Optimizing Robot-to-Human Object Handovers using Vision-based Affordance Information

Lehotský, Daniel; Christensen, Albert Daugbjerg; Chrysostomou, Dimitrios

Published in:
IST 2023 - IEEE International Conference on Imaging Systems and Techniques, Proceedings

DOI (link to publication from Publisher):
10.1109/IST59124.2023.10355704

Publication date:
2023

Document Version
Accepted author manuscript, peer reviewed version

Optimizing Robot-to-Human Object Handovers using Vision-based Affordance Information

Daniel Lehotsky
Dept. of Architecture, Design and Media Technology
Aalborg University
Aalborg, Denmark
dale@create.aau.dk

Albert Christensen
Dept. of Materials & Production
Aalborg University
Aalborg, Denmark
Albert.Christensen@agcocorp.com

Dimitrios Chrysostomou
Dept. of Materials & Production
Aalborg University
Aalborg, Denmark
dimi@mp.aau.dk

***Abstract*—Robotic handovers of objects to humans require selecting appropriate grasp poses and orientations to enable efficient subsequent use. We present two methods to compute suitable handover orientations based solely on object affordances rather than object categories or predefined object-specific rules. The first uses human demonstration data to learn average handover orientations per object directly from affordances. The second is a rule-based method that orients graspable affordances towards the receiver. We integrated both approaches into a robotic system performing task-oriented grasping and handovers based on affordance segmentation. A user study indicates the rule-based method produces equally comfortable and natural handover orientations compared to learning from demonstration, while being simpler to implement. Further experiments demonstrate the robot's ability to successfully hand over objects with proper orientations. This is the first prototype deriving handover orientations solely from affordances treated as pixel-wise semantic segmentation, providing a practical approach without per-object datasets. https://bit.ly/RobotHandovers**

***Index Terms*—Robotic handover, Handover orientation, Object affordances, Industry 4.0, Human-robot interaction**

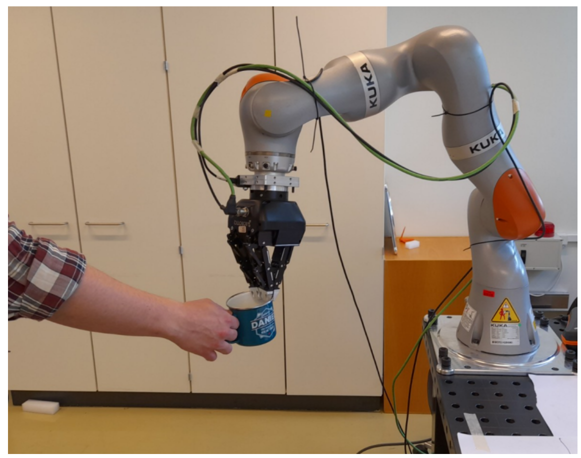

Fig. 1. In our real-world handover experiment, the Kuka LBR iiwa robot arm delivers an object after computing appropriate handover orientations using pixel-wise affordance segmentation.

I. Introduction

As robots become more integrated into human workspaces, the ability to smoothly transfer objects between robots and humans is critical for seamless collaboration. Unlike humans, who excel at fluid object handovers, robots require systematic approaches to perform such dexterous tasks successfully. The execution of robot-to-human object transfers necessitates overcoming complex challenges such as object detection, precise grasping, and contextual trajectory planning across diverse domains, such as robotic-assisted surgery [1], disassembly tasks in restricted industrial environments [2] and package delivery using language interfaces [3], [4]. Object handovers can be categorized into two types: task-agnostic and task-oriented [5]. In task-agnostic handovers, the key concern is simply the success rate of the physical exchange itself. However, for more natural human-robot collaboration (HRC), we must also consider the user's subsequent task after completing the handover [6]. This involves computing appropriate grasps and object orientations to enable efficient post-handover object use, known as task-oriented handovers [7], [8]. For example, a cup should be grasped and oriented with the handle facing the human recipient, as illustrated in Fig. 1.

In task-oriented handovers, computing suitable object orientations depends on the object's inherent functionalities or affordances [9]. However, current methods for determining proper orientations rely on predefined rules [10], [11] or are defined categorically for specific objects [12], [13]. Such methods scale poorly to a larger set of objects, and furthermore they either do not consider the inherent functionalities of the grasped object or only consider them implicitly by knowing the object class.

In our previous work, we demonstrated that affordance theory framework was well suited for solving task-oriented grasping [14]. By the term object affordances we defined the inherent functionalities of an object irrespective of its current state as similarly defined by Humphreys in [15]. Finding object affordances was treated as a pixel-wise segmentation problem based on the visual inputs of RGB cameras capturing the working environment. In this work, we present two methods for computing appropriate handover orientations based solely on detected object affordances rather than categories or heuristics. First, a data-driven approach uses human demonstration to learn average handover orientations directly from visual affordances. Second, a rule-based method orients graspable affordances towards the receiver. We integrate both techniques into a robotic system performing grasping and handovers based on affordance segmentation. Experiments indicate that the rule-based approach produces equally comfortable handover orientations compared to learning while being simpler to implement. The quantitative testing resulted in a 75% success rate for the rule-based and only a 50% success rate for the observation-based method regarding the ability to handover items with the appropriate orientations. We also demonstrate the robot successfully handing over objects with proper orientations. Additionally, our open source code used in this study is available in our GitHub repository[1].

II. Related work

Early methods for computing handover orientations were object-specific. These methods required prior knowledge of objects, and appropriate handover orientations were assigned manually per object basis [10], [11], [16]. As such, not only do these methods not scale well, but handover orientations proposed by a loss function often appear unnatural to humans as the orientations do not account for the object's affordances [10]. On the other hand, orientations learned from human-to-human handover observations account for an object's affordances as humans use their understanding of the object's functional properties to perform task-oriented grasping [17].

Chan et al. [18] studied handover orientations by observing natural human handovers. They built a knowledge base by observing how ten different objects were used and extracted features related to various affordances, i.e., *cut*, *screw*, *translate*, and *slide*. A database of appropriate object orientations for the receiver was formed. Using the extracted features, they could generalize observed orientations to unseen objects. In their subsequent studies [6], [7], they presented a method to compute handover orientations without relying on predefined features or a database. Instead, they directly observed human handovers, using an 'affordance axis' for each object at handover time based on the object's inherent affordances.

Razalli and Demiris [19] applied a multitask variational autoencoder to model both the receiver's and giver's handover pose and an appropriate object handover orientation. This model required three inputs – the initial pose of the receiver, the initial pose of the giver, and the object label, where the giver's and receiver's initial poses were captured from motion-captured data.

Ardon et al. [20] proposed a method for computing appropriate handover location and robot configuration for people with limited arm mobility. By doing so, they addressed the handover orientation problem implicitly. Task-oriented grasps were selected using object affordances, and the best grasp was selected on the criteria of *appropriateness*, *safety*, and *reachability*. In order to maximize these requirements, they inadvertently addressed the issue of the appropriate handover orientations for people with limited arm mobility.

As shown, most of the research into object handover orientations is object-specific. While some approaches utilize object affordance to optimize grasp selection, computing appropriate handover orientations from object affordances rather than object categories is still an open problem. Furthermore, most of the related work focused on learning orientations from observations. However, collecting datasets of human handover examples is time-consuming and expensive. A set of rules that are easy to integrate into a robotic handover system and consider object affordances has not been thoroughly investigated. Therefore, our work introduces a data-driven method to learn average handover orientations directly from visual affordances and a rule-based method that orients graspable affordances towards the receiver.

III. Methodology

A. Computing object handover orientations

Two methods were implemented for computing object handover orientations. The observation-based method follows the dataset collection procedure proposed by Chan et al. [6], [7]. However, collecting and annotating datasets is a tedious process. Therefore, the affordance-aware rule-based method was also implemented. This method was inspired by the work of Ardon et al. [20], but instead of optimizing for their proposed criteria, we simplify the computational part of the handover orientation into a single if statement depending on the detected affordances.

For both implemented methods, twelve object categories were considered - a bottle, a bowl, a cup, a hammer, a knife, a ladle, a mallet, a mug, scissors, a scoop, a spatula, and a spoon. Regarding affordances of these objects, our AffNet-DR detector [14] can detect seven object affordances, all defined in the UMD dataset [21]. These are *grasp*, *wide-grasp*, *cut*, *scoop*, *contain*, *pound*, and *support*.

1) Observation-based method: In order to learn object orientations from human-to-human handovers, an annotated dataset was compiled. Eight participants were asked to hand over twelve objects to one of the authors. Participants were instructed to hand over objects with the perceived comfort of the receiver in mind and using only their right hand. The handovers were recorded with Kinect V2 as a single clip. RGB, depth, point cloud and skeleton tracking data were saved with every clip.

For each video clip, the handover frame was manually annotated using labelCloud[2]. The handover frame was identified as the point when the distance between the giver and receiver was minimal, calculated using the skeleton tracking data. Prior to data collection, each object was assigned a predefined coordinate frame. The handover orientation was annotated by aligning a 3D bounding box to the object's coordinate frame in the handover frame, as illustrated in Fig. 2. This process

[1]https://bit.ly/RobotHandovers

[2]https://github.com/ch-sa/labelCloud

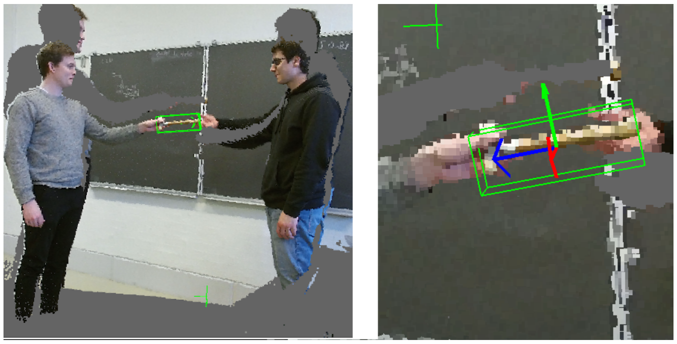

Fig. 2. The annotation procedure of the handover frame, where a 3D bounding box is aligned with the object.

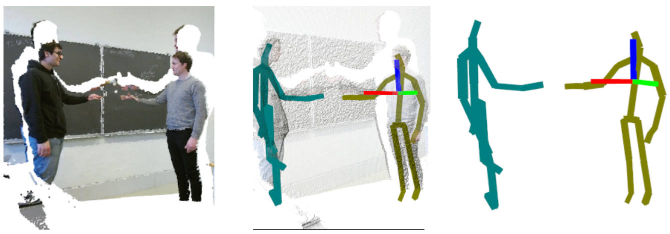

Fig. 3. A giver is assigned a coordinate frame with the x-axis (red) pointing towards the receiver and the z-axis (blue) pointing from the giver's torso to the giver's head. The y-axis (green) is inserted to complete the right-handed coordinate system. Afterwards, the recorded handover orientations are transformed from the Kinect's recording frame to the giver's frame.

was repeated for every video to create a dataset of handover orientations.

After annotating the orientation data, the coordinates were transformed from the Kinect sensor's reference frame into a standardized "giver's frame" to enable more straightforward implementation on the robot system. The giver's frame originates at the torso with the x-axis pointing towards the receiver, the z-axis pointing upwards along the spine towards the head, and the y-axis determined by the cross product to complete a right-handed coordinate system, as visualized in Fig. 3. By converting all annotations into this standardized frame, the robotic system can more efficiently utilize the handover orientation data regardless of the sensor's original reference frame.

With all the clips annotated, a mean handover orientation $\bar{q}$ was computed for each object. We cast the computation of $\bar{q}$ as a minimisation problem defined as follows:

$$\bar{q} = \arg\min_{q'} \sum_i dist(q' - q_i)$$

where

$$dist(q' - q_i) = min\{||q' - q_i||_2, ||q' + q_i||_2\}$$

where q' is the initial solution being minimised and q_i are the observed orientations for the given object. q' is defined as $q' = [x, y, z, w]$, where $\{x, y, z, w\}$ are randomly initiated such that q' is a valid quaternion. The distance function was chosen as suggested by Hartley et al. [22].

The minimisation problem was implemented in Python using the *scipy*[3] library. The function was restarted 50 times, each time with a new initial solution q'. All computed solutions were normalized. After 50 iterations, the solution that produced the minimum sum of distances to all observations for a given object was selected, as the true mean handover orientation $\bar{q}$.

2) Rule-based method: The rule-based method computes object handover orientations using only the knowledge of object affordances. Our solution is based on two rules. The first one applies to objects with the *grasp* affordance, which is associated with the handle part of the object. Therefore, if the grasp affordance is detected, the robot should grasp the non-handle parts associated with the utility affordance and orient the handle towards the receiver. The utility affordance is a general term that describes the object's functionality and is associated with parts not meant to be grasped, like a knife blade or hammerhead. This rule applies to objects with distinct handles, including hammers, knives, ladles, mallets, scissors, scoops, spatulas, and spoons. This rule aligns with findings from Ray et al. [23] and Cini et al. [24] who showed that humans tend to orient handles towards receivers during human-to-human handovers. The second rule applies to cups, mugs, bowls, and bottles. These objects offer *wide-grasp* affordance for containing liquids or other objects. These objects should be oriented vertically when handed to the receiver to avoid spillage.

Analysis of computed mean handover orientations confirmed that these rules match human behavior. Objects with handles showed minor variations but were consistently oriented with the handle pointed towards the receiver. Objects with *wide-grasp* affordance were oriented with their openings upwards. This vertical orientation minimises spillage risk, aligning with the second proposed rule.

B. A robotic system capable of task-oriented handovers

We implemented a robotic system designed to facilitate task-oriented handovers while optimizing both the receiver's ease of grasping and the object's usability. Illustrated in Fig. 4, the system is designed to ensure an unhindered grip for the receiver while orienting the object appropriately according to either the rule-based or observation-based method introduced earlier. The system was implemented on a KUKA LBR iiwa 7 R800 7-DOF manipulator equipped with a Robotic 3-finger gripper combined with an Intel RealSense D435i RGB-D sensor and a Hokuyo URG-04LX-UG01 2D laser scanner.

Expanding on our earlier work in task-oriented robotic grasping based on object affordance segmentation with our synthetically-trained affordance predictor named AffNet-DR [14], the current approach further incorporates computing appropriate handover orientations. As shown in Fig. 4, the system uses an RGB-D sensor to capture scene images, which are analyzed by AffNet-DR to generate a detailed affordance segmentation map (Fig. 4 - Steps 1 and 2). A suitable grasp is

[3]https://docs.scipy.org/doc/scipy/reference/generated/scipy.optimize.minimize.html

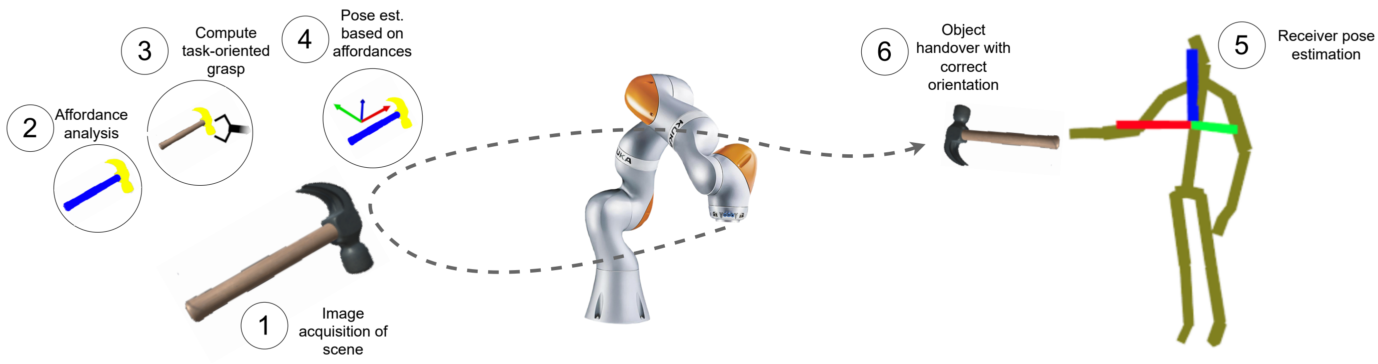

Fig. 4. An overview of the proposed robot-to-human handover system. Step 1: An RGB sensor captures the work environment. Step 2: Pixel-wise object affordances are found in the RGB image using AffNet-DR. Step 3: Grasps that facilitate smooth transition of the object from the robot to an operator are computed from the affordance segmentation map. Step 1-3 are covered more in depth in our previous work [14]. Step 4 computes a handover orientation based on the observation-based or rule-based method, before performing the robot-to-human handover with the appropriate orientation.

calculated to make the object's handle accessible to the human receiver (Fig. 4 - Step 3) and the object's pose is determined based on a modified ICP algorithm that computes the nearest neighbour only within same affordances (Fig. 4 - Step 4).

The affordances segmented by our AffNet-DR are used as features for selecting a source point cloud suitable as an input to the ICP algorithm. Once the ICP algorithm converges, object rotation is computed using singular value decomposition. With the affordance and pose of the object computed, it is possible to grasp the object by a specific affordance and move it to the handover location with the desired orientation.

Next, the position of the receiver in relation to the robot must be determined. The receiver's 2D position is derived from the laser line scan data using a basic pass through filter. Using the polar coordinates of the receiver, a "giver's frame" at the robot's base rotated to face the receiver is computed. Both the handover location and orientation are defined within this frame (Fig. 4 - Step 5). The handover location is identified as the midpoint between the giver and receiver, as noted by [25]. Finally, the system can efficiently estimate and redetermine the most suitable handover orientation by relying solely on affordance segmentation (Fig. 4 - Step 6). This orientation is computed using either the proposed rule-based or observation-based methods.

The key distinction between the rule-based and observation-based methods is in data requirements. The rule-based method requires only two source point clouds - a generic cup point cloud and a generic grasp point cloud. The observation-based method needs seven point clouds to encompass all object-specific affordance combinations AffNet-DR can detect. The second distinction lies in how affordances are used to compute appropriate handover orientations. With the observation-based method, affordances are encoded in a one-hot vector, which is then used as the input to the k-nearest algorithm that identifies the closest computed mean handover orientation. With the rule-based method, the segmented affordances are used to distinguish which of the two proposed rules should be applied.

IV. Experiments

To evaluate the two different approaches of robot-to-human handovers we conducted a series of experiments. First, we present a user study with six participants aimed at understanding preferences between the observation-based and rule-based handover orientation methods. The participants were researchers and students recruited from the university campus - 4 males, 2 females, ages 20-40. Some had experience with robotic applications, but none had prior experience with robot-to-human handovers. The goal was to evaluate their subjective impressions of handovers generated using each orientation method.

The second experiment was a full system test of the complete robotic handover pipeline with the aim of evaluating its overall performance and ability to achieve appropriate object orientations. This system-focused test provided an objective assessment to complement the user study's subjective ratings. The full pipeline was systematically tested under controlled conditions rather than with human participants.

A. User study on handover orientations preference

During the experiments, objects were always placed in their pre-determined poses on the work table. Using the pre-determined poses allowed for computing grasp poses, affordance segmentation, and object pose estimation in advance, which in turn allowed participants to concentrate and evaluate purely the effect of the handover orientation. The participants were not notified that certain parts of the pipeline were pre-computed.

In the study, participants experienced five handovers - one each for a cup, bowl, mallet, spoon, and spatula. The order of objects was consistent across users. Three orientation methods were tested: *random baseline*, *rule-based*, and *observation-based*. The methods were presented as Method A, B, or C to avoid bias. After all five objects were handed to the participant, they were asked to evaluate the handover experience by completing a 7-item questionnaire rating safety, comfort, agreement, understanding, naturalness, appropriateness and preference on 5-point Likert scales. This was repeated for

TABLE I

RESULTS OF THE USER STUDY WHERE THE AVERAGE RATINGS ARE SHOW. THE STUDY WAS EVALUATED ON THE 5-POINT LIKERT SCALE WITH 1 BEING "STRONG AGREE" AND 5 BEING "STRONG DISAGREE". RANDOM-BASED METHOD WAS USED AS A BASELINE. BOLD INDICATE THE BEST SCORE.

Question	Random-based	Rule-based	Observations-based
How safe did you feel during the handover?	2	**1.83**	2.67
The objects were oriented in a way that was comfortable for me when grasping.	2.5	**2**	**2**
I would orient the objects similarly when performing handovers.	3.33	**2**	**2**
I agree with the way the objects were oriented.	3.5	**1.83**	**1.83**
How well would you say you understood the handover process?	1.5	**1.33**	1.67
How natural did you find the object orientations when you grasped the objects?	3.17	**1.83**	**1.83**
How appropriate did you find the object orientations when you grasped the objects?	3.5	**1.83**	**1.83**

TABLE II

INDICATED USER PREFERENCES FOR HANDOVER METHODS

Random-based	Rule-based	Observation-based
0%	**66.67%**	33.33%

TABLE III

FULL HANDOVER SYSTEM SUCCESS RATE

	Rule-based	Observation-based
Failure to grasp	8.33%	**6.67%**
Failure to gen. trajectory	**0.00%**	15.00%
Success rate	**91.67%**	78.33%
Success rate w. correct orientation	**75.00%**	50.00%

all three methods. Finally, participants selected their preferred overall method. The questions of the questionnaire along with the average scores of the responses can be seen in Table I and the respective user preference are shown in Table II.

B. Robotic task-oriented handover system

A systematic test was conducted to evaluate the end-to-end performance of the robotic handover system using both the rule-based and observation-based orientation methods. For each configuration, the robot performed a total of 60 handovers, attempting ten trials for each of the six household objects: mallet, ladle, spatula, knife, bowl, and mug.

During each handover trial, the robot autonomously executed object grasping, orientation, and handing to a human receiver. Trials were conducted in a controlled lab environment with consistent conditions. The handover success rate was recorded based on whether the human receiver was able to obtain the object. The ability to perform the handover with the appropriate object orientation is reported along with the overall handover success rates in Table III.

A chi-squared test found no significant difference in overall handover success rate between the rule-based (91.7%) and observation-based (78.3%) configurations (p=0.057). This robot setup demonstrated high handover success rates with both methods. However, the rule-based orientation approach enabled significantly higher orientation accuracy (75.0%) compared to the observation-based method (50.0%) while maintaining comparable handover success.

V. DISCUSSION

The user study found that rule-based and observation-based methods produced equally comfortable, appropriate, and human-like handover orientations. This indicates orientations can be generated from object affordances rather than human demonstrations, aligning with previous findings [7].

As Table II shows, the participants of the user study preferred the rule-based and observation-based methods over the random-based method, Interestingly, random-based method would occasionally provide the user with the properly oriented object. Despite that, none of the participants preferred the random-based method, suggesting that orientation consistency can affect user perception and overall handover experience.

When comparing the rule-based and observation-based methods, several findings were evident. Firstly, the rule-based method outperformed its observation-based counterpart. However, both methods exhibited issues related to grasping, primarily due to faulty pose estimation. This suggests the need for either refining the ICP algorithm further or substituting it with another technique. Uniquely for the observation-based method was a trajectory generation failure. While both methods employ XYZ coordinates to define the goal handover pose, their approaches to orientation differ. The observation-based method precisely defines the target orientation, whereas the rule-based method offers flexibility in orientation as long as the given rule is met. This advantage arises because certain rotations do not alter the handle's orientation in relation to the receiver or the container's orthogonality relative to the ground.

Furthermore, while both the rule-based and observation-based methods scored almost identically on the 5-point Likert scale, the rule-based method still shows as the preferred method. We theorize this clear preference relates to the way robot achieved handover orientations computed by the observation-based method, particularly the robot's trajectory.

VI. Conclusion

This work represents the first derivation of handover orientations directly from object affordances treated as pixel-wise segmentation. We introduced and evaluated two affordance-based methods for robot-to-human handover orientation generation. User studies demonstrated a strong preference for the rule-based method, while system testing showed it achieved higher orientation accuracy compared to the observation-based approach. By linking visual affordances to orientation rules, consistent and natural human-like handovers were achieved. The practical viability of the rule-based method was confirmed in a full robot-to-human handover system that achieved a success rate with correct object orientations of 75.0%.

Future work could explore more robust pose estimation methods like dense fusion or point cloud registration for improved accuracy. Additionally, as the current receiver detection approach lacks robustness, predicting the receiver pose through skeleton tracking or wearable movement tracking suit could optimize handover locations. Finally, further training of the AffNet-DR to handle grasping of unseen objects will certainly extend the capabilities of this work and move a step closer to enable seamless HRC in everyday environments.

References

[1] A. Wilcox, J. Kerr, B. Thananjeyan, J. Ichnowski, M. Hwang, S. Paradis, D. Fer, and K. Goldberg, "Learning to localize, grasp, and hand over unmodified surgical needles," in *2022 International Conference on Robotics and Automation (ICRA)*, 2022, pp. 9637–9643.

[2] S. Hjorth, J. Lachner, S. Stramigioli, O. Madsen, and D. Chrysostomou, "An energy-based approach for the integration of collaborative redundant robots in restricted work environments," in *2020 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*, 2020, pp. 7152–7158.

[3] C. Li, J. Park, H. Kim, and D. Chrysostomou, "How can i help you? an intelligent virtual assistant for industrial robots," in *Companion of the 2021 ACM/IEEE International Conference on Human-Robot Interaction*, ser. HRI '21 Companion. New York, NY, USA: Association for Computing Machinery, 2021, p. 220–224. [Online]. Available: https://doi.org/10.1145/3434074.3447163

[4] C. Li, A. K. Hansen, D. Chrysostomou, S. Bøgh, and O. Madsen, "Bringing a natural language-enabled virtual assistant to industrial mobile robots for learning, training and assistance of manufacturing tasks," in *2022 IEEE/SICE International Symposium on System Integration (SII)*, 2022, pp. 238–243.

[5] V. Ortenzi, A. Cosgun, T. Pardi, W. P. Chan, E. Croft, and D. Kulić, "Object handovers: A review for robotics," *IEEE Transactions on Robotics*, vol. 37, no. 6, pp. 1855–1873, 2021.

[6] W. P. Chan, M. K. Pan, E. A. Croft, and M. Inaba, "Characterization of handover orientations used by humans for efficient robot to human handovers," in *2015 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*, 2015, pp. 1–6.

[7] ——, "An affordance and distance minimization based method for computing object orientations for robot human handovers," *International Journal of Social Robotics*, vol. 12, pp. 143–162, 2020.

[8] R. Ye, W. Xu, Z. Xue, T. Tang, Y. Wang, and C. Lu, "H2o: A benchmark for visual human-human object handover analysis," in *2021 IEEE/CVF International Conference on Computer Vision (ICCV)*. Los Alamitos, CA, USA: IEEE Computer Society, oct 2021, pp. 15 742–15 751. [Online]. Available: https://doi.ieeecomputersociety.org/10.1109/ICCV48922.2021.01547

[9] V. Ortenzi, F. Cini, T. Pardi, N. Marturi, R. Stolkin, P. Corke, and M. Controzzi, "The grasp strategy of a robot passer influences performance and quality of the robot-human object handover," *Frontiers in Robotics and AI*, vol. 7, 2020. [Online]. Available: https://www.frontiersin.org/articles/10.3389/frobt.2020.542406

[10] M. Cakmak, S. S. Srinivasa, M. K. Lee, J. Forlizzi, and S. Kiesler, "Human preferences for robot-human hand-over configurations," in *2011 IEEE/RSJ International Conference on Intelligent Robots and Systems*, 2011, pp. 1986–1993.

[11] J. Aleotti, V. Micelli, and S. Caselli, "Comfortable robot to human object hand-over," in *2012 IEEE RO-MAN: The 21st IEEE International Symposium on Robot and Human Interactive Communication*. IEEE, 2012, pp. 771–776.

[12] Y.-W. Chao, C. Paxton, Y. Xiang, W. Yang, B. Sundaralingam, T. Chen, A. Murali, M. Cakmak, and D. Fox, "Handoversim: A simulation framework and benchmark for human-to-robot object handovers," in *2022 International Conference on Robotics and Automation (ICRA)*, 2022, pp. 6941–6947.

[13] R. Newbury, A. Cosgun, T. Crowley-Davis, W. P. Chan, T. Drummond, and E. A. Croft, "Visualizing robot intent for object handovers with augmented reality," in *2022 31st IEEE International Conference on Robot and Human Interactive Communication (RO-MAN)*, 2022, pp. 1264–1270.

[14] A. D. Christensen, D. Lehotsky, M. W. Jørgensen, and D. Chrysostomou, "Learning to segment object affordances on synthetic data for task-oriented robotic handovers," in *The 33rd British Machine Vision Conference*. British Machine Vision Association, 2022.

[15] G. Humphreys, "Objects, affordances… action!" *The Psychologist*, 2001.

[16] J. Aleotti, V. Micelli, and S. Caselli, "An affordance sensitive system for robot to human object handover," *International Journal of Social Robotics*, vol. 6, no. 4, pp. 653–666, 2014. [Online]. Available: https://doi.org/10.1007/s12369-014-0241-3

[17] S. H. Creem and D. R. Proffitt, "Grasping objects by their handles: a necessary interaction between cognition and action." *Journal of experimental psychology: Human Perception and Performance*, vol. 27, no. 1, p. 218, 2001.

[18] W. P. Chan, Y. Kakiuchi, K. Okada, and M. Inaba, "Determining proper grasp configurations for handovers through observation of object movement patterns and inter-object interactions during usage," in *2014 IEEE/RSJ International Conference on Intelligent Robots and Systems*. IEEE, 2014, pp. 1355–1360.

[19] H. Razali and Y. Demiris, "Multitask variational autoencoding of human-to-human object handover," in *2021 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*. IEEE, 2021, pp. 7315–7320.

[20] P. Ardón, M. E. Cabrera, E. Pairet, R. P. Petrick, S. Ramamoorthy, K. S. Lohan, and M. Cakmak, "Affordance-aware handovers with human arm mobility constraints," *IEEE Robotics and Automation Letters*, vol. 6, no. 2, pp. 3136–3143, 2021.

[21] A. Myers, C. L. Teo, C. Fermüller, and Y. Aloimonos, "Affordance detection of tool parts from geometric features," in *2015 IEEE International Conference on Robotics and Automation (ICRA)*. IEEE, 2015, pp. 1374–1381.

[22] R. Hartley, J. Trumpf, Y. Dai, and H. Li, "Rotation averaging," *International journal of computer vision*, vol. 103, pp. 267–305, 2013.

[23] M. Ray and T. N. Welsh, "Response selection during a joint action task," *Journal of Motor Behavior*, vol. 43, no. 4, pp. 329–332, 2011.

[24] F. Cini, V. Ortenzi, P. Corke, and M. Controzzi, "On the choice of grasp type and location when handing over an object," *Science Robotics*, vol. 4, no. 27, p. eaau9757, 2019.

[25] C. Hansen, P. Arambel, K. Ben Mansour, V. Perdereau, and F. Marin, "Human–human handover tasks and how distance and object mass matter," *Perceptual and motor skills*, vol. 124, no. 1, pp. 182–199, 2017.