c$ is:

$$\begin{aligned} \hbar &= \left(h_x\hat{\boldsymbol{\imath}} + h_y\hat{\boldsymbol{\jmath}} + h_z\hat{\boldsymbol{k}} + h_w\right) \\ \mathbf{v} &= \left(v_x\hat{\boldsymbol{\imath}} + v_y\hat{\boldsymbol{\jmath}} + v_z\hat{\boldsymbol{k}} + 0\right) \\ {}^{p}\mathcal{S}_c &= \left(\hbar + \left(\frac{1}{2}\mathbf{v} \otimes \hbar\right)\boldsymbol{\varepsilon}\right) \end{aligned} \tag{1}$$

where $\hat{\boldsymbol{\imath}}$, $\hat{\boldsymbol{\jmath}}$, $\hat{\boldsymbol{k}}$ are the imaginary elements, with $\hat{\boldsymbol{\imath}}^2 = \hat{\boldsymbol{\jmath}}^2 = \hat{\boldsymbol{k}}^2 = \hat{\boldsymbol{\imath}}\hat{\boldsymbol{\jmath}}\hat{\boldsymbol{k}} = -1$, and $\boldsymbol{\varepsilon}$ is the dual element, with $\boldsymbol{\varepsilon}^2 = 0$ and $\boldsymbol{\varepsilon} \neq 0$.

Chaining transformations corresponds to multiplication of the transformation matrices or the dual quaternion. For a kinematic chain, we must match the child frame of predecessor to parent frame of successor transformations. The result is the transform from the parent of the initial to the child of the final transformation.

$${}^{a}\mathcal{S}_b \otimes {}^{b}\mathcal{S}_c = {}^{a}\mathcal{S}_c \tag{2}$$

We illustrate the kinematics computation for the simple serial manipulator in Fig. 2b. Note that the local frames and relative transforms of the robot in Fig. 2b correspond to those drawn in Fig. 2a.

The kinematic position of a robot is fully determined by its configuration $\boldsymbol{\phi}$, i.e, the vector of joint angles,

$$\boldsymbol{\phi} = [\phi_0, \phi_1, \ldots, \phi_n]^T \tag{3}$$

The relative frame at each joint i is a function of the corresponding configuration: ${}^{i-1}\mathcal{S}_i(\phi_i)$. The frame for the end-effector is the product of all frames in the chain

$${}^{0}\mathcal{S}_e(\boldsymbol{\phi}) = \left({}^{0}\mathcal{S}_1(\phi_0)\right) \otimes \left({}^{1}\mathcal{S}_2(\phi_1)\right) \otimes \ldots \otimes \left({}^{n-1}\mathcal{S}_n(\phi_{n-1})\right) \otimes \left({}^{n}\mathcal{S}_e(\phi_n)\right) \tag{4}$$

Cartesian Control We compute the least-squares solution for Cartesian motion using a singularity-robust Jacobian pseudoinverse:

$$\dot{\mathbf{x}} = \mathbf{J}\dot{\boldsymbol{\phi}} \quad \rightsquigarrow \quad \dot{\boldsymbol{\phi}} = \mathbf{J}^{+}\dot{\mathbf{x}} \tag{5}$$

$$\mathbf{J}^{+} = \sum_{i=0}^{\min(m,n)} \frac{s_i}{\max({s_i}^2, {s_{\min}}^2)} \mathbf{v_i}\mathbf{u_i}^T \tag{6}$$

where $\dot{\mathbf{x}} = [\omega, \dot{v}]$ is the vector of rotational velocity ω and translational velocity $\dot{v}$, $\mathbf{J} = \mathbf{USV}^T$ is the singular value decomposition[2] of Jacobian J, and $s_{\min}$ is a selected constant for the minimum acceptable singular value.

[2] The SVD, while more expensive to compute, provides better accuracy and numerical stability for Cartesian control than the LU decomposition.

doi: 10.1007/978-3-319-91581-4_19

We determine Cartesian velocity $\dot{\mathbf{x}}$ with a proportional gain on position error, computed as the velocity to reach the desired target in unit time, decoupling rotational ω and translational $\dot{v}$ parts to achieve straight-line translations:

$$\hbar_1 = \exp\left(\frac{\omega \Delta t}{2}\right) \otimes \hbar_0 \quad \rightsquigarrow \quad -\omega \Delta t = 2 \ln\left(\hbar_0 \otimes \hbar_1^*\right) \tag{7}$$

$$v_1 = \dot{x} \Delta t + v_0 \quad \rightsquigarrow \quad -\dot{x} \Delta t = v_0 - v_1 \tag{8}$$

In combination, we compute the reference joint velocity as:

$$\dot{\boldsymbol{\phi}} = \mathbf{J}^+ \left(-k \begin{bmatrix} 2 \ln\left({}^0\hbar_e \otimes \left({}^0\hbar_r\right)^*\right) \\ {}^0v_e - {}^0v_r \end{bmatrix} \right) \tag{9}$$

where e is the actual end-effector frame and r is the desired or reference frame.

3.2 Design Patterns

While the method above provides a general approach to enabling mixed-reality deictic gestures, there are a variety of different possible forms of deictic gestures that might be generated using that approach. In this section, we propose three candidate gesture designs enabled by the proposed approach: *Fixed Translation*, *Reaching*, and *Floating*.

Fixed Translation The first proposed design, *Fixed Translation*, is the most straightforward manifestation of the proposed approach. In this design, the visualized arm rotates in place to point to the desired target. To enable this design, we must find a target orientation ${}^0\hbar_r$ for the end-effector. We find the relative rotation ${}^e\hbar_r$ between the current end-effector frame and pointing direction towards the target based on the end-effector's pointing direction vector and the vector from the end-effector to the target (see Fig. 2b).

First, we find the end-effector's global pointing vector $\hat{u}_e$ by rotating the local pointing direction $\hat{a}_e$.

$$\hat{u}_e = {}^0\hbar_e \otimes \hat{a}_e \otimes \left({}^0\hbar_e\right)^* \tag{10}$$

Then, we find the vector from the end-effector to the target by subtracting the end-effector translation ${}^0v_e$ from the target translation ${}^0v_b$ and normalizing to a unit vector.

$$\begin{aligned} {}^0v_e &= 2\ {}^0d_e \otimes \left({}^0\hbar_e\right)^* \\ \hat{u}_b &= \frac{{}^0v_b - {}^0v_e}{\|{}^0v_b - {}^0v_e\|} \end{aligned} \tag{11}$$

Next, we compute the relative rotation between the two vectors $\hat{u}_e$ and $\hat{u}_b$ using the dot product to find the angle θ and cross product to find the axis $\hat{a}$,

$$\theta = \cos^{-1}\left(\hat{u}_e \bullet \hat{u}_b\right) \tag{12}$$

$$\hat{a} = \frac{\hat{u}_e \times \hat{u}_b}{\sin\theta} \tag{13}$$

doi: 10.1007/978-3-319-91581-4_19

The axis $\hat{a}$ and angle θ then give us the rotation unit quaternion ${}^{e}\mathcal{h}_{b}$:

$$ {}^{e}\mathcal{h}_{r} = \left(\hat{a} \sin \frac{\theta}{2} + \cos \frac{\theta}{2} \right) \tag{14} $$

Note that a direct conversion of the vectors to the rotation unit quaternion avoids the need for explicit evaluation of transcendental functions.

Now, we compute the global reference frame for the end-effector using

$$ \begin{aligned} {}^{e}\mathcal{S}_{r} &= \left({}^{e}\mathcal{h}_{b} + 0\boldsymbol{\varepsilon} \right) \\ {}^{0}\mathcal{S}_{r} &= {}^{0}\mathcal{S}_{e} \otimes {}^{e}\mathcal{S}_{b} \end{aligned} \tag{15} $$

Combining, (15) and (9), we compute the joint velocities $\dot{\boldsymbol{\phi}}$ for the robot arm.

Reaching Our second proposed design, *Reaching*, stretches the arm out towards the target, increasing gesture legibility in a way that would not be feasible with a physical arm. To enable this design, we compute the instantanous desired orientation as in the fixed translation case, but now set the desired translation to the target object's translation ${}^{0}\mathcal{v}_{b}$.

$$ \begin{aligned} {}^{0}\mathcal{h}_{r} &= {}^{0}\mathcal{h}_{e} \otimes {}^{e}\mathcal{h}_{b} \\ {}^{0}\mathcal{S}_{r} &= \left({}^{0}\mathcal{h}_{r} + \left(\frac{1}{2} {}^{0}\mathcal{h}_{r} \otimes {}^{0}\mathcal{v}_{b} \right) \boldsymbol{\varepsilon} \right) \end{aligned} \tag{16} $$

Then we combine, (16) and (9) to compute the joint velocities $\dot{\boldsymbol{\phi}}$ for the robot arm.

Floating Translation Diectic information is conveyed primarily by the orientation of the end-effector rather than its translation. Thus, in our final design, *Floating Translation*, we consider a case where the translation can freely float, allowing the arm to point with more natural-looking configurations. First, we remove the translational component from the control law. Second, we center all joints within the Jacobian null space, so centering does not impact end-effector velocity. We update the workspace control law with a weighting matrix and null space projection term:

$$ \dot{\boldsymbol{\phi}} = \mathbf{J}^{+} \mathbf{W} \dot{\mathbf{x}} + \left(\mathbf{I} - \mathbf{J}^{+} \mathbf{J} \right) \dot{\boldsymbol{\phi}}_{N} \tag{17} $$

The weighting matrix $\mathbf{W}$ removes the translational component from the Jacobian $\mathbf{J}$, so only rotational error contributes to the joint velocity $\dot{\boldsymbol{\phi}}$. Structurally, $\mathbf{J}^{+}$ consists of rotational block $\mathbf{j}_{\omega}^{+}$ and translational block $\mathbf{j}_{\dot{v}}^{+}$. We construct $\mathbf{W}$ to remove $\mathbf{j}_{\dot{v}}^{+}$.

$$ \begin{aligned} \mathbf{J}^{+} &= \left[\mathbf{j}_{\omega}^{+} \mid \mathbf{j}_{\dot{v}}^{+} \right] \\ \mathbf{W} &= \begin{bmatrix} \mathbf{I}_{3 \times n} \\ \mathbf{0}_{3 \times n} \end{bmatrix} \\ \mathbf{J}^{+} \mathbf{W} &= \left[\mathbf{j}_{\omega}^{+} \mid \mathbf{0} \right] \end{aligned} \tag{18} $$

doi: 10.1007/978-3-319-91581-4_19

where n is the length of $\boldsymbol{\phi}$, or equivalently the number of rows in $\mathbf{J}^+$.

We use the null space projection to move all joints towards their center configuration, without impacting on end-effector pose:

$$\dot{\boldsymbol{\phi}} = \boldsymbol{\phi}_c - \boldsymbol{\phi}_a \tag{19}$$

where $\boldsymbol{\phi}_c$ is the center configuration and $\boldsymbol{\phi}_a$ is the actual configuration.

The combined workspace control law is

$$\dot{\boldsymbol{\phi}} = \left(\mathbf{J}^+\right) \begin{bmatrix} \mathbf{I}_{3\times n} \\ \mathbf{0}_{3\times n} \end{bmatrix} \dot{\mathbf{x}} + \left(\mathbf{I} - \mathbf{J}^+\mathbf{J}\right)\left(\boldsymbol{\phi}_c - \boldsymbol{\phi}_a\right) \tag{20}$$

In this paper, we have proposed a new form of mixed-reality deictic gesture, and proposed a space of candidate designs for manifesting such gestures. In current and future work, we will implement all three designs using the Microsoft Hololens, and evaluate their performance with respect to both each other, and to the other categories of gesture we have described. In the next section, we turn to methods by which such gestures might be generated by human teleoperators during human-subject experiments.

4 An Interface for Virtual Reality Teleoperation

In the previous sections, we presented a framework for mixed-reality deixis, and a novel form of mixed-reality deictic gesture. But a question remains as to how robots might decide to generate such gestures. While in future work our interests lie in computational approaches for allowing robots to decide for themselves when and how to generate such gestures, in this work we first examine how humans might trigger such gestures, and how novel *virtual reality* technologies might facilitate this process.

Specifically, we examine the use of virtual reality and gesture recognition technologies may be used to control gesture-capable robots used by Human-Robot Interaction (HRI) researchers during human-subject experiments [50]. Manual control of language- and gesture-capable robots is crucial for HRI researchers seeking to evaluate human perceptions of potential autonomous capabilities which either do not yet exist, or are not yet robust enough to work consistently and predictably, as in the *Wizard of Oz* (WoZ) experimental paradigm [51]. For the purposes of such experiments, manual control of dialogue and gestural capabilities is particularly challenging [52]. Not only is it repetitive and time consuming to design WoZ interfaces for such capabilities, but such interfaces are not always effective, as the time necessary for an experimenter to decide to issue a command, click the appropriate button, and have that command take effect on the robot is typically too long to facilitate natural interaction.

What is more, such interfaces typically require experimenters to switch back and forth between monitoring a camera stream depicting the robot's environment and consulting their control interface: a pattern that can decrease robots' situational awareness and harm experiment effectiveness [53]. This is particularly true when the camera stream depicts the robot's environment from a

doi: 10.1007/978-3-319-91581-4_19

third-person perspective, which can lead to serious performance challenges [54]. While some recent approaches have introduced the use of augmented reality for safely teleoperating co-present robots [55,56], robots are not typically co-present with teleoperators during tightly controlled WoZ experiments. For such applications, Virtual Reality (VR) teleoperation provides one possible solution. VR is also beneficial as immersion in the robot's perspective improves depth perception and enhances visual feedback, resulting in an overall more immersive experience [57]. On the other hand, immersive first-person teleoperation comes with its own concerns. Recent researchers have noted safety concerns, as a sufficiently constrained robot perspective may limit the teleoperator's situational awareness [58]. What is more, VR teleoperation in particular raises challenges as the teleoperator may no longer be able see their teleoperation interface.

4.1 Previous Work

There have been a large number of approaches to robot teleoperation through virtual reality, even within only the past year. First, there has been some work on robot teleoperation using touchscreens displaying first- or third-person views of the robot's environment [59, 60]. There have been a number of approaches enabling first-person robot teleoperation using virtual reality displays, using a variety of different control modalities, including joysticks [61], VR hand controllers [62–67], gloves [68, 69], and full-torso exo-suits [70]. There has been less work enabling hands-free teleoperation, with the closest previous work we are aware of being Miner and Stansfield's approach, which allowed gesture-based control in *simulated, third-person* virtual reality. The only approaches we are aware of enabling first-person hands-free control are our own approach (discussed in the next section), and the Kinect-based approach of Sanket et al., which was presented at the same workshop as our own work [66].

4.2 Integrated Approach

(a) The teleoperation interface, and its integration with the Softbank Pepper robot.

(b) Simulated stereoscopic display displayed to the user through the VR headset.

(c) View of gesture tracking data in the Leap Motion SDK Diagnostic Visualizer.

Fig. 3: Multiple views of integrated system

doi: 10.1007/978-3-319-91581-4_19

In our recent work [50], we have proposed a novel teleoperation interface which provides hands-free WoZ control of a robot while providing the teleoperator with an immersive VR experience from the robot's point of view. This interface integrates a VR headset, interfaced directly with the robot's camera to allowing the experimenter to see exactly what the robot sees (Fig. 3b), with a Leap Motion Controller. Translating traditional joystick or gamepad control to robotic arm motions can be challenging, but the Leap Motion Controller can simplify this process by allowing the user to replicate the gesture he/she desires of the robot, making it a powerful hands-free teleoperation device [71]. There has been work on using the Leap Motion for teleoperation *outside* the context of virtual reality [72–74] but to the best of our knowledge our approach is the first to pair it with an immersive virtual-reality display. In our approach, we use the Leap Motion sensor to capture the experimenter's gestures, and then generate analogous gestures on the robot in real time. Specifically, we first extract hand position and orientation data from raw Leap Motion data. Fig. 3c shows the visualization of the tracking data produced by the Leap Motion. Each arrow represents a finger, and each trail represents the corresponding movement of that finger. Changes in this position and orientation data is used to trigger changes in the robot's gestures according to the following equations:

$$robotGesturePitch = \begin{cases} low \;\; \tau_{p_1} < humanGesturePitch < \tau_{p_2} \\ high \; \tau_{p_2} < humanGesturePitch < \tau_{p_3} \end{cases}$$

$$robotGestureRoll = \begin{cases} low \;\; \tau_{r_1} < humanGestureRoll < \tau_{r_2} \\ high \; \tau_{r_2} < humanGestureRoll < \tau_{r_3} \end{cases}$$

Here, parameters $\tau_{p_1} < \tau_{p_2} < \tau_{p_3}$ and $\tau_{r_1} < \tau_{r_2} < \tau_{r_3}$ are manually defined pitch and raw thresholds. While in this work our initial prototype makes use of these simple inequalities, in future work we aim to examine more sophisticated geometric and approximate methods for precisely mapping human gestures to robot gestures, with the aim of enabling a level of control currently seen in suit-based teleoperation systems [70].

All components of the proposed interface are integrated using the Robot Operating System (ROS) [75]. As shown in Fig. 4, the Leap Motion publishes raw sensor data, which is converted into motion commands. These motion commands are then sent to the robot[3]. Similarly, camera data is published by the robot, to a topic subscribed to by the Android VR app which displays it in the VR headset[4].

[3] While in this work we use the Softbank Pepper robot, our general framework is not necessarily specific to this particular robot.

[4] In this work, we use a single camera, as Pepper has a single RGB camera rather than stereo cameras. In the future, we hope to use stereoscopic vision as input for a more immersive VR experience. In addition, images could be just as easily streamed to this app from a ROS simulator (e.g., Gazebo [76]) or from some other source.

doi: 10.1007/978-3-319-91581-4_19

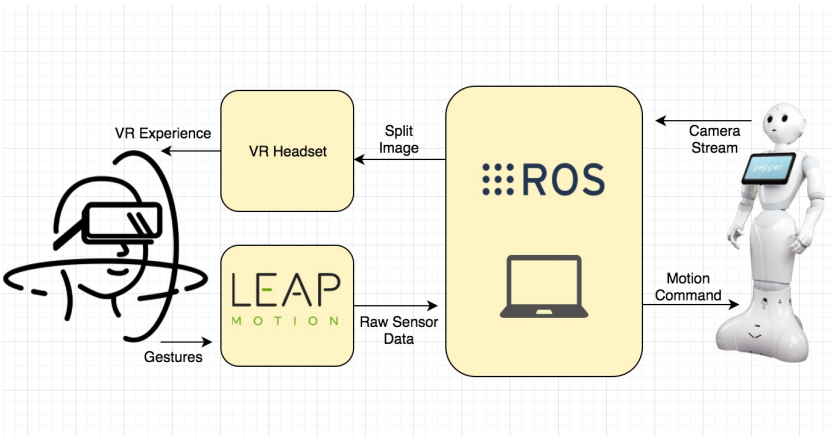


Fig. 4: Architecture diagram: The user interacts directly with a VR headset (e.g., Google Cardboard) and a Leap Motion gesture sensor. These devices send data to and receive data from a humanoid robot (e.g., the Softbank Pepper) using an instance of the ROS architecture whose Master node is run on a standard Linux laptop.

5 Conclusion

Virtual, augmented, and mixed reality stand to enable – and are already enabling – promising new paradigms for human-robot interaction. In this work, we summarized our own recent work in all three of these areas. We see a long, bright avenue for future work in this area for years to come. In our own future work, we plan to focus on exploring the space of different designs for mixed-reality deictic gesture, and integrating these approaches with our existing body of work on natural language generation, thus enabling exciting new ways for robots to express themselves.

References

1. McNeill, D.: Hand and mind: What gestures reveal about thought. University of Chicago press (1992)
2. Fillmore, C.J.: Towards a descriptive framework for spatial deixis. Speech, place and action: Studies in deixis and related topics (1982) 31–59
3. Salem, M., Kopp, S., Wachsmuth, I., Rohlfing, K., Joublin, F.: Generation and evaluation of communicative robot gesture. International Journal of Social Robotics **4**(2) (2012) 201–217
4. Huang, C.M., Mutlu, B.: Modeling and evaluating narrative gestures for humanlike robots. In: Robotics: Science and Systems. (2013) 57–64
5. Holladay, R.M., Dragan, A.D., Srinivasa, S.S.: Legible robot pointing. In: Robot and Human Interactive Communication, 2014 RO-MAN: The 23rd IEEE International Symposium on, IEEE (2014) 217–223

doi: 10.1007/978-3-319-91581-4_19

6. Sauppé, A., Mutlu, B.: Robot deictics: How gesture and context shape referential communication. In: Proceedings of the 2014 ACM/IEEE international conference on Human-robot interaction, ACM (2014) 342–349
7. Gulzar, K., Kyrki, V.: See what i mean-probabilistic optimization of robot pointing gestures. In: Proceedings of the fifteenth IEEE/RAS International Conference on Humanoid Robots (Humanoids), IEEE (2015) 953–958
8. Admoni, H., Weng, T., Scassellati, B.: Modeling communicative behaviors for object references in human-robot interaction. In: Proceedings of the IEEE International Conference on Robotics and Automation (ICRA), IEEE (2016) 3352–3359
9. Williams, T., Szafir, D., Chakraborti, T., Ben Amor, H.: Virtual, augmented, and mixed reality for human-robot interaction. In: Companion of the 2018 ACM/IEEE International Conference on Human-Robot Interaction, ACM (2018) 403–404
10. Williams, T.: A consultant framework for natural language processing in integrated robot architectures. IEEE Intelligent Informatics Bulletin (2017)
11. Williams, T.: Situated Natural Language Interaction in Uncertain and Open Worlds. PhD thesis, Tufts University (2017)
12. Kunze, L., Williams, T., Hawes, N., Scheutz, M.: Spatial referring expression generation for hri: Algorithms and evaluation framework. In: AAAI Fall Symposium on AI and HRI. (2017)
13. Williams, T., Scheutz, M.: Referring expression generation under uncertainty: Algorithm and evaluation framework. In: Proceedings of the 10th International Conference on Natural Language Generation. (2017)
14. Williams, T., Scheutz, M.: Resolution of referential ambiguity in human-robot dialogue using dempster-shafer theoretic pragmatics. In: Proceedings of Robotics: Science and Systems. (2017)
15. Briggs, G., Williams, T., Scheutz, M.: Enabling robots to understand indirect speech acts in task-based interactions. Journal of Human-Robot Interaction (2017)
16. Williams, T., Briggs, G., Oosterveld, B., Scheutz, M.: Going beyond literal command-based instructions: Extending robotic natural language interaction capabilities. In: Proceedings of the 29th AAAI Conference on Artificial Intelligence. (2015)
17. Williams, T., Thames, D., Novakoff, J., Scheutz, M.: thank you for sharing that interesting fact!: Effects of capability and context on indirect speech act use in task-based human-robot dialogue. In: Proceedings of the 13th ACM/IEEE International Conference on Human-Robot Interaction. (2018)
18. Fussell, S.R., Setlock, L.D., Kraut, R.E.: Effects of head-mounted and scene-oriented video systems on remote collaboration on physical tasks. In: Proceedings of the SIGCHI conference on Human factors in computing systems, ACM (2003) 513–520
19. Kraut, R.E., Fussell, S.R., Siegel, J.: Visual information as a conversational resource in collaborative physical tasks. Human-computer interaction **18**(1) (2003) 13–49
20. Datcu, D., Cidota, M., Lukosch, H., Lukosch, S.: On the usability of augmented reality for information exchange in teams from the security domain. In: Intelligence and Security Informatics Conference (JISIC), 2014 IEEE Joint, IEEE (2014) 160–167
21. White, S., Lister, L., Feiner, S.: Visual hints for tangible gestures in augmented reality. In: Mixed and Augmented Reality, 2007. ISMAR 2007. 6th IEEE and ACM International Symposium on, IEEE (2007) 47–50

doi: 10.1007/978-3-319-91581-4_19

22. Fussell, S.R., Setlock, L.D., Yang, J., Ou, J., Mauer, E., Kramer, A.D.: Gestures over video streams to support remote collaboration on physical tasks. Human-Computer Interaction **19**(3) (2004) 273–309
23. Wahlster, W., André, E., Graf, W., Rist, T.: Designing illustrated texts: how language production is influenced by graphics generation. In: Proceedings of the fifth conference on European chapter of the Association for Computational Linguistics, Association for Computational Linguistics (1991) 8–14
24. Wazinski, P.: Generating spatial descriptions for cross-modal references. In: Proceedings of the third conference on Applied natural language processing, Association for Computational Linguistics (1992) 56–63
25. Green, S.A., Billinghurst, M., Chen, X., Chase, J.G.: Human robot collaboration: An augmented reality approacha literature review and analysis. In: Proceedings of the ASME International Design Engineering Technical Conferences and Computers and Information in Engineering Conference, American Society of Mechanical Engineers (2007) 117–126
26. Green, S., Billinghurst, M., Chen, X., Chase, G.: Human-robot collaboration: A literature review and augmented reality approach in design. International Journal of Advanced Robotic Systems (2008)
27. Billinghurst, M., Clark, A., Lee, G.: A survey of augmented reality. Foundations and Trends in Human-Computer Interaction **8**(2-3) (2015) 73–272
28. Sibirtseva, E., Kontogiorgos, D., Nykvist, O., Karaoguz, H., Leite, I., Gustafson, J., Kragic, D.: A comparison of visualisation methods for disambiguating verbal requests in human-robot interaction. arXiv preprint arXiv:1801.08760 (2018)
29. Green, S.A., Chase, J.G., Chen, X., Billinghurst, M.: Evaluating the augmented reality human-robot collaboration system. International journal of intelligent systems technologies and applications **8**(1-4) (2009) 130–143
30. Andersen, R.S., Madsen, O., Moeslund, T.B., Amor, H.B.: Projecting robot intentions into human environments. In: Robot and Human Interactive Communication (RO-MAN), 2016 25th IEEE International Symposium on, IEEE (2016) 294–301
31. Chadalavada, R.T., Andreasson, H., Krug, R., Lilienthal, A.J.: That's on my mind! robot to human intention communication through on-board projection on shared floor space. In: Mobile Robots (ECMR), 2015 European Conference on, IEEE (2015) 1–6
32. Frank, J.A., Moorhead, M., Kapila, V.: Mobile mixed-reality interfaces that enhance human–robot interaction in shared spaces. Frontiers in Robotics and AI **4** (2017) 20
33. Ganesan, R.K.: Mediating human-robot collaboration through mixed reality cues. Master's thesis, Arizona State University (2017)
34. Katzakis, N., Steinicke, F.: Excuse me! perception of abrupt direction changes using body cues and paths on mixed reality avatars. arXiv preprint arXiv:1801.05085 (2018)
35. Rosen, E., Whitney, D., Phillips, E., Chien, G., Tompkin, J., Konidaris, G., Tellex, S.: Communicating robot arm motion intent through mixed reality head-mounted displays. arXiv preprint arXiv:1708.03655 (2017)
36. Walker, M., Hedayati, H., Lee, J., Szafir, D.: Communicating robot motion intent with augmented reality. In: Proceedings of the 2018 ACM/IEEE International Conference on Human-Robot Interaction, ACM (2018) 316–324
37. Williams, T.: A framework for robot-generated mixed-reality deixis. In: Proceedings of the 1st International Workshop on Virtual, Augmented, and Mixed Reality for HRI (VAM-HRI). (2018)

doi: 10.1007/978-3-319-91581-4_19

38. Oosterveld, B., Brusatin, L., Scheutz, M.: Two bots, one brain: Component sharing in cognitive robotic architectures. In: Proceedings of 12th ACM/IEEE International Conference on Human-Robot Interaction Video Contest. (2017)
39. Correia, F., Alves-Oliveira, P., Maia, N., Ribeiro, T., Petisca, S., Melo, F.S., Paiva, A.: Just follow the suit! trust in human-robot interactions during card game playing. In: Robot and Human Interactive Communication (RO-MAN), 2016 25th IEEE International Symposium on, IEEE (2016) 507–512
40. Bethel, C.L., Carruth, D., Garrison, T.: Discoveries from integrating robots into swat team training exercises. In: Safety, Security, and Rescue Robotics (SSRR), 2012 IEEE International Symposium on, IEEE (2012) 1–8
41. Goldfine, S.: Assessing the prospects of security robots (October 2017)
42. Kotchetkov, I.S., Hwang, B.Y., Appelboom, G., Kellner, C.P., Connolly Jr, E.S.: Brain-computer interfaces: military, neurosurgical, and ethical perspective. Neurosurgical focus **28**(5) (2010) E25
43. Dragan, A.D., Lee, K.C., Srinivasa, S.S.: Legibility and predictability of robot motion. In: Human-Robot Interaction (HRI), 2013 8th ACM/IEEE International Conference on, IEEE (2013) 301–308
44. Şucan, I.A., Chitta, S.: MoveIt! (2015) http://moveit.ros.org.
45. Dantam, N.T., Chaudhuri, S., Kavraki, L.E.: The task motion kit (accepted). Robotics and Automation Magazine (2018)
46. Diankov, R.: Automated Construction of Robotic Manipulation Programs. PhD thesis, Carnegie Mellon University, Robotics Institute (August 2010)
47. Hartenberg, R.S., Denavit, J.: Kinematic synthesis of linkages. McGraw-Hill (1964)
48. Smits, R., Bruyninckx, H., Aertbeliën, E.: KDL: Kinematics and dynamics library (2011) http://www.orocos.org/kdl.
49. Willow Garage: URDF XML (2013) http://wiki.ros.org/urdf/XML.
50. Tran, N., Rands, J., , Williams, T.: A hands-free virtual-reality teleoperation interface for wizard-of-oz control. In: Proceedings of the 1st International Workshop on Virtual, Augmented, and Mixed Reality for HRI (VAM-HRI). (2018)
51. Riek, L.D.: Wizard of oz studies in hri: a systematic review and new reporting guidelines. Journal of Human-Robot Interaction **1**(1) (2012)
52. Bonial, C., Marge, M., Foots, A., Gervits, F., Hayes, C.J., Henry, C., Hill, S.G., Leuski, A., Lukin, S.M., Moolchandani, P., et al.: Laying down the yellow brick road: Development of a wizard-of-oz interface for collecting human-robot dialogue. arXiv preprint arXiv:1710.06406 (2017)
53. Chen, J.Y.C., Haas, E.C., Barnes, M.J.: Human performance issues and user interface design for teleoperated robots. IEEE Transactions on Systems, Man, and Cybernetics, Part C (Applications and Reviews) **37**(6) (Nov 2007) 1231–1245
54. Nawab, A., Chintamani, K., Ellis, D., Auner, G., Pandya, A.: Joystick mapped augmented reality cues for end-effector controlled tele-operated robots. In: 2007 IEEE Virtual Reality Conference. (March 2007) 263–266
55. Gong, L., Gong, C., Ma, Z., Zhao, L., Wang, Z., Li, X., Jing, X., Yang, H., Liu, C.: Real-time human-in-the-loop remote control for a life-size traffic police robot with multiple augmented reality aided display terminals. In: 2017 2nd International Conference on Advanced Robotics and Mechatronics (ICARM). (Aug 2017) 420–425
56. Hedayati, H., Walker, M., Szafir, D.: Improving collocated robot teleoperation with augmented reality. In: Proceedings of the 2018 ACM/IEEE International Conference on Human-Robot Interaction, ACM (2018) 78–86

doi: 10.1007/978-3-319-91581-4_19

57. Miner, N.E., Stansfield, S.A.: An interactive virtual reality simulation system for robot control and operator training. In: Proceedings of the 1994 IEEE International Conference on Robotics and Automation. (May 1994) 1428–1435 vol.2
58. Rakita, D., Mutlu, B., Gleicher, M.: An autonomous dynamic camera method for effective remote teleoperation. In: Proceedings of the 2018 ACM/IEEE International Conference on Human-Robot Interaction, ACM (2018) 325–333
59. Quintero, C.P., Ramirez, O.A., Jägersand, M.: Vibi: Assistive vision-based interface for robot manipulation. In: ICRA. (2015) 4458–4463
60. Hashimoto, S., Ishida, A., Inami, M., Igarashi, T.: Touchme: An augmented reality based remote robot manipulation. In: The 21st International Conference on Artificial Reality and Telexistence, Proceedings of ICAT2011. Volume 2. (2011)
61. Pereira, A., Carter, E.J., Leite, I., Mars, J., Lehman, J.F.: Augmented reality dialog interface for multimodal teleoperation. In: 26th IEEE International Symposium on Robot and Human Interactive Communication (RO-MAN), 2017. (2017)
62. Haring, K.S., Finomore, V., Muramato, D., Tenhundfeld, N.L., Redd, M., Wen, J., Tidball, B.: Analysis of using virtual reality (vr) for command and control applications of multi-robot systems. In: Proceedings of the 1st International Workshop on Virtual, Augmented, and Mixed Reality for HRI (VAM-HRI). (2018)
63. Lipton, J.I., Fay, A.J., Rus, D.: Baxter's homunculus: Virtual reality spaces for teleoperation in manufacturing. IEEE Robotics and Automation Letters **3**(1) (2018) 179–186
64. Oh, Y., Parasuraman, R., McGraw, T., Min, B.C.: 360 vr based robot teleoperation interface for virtual tour. In: Proceedings of the 1st International Workshop on Virtual, Augmented, and Mixed Reality for HRI (VAM-HRI). (2018)
65. Rosen, E., Whitney, D., Phillips, E., Ullman, D., Tellex, S.: Testing robot teleoperation using a virtual reality interface with ros reality. In: Proceedings of the 1st International Workshop on Virtual, Augmented, and Mixed Reality for HRI (VAM-HRI). (2018)
66. Gaurav, S., Al-Qurashi, Z., Barapatre, A., Ziebart, B.: Enabling effective robotic teleoperation using virtual reality and correspondence learning via neural network. In: Proceedings of the 1st International Workshop on Virtual, Augmented, and Mixed Reality for HRI (VAM-HRI). (2018)
67. Whitney, D., Rosen, E., Phillips, E., Konidaris, G., Tellex, S.: Comparing robot grasping teleoperation across desktop and virtual reality with ros reality. In: Proceedings of the International Symposium on Robotics Research. (2017)
68. Allspaw, J., Roche, J., Lemiesz, N., Yannuzzi, M., Yanco, H.A.: Remotely teleoperating a humanoid robot to perform fine motor tasks with virtual reality–. In: Proceedings of the 2018 Conference on Waste Management. (2018)
69. Allspaw, J., Roche, J., Norton, A., Yanco, H.A.: Remotely teleoperating a humanoid robot to perform fine motor tasks with virtual reality–. In: Proceedings of the 1st International Workshop on Virtual, Augmented, and Mixed Reality for HRI (VAM-HRI). (2018)
70. Bennett, M., Williams, T., Thames, D., , Scheutz, M.: Differences in interaction patterns and perception for teleoperated and autonomous humanoid robots. In: Proceedings of the IEEE/RSJ International Conference on Intelligent Robots and Systems. (2017)
71. Pititeeraphab, Y., Choitkunnan, P., Thongpance, N., Kullathum, K., Pintavirooj, C.: Robot-arm control system using leap motion controller. In: 2016 International Conference on Biomedical Engineering (BME-HUST). (Oct 2016) 109–112

doi: 10.1007/978-3-319-91581-4_19

72. Bassily, D., Georgoulas, C., Guettler, J., Linner, T., Bock, T.: Intuitive and adaptive robotic arm manipulation using the leap motion controller. In: ISR/Robotik 2014; 41st International Symposium on Robotics. (June 2014) 1–7
73. Lin, Y., Song, S., Meng, M.Q.H.: The implementation of augmented reality in a robotic teleoperation system. In: Real-time Computing and Robotics (RCAR), IEEE International Conference on, IEEE (2016) 134–139
74. Weichert, F., Bachmann, D., Rudak, B., Fisseler, D.: Analysis of the accuracy and robustness of the leap motion controller. Sensors **13**(5) (2013) 6380–6393
75. Quigley, M., Faust, J., Foote, T., Leibs, J.: Ros: an open-source robot operating system. In: ICRA workshop on open source software. (2009)
76. Koenig, N., Howard, A.: Design and use paradigms for gazebo, an open-source multi-robot simulator. In: Proceedings of the IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS), IEEE (2004) 2149–2154

doi: 10.1007/978-3-319-91581-4_19