RoGuE : Robot Gesture Engine

Rachel M. Holladay and Siddhartha S. Srinivasa
Robotics Institute, Carnegie Mellon University
Pittsburgh, Pennsylvania 15213

Abstract

We present the Robot Gesture Library (RoGuE), a motion-planning approach to generating gestures. Gestures improve robot communication skills, strengthening robots as partners in a collaborative setting. Previous work maps from environment scenario to gesture selection. This work maps from gesture selection to gesture execution. We create a flexible and common language by parameterizing gestures as task-space constraints on robot trajectories and goals. This allows us to leverage powerful motion planners and to generalize across environments and robot morphologies. We demonstrate RoGuE on four robots: HREB, ADA, CURI and the PR2.

1 Introduction

To create robots that seamlessly collaborate with people, we propose a motion-planning based method for generating gestures. Gesture augment robots' communication skills, thus improving the robot's ability to work jointly with humans in their environment (Breazeal et al. 2005; McNeill 1992). This work maps existing gesture terminology, which is generally qualitative, to precise mathematical definitions as task space constraints.

By augmenting our robots with nonverbal communication methods, specifically gestures, we can improve their communication skills. In human-human collaborations, gestures are frequently used for explanations, teaching and problem solving (Lozano and Tversky 2006; Tang 1991; Reynolds and Reeve 2001; Garber and Goldin-Meadow 2002). Robots have used gestures to improve their persuasiveness and understandability while also improving the efficiency and perceived workload of their human collaborator (Chidambaram, Chiang, and Mutlu 2012; Lohse et al. 2014).

While gestures improve a robot's skills as a partner, their use also positively impacts people's perception of the robot. Robots that use gestures are viewed as more active, likable and competent (Salem et al. 2013; 2011). Gesturing robots are more engaging in game play, storytelling and over long-term interactions (Carter et al. 2014; Huang and Mutlu 2014; Kim et al. 2013).

As detailed in Sec. 2.1 there is no standardized classification for gestures. Each taxonomy describes gestures in a
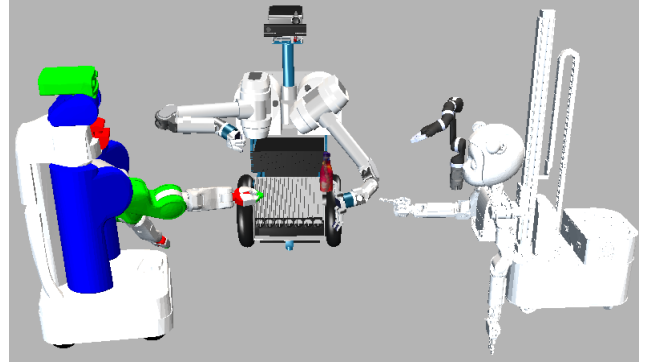
Figure 1: Robots HERB, ADA, CURI and PR2 using RoGuE (Robot Gesture Engine) to point at the fuze bottle.

qualitative manner, as a method for expressing intent and emotion. Previous robot gestures systems, further detailed in Sec. 2.2, often concentrate on transforming ideas, such as inputted text, into speech and gestures.

These gesture systems, some of which use canned gestures, are difficult to use in cluttered environment with object-centric motions. It is critical to consider the environment when executing a gesture. Even given a gesture-object pair, the gesture execution might vary due to occlusions and obstructions in the environment. Additionally, many previous gesture systems are tuned to a specific platform, limiting their generalization across morphologies.

Our key insight is that many gestures can be translated into task-space constraints on the trajectory and the goal, which serves as a common language for gesture expression. This is intuitive to express, and consumable by our powerful motion planning algorithms that plan in clutter. This also enables generalization across robot morphologies, since the robot kinematics are handled by the robot's planner, not the gesture system.

Our key contribution is a planning engine, RoGuE (Robot Gesture Engine), that allows a user to easily specify task-space constraints for gestures. Specifically we formalize gestures as instances of Task Space Regions (TSR), a general constraint representation framework (Berenson, Srinivasa, and Kuffner 2011). This engine has already been deployed to several robot morphologies, specifically HERB, ADA, CURI and the PR2, seen in Fig.1. The source code for

RoGuE is publicly available and detailed in Sec. 6.

RoGuE is presented as a set of gesture primitives that parametrize the mapping from gestures to motion planning. We do not explicitly assume a higher-level system that autonomously call these gesture primitives.

We begin by presenting an extensive view of related work followed by the motivation and implementation for each gesture. We conclude with a description of each robot platform using RoGuE and a brief discussion.

2 Related Work

In creating a taxonomy of collaborative gestures we begin by examining existing classifications and previous gesture systems.

2.1 Gestures Classification

Despite extensive work in gestures, there is no existing common taxonomy or even standardized set of definitions (Wexelblat 1998). Kendon presents a historical view of terminology, demonstrating the lack of agreement, in addition to adding his own classification (Kendon 2004).

Karam examines forty years of gesture usage in human-computer interaction research and established five major categories of gestures (Karam and others 2005). Other methodological gestures classifications name a similar set of five categories (Nehaniv et al. 2005).

Our primary focus is gestures that assist in collaboration and there has been prior work with a similar focus. Clark categorizes coordination gestures into two groups: directing-to and placing-for (Clark 2005). Sauppe presents a series of gestures focused on deictic gestures that a robot would use to communicate information to a human partner (Sauppé and Mutlu 2014). Sauppe's taxonomy includes: pointing, presenting, touching, exhibiting, sweeping and grouping. We implemented four of these, omitting touching and grouping.

2.2 Gesture Systems

Several gesture systems, unlike our own, integrate body and facial features with a verbal system, generating the voice and gesture automatically based on a textual input (Tojo et al. 2000; Kim et al. 2007; Okuno et al. 2009; Salem et al. 2009). BEAT, Behavior Expression Animation Toolkit, is an early system that allows animators to input text and outputs synchronized nonverbal behaviors and synthesized speech (Cassell, Vilhjálmsson, and Bickmore 2004).

Some approach gesture generation as a learning problem, either learning via direct imitation or through Gaussian Mixture Models (Hattori et al. 2005; Calinon and Billard 2007). These method focus on being data-driven or knowledge-based as they learn from large quantities of human examples (Kipp et al. 2007; Kopp and Wachsmuth 2000).

Alternatively, gestures are framed as constrained inverse kinematics problem, either concentrating on smooth joint acceleration or on synchronizing head and eye movement (Bremner et al. 2009; Marjanovic, Scassellati, and Williamson 1996).

Other gesture system focus on collaboration and embodiment or attention direction(Fang, Doering, and Chai 2015; Sugiyama et al. 2005). Regardless of the system,

its clear that gestures are an important part of an engaging robot's skill set (Severinson-Eklundh, Green, and Hüttenrauch 2003; Sidner, Lee, and Lesh 2003).

3 Library of Collaborative Gestures

As mentioned in Sec. 2.1, in creating our list we were inspired by Sauppe's work (Sauppé and Mutlu 2014). Zhang studies a teaching scenario and found that only five gestures made up 80% of the gestures used and of this, pointing dominated use (Zhang et al. 2010). Motivated by both these works we focus on four key features: pointing, presenting, exhibiting and sweeping. We further define and motivate the choice of each of these gestures.

3.1 Pointing and Presenting

Across language and culture we use pointing to refer to objects on a daily basis (Kita 2003). Simple deictic gestures ground spatial references more simply than complex referential description (Kirk, Rodden, and Fraser 2007). Robots can, as effectively as human agents, use pointing as a referential cue to direct human attention (Li et al. 2015).

Previous pointing systems focus on understandability, either by simulating cognition regions or optimizing for a legibility (Hato et al. 2010; Holladay, Dragan, and Srinivasa 2014). Pointing in collaborative virtual environments concentrates on improving pointing quality by verifying that the information was understood correctly (Wong and Gutwin 2010). Pointing is also viewed as a social gesture that should balance social appropriateness and understandability (Liu et al. 2013).

The natural human end effector shape when pointing is to close all but the index finger (Gokturk and Sibert 1999; Cipolla and Hollinghurst 1996) which serves as the pointer. Kendon formally describes this position as 'Index Finger Extended Prone'. We adhere to this style as closely as the robot's morphology will allow. Presenting achieves a similar goal, but is done with, as Kendon describes, an 'Open Hand Neutral' and 'Open Hand Supine' hand position.

3.2 Exhibiting

While pointing and presenting can refer to an object or spatial region, exhibiting is a gesture used to show off an object. Exhibiting involves holding an object and bringing emphasis to it by lifting it into view. Specifically, exhibiting deliberating displays the object to an audience (Clark 2005). Exhibiting and pointing are often used in conjunction in teaching scenarios (Lozano and Tversky 2006).

3.3 Sweeping

A sweeping gesture involves making a long, continuous curve over the objects or spatial area being referred to. Sweeping is a common technique used by teachers, where they sweep across various areas to communicate abstract ideas or regions (Alibali, Flevares, and Goldin-Meadow 1997).

4 Gesture Formulation as TSRs

Each of the main four gestures are implemented with a base of a task space region, described below. The implementa-

tion for each, as described, is nearly identical for each of the robots listed in Sec. 5, with minimal changes made necessary by each robot's morphology.

The gestures pointing, presenting and sweeping can reference either a point in space or an object while exhibiting can only be applied to an object, since something must be physically grasped. For ease, we will define R as the spatial region or object being referred to. Within our RoGuE framework, R can be an object pose, a 4x4 transform or a 3-D coordinate in space.

4.1 Task Space Region

A task space region (TSR) is a constraint representation proposed in (Berenson, Srinivasa, and Kuffner 2011) and a summary is given here. We consider the pose of a robot's end effector as a point in three-dimensional space, SE(3). Therefore we can describe end effector constraint sets as subsets of our space, SE(3).

A TSR is composed of three components: T_0^w, T_e^w, B_w. T_0^w describes the reference transform for the TSR, given the world coordinates in the TSR's frame. T_e^w is the offset transform that gives the end effector's pose in the world coordinates. Finally, B_w is a 6x2 matrix that bounds the coordinates of the world. The first three rows declare the allowable translations in our constrained world in terms of x, y, z. The last three rows give the allowable rotation, in roll pitch yaw notion. Hence our B_w, which describes the end effector's constraints, is of the form:

$$\begin{bmatrix} x_{\min} & y_{\min} & z_{\min} & \psi_{\min} & \theta_{\min} & \phi_{\min} \\ x_{\max} & y_{\max} & z_{\max} & \psi_{\max} & \theta_{\max} & \phi_{\max} \end{bmatrix}^{\mathsf{T}} \quad (1)$$

For more detailed description of TSRs please refer to (Berenson, Srinivasa, and Kuffner 2011). Each of the main four gestures is composed of one or more TSRs and some wrapping motions. In describing each gesture below, we will first describe the necessary TSR(s) followed by the actions taken.

4.2 Pointing

Our pointing TSR is composed of two TSRs chained together. The details of TSR chains can be found in (Berenson et al. 2009). The first TSR uses a T_e^w that aligns the z axis of the robot's end effector with R. T_e^w varies from robot to robot. We next want to construct our B_w. With respect to rotation, R can be pointed at from any angle. This creates a sphere of pointing configuration where each configuration is a point on the surface of the sphere with the robot's z-axis, or the axis coming out of the palm, aligned with R. We constrain the pointing to not come from below the object, thus yielding a hemisphere, centered at the object, of pointing configurations. The corresponding B_w is therefore:

$$\begin{bmatrix} 0 & 0 & 0 & -\pi & 0 & -\pi \\ 0 & 0 & 0 & \pi & \pi & -\pi \end{bmatrix}^{\mathsf{T}} \quad (2)$$

While this creates a hemisphere, a pointing configuration is not constrained to be some minimum or maximum distance from the object. This translates into an allowable range in our z-axis, the axis coming out of the end effector's palm,
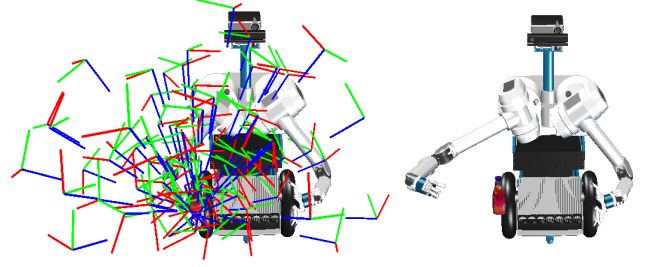

Figure 2: Pointing. The leftside shows many of the infinite solutions to pointing. The rightside shows the execution of one.

(a) No other objects occlude the bottle, thus this point is occlusion free.

(b) Other objects occlude the bottle such that only 27% of the bottle is visible from this angle.

Figure 3: Demonstrating how the offscreen render tool can be used to measure occlusion. For both scenes the robot is pointing at the bottle shown in blue.

in the hand frame. We achieve this with a second TSR that is chained to the intial rotation TSR.

We have an allowable range in our z-axis. Therefore, our pointing region as defined by the TSR is a hemisphere centered on the object of varying radius. This visualized TSR can be seen in Fig.2.

In order to execute a point we plan to a solution of the TSR and change our robot's end effector's preshape to what is closest to the Index Finger Point. This also varies from robot to robot, depending on the robot's end effector.

Accounting for Occlusion

As described in (Holladay, Dragan, and Srinivasa 2014) we want to ensure that our pointing configuration not only refers to the desired R but also does not refer to any other candidate R. Stating this another way, we do not want R to be occluded by other candidate items. The intuitive idea is that you cannot point at a bottle if there is a cup in the way, since it would look as though you were pointing at the cup instead. We use a tool, offscreen render [1], to measure occlusion.

A demonstration of its use can be seen in Fig.3. In this

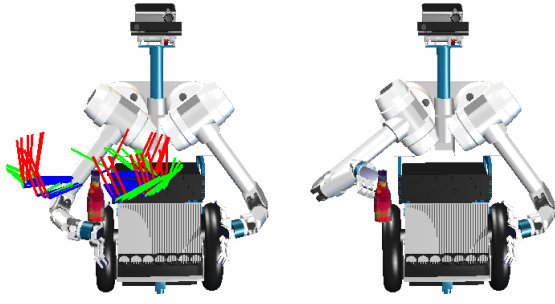[1]`https://github.com/personalrobotics/offscreen_render`

Figure 4: Presenting. The leftside shows a visualization of the TSR solutions. In the rightside, the robot has picked one solution and executed it.
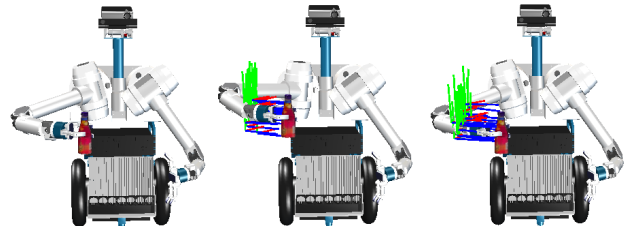

Figure 5: Exhibiting. The robot grasps the object, lifts it up for display and places it back down.
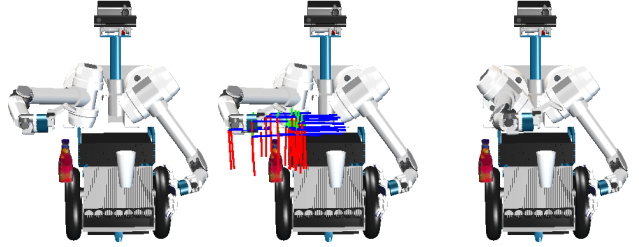

Figure 6: Sweeping. The robot starts on one end of the reference, moving its downward facing falm to the other end of the area.

case there is a simple scene with a glass, plate and a bottle that the robot wants to point at. There are many different ways to achieve this point and here we show two perspectives from two different end effector locations that point at the bottle, Fig.3a and Fig.3b.

In the first column we see the robot's visualization of the scene. The target object, the bottle, is in blue, while the other clutter objects, the plate and glass, are in green. Next, seen in second column, we remove the clutter, effectively measuring how much the clutter blocks the bottle. The third column shows how much of the bottle we would see if the clutter did not exist in the scene. We can then take a ratio, in terms of pixels, of the second column with respect to the third column to see what percentage of the bottle is viewable given the clutter.

For the pointing configuration shown in Fig.3a, the bottle is entirely viewable, which represents a good pointing configuration. In the pointing configuration shown in Fig.3b, the clutter occludes the bottle significantly, such that only 27% is viewable. Thus we would prefer the pointing configuration of Fig.3a over that of Fig.3b since more of the bottle is visible.

The published RoGuE implementation, as detailed in Sec. 6 does not include occlusion checking, however, we have plans to incorporate it in the coming future, with more details in Sec. 7.1.

4.3 Presenting

Presenting is a much more constrained gesture than pointing. While we can refer to R from any angle, we constrain our gesture to be level with the object and within a fixed radius. Thus our B_w is

$$\begin{bmatrix} 0 & 0 & 0 & 0 & 0 & -\pi \\ 0 & 0 & 0 & 0 & 0 & \pi \end{bmatrix}^{\mathsf{T}} \quad (3)$$

This visualized TSR can be seen in Fig.4. Once again we execute presenting by planning to a solution of the TSR and changing the preshape to the Open Hand Neutral pose.

4.4 Exhibiting

Exhibiting involves grasping the reference object, lifting it to be displayed and placing the object back down. Therefore, first the object is grasped using any grasp primitive, in our case a grasp TSR. We then use the lift TSR, which constrains the motion to have epsilon rotation in any direction while allowing it to translate upwards to the desired height. The system then pauses for the desired wait time before performing the lift action in reverse, therefore placing the object back at its original location. This process can be seen in Fig.5. Optionally, the system can un-grasp the object and return its end effector the pose it was in before the exhibiting gesture occurred.

4.5 Sweeping

Sweeping is motioning from one area across to another area in a smooth curve. We first plan our end effector to hover over the starting position of the curve. We then modify our hand to a preshape that faces the end effector's palm downwards.

Next we want to translate to the end of the curve smoothly along a plane. We execute this via two TSRs that are chained together. We define one TSR above the end position, allowing for epsilon movement in the x, y, z and pitch. This constrains our curve to end at the end of our bounding curve. Our second TSR constrains our movement, allowing us to translate in the x and y plane to our end location and giving some leeway by allowing epsilon movement in all other directions. By executing these TSRs we smoothly curve from the starting position to the ending position, thus sweeping out our region. This is displayed in Fig.6.
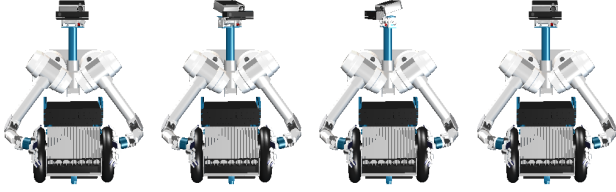
4.6 Additional Gestures

In addition to the four key gestures described above, on some of the robot platforms described in Sec. 5 we implement nodding and waving.
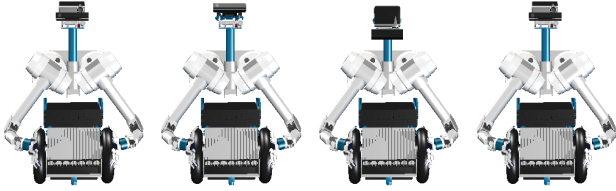
Wave: We created a pre-programmed wave. We first record a wave trajectory and then execute it by playing this

Figure 7: Wave

(a) Noding No by shaking from side to side

(b) Noding Yes, by shaking up and down

trajectory back. While this is not customizable, we have find people react very positively to this simple motion.

Nod: Nodding is useful in collaborative settings (Lee et al. 2004). We provide the ability to nod yes, by servoing the head up and down, and to nod no, by servoing the head side to side as seen in Fig.8a and Fig.8b.

5 Systems Deployed

RoGuE was deployed onto four robots, listed below. The details of each robot are given. At the time of this publication, the authors only have physical access to HERB and ADA, thus the gestures for CURI and the PR2 were tested only in simulation. All of the robot can be seen in Fig.1, pointing at a fuze bottle.

HERB HERB (Home Exploring Robot Bulter) is a bi-manual mobile manipulator with two Barrett WAMs, each with 7-degrees of freedom (Srinivasa et al. 2012). HERB's Barrett hands have three fingers, one of which is a thumb, allowing for the canonical index out-stretched pointing configuration. While HERB does not specifically look like a human, the arm and head configuration is human-like.

ADA ADA (Assistive Dexterous Arm) is a Kinova MICO[2] 6-degree of freedom commercial wheelchair mounted or workstation-mounted arm with an actuated gripper. ADA's gripper has only two fingers so the pointing configuration consists of the two fingers closed together. Since ADA is a free standing arm there is little anthropomorphism.

CURI Curi is a humanoid mobile robot with two 7-degree of freedom arms, an omnidirectional mobile base, and a socially expressive head. Curi has human-like hands and therefore can also point through the outstretched index finger con-

figuration. Curi's head and hands are particularly human-like.

PR2 The PR2 is a bi-manual mobile robot with two 7-degree of freedom arms and gripper hands (Bohren et al. 2011). Like ADA, the PR2 has only two fingers and therefore points with the fingers closed. The PR2 is similar to HERB in that it is not a humanoid but has a human-like configuration.

6 Source Code and Reproducibility

RoGuE is composed largely of two components, the TSR Library and Action Library. The TSR library holds all the TSRs that are then used by the Action Library, which executes the gestures. Between the robots the files are nearly identical, with minor morphology differences, such as different wrist to palm distance offsets.

For example, below is a comparison of the simplified source code for the present gesture on HERB and ADA. The focus is the location of the object being presented. The only difference is that HERB and ADA have different hand pre-shapes.

```
1  def Present_HERB(robot, focus, arm):
2      present_tsr = robot.tsrlibrary('present', focus, arm)
3      robot.PlanToTSR(present_tsr)
4      preshape = {finger1=1, finger2=1,
5                  finger3=1, spread=3.14}
6      robot.arm.hand.MoveHand(preshape)
```

```
1  def Present_ADA(robot, focus, arm):
2      present_tsr = robot.tsrlibrary('present', focus, arm)
3      robot.PlanToTSR(present_tsr)
4      preshape = {finger1=0.9, finger2=0.9}
5      robot.arm.hand.MoveHand(preshape)
```

Referenced are the TSR[2] and Action Library[3] for HERB.

7 Discussion

We presented RoGuE, the Robot Gesture Engine, for executing several collaborative gestures across multiple robot platforms. Gestures are an invaluable communicate tool for our robots as they become engaging and communicative partners.

While existing systems map from situation to gesture selection we focused on the specifics of executing those gestures on robotic systems.

We formalized a set of collaborative gestures and parametrized them as task space constraints on the trajectory and goal location. This insight allowed us to use powerful existing motion planners that generalize across environments, as seen in Fig.9, and robots, as seen in Fig.1.

---

[2]https://github.com/personalrobotics/
herbpy/blob/master/src/herbpy/tsr/generic.py
[3]https://github.com/personalrobotics/
herbpy/blob/master/src/herbpy/action/rogue.
py

Figure 9: Four pointing scenarios with varying clutter. In each case HERB the robot is pointing at the fuze bottle.

7.1 Future Work

The gestures and their implementations serve as primitives that can be used as stand-alone features or as components to a more complicated robot communication system. We have not incorporated these gestures in conjunction with head motion or speech. Taken all together speech, gestures, head movement and gaze should be synchronized. Hence the gestures system would have to account for timing.

In the current implementation of RoGuE, the motion planner randomly samples from the TSR. Instead of picking a random configuration, we want to bias our planner to picking pointing configurations that are more clear and natural. As discussed towards the end of Sec. 4.2, we can score our pointing configuration based on occlusion, therefore picking a configuration that minimizes occlusion.

RoGuE serves as a starting point for formalizing gestures as motion. Continuing this line of work we can step closer towards effective, communicative robot partners.

Acknowledgments

References

Alibali, M. W.; Flevares, L. M.; and Goldin-Meadow, S. 1997. Assessing knowledge conveyed in gesture: Do teachers have the upper hand? *Journal of Educational Psychology* 89(1):183.

Berenson, D.; Chestnutt, J.; Srinivasa, S. S.; Kuffner, J. J.; and Kagami, S. 2009. Pose-constrained whole-body planning using task space region chains. In *Humanoid Robots, 2009. Humanoids 2009. 9th IEEE-RAS International Conference on*, 181–187. IEEE.

Berenson, D.; Srinivasa, S. S.; and Kuffner, J. 2011. Task space regions: A framework for pose-constrained manipulation planning. *The International Journal of Robotics Research* 0278364910396389.

Bohren, J.; Rusu, R. B.; Jones, E. G.; Marder-Eppstein, E.; Pantofaru, C.; Wise, M.; Mösenlechner, L.; Meeussen, W.; and Holzer, S. 2011. Towards autonomous robotic butlers: Lessons learned with the pr2. In *Robotics and Automation (ICRA), 2011 IEEE International Conference on*, 5568–5575. IEEE.

Breazeal, C.; Kidd, C. D.; Thomaz, A. L.; Hoffman, G.; and Berlin, M. 2005. Effects of nonverbal communication on efficiency and robustness in human-robot teamwork. In *Intelligent Robots and Systems, 2005.(IROS 2005). 2005 IEEE/RSJ International Conference on*, 708–713. IEEE.

Bremner, P.; Pipe, A.; Melhuish, C.; Fraser, M.; and Subramanian, S. 2009. Conversational gestures in human-robot interaction. In *Systems, Man and Cybernetics, 2009. SMC 2009. IEEE International Conference on*, 1645–1649. IEEE.

Calinon, S., and Billard, A. 2007. Incremental learning of gestures by imitation in a humanoid robot. In *Proceedings of the ACM/IEEE international conference on Human-robot interaction*, 255–262. ACM.

Carter, E. J.; Mistry, M. N.; Carr, G. P. K.; Kelly, B.; Hodgins, J. K.; et al. 2014. Playing catch with robots: Incorporating social gestures into physical interactions. In *Robot and Human Interactive Communication, 2014 RO-MAN: The 23rd IEEE International Symposium on*, 231–236. IEEE.

Cassell, J.; Vilhjálmsson, H. H.; and Bickmore, T. 2004. Beat: the behavior expression animation toolkit. In *Life-Like Characters*. Springer. 163–185.

Chidambaram, V.; Chiang, Y.-H.; and Mutlu, B. 2012. Designing persuasive robots: how robots might persuade people using vocal and nonverbal cues. In *Proceedings of the seventh annual ACM/IEEE international conference on Human-Robot Interaction*, 293–300. ACM.

Cipolla, R., and Hollinghurst, N. J. 1996. Human-robot interface by pointing with uncalibrated stereo vision. *Image and Vision Computing* 14(3):171–178.

Clark, H. H. 2005. Coordinating with each other in a material world. *Discourse studies* 7(4-5):507–525.

Fang, R.; Doering, M.; and Chai, J. Y. 2015. Embodied collaborative referring expression generation in situated human-robot interaction. In *Proceedings of the Tenth Annual ACM/IEEE International Conference on Human-Robot Interaction*, 271–278. ACM.

Garber, P., and Goldin-Meadow, S. 2002. Gesture offers insight into problem-solving in adults and children. *Cognitive Science* 26(6):817–831.

Gokturk, M., and Sibert, J. L. 1999. An analysis of the index finger as a pointing device. In *CHI'99 Extended Abstracts on Human Factors in Computing Systems*, 286–287. ACM.

Hato, Y.; Satake, S.; Kanda, T.; Imai, M.; and Hagita, N. 2010. Pointing to space: modeling of deictic interaction referring to regions. In *Proceedings of the 5th ACM/IEEE*

international conference on Human-robot interaction, 301–308. IEEE Press.

Hattori, Y.; Kozima, H.; Komatani, K.; Ogata, T.; and Okuno, H. G. 2005. Robot gesture generation from environmental sounds using inter-modality mapping.

Holladay, R. M.; Dragan, A. D.; and Srinivasa, S. S. 2014. Legible robot pointing. In *Robot and Human Interactive Communication, 2014 RO-MAN: The 23rd IEEE International Symposium on*, 217–223. IEEE.

Huang, C.-M., and Mutlu, B. 2014. Multivariate evaluation of interactive robot systems. *Autonomous Robots* 37(4):335–349.

Karam, et al. 2005. A taxonomy of gestures in human computer interactions.

Kendon, A. 2004. *Gesture: Visible action as utterance*. Cambridge University Press.

Kim, H.-H.; Lee, H.-E.; Kim, Y.-H.; Park, K.-H.; and Bien, Z. Z. 2007. Automatic generation of conversational robot gestures for human-friendly steward robot. In *Robot and Human interactive Communication, 2007. RO-MAN 2007. The 16th IEEE International Symposium on*, 1155–1160. IEEE.

Kim, A.; Jung, Y.; Lee, K.; and Han, J. 2013. The effects of familiarity and robot gesture on user acceptance of information. In *Proceedings of the 8th ACM/IEEE international conference on Human-robot interaction*, 159–160. IEEE Press.

Kipp, M.; Neff, M.; Kipp, K. H.; and Albrecht, I. 2007. Towards natural gesture synthesis: Evaluating gesture units in a data-driven approach to gesture synthesis. In *Intelligent Virtual Agents*, 15–28. Springer.

Kirk, D.; Rodden, T.; and Fraser, D. S. 2007. Turn it this way: grounding collaborative action with remote gestures. In *Proceedings of the SIGCHI conference on Human Factors in Computing Systems*, 1039–1048. ACM.

Kita, S. 2003. *Pointing: Where language, culture, and cognition meet*. Psychology Press.

Kopp, S., and Wachsmuth, I. 2000. A knowledge-based approach for lifelike gesture animation. In *ECAI 2000-Proceedings of the 14th European Conference on Artificial Intelligence*.

Lee, C.; Lesh, N.; Sidner, C. L.; Morency, L.-P.; Kapoor, A.; and Darrell, T. 2004. Nodding in conversations with a robot. In *CHI'04 Extended Abstracts on Human Factors in Computing Systems*, 785–786. ACM.

Li, A. X.; Florendo, M.; Miller, L. E.; Ishiguro, H.; and Saygin, A. P. 2015. Robot form and motion influences social attention. In *Proceedings of the Tenth Annual ACM/IEEE International Conference on Human-Robot Interaction*, 43–50. ACM.

Liu, P.; Glas, D. F.; Kanda, T.; Ishiguro, H.; and Hagita, N. 2013. It's not polite to point: generating socially-appropriate deictic behaviors towards people. In *Proceedings of the 8th ACM/IEEE international conference on Human-robot interaction*, 267–274. IEEE Press.

Lohse, M.; Rothuis, R.; Gallego-Pérez, J.; Karreman, D. E.; and Evers, V. 2014. Robot gestures make difficult tasks easier: the impact of gestures on perceived workload and task performance. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, 1459–1466. ACM.

Lozano, S. C., and Tversky, B. 2006. Communicative gestures facilitate problem solving for both communicators and recipients. *Journal of Memory and Language* 55(1):47–63.

Marjanovic, M. J.; Scassellati, B.; and Williamson, M. M. 1996. *Self-taught visually-guided pointing for a humanoid robot*. From Animals to Animats: Proceedings of.

McNeill, D. 1992. *Hand and mind: What gestures reveal about thought*. University of Chicago Press.

Nehaniv, C. L.; Dautenhahn, K.; Kubacki, J.; Haegele, M.; Parlitz, C.; and Alami, R. 2005. A methodological approach relating the classification of gesture to identification of human intent in the context of human-robot interaction. In *Robot and Human Interactive Communication, 2005. ROMAN 2005. IEEE International Workshop on*, 371–377. IEEE.

Okuno, Y.; Kanda, T.; Imai, M.; Ishiguro, H.; and Hagita, N. 2009. Providing route directions: design of robot's utterance, gesture, and timing. In *Human-Robot Interaction (HRI), 2009 4th ACM/IEEE International Conference on*, 53–60. IEEE.

Reynolds, F. J., and Reeve, R. A. 2001. Gesture in collaborative mathematics problem-solving. *The Journal of Mathematical Behavior* 20(4):447–460.

Salem, M.; Kopp, S.; Wachsmuth, I.; and Joublin, F. 2009. Towards meaningful robot gesture. In *Human Centered Robot Systems*. Springer. 173–182.

Salem, M.; Rohlfing, K.; Kopp, S.; and Joublin, F. 2011. A friendly gesture: Investigating the effect of multimodal robot behavior in human-robot interaction. In *RO-MAN, 2011 IEEE*, 247–252. IEEE.

Salem, M.; Eyssel, F.; Rohlfing, K.; Kopp, S.; and Joublin, F. 2013. To err is human (-like): Effects of robot gesture on perceived anthropomorphism and likability. *International Journal of Social Robotics* 5(3):313–323.

Sauppé, A., and Mutlu, B. 2014. Robot deictics: How gesture and context shape referential communication.

Severinson-Eklundh, K.; Green, A.; and Hüttenrauch, H. 2003. Social and collaborative aspects of interaction with a service robot. *Robotics and Autonomous systems* 42(3):223–234.

Sidner, C. L.; Lee, C.; and Lesh, N. 2003. Engagement rules for human-robot collaborative interactions. In *IEEE International Conference On Systems Man And Cybernetics*, volume 4, 3957–3962.

Srinivasa, S. S.; Berenson, D.; Cakmak, M.; Collet, A.; Dogar, M. R.; Dragan, A. D.; Knepper, R. A.; Niemueller, T.; Strabala, K.; Vande Weghe, M.; et al. 2012. Herb 2.0: Lessons learned from developing a mobile manipulator for the home. *Proceedings of the IEEE* 100(8):2410–2428.

Sugiyama, O.; Kanda, T.; Imai, M.; Ishiguro, H.; and Hagita, N. 2005. Three-layered draw-attention model for humanoid robots with gestures and verbal cues. In *Intelligent Robots and Systems, 2005.(IROS 2005). 2005 IEEE/RSJ International Conference on*, 2423–2428. IEEE.

Tang, J. C. 1991. Findings from observational studies of collaborative work. *International Journal of Man-machine studies* 34(2):143–160.

Tojo, T.; Matsusaka, Y.; Ishii, T.; and Kobayashi, T. 2000. A conversational robot utilizing facial and body expressions. In *Systems, Man, and Cybernetics, 2000 IEEE International Conference on*, volume 2, 858–863. IEEE.

Wexelblat, A. 1998. Research challenges in gesture: Open issues and unsolved problems. In *Gesture and sign language in human-computer interaction*. Springer. 1–11.

Wong, N., and Gutwin, C. 2010. Where are you pointing?: the accuracy of deictic pointing in cves. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, 1029–1038. ACM.

Zhang, J. R.; Guo, K.; Herwana, C.; and Kender, J. R. 2010. Annotation and taxonomy of gestures in lecture videos. In *Computer Vision and Pattern Recognition Workshops (CVPRW), 2010 IEEE Computer Society Conference on*, 1–8. IEEE.