

Hands-On Robotics: Enabling Communication Through Direct Gesture Control

Max Pascher
max.pascher@udo.edu
TU Dortmund University
Dortmund, Germany

Roman Heger
roman.heger@uni-due.de
University of Duisburg-Essen
Essen, Germany

Alia Saad
alia.saad@uni-due.de
University of Duisburg-Essen
Essen, Germany

Jens Gerken
jens.gerken@udo.edu
TU Dortmund University
Dortmund, Germany

Jonathan Liebers
jonathan.liebers@uni-due.de
University of Duisburg-Essen
Essen, Germany

Stefan Schneegass
stefan.schneegass@uni-due.de
University of Duisburg-Essen
Essen, Germany

Uwe Gruenefeld
uwe.gruenefeld@uni-due.de
University of Duisburg-Essen
Essen, Germany

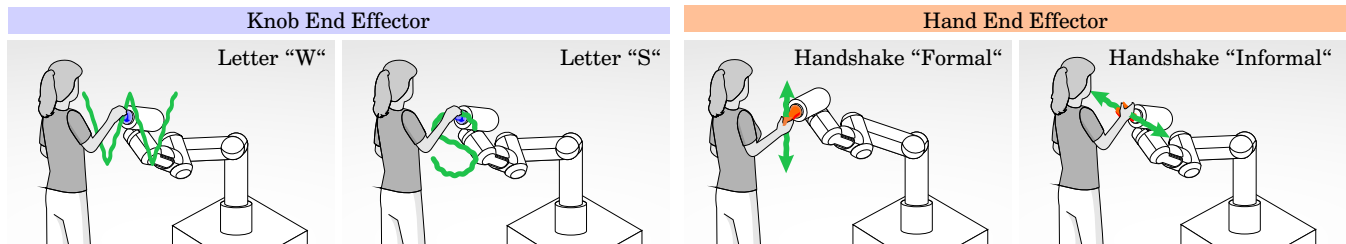


Figure 1: We investigated four different user performed gestures by directly manipulating a robot; two letter-based gestures using a knob as an end effector; and two handshake gestures with a hand as an end effector.

ABSTRACT

Effective Human-Robot Interaction (HRI) is fundamental to seamlessly integrating robotic systems into our daily lives. However, current communication modes require additional technological interfaces, which can be cumbersome and indirect. This paper presents a novel approach, using direct motion-based communication by moving a robot's end effector. Our strategy enables users to communicate with a robot by using four distinct gestures – two handshakes ('formal' and 'informal') and two letters ('W' and 'S'). As a proof-of-concept, we conducted a user study with 16 participants, capturing subjective experience ratings and objective data for training machine learning classifiers. Our findings show that the four different gestures performed by moving the robot's end effector can be distinguished with close to 100% accuracy. Our research offers implications for the design of future HRI interfaces, suggesting that motion-based interaction can empower human operators to communicate directly with robots, removing the necessity for additional hardware.

CCS CONCEPTS

• Human-centered computing → Gestural input; • Computer systems organization → Robotics.

KEYWORDS

embodied interaction, gesture recognition, human-robot interaction, motion-based communication

ACM Reference Format:

Max Pascher, Alia Saad, Jonathan Liebers, Roman Heger, Jens Gerken, Stefan Schneegass, and Uwe Gruenefeld. 2024. Hands-On Robotics: Enabling Communication Through Direct Gesture Control. In *Companion of the 2024 ACM/IEEE International Conference on Human-Robot Interaction (HRI '24 Companion)*, March 11–14, 2024, Boulder, CO, USA. ACM, New York, NY, USA, 6 pages. <https://doi.org/10.1145/3610978.3640635>

1 INTRODUCTION

In recent years, rapid advancements in robotic technologies have led to their growing integration in our daily lives, acting as versatile assistants in workplaces and homes [2, 28, 34]. These robotic companions enhance human capabilities and efficiency and, as such, substantially change how we interact with the world [19]. As robotic solutions evolve and diversify, their capacity of autonomous actions increases – with seamless close-contact interactions between humans and robots becoming a reality [15].



This work is licensed under a Creative Commons Attribution International 4.0 License.

HRI '24 Companion, March 11–14, 2024, Boulder, CO, USA
© 2024 Copyright held by the owner/author(s).
ACM ISBN 979-8-4007-0323-2/24/03.
<https://doi.org/10.1145/3610978.3640635>

However, for successful Human-Robot Collaboration (HRC), effective communication channels must be established for accurate transmission of intent and coordination of respective actions [35]. Traditional modes of HRI, such as voice commands [6] or touch interfaces [3] can effectively convey instructions to robots. Yet, as robots take on more complex tasks and work in proximity to humans – sometimes working *hand in hand* – the need for a more natural and intuitive communication approach becomes apparent. *Motion-based communication*, where humans actively manipulate the robot’s end effector, can be a viable solution to this challenge. By mimicking how we naturally interact with one another, *motion-based communication* bridges the gap between humans and robots, promoting a more intuitive and seamless exchange of information and – consequently – improved collaboration.

Therefore, tracking and measuring movements is a crucial aspect to consider. Two types of tracking approaches – *relative* and *absolute* – exist. Inertial sensors (e.g., accelerometers [7]) provide relative tracking information, while optical tracking methods (e.g., cameras [1]) offer more exact positional data. Previous studies, like those using a *Nintendo Wii* controller [37], have successfully demonstrated gesture recognition using inertial sensors with limited training samples. In contrast, our approach focuses on mechanical tracking – positioned between *relative* and *absolute* tracking – using a seven Degrees-of-Freedom (DoFs) robotic arm for precise movement data. Our research uses the robot’s joint motion data to enable accurate and efficient *motion-based communication*.

Gesture recognition has long been recognized as a challenging yet crucial aspect of enabling natural and intuitive communication between humans and machines. Recognizing gestures using pattern recognizers from the δ -family has been one of the earlier approaches in the field [4, 5, 40–42, 45], leading to numerous follow-ons by other researchers [17, 24, 26, 27]. These recognizers utilize a predefined set of gesture templates to match and identify user motions. Template-based solutions showed robust performance both in 2D and 3D spaces [13].

Machine Learning (ML) techniques are potent tools for gesture recognition [33], allowing systems to learn and adapt to a wide range of gestures. Although it may seem intuitive that existing ML approaches will work for recognizing gestures directly performed with a robot’s end effector, so far, this remains unproven. When designing gestures, *affordance* plays a significant role in determining the gesture vocabulary. An artificial hand on a robot may prompt gestures like a handshake, while the affordance of a knob encourages other gestures.

This work uses the robot’s sensor data to study recognizing and distinguishing gestures performed by direct interaction with a robotic arm. We performed a laboratory study (N=16) to assess the feasibility of four distinct gestures and determine their respective recognition accuracy, resulting in a f1-score up to 0.99. Here, we demonstrate the feasibility of direct interactions’ usage for robust gesture recognition in HRI.

2 RELATED WORK

Our research integrates insights from collaborative robots in close-contact interactions and embodied user input for gesture recognition and classification.

In HRC, collaborative robots – known as cobots – are increasingly common in various settings, including domestic care [9, 34]. They are categorized based on environment sharing [30, 36, 38] and types of cooperation [8, 21]. Previous research focused on cobots adapting to human movements and behavior, while maintaining appropriate distance [31], avoiding collisions [22], and customizing assistance based on skills [10] and comfort [14]. Supporting this, Drolshagen et al. indicated no adverse effects on collaboration when safety aspects are met [16], while Maurtua et al. questioned study participants that anticipate increased interaction with cobots in the future [29]. Efficiency in HRC can be enhanced with suitable techniques and sensors [11]. The embodiment of cobots positively affects perception and trust [47], while touch-based interactions improve non-verbal communication and reduce human stress responses [44]. Interactive perception combines physical and traditional methods but may be limited by occlusion [25].

Gesture recognition plays a pivotal role in HRI. Utilizing the acceleration sensor of the *Nintendo Wii mote*, Schlömer et al. implemented gesture recognition by employing a hidden Markov model (HMM) for training and recognizing user-selected gestures [37]. Despite the small training set, their evaluation demonstrated an accuracy between 0.85 to 0.95. Wu et al. proposed a similar approach using acceleration-based movement data, achieving an accuracy of almost 0.99 for four gestures using the *FDSVM* method. Cabrera and Wachs used the *Microsoft Kinect* sensor to use skeleton data for one-shot gesture recognition [12], comparing three classifiers with accuracy between 0.81 and 0.86. Using a WiFi-based approach, gesture recognition for multi-user applications was demonstrated by Venkatnarayan et al. [43]. Their system identifies concurrent gestures, quantifies the total gesture number, and creates virtual samples for different combinations using training data from a single user. Achieving an accuracy of over 0.90, it can recognize up to eight gestures performed simultaneously.

By leveraging knowledge and techniques from collaborative robotics and gesture recognition, advanced systems capable of accurate gesture recognition – while adapting and responding to human movements and behavior – become possible. These capabilities result in smoother and more efficient HRC, fostering overall safer and more productive interactions.

3 GESTURE INPUT INTERACTIONS

We investigate the use of a robotic arm for effective embodied gesture recognition, specifically focusing on movements that directly manipulate the robot’s end effector through physical interaction. In domestic care, assistive robotic arms often use multi-finger end effectors comparable to a human hand. For this kind of gripper, we use two types of handshakes as natural gestures. The robot’s flanch also supports adding a grasping object (e.g., a protrusion).

Several gestures have been introduced in prior works [23]. Here, we selected two **letter-based gestures** and two **handshake gestures**, simulating a natural human-like interaction with the robotic arm (see Figure 1). To accommodate the differing movements and grips between the gestures (letters vs. hand), we attached a spherical knob for the letters, enabling a firmer grip and smoother motion in 3D space. A hand model resembling a human hand was used for a more lifelike interactions.

Letter Using the knob end effector, we investigated two types of motions: *circular*, represented by the curved letter “S” fostering a smooth transition in performing the gesture, and *linear*, represented by the sharp-edged letter “W” consisting of several stops for direction changing during the motion.

Handshake This gesture imitates the traditional human-to-human handshake and, as such, was performed with the hand end effector. Potential applications include an introductory interaction to initiate communication with a cobot to start a procedure. We examined both the *formal* handshake, involving grasping the hand followed by an up-and-down movement, and the *informal* one, also known as the *G-lock handshake*.

4 STUDY: GESTURE RECOGNITION USING A ROBOTIC ARM

In this study, we introduce a novel approach in human-to-robot communication through gesture-based mechanical manipulation of a robotic arm. Our method centers on extracting movement values from the robotic arm’s joints and does not depend on additional tracking requirements. We investigate how mechanical manipulation of a robotic arm can achieve accurate gesture recognition.

4.1 Study Design

We conducted a within-subject controlled laboratory study to assess the accuracy of gesture recognition. The independent variable were *gestures*, with two pre-selected letter gestures (denoted *LS* and *LW*) and two handshake gestures (denoted *HS* and *GL*). The *gesture recognition accuracy* serves as the dependent variable.

4.2 Participants and Procedure

We recruited 16 participants (6 females, 10 males), aged between 22 and 35 years ($M=27.75$, $SD=3.96$) via mailing lists and social media. All participants were right-handed and reported no motor functions limitations or injuries. The study received approval from our institution’s ethics committee, and each participant received a 10 Euro remuneration upon task completion.

Study sessions for each participant began with the experimenter explaining the study’s purpose and demonstrating the interaction process. Subsequently, participants completed demographic and consent forms. At the start of each run, the robot arm end effector was automatically positioned to a predefined starting point to maintain consistency across all participants. The study conductor briefed participants on the interaction pattern, allowing them to start and stop freely. The conductor then started the recording of the robot’s movement. The corresponding movement data were recorded while participants were moving the robot arm end effector in the desired gesture. Each gesture was recorded five times consecutively, totaling 20 recordings per participant. Gesture order was counterbalanced using a Latin Square design. Following the completion of all recordings, a post-study questionnaire was administered. On average, participants finished the entire task within 40 minutes.

4.3 Apparatus

For our study, we used a Franka Emika robot placed on a fixed table at a height of 61 cm [18]. Both types of end effectors were 3D printed by the research team; the models were acquired from an open-source library [32]. Manipulating the robotic arm required setting the robot status to the *free guiding* mode for a safe interaction. We affixed a clamp to engage the *free guiding* mode buttons, allowing users to freely execute gestures. This action also deactivated any ongoing autonomous movements by the robot. While performing the gestures, participants did not receive any additional feedback.

4.4 Results

In our analysis, we collected *objective* and *subjective* measures. We applied non-parametric Friedman tests to detect significant main effects between gestures. Post-hoc, we conducted Wilcoxon signed-rank tests (Bonferroni corrected) for pairwise comparisons. The effect sizes of the Wilcoxon tests are reported as r ($r > 0.1$ small, > 0.3 medium, and > 0.5 large effect).

4.4.1 Objective Measures. For objective measures, we report the median (interquartile range) of the *temporal* and *spatial* dimensions of each gesture in ascending order (see Figure 2 left).

Temporal: The duration per gesture are: *LS*=1.78s (IQR=1.05s), *LW*=1.80s (IQR=1.12s), *GL*=6.68s (IQR=1.12s), and *HS*=6.83s (IQR=0.63s). A Friedman test revealed significant differences between the conditions ($\chi^2(3)=39.00$, $p \leq 0.001$, $N=16$). Post-hoc tests showed significant differences between *GL* and *LW* ($W=136$, $Z=3.52$, $p \leq 0.001$, $r=0.62$), *GL* and *LS* ($W=136$, $Z=3.52$, $p \leq 0.001$, $r=0.62$), *HS* and *LW* ($W=136$, $Z=3.52$, $p \leq 0.001$, $r=0.62$), and *HS* and *LS* ($W=136$, $Z=3.52$, $p \leq 0.001$, $r=0.62$), but no significant differences between *GL* and *HS* ($W=62$, $Z=-0.31$, $p=1$) and *LW* and *LS* ($W=62$, $Z=-0.31$, $p=1$).

Spatial: The maximum distance per gesture are: *HS*=0.16m (IQR=0.06m), *GL*=0.28 (IQR=0.18m), *LW*=0.31 (IQR=0.20m), and *LS*=0.34m (IQR=0.27m). A Friedman test revealed significant differences between the conditions ($\chi^2(3)=21.38$, $p \leq 0.001$, $N=16$). Post-hoc tests showed significant differences between *HS* and *LW* ($W=0$, $Z=-3.52$, $p \leq 0.001$, $r=0.62$) and *HS* and *LS* ($W=0$, $Z=-3.52$, $p \leq 0.001$, $r=0.62$), but no significant differences between *GL* and *HS* ($W=116$, $Z=2.48$, $p=0.066$), *GL* and *LW* ($W=37$, $Z=-1.60$, $p=0.700$), *GL* and *LS* ($W=36$, $Z=-1.65$, $p=0.627$), and *LW* and *LS* ($W=48$, $Z=-1.03$, $p=1$).

4.4.2 Subjective Measures. We used the Likert-items for *Mental Demand* and *Physical Demand* of the NASA Raw-TLX questionnaire [20] to assess subjective measures regarding task load. The median task load scores (with interquartile ranges) for each gesture, in ascending order, are as follows: For **Mental Demand** we found *LW*=10.00 (IQR=5.00), *LS*=10.00 (IQR=11.25), *HS*=5.00 (IQR=5.00), and *GL*=7.50 (IQR=5.00). A Friedman test revealed no significant differences between the conditions ($\chi^2(3)=5.23$, $p=0.156$, $N=16$). For **Physical Demand** we obtained *LW*=10.00 (IQR=1.25), *LS*=10.00 (IQR=7.50), *HS*=10.00 (IQR=7.50), and *GL*=12.50 (IQR=10.00). A Friedman test revealed no significant differences between the conditions ($\chi^2(3)=1.04$, $p=0.792$, $N=16$).

Suggested Gestures: Participants were invited to suggest new gestures, and most ($n=10$) recommended knob-based gestures. These suggestions ranged from additional letters such as “I”, “L”, or “U”, to straightforward non-letter gestures like a straight line “—”, a

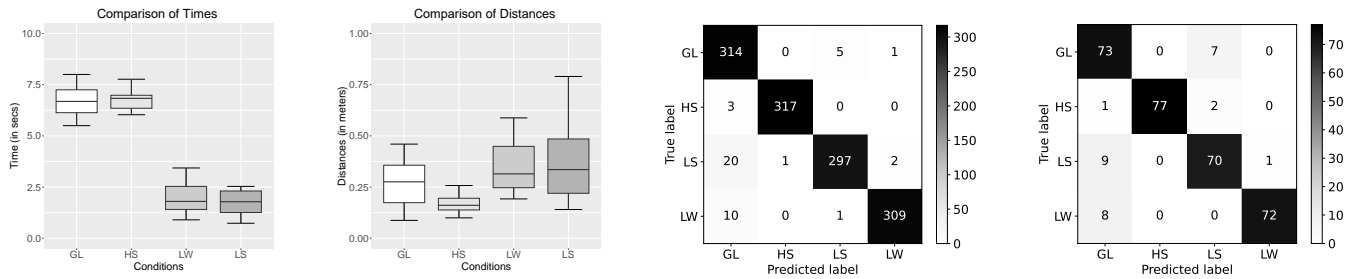


Figure 2: Measures of gestures, for (first) temporal, (second) spatial dimensions, and confusion matrices, for (third) inverse classification (cross-validation), the training sample = 0.2, and (fourth) user-independent (cross-subject), with training sample = 0.5.

checkmark “√”, and more intricate gestures like a question mark “?” or a square “□”. Certain suggestions leaned towards human-like interactions, including fist bumps, holding hands, or shaking objects. This aligns with participants’ impressions of the interaction, describing the handshake as “natural” (P5, P6).

5 GESTURE RECOGNITION

Applying ML-based gesture recognition, we conducted a cross-validated classification using an 80-20 and 20-80 training-testing data set and evaluated a user-independent classification.

Data Analysis. For the training process, we computed four low-level descriptive statistical values (minimum, maximum, mean, and standard deviation) for all positions, as well as velocity and effort values per joint for each of the seven joints. This combination resulted in a set of 84 features per gesture, per participant. As a classifier, we used the Random Forest (RF) implementation of *scikit-learn* with default parameters (e.g., 100 estimators) [39].

Cross-Validation Classification. To analyze the effect of differences *between gestures*, we calculated the mean accuracy for each individual gesture using k-fold cross-validation (k=5). We performed the classical 80-20 split of training and test set, and achieved an average f1-score of 0.99. To further test the robustness of distinguishing gestures across participants, we performed an inverse classification, where the train and test sets were 20% and 80%, respectively. An average f1-score of 0.96 was achieved. The confusion matrix of the 20-80 classifier is illustrated in Figure 2 (right).

User-Independent Classification. To test the generalizability of our approach, i.e., its performance against new and unknown users, the data set was equally divided into disjoint training and test sets (i.e., with no overlap of participants). Classifying two folds achieved a mean f1-score accuracy of 0.91 (0.85 and 0.97 respectively).

6 DISCUSSION

We investigated gesture recognition by using motion data of a robotic arm manipulated by the user, by incorporating *relative* and *absolute* tracking data of the robot’s end effector. A critical consideration is the robot’s inverse kinematics (IK), which may deviate from natural human motion due to joint angle constraints.

Validity of Mechanical Interaction with a Robot. Cross-validation produced promising results, with varying train and test set sizes

achieving f1-scores between 0.96 and 0.99. These outcomes highlight the potential for solid performance even with limited training data. Furthermore, user-independent classification demonstrated robustness in recognizing gestures, irrespective of users, achieving an accuracy rate of 0.91. These findings underscore the practicality of our solution across diverse users and scenarios, particularly in systems requiring gesture recognition from different individuals.

Subjective Feedback. Participants perceived the gestures as neither *mentally* nor *physically* demanding, indicating a convenient and natural interaction. However, this perception might be influenced by the limited number of repetitions (N=5/gesture). Possible applications of this method include interrupting an undesired ongoing task or altering direction through a single, clear, and unambiguous command.

Limitations & Future Work. We selected and assessed various gestures with different handles for the robot’s end effector, though not exhaustively exploring every combination of gesture and handle. Furthermore, our repertoire of gestures was limited to just four distinct ones.

Future work will involve expanding the number of repetitions per gesture and incorporating additional gestures. We anticipate conducting further studies based on this work, potentially involving a more diverse range of user groups. This may also encompass studies exploring the impact of altering robot tasks through mechanical input.

7 CONCLUSION

We highlight mechanical-based gestures in HRI by manipulating two robot end effectors (knob/hand) with four gestures (two letters / two handshakes). Classification achieved 0.91 for user-independent and 0.99 using 80% of the collected data for training. This robust gesture recognition establishes mechanical interaction with the robotic arm as a feasible, immediate, and intuitive user input.

ACKNOWLEDGMENTS

This work is funded by the *Deutsche Forschungsgemeinschaft (DFG, German Research Foundation)* – Project ID 425869382 & 426052422 and the *German Federal Ministry of Education and Research (BMBWF, FKZ: 16SV8565)*.

REFERENCES

- [1] D.F. Abawi, J. Bienwald, and R. Dorner. 2004. Accuracy in optical tracking with fiducial markers: an accuracy function for ARToolKit. In *Third IEEE and ACM International Symposium on Mixed and Augmented Reality*. 260–261. <https://doi.org/10.1109/ISMAR.2004.8>
- [2] Arash Ajoudani, Andrea Maria Zanchettin, Serena Ivaldi, Alin Albu-Schäffer, Kazuhiro Kosuge, and Oussama Khatib. 2017. Progress and prospects of the human–robot collaboration. *Autonomous Robots* 42, 5 (Oct. 2017), 957–975. <https://doi.org/10.1007/s10514-017-9677-2>
- [3] Rebecca Andreasson, Beatrice Alenljung, Erik Billing, and Robert Lowe. 2017. Affective Touch in Human–Robot Interaction: Conveying Emotion to the Nao Robot. *International Journal of Social Robotics* 10, 4 (Dec. 2017), 473–491. <https://doi.org/10.1007/s12369-017-0446-3>
- [4] Lisa Anthony and Jacob O. Wobbrock. 2010. A Lightweight Multistroke Recognizer for User Interface Prototypes. In *Proceedings of Graphics Interface 2010* (Ottawa, Ontario, Canada) (*GI '10*). Canadian Information Processing Society, CAN, 245–252.
- [5] Lisa Anthony and Jacob O. Wobbrock. 2012. \$N\$-Protractor: A Fast and Accurate Multistroke Recognizer. In *Proceedings of Graphics Interface 2012* (Toronto, Ontario, Canada) (*GI '12*). Canadian Information Processing Society, CAN, 117–120.
- [6] Ameer Badr and Alia Abdul-Hassan. 2020. A Review on Voice-based Interface for Human-Robot Interaction. *Iraqi Journal for Electrical and Electronic Engineering* 16, 2 (Nov. 2020), 1–12. <https://doi.org/10.37917/ijeee.16.2.10>
- [7] Ling Bao and Stephen S. Intille. 2004. Activity Recognition from User-Annotated Acceleration Data. In *Pervasive Computing*, Alois Ferscha and Friedemann Mattern (Eds.). Springer Berlin Heidelberg, Berlin, Heidelberg, 1–17.
- [8] Wilhelm Bauer, Manfred Bender, Martin Braun, Peter Rally, and Oliver Scholtz. 2016. Lightweight robots in manual assembly—best to start simply. *Fraunhofer-Institut für Arbeitswirtschaft und Organisation IAO, Stuttgart* (2016).
- [9] Roger Bemelmans, Gert Jan Gelderblom, Pieter Jonker, and Luc de Witte. 2012. Socially assistive robots in elderly care: a systematic review into effects and effectiveness. *Journal of the American Medical Directors Association* 13, 2 (2012), 114–120. <https://doi.org/10.1016/j.jamda.2010.10.002>
- [10] Kathleen Blanchet, Amel Bouzeghoub, Selma Kchir, and Olivier Lebec. 2020. How to Guide Humans Towards Skills Improvement in Physical Human-Robot Collaboration Using Reinforcement Learning?. In *2020 IEEE International Conference on Systems, Man, and Cybernetics (SMC)*. 4281–4287. <https://doi.org/10.1109/SMC42975.2020.9283469>
- [11] Andrea Bonci, Pangcheng David Cen Cheng, Marina Indri, Giacomo Nabissi, and Fiorella Sibona. 2021. Human-robot perception in industrial environments: A survey. *Sensors* 21, 5 (2021), 1571. <https://doi.org/10.3390/s21051571>
- [12] Maria E. Cabrera and Juan P. Wachs. 2016. Embodied gesture learning from one-shot. In *2016 25th IEEE International Symposium on Robot and Human Interactive Communication (RO-MAN)*. IEEE. <https://doi.org/10.1109/roman.2016.7745244>
- [13] Baptiste Caramiaux, Nicola Montecchio, Atsu Tanaka, and Frédéric Bevilacqua. 2014. Adaptive Gesture Recognition with Variation Estimation for Interactive Systems. *ACM Transactions on Interactive Intelligent Systems* 4, 4 (Dec. 2014), 1–34. <https://doi.org/10.1145/2643204>
- [14] Lipeng Chen, Luis F C Figueredo, and Mehmet R. Dogar. 2018. Planning for Muscular and Peripersonal-Space Comfort During Human-Robot Forceful Collaboration. In *2018 IEEE-RAS 18th International Conference on Humanoid Robots (Humanoids)*. 1–8. <https://doi.org/10.1109/HUMANOIDS.2018.8624978>
- [15] Andrea Cherubini, Robin Passama, André Crosnier, Antoine Lasnier, and Philippe Fraisse. 2016. Collaborative manufacturing with physical human–robot interaction. *Robotics and Computer-Integrated Manufacturing* 40 (Aug. 2016), 1–13. <https://doi.org/10.1016/j.rcim.2015.12.007>
- [16] Sandra Drolshagen, Max Pfingsthorn, Pascal Gliesche, and Andreas Hein. 2021. Acceptance of Industrial Collaborative Robots by People With Disabilities in Sheltered Workshops. *Frontiers in Robotics and AI* 7 (2021), 173. <https://doi.org/10.3389/frobt.2020.541741>
- [17] Florian Ehtler and Andreas Butz. 2012. GISpL. In *Proceedings of the Sixth International Conference on Tangible, Embedded and Embodied Interaction*. ACM. <https://doi.org/10.1145/2148131.2148181>
- [18] Franka Emika GmbH. 2023. <https://frankaemika.github.io/docs/overview.html> last accessed: January 16, 2024.
- [19] Rinat R. Galin and Roman V. Meshcheryakov. 2020. *Human-Robot Interaction Efficiency and Human-Robot Collaboration*. Springer International Publishing, Cham, 55–63. https://doi.org/10.1007/978-3-030-37841-7_5
- [20] Sandra G Hart and Lowell E Staveland. 1988. Development of NASA-TLX (Task Load Index): Results of empirical and theoretical research. In *Advances in psychology*. Vol. 52. Elsevier, 139–183.
- [21] Evert Helms, Rolf Dieter Schraft, and M Hagele. 2002. rob@work: Robot assistant in industrial environments. In *Proceedings. 11th IEEE International Workshop on Robot and Human Interactive Communication*. IEEE, 399–404.
- [22] Khoi Hoang Dinh, Ozgur S Oguz, Mariam Elsayed, and Dirk Wollherr. 2019. Adaptation and transfer of robot motion policies for close proximity human-robot interaction. *Frontiers in Robotics and AI* (2019), 69.
- [23] Long Huang, Zhen Meng, Zeyu Deng, Chen Wang, Liying Li, and Guodong Zhao. 2021. Robot Behavior-Based User Authentication for Motion-Controlled Robotic Systems. In *IEEE INFOCOM 2021-IEEE Conference on Computer Communications Workshops (INFOCOM WKSHPS)*. IEEE, 1–6.
- [24] Carles F. Julià, Nicolas Earnshaw, and Sergi Jordà. 2013. GestureAgents. In *Proceedings of the 7th International Conference on Tangible, Embedded and Embodied Interaction*. ACM. <https://doi.org/10.1145/2460625.2460658>
- [25] Danica Kragic, Joakim Gustafson, Hakan Karaoguz, Patric Jensfelt, and Robert Krug. 2018. Interactive, Collaborative Robots: Challenges and Opportunities.. In *IJCAI*. 18–25.
- [26] Yang Li, Jin HUANG, Feng TIAN, Hong-An WANG, and Guo-Zhong DAI. 2019. Gesture interaction in virtual reality. *Virtual Reality & Intelligent Hardware* 1, 1 (Feb. 2019), 84–112. <https://doi.org/10.3724/sp.j.2096-5796.2018.0006>
- [27] Nathan Magrofuoco, Paolo Roselli, and Jean Vanderdonck. 2021. Two-dimensional Stroke Gesture Recognition. *Comput. Surveys* 54, 7 (July 2021), 1–36. <https://doi.org/10.1145/3465400>
- [28] Hamza Mahdi, Sami Alperen Akgun, Shahed Saleh, and Kerstin Dautenhahn. 2022. A survey on the design and evolution of social robots – Past, present and future. *Robotics and Autonomous Systems* 156 (2022), 104193. <https://doi.org/10.1016/j.robot.2022.104193>
- [29] Inaki Maurtua, Aitor Iburguren, Johan Kildal, Loreto Susperregi, and Basilio Sierra. 2017. Human–robot collaboration in industrial applications: Safety, interaction and trust. *International Journal of Advanced Robotic Systems* 14, 4 (2017), 172988141716010.
- [30] George Michalos, Sotiris Makris, Panagiota Tsarouchi, Toni Guasch, Dimitris Kontovrakis, and George Chrysolouris. 2015. Design considerations for safe human-robot collaborative workplaces. *Procedia CIRP* 37 (2015), 248–253.
- [31] Jonathan Mumm and Bilge Mutlu. 2011. Human-Robot Proxemics: Physical and Psychological Distancing in Human-Robot Interaction. In *Proceedings of the 6th International Conference on Human-Robot Interaction* (Lausanne, Switzerland) (*HRI '11*). Association for Computing Machinery, New York, NY, USA, 331–338. <https://doi.org/10.1145/1957656.1957786>
- [32] My Human Kit on Thingiverse.com (The Humanlab). 2023. <https://www.thingiverse.com/thing:1755922> last accessed: January 15, 2024.
- [33] Rubén E. Nogales and Marco E. Benalcázar. 2021. Hand gesture recognition using machine learning and infrared information: a systematic literature review. *International Journal of Machine Learning and Cybernetics* 12, 10 (July 2021), 2859–2886. <https://doi.org/10.1007/s13042-021-01372-y>
- [34] Max Pascher, Annalies Baumeister, Stefan Schneegass, Barbara Klein, and Jens Gerken. 2021. Recommendations for the Development of a Robotic Drinking and Eating Aid - An Ethnographic Study. In *Human-Computer Interaction - INTERACT 2021* (2021-09-01), Carmelo Ardito, Rosa Lanzilotti, Alessio Malizia, Helen Petrie, Antonio Piccinno, Giuseppe Desolda, and Kori Inkpen (Eds.). Springer, Cham. https://doi.org/10.1007/978-3-030-85623-6_21
- [35] Max Pascher, Uwe Gruenefeld, Stefan Schneegass, and Jens Gerken. 2023. How to Communicate Robot Motion Intent: A Scoping Review. In *Proceedings of the 2023 CHI Conference on Human Factors in Computing Systems - CHI '23*. <https://doi.org/10.1145/3544548.3580857>
- [36] Fabio Pini, Francesco Leali, and Matteo Ansaloni. 2015. A systematic approach to the engineering design of a HRC workcell for bio-medical product assembly. In *2015 IEEE 20th Conference on Emerging Technologies & Factory Automation (ETFA)*. IEEE, 1–8.
- [37] Thomas Schlömer, Benjamin Poppinga, Niels Henze, and Susanne Boll. 2008. Gesture Recognition with a Wii Controller. In *Proceedings of the 2nd International Conference on Tangible and Embedded Interaction* (Bonn, Germany) (*TEI '08*). Association for Computing Machinery, New York, NY, USA, 11–14. <https://doi.org/10.1145/1347390.1347395>
- [38] Jane Shi, Glenn Jimmerson, Tom Pearson, and Roland Menassa. 2012. Levels of human and robot collaboration for automotive manufacturing. In *Proceedings of the Workshop on Performance Metrics for Intelligent Systems*. 95–100.
- [39] The contributors of scikit-learn. 2023. API documentation of sklearn.ensemble.RandomForestClassifier. <https://scikit-learn.org/stable/modules/generated/sklearn.ensemble.RandomForestClassifier.html> last accessed: January 14, 2024.
- [40] Radu-Daniel Vatavu. 2017. Improving Gesture Recognition Accuracy on Touch Screens for Users with Low Vision. In *Proceedings of the 2017 CHI Conference on Human Factors in Computing Systems* (Denver, Colorado, USA) (*CHI '17*). Association for Computing Machinery, New York, NY, USA, 4667–4679. <https://doi.org/10.1145/3025453.3025941>
- [41] Radu-Daniel Vatavu, Lisa Anthony, and Jacob O. Wobbrock. 2012. Gestures as Point Clouds: A \$P\$ Recognizer for User Interface Prototypes. In *Proceedings of the 14th ACM International Conference on Multimodal Interaction* (Santa Monica, California, USA) (*ICMI '12*). Association for Computing Machinery, New York, NY, USA, 273–280. <https://doi.org/10.1145/2388676.2388732>
- [42] Radu-Daniel Vatavu, Lisa Anthony, and Jacob O. Wobbrock. 2018. \$Q\$: A Super-Quick, Articulation-Invariant Stroke-Gesture Recognizer for Low-Resource Devices. In *Proceedings of the 20th International Conference on Human-Computer Interaction with Mobile Devices and Services* (Barcelona, Spain) (*MobileHCI '18*).

- Association for Computing Machinery, New York, NY, USA, Article 23, 12 pages. <https://doi.org/10.1145/3229434.3229465>
- [43] Raghav H. Venkatnarayan, Shakir Mahmood, and Muhammad Shahzad. 2021. WiFi based Multi-User Gesture Recognition. *IEEE Transactions on Mobile Computing* 20, 3 (2021), 1242–1256. <https://doi.org/10.1109/TMC.2019.2954891>
- [44] Christian JAM Willemsse and Jan BF Van Erp. 2019. Social touch in human-robot interaction: Robot-initiated touches can induce positive responses without extensive prior bonding. *International journal of social robotics* 11, 2 (2019), 285–304.
- [45] Jacob O. Wobbrock, Andrew D. Wilson, and Yang Li. 2007. Gestures without Libraries, Toolkits or Training: A \$1 Recognizer for User Interface Prototypes. In *Proceedings of the 20th Annual ACM Symposium on User Interface Software and Technology* (Newport, Rhode Island, USA) (*UIST '07*). Association for Computing Machinery, New York, NY, USA, 159–168. <https://doi.org/10.1145/1294211.1294238>
- [46] Jiahui Wu, Gang Pan, Daqing Zhang, Guande Qi, and Shijian Li. 2009. Gesture Recognition with a 3-D Accelerometer. In *Ubiquitous Intelligence and Computing*. Springer Berlin Heidelberg, 25–38. https://doi.org/10.1007/978-3-642-02830-4_4
- [47] Sean Ye, Karen Feigh, and Ayanna Howard. 2020. Learning in Motion: Dynamic Interactions for Increased Trust in Human-Robot Interaction Games. In *2020 29th IEEE International Conference on Robot and Human Interactive Communication (RO-MAN)*. 1186–1189. <https://doi.org/10.1109/RO-MAN47096.2020.9223437>