

Hand in Hand with Robots: Differences Between Experienced and Naive Users in Human-Robot Handover Scenarios

Sebastian Meyer zu Borgsen^(✉), Jasmin Bernotat, and Sven Wachsmuth

Exzellenzcluster Cognitive Interaction Technology (CITEC),
Bielefeld University, Inspiration 1, 33619 Bielefeld, Germany
{semeyerz,jbernotat,swachsmu}@cit-ec.uni-bielefeld.de

Abstract. Service robots are expected to closely interact with humans in the near future. Their tasks often include delivering and taking objects. Thus, handover scenarios play an important role in human-robot-interaction. A lot of work in this field of research focuses on speed, accuracy and predictability of the robot's movement during object handover. Those robots need to closely interact with naive users and not only experts. In order to evaluate handover interaction performance between human and robot a force measurement based approach was implemented on the humanoid robot Floka. Different gestures with the second arm were added to analyze the influence on synchronization, predictability, and human acceptance. In this paper we present a study where users with different levels of experience were asked to help the robot to learn new objects. We evaluated the impact of previous knowledge with robots on handover interactions. Disparities in timing, distance, and applied force during handover could be observed. We present an automated annotation pipeline for human-robot-interaction that will be used in future studies. While the commonly used force measurement based approach proved to be a valid starting point, our results show that naive user interaction could benefit from better anticipation.

1 Introduction

Handing over objects is a complex task that entails collaboration and precise synchronization in space and time. Object handovers take place everywhere in our daily lives [25] for instance, delivering a drink or helping with the dishes. Therefore, the ability to exchange objects with humans is mandatory for socially accepted interaction with service robots.

Such close interactions demand a social behavior of the robot. One aspect for collaboration is communication which helps to create joint action understanding. It has been shown that integration of non-verbal cues like gaze and head orientation improves robot-to-human object handover [11]. Socially accepted approach directions and distances for approaching someone to handover an object has been analyzed by Koay et al. [15]. Integrated systems on a mobile service robot have been studied with the result that adaptivity and complementary skills of human

and robot allow to hand objects to one another [10,24]. Even though there was progress in this field the robots still need improvement to transfer cognitive and physical load from human to robot. This is especially important for non-expert users, which is one aspect we want to show in this work.

There are multiple areas of research in the field of handover. Including generating and optimizing robotic arm trajectories, detecting and handling the object transfer, positioning of mobile service robots, and verbal and non-verbal communication during the interaction. Acceptance and predictability of robots can be improved by generating smooth and human-like motions: Legible trajectories during collaboration help to decrease the coordination time [9]. An approach of synthesizing object receiving motions of humanoid robots based on a human motion database might create legible movements but might be hard to adapt during execution [27]. Orientations of objects during handover used by humans were tracked and analyzed for efficient robot to human handovers [7]. Maximizing user comfort with an affordance-sensitive system that helps to align objects before handover is perceived more human-like [1,6]. Dynamic movement primitives proved to generate predictable as well as reactive trajectories [21,22]. A notable advantage of this approach is the adaptivity of the trajectory during execution. In a later study the others showed that timing might be even more important than position [16]. Timing had more influence on the perceived safety than the actual trajectories. Adaptive coordination strategies such as waiting for the human are often a trade-off between team performance and user experience [13]. The results of Basili et al. show that approaching and handover are smooth and dynamic actions with the different parts blending into each other [3]. Huber et al. investigated the timing in human-human and human-robot interaction and split handovers into three phases: reaction, manipulation, and posthandover [14].

In the manipulation phase, the object is transferred between subjects. Hence, the robot needs some kind of sensing when to release/grasp the object. Most existing approaches make use of force-torque sensors in the wrist to sense when the human applies force to the robots end-effector either by pulling on the object or by pushing it into the hand of the robot [4,8,12]. More advanced approaches add optical or tactile sensors in the gripper to optimize grip-force control and contact detection [18,20]. All these approaches expect the user to actually apply force above a certain threshold. In a pre-study Chan et al. discovered that this is not the case for all interactions and decided to instruct the participants to pull on the object until it is released [8]. Although this is a good approach to validate and compare algorithms the need to get handling instructions contradicts our anticipation of natural human-robot handover.

2 Human-Robot Handover Experiment

A handover incorporates a lot of communication to synchronize between the interaction partners. Thus, we wanted to test in which way gestures with the second arm of the robot help to indicate the state of the robot. As of now

robots are either not able to move and react with the speed and acceleration of humans or safety concerns lead to a limitation of those. Thus, humans can not easily apply the same patterns and expectations they have from human-human-handovers to the human-robot-case. To overcome this difficulty we designed an experiment to test for the following hypothesis:

H1: *Additional gestures with the second arm help to synchronize between human and robot.*

Approaches discussed in Sect. 1 often evaluate interaction with participants that already have experience with robots or instruct them to test for a distinct behavior. From fairs and events like the RoboCup@home [5] we had the impression that there might be a significant difference in interaction from users that have no experience with robots. This led to the following second hypothesis that we investigated in the study:

H2: *Naive and (robot-)experienced users handle object handovers differently.*

2.1 Experiment Procedure and Design

We chose the robot Floka [5] to study the interaction with the human in a one factorial between-subjects interaction. The human-like torso with two arms allowed us to design gestures that are equal to humans' and thus should be easily recognizable by the participants.

The goal of the user-study was to record the interactions with Floka as natural as possible. We decided to go without a tracking system that depends on markers or sensors attached to the participants to prevent interference. The movements were post-annotated by means of automatic extraction from an external camera. In order to inhibit the emergence of artifacts by participants concentrating on the handover itself, a distractor-task was placed. The participants were instructed to help the robot to learn the shape of new objects.

As gaze might improve turn taking during handover [17] we implemented a turn-taking gaze scheme on Floka. The robot looked at the object while moving towards the participant. After that it gazed into the face of the participant when it was ready to hand/receive the object. These head movements are the same for all the interaction runs. The robot was always rising the right arm for handover without telling the participants before. Figure 1d shows the movements when Floka is learning an object as distractor-task. To test for H1 we designed two different gestures for the left arm, which was not involved in the object transfer, to signal the state of the robot. The first one (C_{low}) was turning the hand in a presenting manner below the object to signalize readiness. Figure 1b shows that this gesture made only use of small movements to be less intrusive. The second gesture (C_{high}) depicted in Fig. 1c started with a protecting movement of the object with the goal to signalize that the robot is not yet ready to hand the object. The trajectory ended as well in a presenting gesture but in a more distinct fashion. Both gestures were synchronized with the handover trajectory. As control condition ($C_{control}$) Floka did not move the left arm during

the handover. The arm was kept in a neutral posture as can be seen in Fig. 1a. Each participant was assigned randomly to one of the three conditions. In these conditions the gesture was only activated for odd-numbered runs to allow measurements of within-subject differences. The interaction consisted of nine gives and receives.

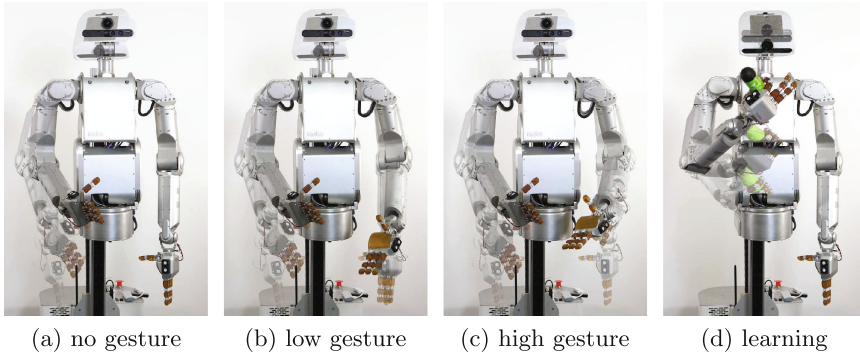


Fig. 1. Blended pictures of Floka during the interaction. Each figure is rendered out of three frames during the movement. Viewpoint is similar to the participant's.

To only analyze the effects of non-verbal communication Floka did neither speak nor respond to speech input during the study. For security reasons, the experimenter stayed next to the external camera with a wireless emergency stop. This e-stop was also programmed to start the experiment. For detection in the manipulation phase we used an *ATI F/T Sensor: Mini40* in the wrist to measure forces applied to the robot similar to related work discussed in Sect. 1. The robot is only able to detect contact after the arm trajectory finished as the trajectory itself applies higher force on the sensor than the interaction itself.

After the interaction the participants had to answer a survey. Besides age and gender the participants gave self-assessment on experience with technology like computers and robots. The attitude towards robots was investigated using the NARS [19] ($\alpha = .64$) question-set. We collected information on how the robot was perceived during the handovers with the Godspeed [2] items ($\alpha = .90$). In the conclusion of the survey the participants were asked whether they noticed different behavior patterns during the interaction. This data was collected by means of a freetext field.

The exact setup can be seen in Fig. 2 from the view of the external camera (Fig. 2b). Figure 2a shows a schematic of the room-setup from top. Floka is positioned such that the participant can freely chose a position in front of the robot. Three objects are placed on a small table near the interaction area. The external camera is placed on a table on the other side of the room to have a complete view on the interaction.

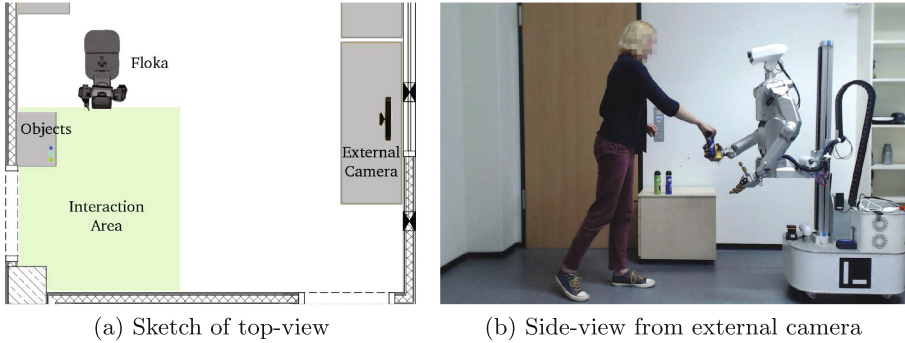


Fig. 2. Setup of the experiment from view of the external camera and as a schematic top view. Our robot Floka is receiving an object from one of the participants. The other two objects are still placed on a small table next to the interaction.

In total $N = 40$ participants took part in our experiment with Floka. Eight runs were not used in the following evaluation because of technical dropouts during the recording. The remaining 32 participants (17 male and 15 female aged between 18–53 years) were randomly assigned to the conditions with the following distribution: 10 $C_{control}$, 10 C_{low} , and 12 C_{high} . We split the participants in three groups based on experience on their self-assessment stated in the survey. The group of naive users only contains participants that stated they have no experience on interaction with robots. This is the first group with 12 persons. The participants that answered to this question with 4–7 form the group of experts containing 11 persons. The remaining participants form the group of semi-experienced with 9 participants.

For each participant the procedure was: enter the room, read and sign the consent form on a designated table, after that they were asked to come to the interaction area. In this area they were introduced to Floka and were instructed for the experiment. After all nine runs they were asked to answer the survey.

2.2 Experiment Annotation

In total 725 experiment recordings were created during this study. Each recording contains the positions of the robot’s joints, the forces applied to the force-torque-sensors in the robot’s wrists, the timing and state of the handover control-system. Video-streams of the robots internal face-camera and the external camera were stored synchronously as-well. A marker on the robot helped to exactly determine the position of the external camera in relation to the robot. This allows mapping internal robot data like forces, torques and positions into the video as depicted in Fig. 3.

A pipeline that loads all files and automatically annotates the recordings was implemented in order to extract and compare positions and velocities of the human and the robot. Convolutional Pose Machines (CPMs) [26] were used to

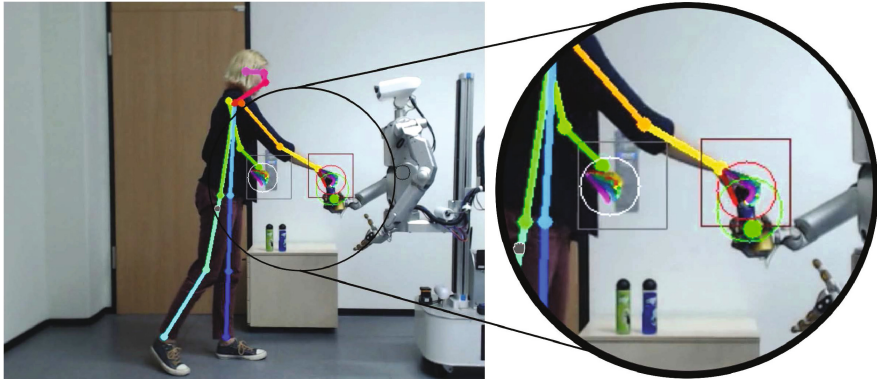


Fig. 3. Visualization of the results from post-processing with help of OpenPose [26]. Each joint is visualized in a different color. The position of Floka's hand is marked with a green dot. Bounding boxes of possible participant hand positions are red for the right and gray for the left hand. Accordingly, the center of hand joints is surrounded by light red and light gray, respectively. When contact between human and robot is detected a green circle is drawn around the hands in contact. (Color figure online)

extract the position of the human in the videos. To precisely annotate the hands, the pose detections were enhanced with hand keypoint detection [23]. The resulting annotation can be seen in Fig. 3. The processing pipeline generated a log of the positions extracted by the CPM and hand tracking as well as a video with all data visualized for each recording. In addition, timestamps for the different phases of the handovers were logged for approaching, contact and retraction phase. Figure 4 shows the trajectories of the robot and participants during handover.

2.3 Findings and Results

The analysis of the survey showed that in total 17 of the participants stated that they experienced differences in the behavior of the robot in between the runs. Although only seven were able to describe the differences correctly in the way that the second arm did support the handover with a gesture. Some of them stated in the free-response that movements with the gesture looked more natural. Further analysis of the ratings from the survey, timing, and position data to confirm H1 did not show a statistically measurable effect ($p > .05$). The introduced within person condition that switched between gesture and control condition every second run did not result in measurable differences in behavior of the participant either. This alternation might have reduced the measurability of the overall effect of the gesture in combination with the survey being answered after seeing both behaviors.

For some participants the high gesture looked like Floka was offering the left hand for receiving an object. Although only it's right hand was able to detect and grasp objects. This led to confusion and created huge offsets in timing until they continued to give the object into the right hand. These offsets influenced the overall statistics.

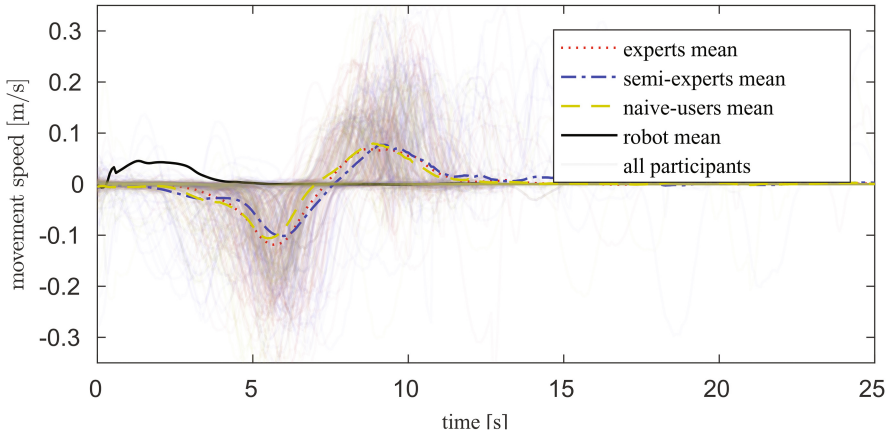


Fig. 4. Velocity profile of the handover runs for right-handed interaction. Floka moves his hand towards the participant. Some participants start to move right after the robot started. Most of them wait until it finished. Then hand the object and move back. The colored lines represent the average for each of the three groups.

One of the analyzed aspects was the reaction time to see how well the movements of the robot and participants aligned. The alignment was calculated as the difference between the time the robot was ready and the time the person’s hand getting close to Floka’s hand. A perfect alignment would be a 0.0 s result. Table 1 shows that the participants took an average of 0.29 s. Meaning they gave a little more time for the robot to finish moving. Negative differences are cases in which the participants tried to hand the object while Floka was still executing the trajectory. Naive users show the least standard variation here as they actively tried to align well with the system and actually helped it to fulfill the task of learning objects most efficiently. In the analysis of recordings we could observe the experts and semi-experienced participants testing the robot. They actively introduced delays to see how the robot would react to them. The testing went as far as giving the object in the hand of the robot and pulling it away as the robot closes its hand.

Table 1. Mean and standard deviation for measured duration grouped by the experience with robots.

	<i>naive</i>	<i>semi</i>	<i>expert</i>	<i>total</i>		<i>naive</i>	<i>semi</i>	<i>expert</i>	<i>total</i>
\bar{x} [s]	0.36	-0.30	0.76	0.29	\bar{x} [s]	1.60	2.68	1.27	1.77
σ [s]	1.09	1.38	1.86	1.49	σ [s]	2.45	3.89	1.62	2.74

(a) reaction times: the time between the robot was ready to receive or hand an object and the participant reaching for it.

(b) manipulation times: the time between first contact between robot and human and contact detected by the robot.

Secondly we calculated the time needed to transfer the object. Which mainly tests how well the force-based approach succeeds in detecting a stable hand-over. The object was dropped in only one run and only two runs timed out after the robot was waiting 30s for the person to pull the object strong enough to trigger the force threshold for releasing the object. A major problem with this approach is that some participants, mostly naive, did not apply force at all on the first tries and expected the robot to see that they hand the object. This was especially happening for handing an object to the robot and happened in more cases for the naive and semi-experienced users. When giving an object to the robot applying pressure seems to be less intuitive than pulling when taking it from the robot. Table 1 visualizes how experts have lower mean and standard deviation of time when exchanging the object with Floka. They seem to be already used to trigger force thresholds to make a robot react.

3 Conclusion

We presented a study on natural human-robot handover with the robot Floka. Therefore, we used an implementation of wrist-force based handover detection. There was no artificial tracking system and only minimal instructions for the participants to observe the interaction without interference. A novel annotation system that does not interfere with the participants was created to evaluate human-robot interaction by making use of deep-learning techniques. This low-cost and easy deployable system allows fully automatic annotation of human motion without time-consuming manual annotation of video data. Furthermore, it can replace other intrusive marker-based tracking-solutions and will be used in future HRI-studies. We could not statistically prove that a gesture with the second arm helps to improve the synchronization between human and robot (H1). The effects of other phenomena appear stronger in the data and have to be addressed beforehand. Although, participants that consciously perceived the gesture stated in the survey that they experienced the robot more human-like when the gesture was part of the interaction. We noticed significant differences in behavior with varying levels of prior knowledge in regard to HRI (H2). While naive users expect the robot to visually perceive the environment and react accordingly, experienced users know that they need to pull and push objects for the robot to perceive their intention. This leads to the assumption that future implementations can not only rely on force measurements to prevent a social gap between users of service robots. Especially with the elderly and disabled in mind, handover of robots needs to be more adaptive to cope with the huge variance in observed handovers and to adapt better to the human expectations. The results of our study contributes to robots with less preconceptions on handover interactions.

Based on these results we will continue on improving interaction experience for inexperienced users as we believe that this is the group of users that need supporting robots the most. One of our goals is to make use of the system we created to post-annotate the videos without manual interference. The visual

perception of the interaction partner appears to be crucial to naturally interact during handovers and create a socially accepted robot. As robots are not expected to move with the same speeds and thus timings as humans in the near future because of security concerns in such close interactions, other methods like reactive movements and gestures are needed to overcome that gap. Another study that further investigates not only the trajectory of the end-effector that is transferring the object but the body-language on the whole could give deeper insights in the non-verbal communication cues. Synchronization by incorporating a second arm, the gaze and even the base could help to clearer communicate the internal state of the robot.

Acknowledgments. This research/work was supported by the Cluster of Excellence Cognitive Interaction Technology ‘CITEC’ (EXC 277) at Bielefeld University, which is funded by the German Research Foundation (DFG).

References

1. Aleotti, J., Micelli, V., Caselli, S.: An affordance sensitive system for robot to human object handover. *Int. J. Soc. Robot.* **6**(4), 653–666 (2014)
2. Bartneck, C., Kulić, D., Croft, E., Zoghbi, S.: Measurement instruments for the anthropomorphism, animacy, likeability, perceived intelligence, and perceived safety of robots. *Int. J. Soc. Robot.* **1**(1), 71–81 (2009)
3. Basili, P., Huber, M., Brandt, T., Hirche, S., Glasauer, S.: Investigating human-human approach and hand-over. In: Ritter, H., Sagerer, G., Dillmann, R., Buss, M. (eds.) *Human Centered Robot Systems*, vol. 6, pp. 151–160. Springer, Heidelberg (2009). doi:[10.1007/978-3-642-10403-9_16](https://doi.org/10.1007/978-3-642-10403-9_16)
4. Bdiwi, M., Suchy, J., Winkler, A.: Handing-over model-free objects to human hand with the help of vision/force robot control. In: *10th International Multi-Conferences on Systems, Signals and Devices (SSD 2013)*, pp. 1–6. IEEE (2013)
5. Meyer zu Borgsen, S., Korthals, T., Lier, F., Wachsmuth, S.: ToBI team of Bielefeld: enhancing robot behaviors and the role of multi-robotics in RoboCup@Home. In: Behnke, S., Raymond, S., Sariel, S., Lee, D.D. (eds.) *RoboCup 2016: Robot World Cup XX*. LNCS (LNAI), vol. 9776. Springer, Heidelberg (2016). doi:[10.1007/978-3-319-68792-6](https://doi.org/10.1007/978-3-319-68792-6)
6. Cakmak, M., Srinivasa, S.S., Lee, M.K., Forlizzi, J., Kiesler, S.: Human preferences for robot-human hand-over configurations. In: *IEEE International Conference on Intelligent Robots and Systems*, pp. 1986–1993 (2011)
7. Chan, W.P., Pan, M.K.X.J., Croft, E.A., Inaba, M.: Characterization of handover orientations used by humans for efficient robot to human handovers. In: *IEEE International Conference on Intelligent Robots and Systems*, vol. 2015, pp. 1–6, December 2015
8. Chan, W.P., Parker, C.A.C., Van der Loos, H.F.M., Croft, E.A.: A human-inspired object handover controller. *Int. J. Robot. Res.* **32**(8), 971–983 (2013)
9. Dragan, A.D., Bauman, S., Forlizzi, J., Srinivasa, S.S.: Effects of robot motion on human-robot collaboration. In: *Proceedings of the Tenth Annual ACM/IEEE International Conference on Human-Robot Interaction - HRI 2015*, vol. 1, pp. 51–58 (2015)
10. Edsinger, A., Kemp, C.C.: Human-robot interaction for cooperative manipulation: handing objects to one another. In: *Proceedings of IEEE International Workshop on Robot and Human Interactive Communication*, pp. 1167–1172 (2007)

11. Grigore, E.C., Eder, K., Pipe, A.G., Melhuish, C., Leonards, U.: Joint action understanding improves robot-to-human object handover. In: *IEEE International Conference on Intelligent Robots and Systems*, pp. 4622–4629 (2013)
12. He, W., Sidobre, D.: Improving human-robot object exchange by online force classification. *J. Hum. Robot Interact.* **4**(1), 75 (2015)
13. Huang, C.M., Cakmak, M., Mutlu, B.: Adaptive coordination strategies for human-robot handovers. In: *Robotics: Science and Systems* (2015)
14. Huber, M., Rickert, M., Knoll, A., Brandt, T., Glasauer, S.: Human-robot interaction in handing-over tasks. In: *Proceedings of the 17th IEEE International Symposium on Robot and Human Interactive Communication, RO-MAN* (2008)
15. Koay, K., Sisbot, E., Syrdal, D., Walters, M.: Exploratory study of a robot approaching a person in the context of handing over an object. In: *AAAI Spring Symposium* (2007)
16. Koene, A., Remazeilles, A., Prada, M., Garzo, A., Puerto, M., Endo, S., Wing, A.M.: Relative importance of spatial and temporal precision for user satisfaction in human-robot object handover interactions. In: *Third International Symposium on New Frontiers in Human-Robot Interaction*, p. 14 (2014)
17. Moon, A., Troniak, D.M., Gleeson, B., Pan, M.K., Zeng, M., Blumer, B.A., MacLean, K., Croft, E.A.: Meet me where i'm gazing. In: *Proceedings of the 2014 ACM/IEEE International Conference on Human-Robot Interaction - HRI 2014*, pp. 334–341. ACM, New York (2014)
18. Nagata, K., Oosaki, Y., Kakikura, M., Tsukune, H.: Delivery by hand between human and robot based on fingertip force-torque information. In: *Proceedings of 1998 IEEE/RSJ International Conference on Intelligent Robots and Systems. Innovations in Theory, Practice and Applications (Cat. No. 98CH36190)*, vol. 2, pp. 750–757. IEEE (1998)
19. Nomura, T., Suzuki, T., Kanda, T., Kato, K.: Measurement of negative attitudes toward robots. *Interact. Stud.* **7**(3), 437–454 (2006)
20. Parastegari, S., Noohi, E., Abbasi, B., Žefran, M.: A fail-safe object handover controller. In: *Proceedings of IEEE International Conference on Robotics and Automation*, pp. 2003–2008 (2016)
21. Prada, M., Remazeilles, A., Koene, A., Endo, S.: Dynamic Movement Primitives for Human-Robot interaction: comparison with human behavioral observation. In: *IEEE International Conference on Intelligent Robots and Systems* (2013)
22. Prada, M., Remazeilles, A., Koene, A., Endo, S.: Implementation and experimental validation of Dynamic Movement Primitives for object handover. In: *IEEE International Conference on Intelligent Robots and Systems*, pp. 2146–2153 (2014)
23. Simon, T., Joo, H., Matthews, I., Sheikh, Y.: Hand keypoint detection in single images using multiview bootstrapping. In: *Conference on Computer Vision and Pattern Recognition* (2017)
24. Sisbot, E.A., Alami, R.: A human-aware manipulation planner. *IEEE Trans. Robot.* **28**(5), 1045–1057 (2012)
25. Strabala, K., Lee, M.K., Dragan, A., Forlizzi, J., Srinavasa, S.S., Cakmak, M., Micelli, V.: Towards seamless human-robot handovers. *J. Hum. Robot Interact.* **1**(1), 112–132 (2013)
26. Wei, S.E., Ramakrishna, V., Kanade, T., Sheikh, Y.: Convolutional pose machines. In: *Conference on Computer Vision and Pattern Recognition* (2016)
27. Yamane, K., Revfi, M., Asfour, T.: Synthesizing object receiving motions of humanoid robots with human motion database. In: *2013 IEEE International Conference on Robotics and Automation*, pp. 1629–1636. IEEE (2013)