Biomimetic Intelligence and Robotics 4 (2024) 100145

Contents lists available at ScienceDirect

Biomimetic Intelligence and Robotics

journal homepage: www.elsevier.com/locate/birob

Review

Human–robot object handover: Recent progress and future direction

Haonan Duan [a,b], Yifan Yang [a,b], Daheng Li [a,b], Peng Wang [a,b,c,*]

[a] *Institute of Automation, Chinese Academy of Sciences, Beijing 100190, China*
[b] *School of Artificial Intelligence, University of Chinese Academy of Sciences, Beijing 100049, China*
[c] *Centre for Artificial Intelligence and Robotics, Hong Kong Institute of Science and Innovation, Chinese Academy of Sciences, Hong Kong 999077, China*

ARTICLE INFO

Article history:
Received 28 November 2023
Revised 4 January 2024
Accepted 21 January 2024
Available online 2 February 2024

Keywords:
Human–robot interactions
Object handover

ABSTRACT

Human–robot object handover is one of the most primitive and crucial capabilities in human–robot collaboration. It is of great significance to promote robots to truly enter human production and life scenarios and serve human in numerous tasks. Remarkable progressions in the field of human–robot object handover have been made by researchers. This article reviews the recent literature on human–robot object handover. To this end, we summarize the results from multiple dimensions, from the role played by the robot (receiver or giver), to the end-effector of the robot (parallel-jaw gripper or multi-finger hand), to the robot abilities (grasp strategy or motion planning). We also implement a human–robot object handover system for anthropomorphic hand to verify human–robot object handover pipeline. This review aims to provide researchers and developers with a guideline for designing human–robot object handover methods.

1. Introduction

With the continuous progress of modern society and science, human production and life have entered the era of big data with the improvement and development of software and hardware technology. Driven by the boom of chip manufacturing, the assistance of software design, and the accumulation of natural disciplines, applications such as large-scale computing, high-performance computing, and cloud computing are changing with each passing day, which in turn is promoting the development of the field of artificial intelligence. Intelligent scientific research such as computer vision and natural language processing not only brings convenience to humans by deploying their algorithms in real-world scenarios, but also provides new directions for the development of intelligent robots [1]. As the "pearl" in the field of advanced manufacturing and the key development direction of the Industry 4.0 revolution, intelligent robots have demonstrated their important research value and application potential in various scenarios in daily life and industrial production. In the past year, the proposal and development of the concepts of embodied artificial intelligence (Embodied AI) and humanoid robots have also made researchers and the public expect robots with universal forms to serve human production and life.

In recent years, cooperation between humans and robots has developed in a closer and more direct way [2–6]. Industry 4.0 envisions a fully shared environment, where robots are driven by advances in robotic hardware technology [7] and able to interact with their surroundings and other agents [8,9]. Human–robot interaction (HRI) has occupied a prominent position in the robot development strategies of various countries. The core advantage of human–robot coexistence teams is assigning humans to focus on more complex tasks, while transfer repetitive and low-skill tasks to robotic assistants. Such successful deployment is beneficial for both humans and robots. The structured nature of traditional industrial environments promotes the use of robots in such scenarios. However, there have been no similar successful applications in unstructured environments (homes, hospitals, etc.). For such an environment, the robot needs to better understand the tasks to be performed, which in turn requires a powerful perception system to detect and track changes in the surrounding dynamic environment, as well as an intelligent action decision-making and motion planning system [10].

Enabling robots to complete object handover with humans is the basic condition and key technology for successfully realizing human–robot interaction and collaboration. As shown in Fig. 1, the action of transferring objects is called object handover. Object handover is defined as the continuous joint action of a giver transferring an object to a receiver. Such frequent collaborative action between humans requires the joint efforts of both parties in the abilities of prediction, perception, action and adjustment. Achieving human–robot object handover as efficient and smooth as it between humans is a major challenge for the development of intelligent robots.

* Corresponding author.
E-mail address: peng_wang@ia.ac.cn (P. Wang).

https://doi.org/10.1016/j.birob.2024.100145

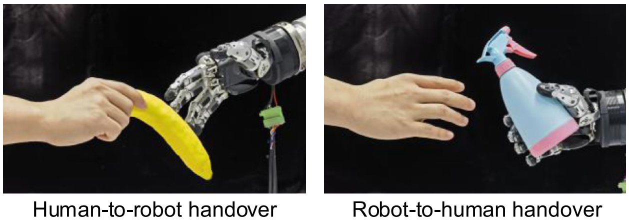

Fig. 1. Human-Robot object handover.

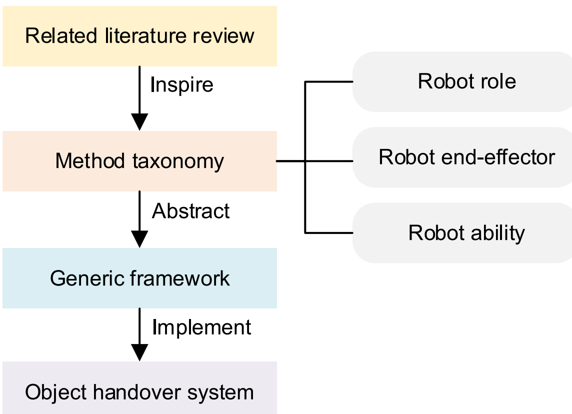

Fig. 2. The construction of this paper.

In this paper, we investigate methods for human–robot object handover mainly after 2021. In order to give readers a more comprehensive perspective on the different approaches, we summarize literature from different dimensions to form method taxonomy. We first focus on the role of robots at the macro level, as receiver taking objects from humans, or as giver delivering objects to humans. We then discuss different end-effectors used by robots, parallel-jaw grippers or multi-finger hands. Finally, the methods to improve the robot grasp strategy or motion planning in object handover are analyzed. We state that the object handover method depends on diverse sensory information or devices, so the perception type is not regarded as a classification dimension. Based on the outlined method pipeline, a general framework is abstracted and an object handover system is implemented for verification. The main contributions can be summarized as:

- We review the human–robot handover literature mainly after 2021, and propose a method taxonomy from three dimensions for comprehensive description.
- We abstract a general framework for human–robot handover to help researchers and developers construct their own human–robot interaction system in an easier way.
- We implement a human–robot object handover system for anthropomorphic hand to verify our framework and list potential directions in the future.

As shown in Fig. 2, this paper is organized as follows. Related literature reviews are outlined in Section 2. Such works inspire us to formulate the method taxonomy from three dimensions in Section 3. Sections 4–6 describe human–robot object handover studies after 2021 in detail. Section 7.1 abstracts a generic framework based on investigated works. A human–robot object handover system is implemented by following the proposed framework in Section 7.2. Section 8 provides the insights of potential development in the future. Section 9 concludes this review.

2. Previous surveys

There are some previous work survey human–robot object handover methods. A complete review of human–robot object handover methods before 2021 is presented in [2], with the perspective of the object handover process. Literatures are divided into pre-handover stage and physical handover stage. Inspired by [2], human–robot collaborations in industry scenario are discussed in [11], with special concentration on communication channels between human and robot. [12,13] focus on visual aspect of the handover to dig out how robot vision promote and affect the success of human–robot handover. [13] further deliberates proactive human–robot collaboration [14] boosted by visual perception. Learning-based methods are summarized in [15], which uncovers the mechanism of building cognitive model and behavioral block based on machine learning [16] in human–robot collaboration. Except from vision, those approaches enable robot to execute smooth motions learned by human demonstrations are outlined in [17]. [18] aims to impartially evaluate the research around human factors that are involved in studies concerning physical interactions and robust manipulation capabilities. They pinpoint the dominant human factors in physical human–robot interaction (pHRI), recognize the elements often tackled in correlation, and establish the commonly used assessment methodologies.

In addition, there are some articles comparing various methods. [19] distinguishes the effects of adaptive and non-adaptive transport methods. They find the adaptive transport approach does not yield substantially longer average physical handover durations compared to the non-adaptive transport approach. The non-adaptive transport method does not prompt a significantly sooner handover intervention during the runs compared to its adaptive counterpart. The adaptive transport method is associated with significantly lower ratings for trust in the robot and perceived safety compared to the non-adaptive transport method. [20] contrasts four trajectory generation methods for robot-to-human object handover scenario. [21] engineers and conducts a comprehensive evaluation of a pair of controllers designed for human–robot handover interactions. These systems are constructed with the capacity to allow end-users to delineate timing parameters for the robot reach motion and deliver feedback in instances where the robotic system is unable to fulfill the established constraints.

Different from the above, this paper focuses on reviewing the human–robot object handover literature after 2021. Due to the rapid changes in the current needs and scenarios of human–robot collaboration, old methods may no longer be suitable for future applications of robots. We adopt the classification in [2], but the method taxonomy we propose characterizes each method along three dimensions. This review aims to provide researchers and developers with a guideline for designing human–robot handover or collaboration methods and systems.

3. Method taxonomy

Our method taxonomy for human–robot object handover is illustrated in Fig. 3. We classify the existing literature from three dimensions and focus on robot ontology and capabilities, rather than from the perspective of deep learning methods. We observe that in the field of robotics, methods differences largely depend on the agent and task, our taxonomy describes human–robot object handover research in a more comprehensive fashion.

Robot Role. From a macro perspective, robots can play two roles in human–robot object handover tasks, receiver and giver. Specifically, in human-to-robot object handover, the robot acts as a receiver and takes the object from the human hand, and vice versa. Different roles bring various problems and challenges.

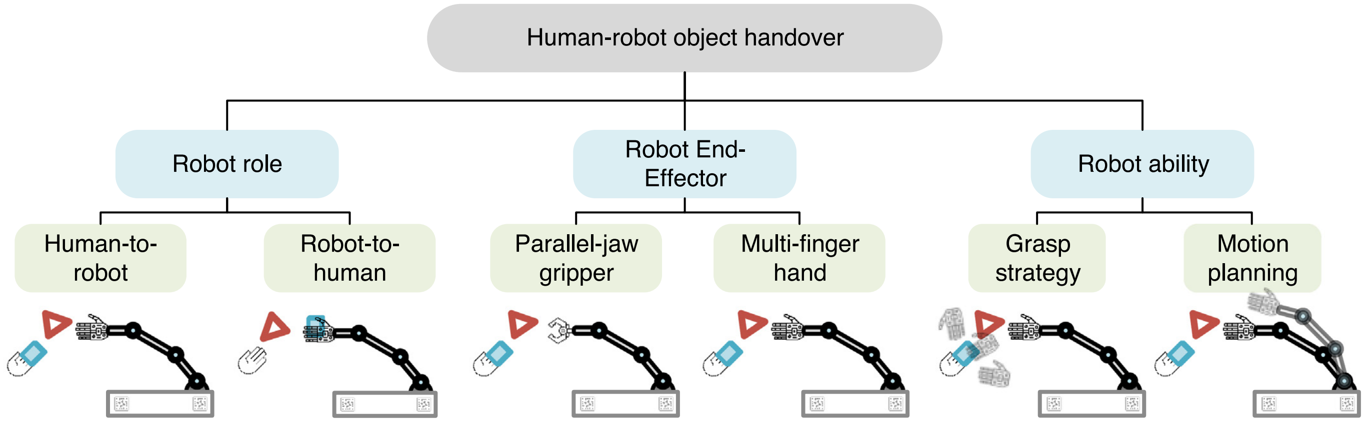

Fig. 3. Method taxonomy for human–robot object handover.

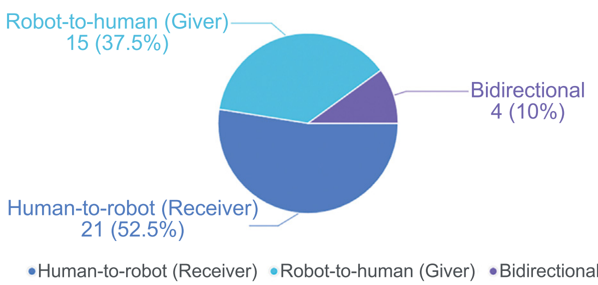

Fig. 4. Robot role distribution in human–robot object handover.

Robot End-effector. With the development of software and hardware technology, the morphology of robot end effectors has evolved from low-DOF parallel-jaw grippers to high-DOF anthropomorphic hands. While parallel-jaw grippers are favored by researchers due to simple mechanical structure and low collision probability, there are still certain studies on utilizing multi-finger hand to complete human–robot interaction.

Robot Ability. The process of human–robot object handover can be divided into two stages: pre-handover and physical handover. For the physical handover stage, the robot capabilities are reflected in grasping force adjustment and failure handling. However, the robustness of grasp strategy and motion planning in the pre-handover stage directly affect the success of object handover. Therefore, most research work focuses on such two abilities.

4. Robot role

The proportion of the role of robots in human–robot object handover tasks in the surveyed research work are shown in Fig. 4. Overall, more than half of the work focus on making the robot act as a receiver to complete the task of human-to-robot handover. In the remaining work, in addition to using robot as regular givers, some researchers are committed to solving the robot-to-robot object handover problem. Only one work explored bidirectional handover scenario. We describe research work in this section mainly from the perspective of scene settings or task scenarios.

4.1. Human-to-robot handover (H2R)

Indeed, the role of robots as receivers in H2R object handover tasks is a thorny and incompletely explored problem. Specifically, robots may face the following challenges:

- Human behavior uncertainty: human may behave very differently when handing over objects, such as speed, orientation, force, etc. Robots need to be able to adapt to these uncertainties.
- Real-time response: robots need to have the ability to respond to human motion. This involves efficient sensor data processing and fast action generation.
- Grasping strategy: robot needs to generate feasible and collision-free grasp to safely take the object occluded by human hand.

To this end, researchers have also mainly explored the above difficulties. Some work train robot object handover policies in simulation environments. [22] builds a pseudo-robot arm as a human arm to guide a 7-DOF Franka robot arm to complete the training of reinforcement learning policy [23] for motion control. In contrast, [24] applies a multi-agent generative adversarial imitation learning to directly learn human–human demonstrations collected in the real world. Such method enables robot to adapt to diverse objects in the handover process. However, it is difficult to achieve sim-to-real transfer for the satisfactory policy trained in the simulation environment due to challenging acquisition of fully-observed state in the real world. Therefore, most work try to implement its ideas on real robots.

We divide human–robot object handover scenarios based on the selected object number. Specifically, those works that use only a single or one class of objects demonstrate the effectiveness of their methods in specific scenarios, while those test multiple classes of objects model real human–robot interaction scenarios in the future.

4.1.1. Single object

[25] feeds human hand RGB-D information into a long short-term memory (LSTM) [26] network to predict final handover position. [27] generates predictions that encapsulate how future human motion could influence the cost of a robot strategy to receive the object from human giver. In order to eliminate the error caused by visual perception, [28,29] rely on inertial measurement unit (IMU)-type sensors to obtain high-precision human hand poses in real time, thereby providing input for the proposed method of generating flexible motion of the robot arm. Differently, other works discuss more specific handover tasks or scenarios. [30–32] construct a scene in which a human hands a cup to a robot, and the amount of water in the cup is unknown. The robot needs to determine whether the cup is empty or filled with water through human behavior, and plan its motion from human arm movement. [33] focuses on human-to-cobot dualarm handover operations for large box-type objects. [34] introduces a

prediction-planning pipeline for preemptive human robot placement handovers in a indirect paradigm. Although these work promote the ability of robots to interact with human in specific scenarios, they are difficult to meet the requirements of real-world human–robot collaboration due to the lack of generalizability.

4.1.2. Diverse objects

In order to restore the real human–robot object handover scene, some literature is dedicated to testing methods on diverse objects. Compared with using a single object, such task further requires the robot generalization ability. By employing single-view visual perception, [35–40] improve smoothness and flexibility of robot arm motion. Different grasp networks are adopted to guarantee robot grasp generalizability. [41] imagines a scenario where a mobile robot cares for patients in the ward, and enables the robot to choose reasonable grasp to take over the object from the patient hand. From the perspective of user privacy, a federated learning scheme is used in [42] to distributedly train the grasp network on multiple clients, making robot adapt to objects with various shapes and sizes. In the hardware setup of the above work, the parallel-jaw gripper is mounted in the tool center point (TCP) of the robot arm as the end effector. However, few works have explored the object handover scenario on multi-finger hand robots, and the high degree of freedom of planning space makes research in this direction predictably more challenging. [43] inherit the idea of learning-based grasp generation, and propose an end-to-end grasp detection network to predict diverse and dense grasp proposals on object point clouds based with 5 predefined grasp taxonomies. This work empowers the multi-finger hand robot grasp generalizability, and combines the mechanical structural characteristics of the multi-finger hand to boost the diversity of grasp configurations. [44] relies on tactile gloves worn on human hand to obtain a classification of the rigid or soft properties of objects before transferring it, allowing multi-finger hand to perform predefined torque-based grasp for objects of different physical properties.

4.2. Robot-to-human handover (R2H)

Compared with human-to-robot handover, robots need to learn dissimilar ability when acting as givers in robot-to-human handover:

- Human comfort: robots need to understand and follow human comfort and expectations, such as where, how quickly, and in which direction objects should be handed to humans.
- Initiative: robots need to be more proactive in determining the timing and method of handover, which may require complex predictions and decisions.
- Release strategy: robots needs to determine when and how to release the object to ensure that the human can safely and comfortably receive the object.

We describe the scenarios targeted by research work into two categories, namely general scenarios and specific tasks.

4.2.1. General scenario

A basic robot-to-human handover is tested in [45] to enhance smoothness of robot motion. [46,47] set up two handover scenarios to verify their reactive object handover method. The straightforward handover is the user approaching and grasping the object from the robot giver. Unlikely, the user is asked to perform a secondary task before engaging in the handover in the perturbed handover. With the same thoughts, [48] treats human and robot as a cohesive entity to simultaneously improve the experience of both during the object handover process. [49–53] explore how robot can grasp objects with the ability of detecting affordances to promote human receiver convenience. Human body poses are predicted in [54] to help robot react human and deliver objects in a proper location. The above studies are all aimed at improving the capabilities of robots in the pre-handover stage.

4.2.2. Specific task

There are some work develops object handover systems for specific tasks, providing inspiration for the deployment of robots in certain operating environments. [55] targets robot to adaptively execute the task of transferring a box to its human counterpart stands on a ladder, by taking into consideration the preferred interaction style of the human partner. Similarly, the small box transfer is conducted in [56]. [57] presents a human–robot collaborative assembly system, which enables robot to recognize human action and deliver necessary object to provide an intuitive experience for the human. [58] asks robot to serve older adults in a nursing room. The handover transfer point should guarantee manipulability and limit the effects of gravity forces. [59] estimates the position of the human hand through a smartwatch and a smartphone on a human partner to provide a target for robot motion control.

4.3. Bidirectional handover

Bidirectional object handover refers to the robot ability to act as both a receiver and a giver, which remains a open challenge. Specifically, when a robot only acts as a receiver, it can usually acquire the status of human hand and object through a single visual perception device. However, when acting as a giver, the initial position of the object may be far away from the human hand, which may place more requirements on the performance, quantity, and installation angle of the sensors, resulting in different hardware settings for the two roles. Despite this, there are some studies probe into bidirectional handover on specific task scenarios.

[60] establishes a human body comfort model and introduce human intention recognition to predict object transfer points in H2R and R2H handovers. [61] facilitates the robot capability to interactively receive or give a large planar object utilizing a vertical grasp posture with human partner. A supermarket-like environment is set in [62], where experiments involving human–robot interaction exclusively through haptic signals are conducted. The findings underscore the significance of force and tactile feedback in successfully executing handover procedures within a collaborative task. [63] innovatively utilizes an on-shoulder supernumerary robotic limb. They consider a scenario involves a situation where a human user requires a tool that is beyond their reach. In this context, the robot is programmed to fetch and deliver the tool to the human user, and subsequently replace the tool once the user has completed their work.

5. Robot end-effector

Robots consist of end-effectors and robot arms. For end-effectors, they are divided into vacuum suction cups, parallel-jaw

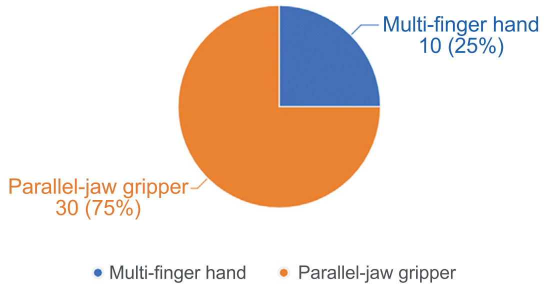

Fig. 5. Robot end-effector distribution in human–robot object handover.

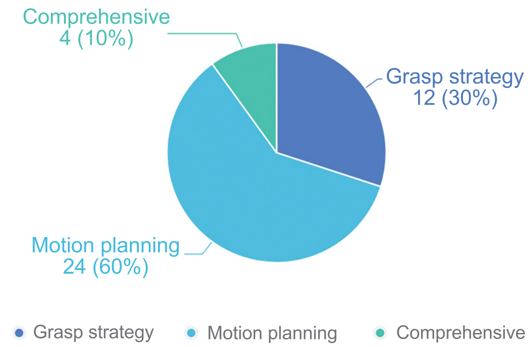

Fig. 6. Robot ability distribution in human–robot object handover.

grippers, and multi-finger hands. Vacuum suction cups are often used in industrial scenarios to execute pick-and-place task for heavy objects, yet not employed in human–robot collaboration tasks. The distinction between parallel-jaw grippers and multi-finger hands is reflected in the degree of freedom, planning space, ease of use and manipulability. As shown in Fig. 5, most current research is based on parallel-jaw gripper robots. As for the robot arms, they are categorized into limited DOF, full DOF and redundant DOF, which are determined by whether its DOF is less than, equal to and greater than 6. For limited DOF robot arms, they are theoretically unable to reach any position in 3D space, resulting in its adoption being limited to certain research focused on specific scenarios. Both of the latter two have the ability to achieve 6D poses. Compared with full DOF robot arms, the obstacle avoidance and singularity handling capabilities of redundant DOF robot arms are relatively better. We find that even if researchers chose parallel-jaw grippers, there exists apparent divergence in the choice of robot arms. This section reports on robot morphology from research efforts, which we span two categories based on end-effectors. The robot arms are detailed in each category.

5.1. Parallel-jaw gripper

Parallel-jaw grippers are widely used in current research due to their simple mechanical structure and degree of freedom. We observe that while leveraging simple end-effectors, researchers prefer to select redundant DOF robot arms, which is probably related to the fact that most work is oriented towards motion planning (detailed in Section 6.2). We subdivide this subsection into 3 subcategories according to the classification of robot arms.

Redundant DOF Robot Arm. [22,27–29,35–38,40,48] all use Franka Emika robot arm. Thanks to its ease of use and rich open source projects, Franka Emika is one of the most popular robot arms in the robotic research community. [30,31] employ Kinova Gen3 and [39] uses a Flexiv Rizon. Kuka LBR/LWR robot arms are adopted in [56,62]. The Baxter robot is a bimanual robot with 7 DOFs for each arm. Since its inception, it has been favored by a group of researchers. In the human–robot handover task, it is applied in [33,55,61].

Full DOF Robot Arm. [32,57,59] use UR5, the most famous collaborative robot arm. [34,51,63] adopt Rethink Robotics Sawyer, Moying Bimanual Robot and Unitree Z1 robot arm respectively.

Limited DOF Robot Arm. The remaining three studies all employ limited DOF robot arms. We notice that the main goal of these works are grasp strategies (detailed in Section 6.1), the accessibility of the robot arm in any position in 3D space is not vital for the author. Specifically, [49] utilizes a 5-DOF Kuka youBot, [41,42] utilize a 5-DOF Trossen Robotics ViperX 300 and [50] utilizes a 4-DOF PR2 Bimanual Robot. [45] validates their method on Toyota Human Support Robot (HSR), which has a single 4-DOF arm.

5.2. Multi-finger hand

The studies using multi-finger hands can be subdivided according to the above criteria as well. We also describe the multi-finger hand used in each literature.

Redundant DOF Robot Arm. [44] adopts Franka Emika robot arm with an Allegro Hand. The combination of Kuka LBR iiwa 7 and Robotiq 3-finger gripper is used in [52]. By equipping an anthropomorphic hand on 7-DOF arm, IVO humanoid robot [64] is used in [54]. [25] employs Talos Humanoid Robot with 7 DOFs for each arm. The end-effector of Talos Humanoid Robot is a 2-DOF 3-finger gripper.

Full DOF Robot Arm. UR5 is utilized in [43,46,47,53]. [43] evaluates their algorithms on HIT-DLR II Dexterous Hand. [46,47, 53] use IH2 Azzurra Hand (Prensilia SRL) and Robotiq 3-finger gripper respectively.

Limited DOF Robot Arm. Only [58] applies Pepper Robot with 5-DOF arm and 1-DOF hand. They set up a structured task scenario and test with simple object.

6. Robot ability

Robot ability, in the context of human–robot object handover, encompasses various capabilities that enable successful interaction between humans and robots. The handover process can be divided into two distinct stages: the pre-handover stage and the physical handover stage. While the physical handover stage primarily relies on the robot grasping force adjustment and failure handling abilities, the robustness of the grasp strategy and motion planning during the pre-handover stage significantly influence the overall success of object handover. Consequently, current research in this field predominantly focuses on enhancing these two key abilities, aiming to improve the efficiency and reliability of human–robot object handover.

As shown in Fig. 6, improving the motion planning capabilities of robots occupies the mainstream in human–robot object handover study. A substantial portion of the effort is dedicated to enhancing the adaptability and initiative of robot arm movements. Concurrently, pertaining to grasping, such literature endeavors to grasp the object by employing simple approaches. Undeniably, robot arm movements carry significant weight. They profoundly affect the perceptual experiences of human participants, particularly their sense of impending danger. Nevertheless, the rudimentary grasp strategies are unmatched to the fluid robot arm movements. In addition, object handover proficiency should account for more than the mere act of transfer — it should also consider subsequent manipulation or tool-use of objects in post-handover.

We survey some works that simultaneously attempt to solve the above two aspects of the problem. [55,61,63] design specific application scenarios and devise task-specific controllers for the robot to accomplish grasp and motion. As for [40], it adopts RL and trains the policy for grasp and motion generation at the same time. However, the vast majority of work remains predominantly concentrated on enhancing a single specific capability in order to optimize the performance of object handover systems.

6.1. Grasp strategy

Grasp strategy denotes that these studies accentuate the generation of viable 6D poses and joint angles for the end-effector. Within these investigations, robots predominantly function as givers, equipped with parallel-jaw grippers. This is attributable to the fact that task difficulty diminishes when employing simple grippers in unobstructed scenarios. Grasp strategy research can be bifurcated into two main categories: grasp part detection and grasp configuration generation. The distinction lies in that the former tends to analyze which part of the object is suitable for grasping, while the latter is inclined towards the generation of an array of grasp candidates for selection. Although the first method offers a more profound comprehension of object characteristics, the downside is its subpar generalization capability for unseen objects. Conversely, the latter falls under the domain of generic grasping [65] and exhibits superior adaptability.

6.1.1. Grasp part detection

[49,50,52,53] focus on enabling robots to comprehend the object affordance during delivering it, thereby elevating the interaction beyond mere object pick and transfer. By adopting the thoughts of learning from human demonstrations, [49] develops a task-oriented handover system that facilitates robot being aware of tool-affordances and their subsequent application. Analogously, to extend handover behaviors to unseen objects, [50] proposes a heuristic-guided, hierarchically optimized cost structure, the optimization of which adapts object configurations for human with restricted arm mobility. [53] modifies Mask R-CNN [66] and trains the network on synthetic dataset to segment object affordance. Based on [52,53] further explores the object deliver orientation method with being aware of object affordances. They compare techniques learned from human–human handovers with rule-based approaches and find that human partners favor the performance of the latter one. [57] builds an assembly system based on human action recognition to grasp proper parts for human. After generating affordance map for handover scene by utilizing [41,67] formulates the grasp choice as Markov Decision Process (MDP) [68] and adopts Double Q-learning (DDQN) [69] to obtain preferred grasp on reasonable grasp point. [32] constructs a scenario of human giver delivering container with uncertain amount of water in it to robot. Based on visual perception, their method estimates the physical properties of the container and extracts feasible grasp region to take over the object.

6.1.2. Grasp configuration generation

[43] generates diverse grasps with five taxonomies for anthropomorphic hand based on PointNet++ [70], which strives to enhance grasp generalizability when receiving objects from human hand. [51] extracts the pose of the human hand and object through various visual modules, and determines the grasp configuration through analytical ways. [42] follows federated learning scheme to train its grasp network to protect user privacy. Their network takes multi-view images as input and predict 6-DOF poses for the end-effector. [44] considers classifying soft/rigid object with tactile glove before delivering it to robot. The robot perform predefined configurations for soft/rigid to stably grasp it without damage. [62] designs a visual servoing controller and a grip force controller to obtain grasp configuration.

6.2. Motion planning

Compared to the relatively immature stage of works centered on grasp strategy, the field of motion planning is blossoming. Explicitly, endeavors in motion planning aim to make robots safer, smoother, and more proactive in object handover process. [22,24] evaluate their approaches in simulation, however, facing the challenging of sim-to-real transfer. We categorize the studies experimented on real robots into three types: those based on learning, those reliant on control, and those grounded in analysis.

6.2.1. Learning-based motion planning

Currently, the learning-based methods can be further divided into supervised learning (SL) or reinforcement learning (RL) approaches. [25,39] adopt the paradigm of temporal models, training on their constructed datasets to generate the motion trajectory of the robot arm with RGB-D images as input. Precisely, [39] integrate target-invariance thoughts [71] and Transformer [72] to generate robust and flexible trajectories. Similar to [25,73] adopts LSTM to predict handover position. [34] encodes human pose and gaze in a RNN-based network [74] and decodes probabilistic heatmap of object transfer point. [54] applies motion attention [75] to predict human body pose to promote safe and smooth R2H handover. Based on sensory data from IMU and smartwatch, [59] incorporates Differentiable Ensemble Kalman Filter (DEnKF) [76] to facilitate the attainment of equilibrium between less-restricted movements and the realization of stable and effective pose estimations that are conducive for human–robot collaboration. In contrast, [37,38] utilize Twin Delayed Deep Deterministic policy gradient algorithm (TD3) [77], a powerful actor–critic RL methods to accomplish object handover. Notably, [37] extends the ideas of [36], replacing model predictive control (MPC) [78] with RL, thereby enhancing the motion capability of the robot arm.

6.2.2. Control-based motion planning

There are some works that favor control methods to tackle the task of human–robot object handover. For instance, [29–31,45] model the handover system and compute optimal hyperparameters for PID-type controllers. [27,36] go a further step, employing MPC to perform reactive H2R object handover. It is worth mentioning that although [36–38] primarily focus on motion planning, they both use the grasp generation network from their previous work [79], hence exhibiting the grasp generalizability for diverse objects.

6.2.3. Analysis-based motion planning

[35] obtains reach cost and selects reasonable grasp target frame by frame to guarantee temporal consistency. A human comfort model and transfer intention recognition model are established in [58,60], which are applied to calculate object transfer point. This aspect incorporates the target 6D pose of the end-effector. Within the analysis-based motion planning methods, dynamic movement primitives (DMPs) [80] are widely used to enhance robot performance in human–robot object handovers. Given a small amount of human demonstration, DMPs can generate flexible motion trajectories and generalize to dynamic changes in the target position. [28,33,46,47,56,81] adopt DMPs and their variations to obtain the robot arm motion. Due to their adaptability to various tasks, DMPs become the first choice when learning-based methods are difficult to applied.

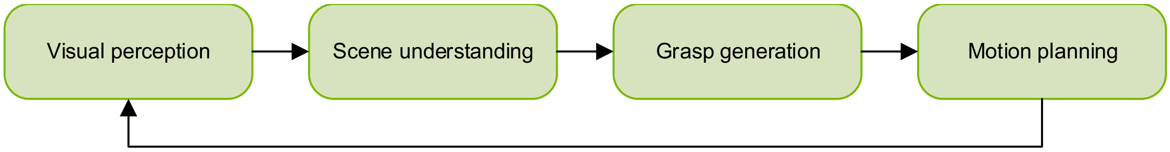

Fig. 7. Generic human–robot object handover framework.

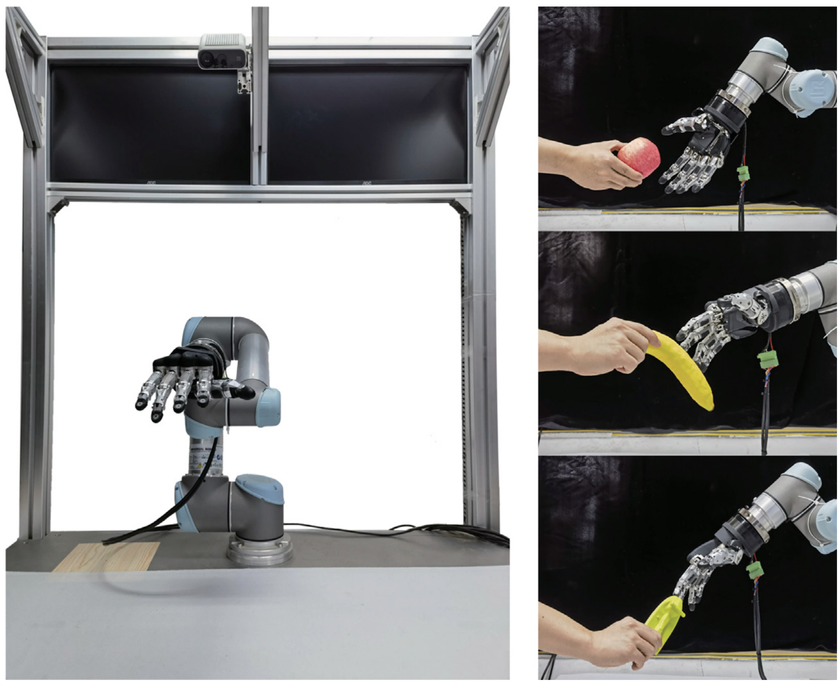

Fig. 8. Human-to-robot handover platform by anthropomorphic hand.

7. Generic framework and implementation for anthropomorphic hand robot

In this section, we first abstract a generic method framework for human–robot object handover based on the literature reviewed above. This framework is versatile, accommodating different robotic roles, diverse end effectors, and varying capabilities. Moreover, it can be flexibly scaled up based on the distinct requirements of differing tasks. Subsequently, leveraging this framework, we implement and verify a human–robot object handover system by anthropomorphic hand robot. Our framework can expedite the design of methods for human–robot interaction and collaboration for researchers and developers.

7.1. Generic framework

We pay special attention to [35–40,43], by virtue of achieving generalizability to grasp diverse objects under the premise of simple system hardware requirements, and enabling the robot arm to execute safe and smooth movement. Although these works all solve the H2R object handover problem, their methods and ideas can be extended to more general human–robot object handovers. Fig. 7 illustrates a generic human–robot object handover method framework that we abstract from these studies.

The framework is composed of four modules, and each module can be flexibly expanded as per the needs of the task. We argue that visual information is still the dominant perceptual input in human–robot object handover tasks, so the basic input module is commonly completed by an RGB-D camera. The scene understanding module takes the visual information obtained by upstream as input to extract information of interest in different tasks. The above work apply different methods (human body key point detection, hand-object detection, or point cloud segmentation) to extract point clouds of human hands and objects. This idea generalizes to R2H object handover. The grasp generation module feeds the object point cloud into the point-based grasp network to predict dense grasp candidates. The motion planning module allows the robot to reactively execute object handover.

Based on this framework, researchers and developers can freely integrate more customized modules to serve the needs of their human–robot collaboration systems. Our proposed framework does not place any restrictions on the form and function of the robot, and can fuse multiple modes of perception such as force, touch or audio, thereby reducing the complexity of designing object handover approaches.

7.2. Object handover system implementation for anthropomorphic hand

Combining the method framework elucidated in Section 7.1, we reproduce the H2R object handover approach proposed in [43]. As shown in Fig. 8, our handover platform consists of a 6-DOF UR5 collaborative robot arm, a 9-DOF Schunk SVH Hand and an Azure Kinect RGB-D camera. AHG-Net based on PointNet++ is adopted as the backbone to generate dense grasp configurations by taking single-view object points as input. The predicted hand joints are remapped from HIT-DLR II Hand to Schunk SVH Hand.

In the experiment, we select 30 novel objects and keep the same criteria as in [43] to determine the success of handover. Each object is delivered 10 times with random poses and human

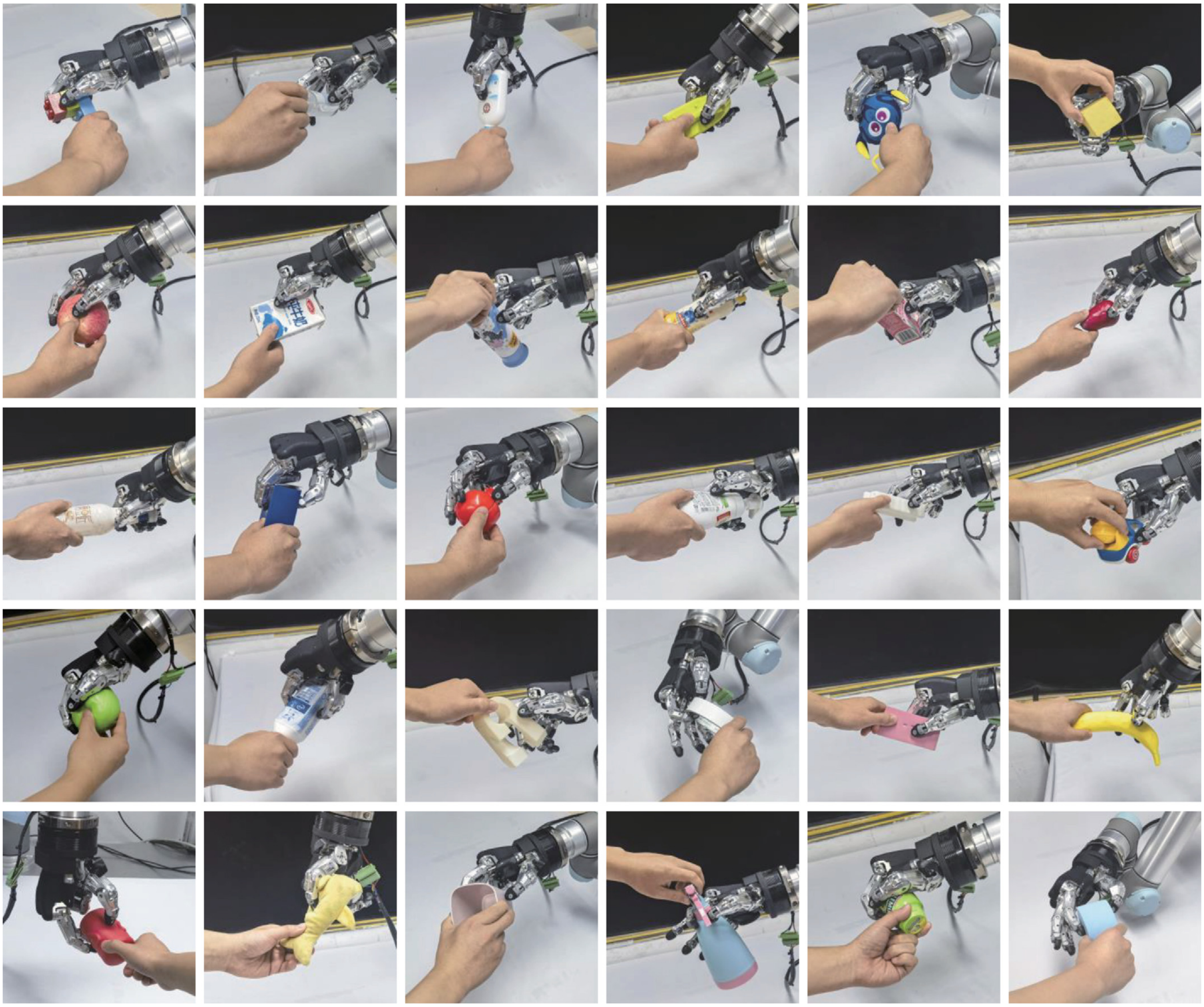

Fig. 9. The qualitative results of our implemented human-to-robot object handover system.

hand occlusion. We only ask the human giver complete the test on the frontal position. The handover time, success rate and number of attempts for each object are listed in Table 1. Overall, our implemented system obtain average success rate of 73.3%. The decrease in success rate is expected because our system only uses a low-precision RGB-D camera to capture the scene, and the joint remapping method is applied to transfer different hand grasp configurations, rather than training a new grasp network with new dataset from scratch. The qualitative results are described in Fig. 9. The robot can not only adapt to diverse objects with various shapes, but also generalize to unpredictable poses of delivered objects occluded by human hands.

Experiment results demonstrate the effectiveness of abstracted framework in Section 7.1, which allow researchers and developers quickly build a human–robot collaboration system by implementing each module with existing methods.

8. Future direction

In this section, we first review some work that does not fall within the scope of proposed method taxonomy. These include robot–robot object handovers, object handover datasets, and those oriented towards the physical handover phase. Inspired by the methodological ideas and implementation approaches of studies after 2021, we put forward some possible future research and development directions.

8.1. Other works

In the field of human–robot object handover, numerous capabilities contribute to the success of this interaction. While research primarily emphasizes improving two abilities in pre-handover stage, there are other aspects that also play a crucial role in enhancing the performance of object handover systems.

One such aspect is robot-to-robot handover (R2R), which explores the transfer of objects between different robot arms or within a single robot hands. Additionally, the creation of comprehensive datasets is essential for advancing human–robot handover research. These datasets provide a valuable resource for training data-driven models and improving the generalization capabilities of handover systems. Furthermore, the physical handover stage, which involves force perception, error handling, and human intention understanding, is another critical aspect.

Table 1
Real-robot experimental results for human-to-robot handover.

Object ID	Times (s)	Success rate	Number of attempts
1	19.11 ± 5.02	80%	1.2
2	19.28 ± 6.89	90%	1.7
3	18.52 ± 7.31	60%	2.0
4	18.90 ± 4.67	70%	1.3
5	20.03 ± 9.46	80%	1.9
6	18.43 ± 4.46	60%	1.6
7	19.89 ± 4.29	70%	1.4
8	18.97 ± 3.69	80%	2.0
9	18.78 ± 3.92	50%	1.1
10	19.66 ± 3.37	90%	1.5
11	18.85 ± 4.88	80%	1.8
12	19.22 ± 2.01	70%	1.3
13	19.15 ± 6.45	60%	2.0
14	19.46 ± 4.51	80%	1.6
15	20.46 ± 7.96	90%	1.4
16	18.91 ± 5.14	70%	1.6
17	19.33 ± 6.12	60%	1.2
18	18.23 ± 7.65	80%	1.8
19	18.83 ± 5.38	70%	1.0
20	19.96 ± 10.48	60%	1.8
21	18.39 ± 4.89	80%	1.6
22	19.78 ± 4.22	90%	1.9
23	18.92 ± 3.54	60%	1.3
24	18.68 ± 3.86	70%	1.7
25	19.58 ± 3.25	80%	1.5
26	18.76 ± 4.18	60%	1.4
27	19.27 ± 1.92	70%	1.9
28	19.18 ± 6.32	80%	1.2
29	19.54 ± 4.77	90%	1.6
30	20.52 ± 7.62	70%	1.5

By considering these different aspects, researchers aim to enhance the adaptability, coordination, and overall performance of human–robot object handover systems.

8.1.1. Robot-to-robot handover (R2R)

Broadly speaking, robot-to-robot object handover includes not only system settings in which objects are transferred between different robot arms, but also scenarios in which a single agent passes an object from one hand to another, known as self-handover. Most research focuses on the scenario where two robot arms cooperate with each other, but the task of self-handover is rarely investigated. This is because the multi-robot arm setup is more in line with the needs of current industrial scenarios, and the low probability of self-collision also reduces the difficulty of algorithm design and development. [73,82–84] evaluate their approaches on two Franka Emika Panda robot arms. [82] equips two eye-in-hand cameras to complete three different industrial parts handover. [73] learns to generate reasonable trajectories from temporal image input. [83] takes attempts to multi-object rearrangement task. [84] tackles the dynamic handover of one robot arm throwing object and another catching it precisely. [81] conducts a hammer self-handover in the experiment. However, the adaptability and coordination capabilities of the robot still need to be improved.

8.1.2. Dataset

Compared with methods based on optimization and analysis, data-driven methods have better generalization capabilities and can effectively handle more corner cases. Data-driven methods typically require large-scale datasets to train a powerful model, and some researchers are committed to building such datasets to promote the progress of related studies. [85] presents H2O, a dataset of 18k video clips involving human-to-human (H2H) object handovers, which is proven to be effective for robot imitation learning on the handover task. Similarly, [86] adopts a markerless approach to capture natural real-world motions and clothing, and employs a multi-camera setup to collect skeletons, fused point clouds, grasp type and handedness labels, object, giver hand, and receiver hand 2D and 3D segmentation, giver and receiver comfort ratings. [87] records bimanual human–human handover with annotation of three handover phase: reach, transfer and retreat. Human preferences are a crucial factor in handover process is considered by [88]. Their dataset contains comfort and transfer point satisfaction ratings. Some other literature generate their datasets for the specific scenarios. [89] arranges three agents (self, other, cobot) to manipulate objects in the kitchen, and investigates how human perceive danger under different situations. [90] constructs a naturalistic collaboration scenario, where a mobile manipulator robot assists a person during a crafting session by providing and retrieving objects used for wooden piece assembly (functional activities) and painting (creative activities). Similar to [32,91] proposes an audio-vision-based dataset and challenge, which estimates object physical properties in five tasks with assorted difficulty levels. [92] continues the cup transferring task in [30,31] and collects the dataset from human–human interaction [93] with eye, head and hand motion data to classify cups with three levels of liquid: empty, half-full, and full of water. [94] discusses how to release grip properly based on the force detection. In addition, [95] proposes a handover benchmark based on PyBullet [96] simulator.

8.1.3. Physical handover

As described in Section 6, the robot ability in the pre-handover stage is indeed the most significant factor related to whether the object handover can be successfully executed. However, in the physical handover stage, the robot force perception and error handling capabilities are undervalued but effective to more precise handover. [97–99] integrate various sensory information to realize reasonable grip-release methods. [93,100–103] concentrate on human intention understanding to enable the robot to execute actions expected by human in advance. The aforementioned methods primarily rely on visual cues to infer human intentions, which are effective when human actions involve significant movement. However, they face challenges in quickly perceiving subtle changes at a finer level. In contrast, electroencephalography (EEG) [104] and electromyography (EMG) [105] offer an alternative perspective for understanding human intention during object handover [106,107]. EEG measures the electrical activity of the brain, while EMG records the electrical signals generated by muscle contractions. These biological signals possess certain advantages that can complement the visual-based approaches mentioned earlier. A handover scenario in which a robot dynamically transfers objects is implemented by [108]. [109] performs an experiment in virtual reality (VR) where a simulated robot offers tools with different degrees of visual occlusion. [110] proposes a method for learning human-to-human handovers observed from motion capture data. Based on multimodal perception, [111] introduces a methodology for fostering collaboration between humans and robots, employing human action recognition, hand gesture recognition, tactile feedback handovers and user personalization. [112] delves into how to cultivate robot to behave and operate in concordance with human user preferences, work habits and task constraints. [113–115] consider psychological and ethical aspects in human–robot handovers, which aim to enhance trust and delight mood for human partner in the collaborative process. [116,117] explore realizing handover not reliant on visual information. Specifically, [116,117] use vibrotactile stimulus and voice respectively. [118] evaluates the handover in aerial case.

8.2. Future directions

We believe that the field of human–robot object handover will gradually develop into a research and application paradigm that is mainly data-driven and supplemented by model guidance. This paradigm shift is expected to become increasingly evident due to the advantages of combining data-driven approaches with model-guided approaches.

The data-driven approach is able to learn from a large number of real handover cases and adapt to a wide variety of objects, human users and environments. This enables robots to provide a more natural and smooth handover experience in future human–robot interactions. In addition, data-driven methods can also improve the accuracy of predictions, allowing robots to more accurately predict and understand the behavior and needs of human users. However, data-driven approaches also have limitations, such as the need for large amounts of high-quality data and performance issues when the data is sparse or unseen. In this case, model-guided approaches can play an important role. Model-guided methods are usually constructed based on physical principles or human behavior patterns and can provide a more general, theoretical understanding. This method can help robots make reasonable predictions and decisions when there is insufficient data or the data cannot be covered.

Taken together, human–robot object handover that is mainly data-driven and supplemented by model guidance will be the main direction of future development, which will bring higher efficiency and better user experience to human–robot interaction. The potential future directions combined with recent trends are listed as follows.

Large Language Model (LLM). With the emergence of ChatGPT, LLM has also attracted the attention of more researchers out of the field of natural language processing [119]. [120,121] successfully accomplish basic human–robot interaction tasks using LLM. However, it is evident that the robot performance is not on par with that of RT-2 [122], Palm-E [123], and other similar systems, which are capable of independently completing tasks within the workspace. This disparity in performance can be attributed to the presence of human partners during the interaction. Integrating LLM's advanced semantic understanding and inference capabilities, and taking human factors into consideration to build a safe and natural human-in-the-loop system for robots can effectively realize more complex object handover tasks.

Multi-finger Hand. Demonstrated by Fig. 5, object handover with multi-finger hands is rarely explored. Combined with Fig. 6, we suggest that researchers should attempt to improve the grasp ability of multi-finger hands.

Long-sequence Task. At present, the vast majority of literature is devoted to the object handover task itself. Even if some researchers outline their methods as a spatial–temporal based paradigm, which is essentially a short-sequence task. How to introduce the ability of object handover into complex long-sequence human–robot collaboration tasks remains an open challenge.

9. Conclusions

This paper provides an in-depth review and analysis of recent developments in human–robot object handover, a fundamental and challenging aspect of human–robot collaboration. We explore this important area from multiple dimensions, including the role, end-effector and ability of robots.

To further contribute to this field, we propose a generic framework for human–robot object handover based on the methods reviewed in this paper, and implemented a handover system for an anthropomorphic hand. Our experiments not only verifies the effectiveness of the proposed framework but also highlight potential directions for future research, such as combining with LLM, focusing on multi-finger hand and long-sequence tasks.

As robots continue to be employed in human production and life, the importance of efficient and intuitive human–robot object handover will inevitably increase. We hope this review and our proposed framework will serve as valuable resources for researchers and developers in designing and implementing their own human–robot collaboration systems. As we move forward in this exciting field, we anticipate that further innovations will bridge the gap between human and robot collaboration, transforming the way we live and work.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgments

This work was supported by the National Natural Science Foundation of China (91748131, 62006229 and 61771471), the Strategic Priority Research Program of Chinese Academy of Science (XDB32050106), and the InnoHK Project.

References

[1] Haonan Duan, Peng Wang, Yayu Huang, Guangyun Xu, Wei Wei, Xiaofei Shen, Robotics dexterous grasping: The methods based on point cloud and deep learning, Front. Neurorobot. 15 (2021) 658280.

[2] Valerio Ortenzi, Akansel Cosgun, Tommaso Pardi, Wesley P Chan, Elizabeth Croft, Dana Kulić, Object handovers: a review for robotics, IEEE Trans. Robot. 37 (6) (2021) 1855–1873.

[3] Francesca Cini, V. Ortenzi, P. Corke, MJSR Controzzi, On the choice of grasp type and location when handing over an object, Science Robotics 4 (27) (2019) eaau9757.

[4] Jianzhi Lyu, Philipp Ruppel, Norman Hendrich, Shuang Li, Michael Görner, Jianwei Zhang, Efficient and collision-free human-robot collaboration based on intention and trajectory prediction, IEEE Trans. Cogn. Dev. Syst. (2022).

[5] Yujiao Cheng, Liting Sun, Changliu Liu, Masayoshi Tomizuka, Towards efficient human-robot collaboration with robust plan recognition and trajectory prediction, IEEE Robot. Autom. Lett. 5 (2) (2020) 2602–2609.

[6] Aaron Edsinger, Charles C. Kemp, Human-robot interaction for cooperative manipulation: Handing objects to one another, in: RO-MAN 2007-the 16th IEEE International Symposium on Robot and Human Interactive Communication, IEEE, 2007, pp. 1167–1172.

[7] Arash Ajoudani, Andrea Maria Zanchettin, Serena Ivaldi, Alin Albu-Schäffer, Kazuhiro Kosuge, Oussama Khatib, Progress and prospects of the human–robot collaboration, Auton. Robots 42 (2018) 957–975.

[8] Esben H. Østergaard, Welcome to industry 5.0, 2018, Retrieved Febr, 5: 2020.

[9] Aude Billard, Danica Kragic, Trends and challenges in robot manipulation, Science 364 (6446) (2019) eaat8414.

[10] Amit Kumar Pandey, Rachid Alami, Towards human-level semantics understanding of human-centered object manipulation tasks for hri: Reasoning about effect, ability, effort and perspective taking, Int. J. Soc. Robot. 6 (2014) 593–620.

[11] Afonso Castro, Filipe Silva, Vitor Santos, Trends of human-robot collaboration in industry contexts: Handover, learning, and metrics, Sensors 21 (12) (2021) 4113.

[12] Nicole Robinson, Brendan Tidd, Dylan Campbell, Dana Kulić, Peter Corke, Robotic vision for human-robot interaction and collaboration: A survey and systematic review, ACM Trans. Hum.-Robot Interact. 12 (1) (2023) 1–66.

[13] Junming Fan, Pai Zheng, Shufei Li, Vision-based holistic scene understanding towards proactive human–robot collaboration, Robot. Comput.-Integr. Manuf. 75 (2022) 102304.

[14] Shufei Li, Pai Zheng, Sichao Liu, Zuoxu Wang, Xi Vincent Wang, Lianyu Zheng, Lihui Wang, Proactive human–robot collaboration: Mutual-cognitive, predictable, and self-organising perspectives, Robot. Comput.-Integr. Manuf. 81 (2023) 102510.

[15] Francesco Semeraro, Alexander Griffiths, Angelo Cangelosi, Human–robot collaboration and machine learning: A systematic review of recent research, Robot. Comput.-Integr. Manuf. 79 (2023) 102432.
[16] Michael I. Jordan, Tom M. Mitchell, Machine learning: Trends, perspectives, and prospects, Science 349 (6245) (2015) 255–260.
[17] Arturo Daniel Sosa-Ceron, Hugo Gustavo Gonzalez-Hernandez, Jorge Antonio Reyes-Avendaño, Learning from demonstrations in human–robot collaborative scenarios: A survey, Robotics 11 (6) (2022) 126.
[18] Nourhan Abdulazeem, Yue Hu, Human factors considerations for quantifiable human states in physical human-robot interaction: A literature review, Sensors 23 (17) (2023).
[19] Marco Käppler, Ilshat Mamaev, Hosam Alagi, Thorsten Stein, Barbara Deml, Optimizing human-robot handovers: the impact of adaptive transport methods, Front. Robot. AI 10 (2023).
[20] Alexander Calvert, Wesley Chan, Tin Tran, Sara Sheikholeslami, Rhys Newbury, Akansel Cosgun, Elizabeth Croft, Comparing subjective perceptions of robot-to-human handover trajectories, 2022, arXiv preprint arXiv:2211.08733.
[21] Alap Kshirsagar, Rahul Kumar Ravi, Hadas Kress-Gazit, Guy Hoffman, Timing-specified controllers with feedback for human-robot handovers, in: 2022 31st IEEE International Conference on Robot and Human Interactive Communication (RO-MAN), IEEE, 2022, pp. 1313–1320.
[22] Alap Kshirsagar, Guy Hoffman, Armin Biess, Evaluating guided policy search for human-robot handovers, IEEE Robot. Autom. Lett. 6 (2) (2021) 3933–3940.
[23] Richard S. Sutton, Andrew G. Barto, Reinforcement Learning: An Introduction, MIT Press, 2018.
[24] Chen Wang, Claudia Pérez-D'Arpino, Danfei Xu, Li Fei-Fei, Karen Liu, Silvio Savarese, Co-gail: Learning diverse strategies for human-robot collaboration, in: Conference on Robot Learning, PMLR, 2022, pp. 1279–1290.
[25] Matija Mavsar, Aleš Ude, RoverNet: Vision-based adaptive human-to-robot object handovers, in: 2022 IEEE-RAS 21st International Conference on Humanoid Robots (Humanoids), IEEE, 2022, pp. 858–864.
[26] Sepp Hochreiter, Jürgen Schmidhuber, Long short-term memory, Neural Comput. 9 (8) (1997) 1735–1780.
[27] Kushal Kedia, Prithwish Dan, Atiksh Bhardwaj, Sanjiban Choudhury, ManiCast: Collaborative manipulation with cost-aware human forecasting, 2023, arXiv preprint arXiv:2310.13258.
[28] Min Wu, Bertram Taetz, Yanhao He, Gabriele Bleser, Steven Liu, An adaptive learning and control framework based on dynamic movement primitives with application to human–robot handovers, Robot. Auton. Syst. 148 (2022) 103935.
[29] Mohamed Djeha, Antonin Dallard, Ahmed Zermane, Pierre Gergondet, Abderrahmane Kheddar, Human-robot handovers using task-space quadratic programming, in: 2022 31st IEEE International Conference on Robot and Human Interactive Communication (RO-MAN), IEEE, 2022, pp. 518–523.
[30] Nuno F. Duarte, Aude Billard, José Santos-Victor, The role of object physical properties in human handover actions: Applications in robotics, IEEE Trans. Cogn. Dev. Syst. (2022).
[31] Linda Lastrico, Nuno Ferreira Duarte, Alessandro Carfì, Francesco Rea, Alessandra Sciutti, Fulvio Mastrogiovanni, José Santos-Victor, Expressing and inferring action carefulness in human-to-robot handovers, 2023, arXiv preprint arXiv:2310.00465.
[32] Yik Lung Pang, Alessio Xompero, Changjae Oh, Andrea Cavallaro, Towards safe human-to-robot handovers of unknown containers, in: 2021 30th IEEE International Conference on Robot & Human Interactive Communication (RO-MAN), IEEE, 2021, pp. 51–58.
[33] Zichen Yan, Wei He, Yuanhang Wang, Liang Sun, Xinbo Yu, Probabilistic motion prediction and skill learning for human-to-cobot dual-arm handover control, IEEE Trans. Neural Netw. Learn. Syst. (2022).
[34] Andrew Choi, Mohammad Khalid Jawed, Jungseock Joo, Preemptive motion planning for human-to-robot indirect placement handovers, in: 2022 International Conference on Robotics and Automation, ICRA, IEEE, 2022, pp. 4743–4749.
[35] Wei Yang, Chris Paxton, Arsalan Mousavian, Yu-Wei Chao, Maya Cakmak, Dieter Fox, Reactive human-to-robot handovers of arbitrary objects, in: 2021 IEEE International Conference on Robotics and Automation, ICRA, IEEE, 2021, pp. 3118–3124.
[36] Wei Yang, Balakumar Sundaralingam, Chris Paxton, Iretiayo Akinola, Yu-Wei Chao, Maya Cakmak, Dieter Fox, Model predictive control for fluid human-to-robot handovers, in: 2022 International Conference on Robotics and Automation, ICRA, IEEE, 2022, pp. 6956–6962.
[37] Sammy Christen, Wei Yang, Claudia Pérez-D'Arpino, Otmar Hilliges, Dieter Fox, Yu-Wei Chao, Learning human-to-robot handovers from point clouds, in: Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, 2023, pp. 9654–9664.
[38] Sammy Christen, Lan Feng, Wei Yang, Yu-Wei Chao, Otmar Hilliges, Jie Song, SynH2R: Synthesizing hand-object motions for learning human-to-robot handovers, 2023, arXiv preprint arXiv:2311.05599.
[39] Gu Zhang, Hao-Shu Fang, Hongjie Fang, Cewu Lu, Flexible handover with real-time robust dynamic grasp trajectory generation, 2023, arXiv preprint arXiv:2308.15622.
[40] Lirui Wang, Yu Xiang, Wei Yang, Arsalan Mousavian, Dieter Fox, Goal-auxiliary actor-critic for 6d robotic grasping with point clouds, in: Conference on Robot Learning, PMLR, 2022, pp. 70–80.
[41] Po-Kai Chang, Jui-Te Huang, Yu-Yen Huang, Hsueh-Cheng Wang, Learning end-to-end 6dof grasp choice of human-to-robot handover using affordance prediction and deep reinforcement learning, in: 2022 IEEE International Conference on Robotics and Automation (ICRA), IEEE, 2022.
[42] Ching-I Huang, Yu-Yen Huang, Jie-Xin Liu, Yu-Ting Ko, Hsueh-Cheng Wang, Kuang-Hsing Chiang, Lap-Fai Yu, Fed-HANet: Federated visual grasping learning for human robot handovers, IEEE Robot. Autom. Lett. (2023).
[43] Haonan Duan, Peng Wang, Yiming Li, Daheng Li, Wei Wei, Learning human-to-robot dexterous handovers for anthropomorphic hand, IEEE Trans. Cogn. Dev. Syst. (2022).
[44] Ayan Mazhitov, Togzhan Syrymova, Zhanat Kappassov, Matteo Rubagotti, Human–robot handover with prior-to-pass soft/rigid object classification via tactile glove, Robot. Auton. Syst. 159 (2023) 104311.
[45] Paramin Neranon, Tanapong Sutiphotinun, A human-inspired control strategy for improving seamless robot-to-human handovers, Appl. Sci. 11 (10) (2021) 4437.
[46] Gojko Perovic, Francesco Iori, Angela Mazzeo, Marco Controzzi, Egidio Falotico, Adaptive robot-human handovers with preference learning, IEEE Robot. Autom. Lett. (2023).
[47] Francesco Iori, Gojko Perovic, Francesca Cini, Angela Mazzeo, Egidio Falotico, Marco Controzzi, DMP-based reactive robot-to-human handover in perturbed scenarios, Int. J. Soc. Robot. 15 (2) (2023) 233–248.
[48] Marta Lagomarsino, Marta Lorenzini, Merryn Dale Constable, Elena De Momi, Cristina Becchio, Arash Ajoudani, Maximising coefficiency of human-robot handovers through reinforcement learning, IEEE Robot. Autom. Lett. (2023).
[49] Meiying Qin, Jake Brawer, Brian Scassellati, Task-oriented robot-to-human handovers in collaborative tool-use tasks, in: 2022 31st IEEE International Conference on Robot and Human Interactive Communication, (RO-MAN), 2022.
[50] Paola Ardón, Maria E Cabrera, Eric Pairet, Ronald PA Petrick, Subramanian Ramamoorthy, Katrin S Lohan, Maya Cakmak, Affordance-aware handovers with human arm mobility constraints, IEEE Robot. Autom. Lett. 6 (2) (2021) 3136–3143.
[51] Chongxi Meng, Tianwei Zhang, Tin lun Lam, Fast and comfortable interactive robot-to-human object handover, in: 2022 IEEE/RSJ International Conference on Intelligent Robots and Systems, IROS, IEEE, 2022, pp. 3701–3706.
[52] Daniel Lehotsky, Albert Christensen, Dimitrios Chrysostomou, Optimizing robot-to-human object handovers using vision-based affordance information, in: IEEE International Conference on Imaging Systems and Techniques (IST 2023), IEEE, 2023.
[53] Albert Daugbjerg Christensen, Daniel Lehotskỳ, Marius Willemoes Jø rgensen, Dimitrios 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.
[54] Javier Laplaza, Albert Pumarola, Francesc Moreno-Noguer, Alberto Sanfeliu, Attention deep learning based model for predicting the 3D human body pose using the robot human handover phases, in: 2021 30th IEEE International Conference on Robot & Human Interactive Communication (RO-MAN), IEEE, 2021, pp. 161–166.
[55] Salih Ertug Ovur, Yiannis Demiris, Naturalistic robot-to-human bimanual handover in complex environments through multi-sensor fusion, IEEE Trans. Autom. Sci. Eng. (2023).
[56] Antonis Sidiropoulos, Yiannis Karayiannidis, Zoe Doulgeri, Human-robot collaborative object transfer using human motion prediction based on cartesian pose dynamic movement primitives, in: 2021 IEEE International Conference on Robotics and Automation, ICRA, IEEE, 2021, pp. 3758–3764.
[57] Zheyu Zhuang, Yizhak Ben-Shabat, Jiahao Zhang, Stephen Gould, Robert Mahony, GoferBot: A visual guided human-robot collaborative assembly system, in: 2022 IEEE/RSJ International Conference on Intelligent Robots and Systems, IROS, IEEE, 2022, pp. 8910–8917.
[58] Jordan Nowak, Philippe Fraisse, Andrea Cherubini, Jean-Pierre Daures, Assistance to older adults with comfortable robot-to-human handovers, in: 2022 IEEE International Conference on Advanced Robotics and Its Social Impacts, ARSO, IEEE, 2022, pp. 1–6.
[59] Fabian C. Weigend, Xiao Liu, Heni Ben Amor, Probabilistic differentiable filters enable ubiquitous robot control with smartwatches, 2023, arXiv preprint arXiv:2309.06606.
[60] Dong Liu, Xianwei Wang, Ming Cong, Yu Du, Qiang Zou, Xiaomin Zhang, Object transfer point predicting based on human comfort model for human-robot handover, IEEE Trans. Instrum. Meas. 70 (2021) 1–11.
[61] Wei He, Jiashu Li, Zichen Yan, Fei Chen, Bidirectional human–robot bimanual handover of big planar object with vertical posture, IEEE Trans. Autom. Sci. Eng. 19 (2) (2021) 1180–1191.
[62] Marco Costanzo, Giuseppe De Maria, Ciro Natale, Handover control for human-robot and robot-robot collaboration, Front. Robot. AI 8 (2021) 672995.

[63] Yuwei Du, Xin Zhang, Mattia Leonori, Pietro Balatti, Jing Jin, Qiang Wang, Arash Ajoudani, Bi-directional human-robot handover using a novel supernumerary robotic system, in: 2023 IEEE International Conference on Advanced Robotics and Its Social Impacts, ARSO, IEEE, 2023, pp. 153–158.
[64] Javier Laplaza, Nicolás Rodríguez, José Enrique Domínguez-Vidal, Fernando Herrero, Sergi Hernández, Alejandro López, Alberto Sanfeliu, Anaís Garrell, IVO robot: A new social robot for human-robot collaboration, in: 2022 17th ACM/IEEE International Conference on Human-Robot Interaction, HRI, IEEE, 2022, pp. 860–864.
[65] J.P. Marques Silva, Karem A. Sakallah, GRASP-a new search algorithm for satisfiability, in: Proceedings of International Conference on Computer Aided Design, IEEE, 1996, pp. 220–227.
[66] Kaiming He, Georgia Gkioxari, Piotr Dollár, Ross Girshick, Mask r-cnn, in: Proceedings of the IEEE International Conference on Computer Vision, 2017, pp. 2961–2969.
[67] Andy Zeng, Shuran Song, Kuan-Ting Yu, Elliott Donlon, Francois R Hogan, Maria Bauza, Daolin Ma, Orion Taylor, Melody Liu, Eudald Romo, et al., Robotic pick-and-place of novel objects in clutter with multi-affordance grasping and cross-domain image matching, Int. J. Robot. Res. 41 (7) (2022) 690–705.
[68] Richard Bellman, A Markovian decision process, J. Math. Mech. (1957) 679–684.
[69] Hado Van Hasselt, Arthur Guez, David Silver, Deep reinforcement learning with double q-learning, in: Proceedings of the AAAI Conference on Artificial Intelligence, Vol. 30, 2016.
[70] Charles Ruizhongtai Qi, Li Yi, Hao Su, Leonidas J Guibas, Pointnet++: Deep hierarchical feature learning on point sets in a metric space, Adv. Neural Inf. Process. Syst. 30 (2017).
[71] Jirong Liu, Ruo Zhang, Hao-Shu Fang, Minghao Gou, Hongjie Fang, Chenxi Wang, Sheng Xu, Hengxu Yan, Cewu Lu, Target-referenced reactive grasping for dynamic objects, in: Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, 2023, pp. 8824–8833.
[72] Ashish Vaswani, Noam Shazeer, Niki Parmar, Jakob Uszkoreit, Llion Jones, Aidan N Gomez, Łukasz Kaiser, Illia Polosukhin, Attention is all you need, Adv. Neural Inf. Process. Syst. 30 (2017).
[73] Matija Mavsar, Barry Ridge, Jun Morimoto, Aleš Ude, et al., Simulation-aided handover prediction from video using recurrent image-to-motion networks, IEEE Trans. Neural Netw. Learn. Syst. (2022).
[74] Jürgen Schmidhuber, Deep learning in neural networks: An overview, Neural Netw. 61 (2015) 85–117.
[75] Wei Mao, Miaomiao Liu, Mathieu Salzmann, History repeats itself: Human motion prediction via motion attention, in: Computer Vision–ECCV 2020: 16th European Conference, Glasgow, UK, August 23–28, 2020, Proceedings, Part XIV 16, Springer, 2020, pp. 474–489.
[76] Xiao Liu, Geoffrey Clark, Joseph Campbell, Yifan Zhou, Heni Ben Amor, Enhancing state estimation in robots: A data-driven approach with differentiable ensemble kalman filters, 2023, arXiv preprint arXiv:2308.09870.
[77] Scott Fujimoto, Herke Hoof, David Meger, Addressing function approximation error in actor-critic methods, in: International Conference on Machine Learning, PMLR, 2018, pp. 1587–1596.
[78] Manfred Morari, Jay H. Lee, Model predictive control: past, present and future, Comput. Chem. Eng. 23 (4–5) (1999) 667–682.
[79] Arsalan Mousavian, Clemens Eppner, Dieter Fox, 6-dof graspnet: Variational grasp generation for object manipulation, in: Proceedings of the IEEE/CVF International Conference on Computer Vision, 2019, pp. 2901–2910.
[80] Matteo Saveriano, Fares J Abu-Dakka, Aljaž Kramberger, Luka Peternel, Dynamic movement primitives in robotics: A tutorial survey, Int. J. Robot. Res. (2021) 02783649231201196.
[81] Yansong Wu, Lingyun Chen, Ignacio Perez Mahiques, Zhenshan Bing, Fan Wu, Alois Knoll, Sami Haddadin, Learning Dynamic Robot-to-Robot Object Handover.
[82] Monica Sileo, Michelangelo Nigro, Domenico D Bloisi, Francesco Pierri, Vision based robot-to-robot object handover, in: 2021 20th International Conference on Advanced Robotics, ICAR, IEEE, 2021, pp. 664–669.
[83] Yunfei Li, Chaoyi Pan, Huazhe Xu, Xiaolong Wang, Yi Wu, Efficient bimanual handover and rearrangement via symmetry-aware actor-critic learning, in: 2023 IEEE International Conference on Robotics and Automation, ICRA, IEEE, 2023, pp. 3867–3874.
[84] Binghao Huang, Yuanpei Chen, Tianyu Wang, Yuzhe Qin, Yaodong Yang, Nikolay Atanasov, Xiaolong Wang, Dynamic handover: Throw and catch with bimanual hands, 2023, arXiv preprint arXiv:2309.05655.
[85] Ruolin Ye, Wenqiang Xu, Zhendong Xue, Tutian Tang, Yanfeng Wang, Cewu Lu, H2O: A benchmark for visual human-human object handover analysis, in: Proceedings of the IEEE/CVF International Conference on Computer Vision, 2021, pp. 15762–15771.
[86] Noah Wiederhold, Ava Megyeri, DiMaggio Paris, Sean Banerjee, Natasha Kholgade Banerjee, HOH: Markerless multimodal human-object-human handover dataset with large object count, 2023, arXiv preprint arXiv:2310.00723.
[87] Alap Kshirsagar, Raphael Fortuna, Zhiming Xie, Guy Hoffman, Dataset of bimanual human-to-human object handovers, Data Brief (2023) 109277.
[88] Noah Wiederhold, Mingjun Li, Nikolas Lamb, DiMaggio Paris, Alaina Tulskie, Sean Banerjee, Natasha Kholgade Banerjee, Studying how object handoff orientations relate to subject preferences on handover, in: 2023 IEEE International Conference on Advanced Robotics and Its Social Impacts, ARSO, IEEE, 2023, pp. 40–45.
[89] Jan Leusmann, Carl Oechsner, Johanna Prinz, Robin Welsch, Sven Mayer, A database for kitchen objects: Investigating danger perception in the context of human-robot interaction, in: Extended Abstracts of the 2023 CHI Conference on Human Factors in Computing Systems, 2023, pp. 1–9.
[90] Leimin Tian, Kerry He, Shiyu Xu, Akansel Cosgun, Dana Kulic, Crafting with a robot assistant: Use social cues to inform adaptive handovers in human-robot collaboration, in: Proceedings of the 2023 ACM/IEEE International Conference on Human-Robot Interaction, 2023, pp. 252–260.
[91] Alessio Xompero, Yik Lung Pang, Timothy Patten, Ahalya Prabhakar, Berk Calli, Andrea Cavallaro, Audio-visual object classification for human-robot collaboration, in: ICASSP 2022-2022 IEEE International Conference on Acoustics, Speech and Signal Processing, ICASSP, IEEE, 2022, pp. 9137–9141.
[92] Nuno Ferreira Duarte, Mirko Raković, José Santos-Victor, Robot learning physical object properties from human visual cues: A novel approach to infer the fullness level in containers, in: 2022 International Conference on Robotics and Automation, ICRA, IEEE, 2022, pp. 10375–10381.
[93] Nuno Ferreira Duarte, Mirko Raković, José Santos-Victor, Learning motor resonance in human-human and human-robot interaction with coupled dynamical systems, in: 2021 IEEE International Conference on Robotics and Automation, ICRA, IEEE, 2021, pp. 3662–3668.
[94] Parag Khanna, Mårten Björkman, Christian Smith, A multimodal data set of human handovers with design implications for human-robot handovers, 2023, arXiv preprint arXiv:2304.02154.
[95] Yu-Wei Chao, Chris Paxton, Yu Xiang, Wei Yang, Balakumar Sundaralingam, Tao Chen, Adithyavairavan Murali, Maya Cakmak, Dieter Fox, Handoversim: A simulation framework and benchmark for human-to-robot object handovers, in: 2022 International Conference on Robotics and Automation, ICRA, IEEE, 2022, pp. 6941–6947.
[96] Erwin Coumans, Yunfei Bai, PyBullet, a Python module for physics simulation for games, robotics and machine learning, 2016–2021, http://pybullet.org.
[97] Ilshat Mamaev, David Kretsch, Hosam Alagi, Björn Hein, Grasp detection for robot to human handovers using capacitive sensors, in: 2021 IEEE International Conference on Robotics and Automation, ICRA, IEEE, 2021, pp. 12552–12558.
[98] Mohammadhadi Mohandes, Behnam Moradi, Kamal Gupta, Mehran Mehrandezh, Robot to human object handover using vision and joint torque sensor modalities, in: International Conference on Robot Intelligence Technology and Applications, Springer, 2022, pp. 109–124.
[99] Parag Khanna, Mårten Björkman, Christian Smith, Human inspired grip-release technique for robot-human handovers, in: 2022 IEEE-RAS 21st International Conference on Humanoid Robots (Humanoids), IEEE, 2022, pp. 694–701.
[100] Afonso Castro, Joel Baptista, Filipe Silva, Vitor Santos, Classification of handover interaction primitives in a COBOT–human context with a deep neural network, J. Manuf. Syst. 68 (2023) 289–302.
[101] Weitian Wang, Rui Li, Yi Chen, Yi Sun, Yunyi Jia, Predicting human intentions in human–robot hand-over tasks through multimodal learning, IEEE Trans. Autom. Sci. Eng. 19 (3) (2021) 2339–2353.
[102] Javier Laplaza, Anaís Garrell, Francesc Moreno-Noguer, Alberto Sanfeliu, Context and intention for 3D human motion prediction: Experimentation and user study in handover tasks, in: 2022 31st IEEE International Conference on Robot and Human Interactive Communication (RO-MAN), IEEE, 2022, pp. 630–635.
[103] Rhys Newbury, Akansel Cosgun, Tysha Crowley-Davis, Wesley P Chan, Tom Drummond, Elizabeth 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), IEEE, 2022, pp. 1264–1270.
[104] Ernst Niedermeyer, F.H. Lopes da Silva, Electroencephalography: Basic Principles, Clinical Applications, and Related Fields, 2005.
[105] Roberto Merletti, Philip J. Parker, Electromyography: Physiology, Engineering, and Non-Invasive Applications, vol. 11, John Wiley & Sons, 2004.
[106] Jianzhi Lyu, Alexander Maỳe, Michael Görner, Philipp Ruppel, Andreas K Engel, Jianwei Zhang, Coordinating human-robot collaboration by EEG-based human intention prediction and vigilance control, Front. Neurorobot. 16 (2022) 1068274.
[107] Sara Cooper, Samuel FP Fensome, Dimitrios Kourtis, Stuart Gow, Mauro Dragone, An EEG investigation on planning human-robot handover tasks, in: 2020 IEEE International Conference on Human-Machine Systems, ICHMS, IEEE, 2020, pp. 1–6.

[108] Kerry He, Pradeepsundar Simini, Wesley P. Chan, Dana Kulić, Elizabeth Croft, Akansel Cosgun, On-the-go robot-to-human handovers with a mobile manipulator, in: 2022 31st IEEE International Conference on Robot and Human Interactive Communication (RO-MAN), IEEE, 2022, pp. 729–734.
[109] Valerio Ortenzi, Maija Filipovica, Diar Abdlkarim, Tommaso Pardi, Chie Takahashi, Alan M Wing, Massimiliano Di Luca, Katherine J Kuchenbecker, Robot, pass me the tool: handle visibility facilitates task-oriented handovers, in: 2022 17th ACM/IEEE International Conference on Human-Robot Interaction, HRI, IEEE, 2022, pp. 256–264.
[110] Haziq Razali, Yiannis 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.
[111] James Male, Gorkem Anil Al, Arya Shabani, Uriel Martinez-Hernandez, Multimodal sensor-based human-robot collaboration in assembly tasks, in: 2022 IEEE International Conference on Systems, Man, and Cybernetics, SMC, IEEE, 2022, pp. 1266–1271.
[112] Vivek Gupte, Dan R. Suissa, Yael Edan, Optometrist's algorithm for personalizing robot-human handovers, 2023, arXiv preprint arXiv:2308.15007.
[113] Pallavi Tilloo, Jesse Parron, Omar Obidat, Michelle Zhu, Weitian Wang, A POMDP-based robot-human trust model for human-robot collaboration, in: 2022 12th International Conference on CYBER Technology in Automation, Control, and Intelligent Systems, CYBER, IEEE, 2022, pp. 1009–1014.
[114] Alexandra Weidemann, Nele Rußwinkel, The role of frustration in human–robot interaction–what is needed for a successful collaboration? Front. Psychol. 12 (2021) 707.
[115] Tair Faibish, Alap Kshirsagar, Guy Hoffman, Yael Edan, Human preferences for robot eye gaze in human-to-robot handovers, Int. J. Soc. Robot. 14 (4) (2022) 995–1012.
[116] Muhammad Akmal Bin Mohammed Zaffir, Takahiro Wada, Presenting human-robot relative hand position using a multi-step vibrotactile stimulus for handover task, in: Companion of the 2023 ACM/IEEE International Conference on Human-Robot Interaction, 2023, pp. 426–430.
[117] Dorothea Langer, Franziska Legler, Philipp Kotsch, André Dettmann, Angelika C Bullinger, I let go now! Towards a voice-user interface for handovers between robots and users with full and impaired sight, Robotics 11 (5) (2022) 112.
[118] Gianluca Corsini, Martin Jacquet, Hemjyoti Das, Amr Afifi, Daniel Sidobre, Antonio Franchi, Nonlinear model predictive control for human-robot handover with application to the aerial case, in: 2022 IEEE/RSJ International Conference on Intelligent Robots and Systems, IROS, IEEE, 2022, pp. 7597–7604.
[119] Wayne Xin Zhao, Kun Zhou, Junyi Li, Tianyi Tang, Xiaolei Wang, Yupeng Hou, Yingqian Min, Beichen Zhang, Junjie Zhang, Zican Dong, et al., A survey of large language models, 2023, arXiv preprint arXiv:2303.18223.
[120] Ceng Zhang, Junxin Chen, Jiatong Li, Yanhong Peng, Zebing Mao, Large language models for human-robot interaction: A review, Biomim. Intell. Robot. (2023) 100131.
[121] Bowen Zhang, Harold Soh, Large language models as zero-shot human models for human-robot interaction, 2023, arXiv preprint arXiv:2303.03548.
[122] Anthony Brohan, Noah Brown, Justice Carbajal, Yevgen Chebotar, Xi Chen, Krzysztof Choromanski, Tianli Ding, Danny Driess, Avinava Dubey, Chelsea Finn, et al., Rt-2: Vision-language-action models transfer web knowledge to robotic control, 2023, arXiv preprint arXiv:2307.15818.
[123] Danny Driess, Fei Xia, Mehdi SM Sajjadi, Corey Lynch, Aakanksha Chowdhery, Brian Ichter, Ayzaan Wahid, Jonathan Tompson, Quan Vuong, Tianhe Yu, et al., Palm-e: An embodied multimodal language model, 2023, arXiv preprint arXiv:2303.03378.