Assisted Human-Robot Interaction for Industry Application Based Augmented Reality

Haonan Fang[✉], Jingqian Wen, XiaoNan Yang, Peng Wang, and Yinqian Li

School of Mechanical Engineering, Beijing Institute of Technology, Beijing, China
ffanghaonan@126.com, {wenjq,yangxn,3220210403}@bit.edu.cn

Abstract. Augmented Reality provides more possibilities for industrial manufacturing to solve the current problems. However, previous researches have been focused on virtual scene visualization and robot trajectory prediction. The investigation of robotic real-time teleoperation using AR is limited, which is still facing great challenges. In this research, a novel method is presented for human-robot interaction for industrial applications based on AR, like assembly. Users can intuitively teleoperate a 6-DOFs industry robot to recognize and locate entities by multi-channel operation via HoloLens2. Augmented reality, as a medium, bridges human zones and robot zones. The system can transform user instructions from virtual multi-channel user interface to robots by communication module in AR environment. The above approaches and related devices are elaborated in this paper. An industrial case is presented to implement and validate the feasibility of the AR-based human-robot interactive system. The results provide evidence to support that AR is one of the efficient methods to realize multi-channel human-robot interaction.

Keywords: Human-robot interaction · Augmented Reality · Manufacturing industry

1 Introduction

With the advances of Industry 4.0, the manufacturing industry calls for a new generation of novel systems and technologies to reduce product development time and cost. As an important part of intelligent mechanical equipment, industrial robot has been the heart of the modern manufacturing plant, which is widely used in many operations, e.g. welding, assembly, machining. Within the usage of industrial robots, human-robot interaction (HRI) is developed as an essential technology. It provides efficient support and becomes an essential part of manufacturing, especially assembly operations [1]. A growing number of novel technologies have been developed for HRI. However, there are still some problems, such as lack of safety, complex operation process and few personnel with rich experience. Augmented reality (AR) is considered a promising method to resolve the above questions. It combines the physical manufacturing environment with the virtuality in cyberspace to provide a new tool to developers and customers of industrial systems. Therefore, it necessitates integrating the AR approach with the HRI application to offer a more natural and convenient method for the operator, which can also enhance the efficiency of human interaction with robot systems [2].

In parallel, there are multiple AR hard devices developed by companies, which can be divided into Head-mounted displays (HMDs), hand-held devices and spatial devices. Meanwhile, HMDs are considered to visualize 3D data most directly so that the users can observe a spatial display of detailed information, which supports understanding and choosing [3]. Besides, the advantage that frees users both hands from manual operations of head-mounted AR, like assembling, was considerably convenient in manufacturing. As a cutting edge of head-mounted AR, the Microsoft HoloLens2 provides a possibility to work remotely without any external interference by using gaze and gesture controllers [4].

Current applications of AR-HRI mainly apply to virtual reality fusion technology, assisting operators to see the virtual information in its real physical surroundings to augment the understanding of the real world by showing information. Previous studies include displaying operating instructions for non-experts, forecasting in advance the motion path, and observing potential pick location during grasping movement by HoloLens2 [5]. Nevertheless, using AR technology to build up real-time HRI has been rarely studied.

According to the above challenges and requirements, this paper presents a system to recognize and locate entities by multi-channel teleoperation via HoloLens2. The research in this paper assists users to teleoperate robots via the multi-channel in a fixed station that is not convenient to move, and have to accomplish a real-time selection of relevant components without interruption. Meanwhile, getting the correct component directly by using the robot vision with receiving the orders from AR, instead of manual measurements and searching, enables saving manpower and decreasing the workload of operation. This research builds a new channel for human-robot interaction with AR systems in the future.

2   Relative Work

2.1   Human-Robot Interaction in Manufacturing

In recent years, the focus of human-robot interaction has shifted from equipment to the demands of humans. Traditional HRI is mainly through an "interactive interface", such as a mouse, keyboard, and operation panel, to operate robots. Workers transfer decisions to robots by tapping the keyboard or clicking the touch screen [6]. After processing information, robots transfer it to users through the visual channel. The major concern of the above process is the accuracy of transmitted contents. However, the connection between the mentioned separate robot workspaces and the robot's physical environment can be unintuitive. Once robots are fixed, the freedom of the relevant operators is limited.

Recently, the advances of Industry 4.0 bring about methods of robotic implementation. The overwhelming transformation put forwards that, a long period of time in the future, eliminating the boundaries of information exchange will become a new development direction of HRI. Robots can actively adapt to the demands of users through multi-channel perception, and produce a more natural and convenient way to collaborate [7]. Working side-by-side with robots in close proximity simultaneously also enables informative and real-time communication [8].

2.2 AR-Based Operator Support System

The burgeoning world of augmented reality provides a new route toward addressing some of the challenges that have been mentioned above. AR can provide a new tool to developers and customers of industrial systems in the manufacturing industry. The combination of digital content with real working spaces creates a fresh environment, which can produce visible feedback during HRI. In the literature, abundant recent works have demonstrated their HRI systems on real industrial manufacturing tasks, where both aspects, control and visualization, are considered (see Fig. 1).

As one of the main advantages of AR, visualization exhibits extensive applications in safety and robotics trajectory planning. In 2019, Faizan Muhammad [9] outlined a system that allows the users to visualize the robot's sensory information and intended path in intuitive modalities overlaid onto the real world. And users can interact with visual objects, such as textual notes, as a means of communication with the robot. Similarly, an AR approach [3] was proposed to visualize the navigation stack. Therefore, relevant data including laser scan, environment map and path planning data are visualized in 3D within Hololens. Then the research was extended to an AR-based safety margins visualization system. Antti Hietanen [10] proposes a shared workspace model for HRI manufacturing and interactive UIs. It focuses on monitoring safety margins and unfolding the virtual zones to users with AR technology.

In the process of robots control and application, AR technology is presented and demonstrates to possess huge potential towards accurate HRI operations. In the process of motion prediction, Dennis Krupke [5] introduced a system in which users can intuitively control a co-located robot arm for pick-and-place tasks. In addition, an AR interface has been developed to show a preview of robot pick selection or trajectory to the operator. Similarly, Kenneth A. Stone [11] detailed an app to edit the code directly and teleoperate robotic motion via virtual keyboard and a virtual panel, which realized the visualization of target, path and visual robot within AR. Then an MR-based distributed system [12] has been proposed to evaluate the robots movements before execution. The planning focus is the use of cloud-based computing and control to find a collision-free path. In addition, Chung Xue Er [13] also proposed a demo to preview and evaluate the motion planned artificially, using a visual robot to detect collision automatically and interact.

During the process of inputting commands, Xi Vincent Wang [1] developed a feasible AR system with a novel closed-loop structure. Within the system, users can manipulate the virtual robot via gesture commands to determine the posture, trajectory and task. After confirming trajectory, the physical robot will duplicate the operation accordingly. What's more, Congyuan Liang [14] also introduced a system to allow operator to teleoperate a robot via capture gestures information.

To recap, despite the significant development of the AR-based HRI application in recent years, there is still a lack of a comprehensive design for AR-based robotics input commands system. But the channel of most previous works is gesture information. Workers using their hands for a long time will be tired and lower productivity. Thus in this research, a novel AR-based HRI system is proposed within multi-channel teleoperation.
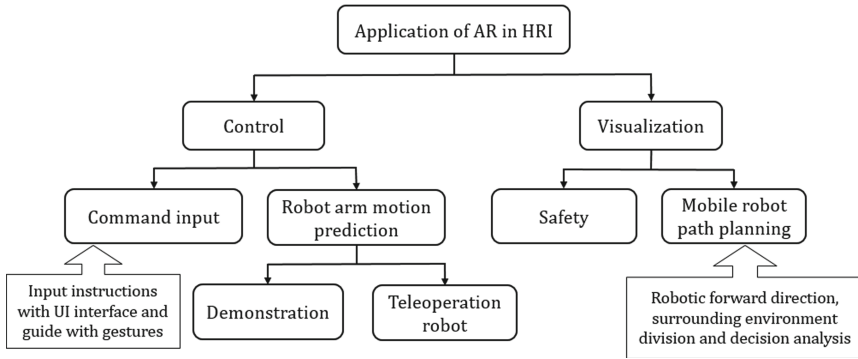
Fig. 1. The framework of AR-based HRI systems.

3 System Description

3.1 System Framework

To improve the efficiency and accuracy of the teleoperation between humans and robots, an AR-based framework is presented for assisted industry operations, as shown in Fig. 2. In this study, two separate workspaces, human zone and robot zone, are combined for assembly and other industrial operations by AR technology.

In the human zone, users, wearing the Microsoft Hololens2, could operate the AR user interface by gesture, voice and gaze. The multi-channel interaction could transform its input data to convey their orders and control the HRI system. In the robot zone, robots could receive the orders that transferred by the hardware device to gain the correct components. As a medium, AR technology connects the physical world and visual world between two zones.
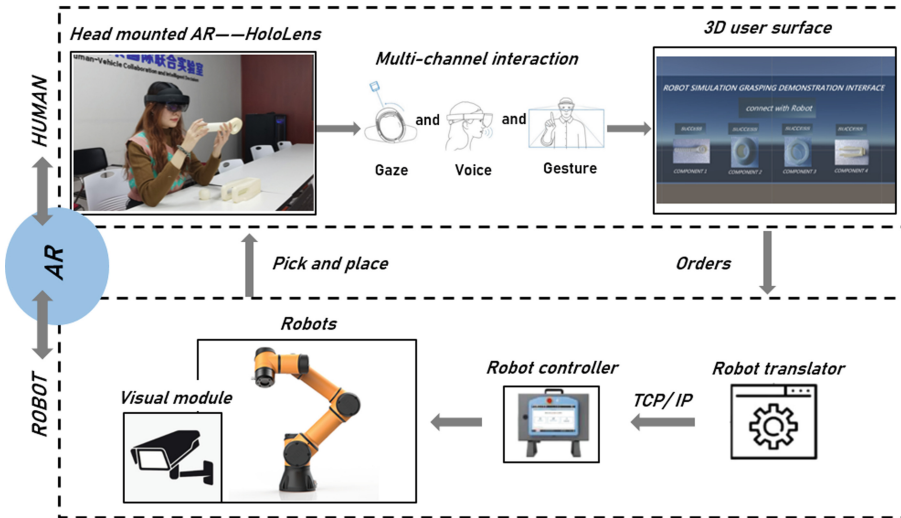

Fig. 2. AR-based framework for HRI.

Hardware Devices. The Hardware equipment contains Microsoft Hololens2 and a collaborative robot with an industrial camera.

With huge potential and advantages, HoloLens can mix images in the real world and related virtual images to display devices, giving users a powerful sense of immediacy. Scanning the real world and building real-time environment meshes, Hololens is capable of realizing world-scale positional tracking and spatial mapping of the virtual environment [15]. In this way, digital contents are allowed to be anchored to objects or surfaces wherever you are.

Hololens supports multiple modals to realize the interaction between humans and the virtual environment. Users remain focused on safely completing tasks error-free with hand tracking, built-in voice commands, eye tracking, spatial mapping, and large field of view. Specifically, Hololens understands precisely where you're looking to adjust its holograms to follow users' eyes in real time. Hololens adapts to both hands to fully execute hand tracking, touch, grasp, and move holograms in a natural way. Concurrently, built-in voice systems allow Hololens respond to the commands when users are occupied with a task and have no hands to quickly operate Hololens [16]. All above functions break through the space limitation of the two-dimensional interface and expand the human-robot interaction interface from two-dimensional to three-dimensional.

AUBO robots, as 6-DOF industrial robots, are used to interact with human. Cooperative robots that are also called workplace assistant robots or just "cobots", can cooperate with the operators on the same assembly line. The "partners" can fully give play to the efficiently of robots and human intelligence.

The type of AUBO robot used in this paper is the i3 series which has the range of motion of a sphere with a radius of 625mm. Meanwhile, as an HRI interface, AUBORPE (AUBO robot programming environment) displays on the touch panel of a teach pendant, which provides a visual operation interface to program and simulate the robot [17]. The system sets up two types of coordinate system for position control, including based on base or flange center coordinates. Equipped with an industrial camera of HIKROBOT, the robot will perform autonomously the operation of distinguishing, picking and handling[18]. Among the process, path planning is calculated and selected voluntarily through built-in scripts of the robot.

Software Device. The software setup contains unity to produce visual world and Vision Master to be used in visual positioning.

Unity3D is a comprehensive large-scale development tool for making interactive content such as scene visualization and 3D dynamic model. C# is used in Unity to be its primary language for developer scripting, which is a secure but straightforward language for beginners to cross the threshold [19]. The convenience of operation improves the productivity and problem-solving skills of experts and engineers.

Microsoft builds, without doubt, an open source cross-platform development kit, called Mixed-Reality Toolkit, to accelerate development of applications. Unity3D offers a sensible method for the development of the target hardware in Hololens. Building the virtual environment and virtual objects using Unity3D forms a portion of AR, such as research on collision detection and ergonomics analysis. The paper is mainly used for the establishment of the virtual user interface, as a way to interact with operators and transform native think into digital signals.

3.2   System Approach

The main goal of this project is to provide an example of the potential of AR technology applied to the world of industrial robots. The paper establishes communication and transfers information between users and robots via two 'middlemen', a multi-channel interactive user interface presented by HoloLens2 and the computer vision module of robots. The research emphases in this paper are the establishment of multi-channel AR user interface, computer vision recognition technology and the communication module.

Multi-channel AR User Interface. In this system, the user interface is used to complete interactive operation with special virtual digital contents through gaze, voice and gesture within the Microsoft HoloLens2 [20].

Users can interact with the menu page by multi-channel interactive operation, e.g. air-tap, bloom, hold and manipulation, to click, exit and flip the menu page. The above application enables to implement the several parts of complex interactive operation, setting buttons for users to click and activate. By the way, in the center of the user's field of vision is a white dot called the "gaze", which is emitted through the user's eyes [5]. Set colliders covering with cubes to collide with the above visual rays, and users can place their gaze on any cube they wish to interact with. About the voice, editing the setting and adding codes of the interactive cube can take actions after users say the corresponding words. The text on the button can be activated as a language instruction by a module called "SEE IT SAY IT" in Hololens.

The specific interaction methods and their combination modes can be formulated according to the actual scene. In this topic, the first step is to connect the robot with Hololens. Users gaze at the cube with "connect with Robot" for three seconds or say "connect" while gazing. The cube will highlight after the connection is successful. Then users perform an action of "hold" by two fingers toward the needed component, and the instruction can be transmitted to the robot. The cube with "SUCCESS" will highlight after successful choice. Communication technology will be commended in the next section (see Fig. 3).

Following these steps, users can select designated components and gain the corresponding information they need.
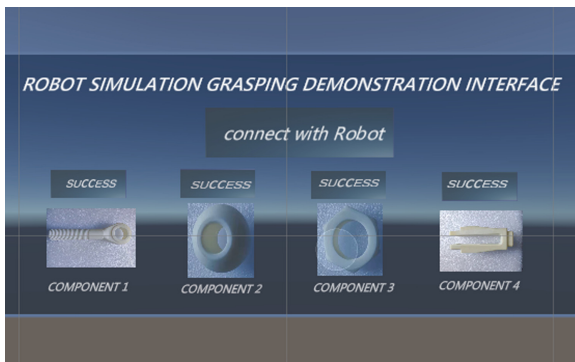

Fig. 3. Multi-channel AR user interface in unity

Communication Module. The communication module likes a "gap" to connect Hololens2 and the robot through the above multi-channel UI. Its framework and logical description are shown in Fig. 4.

The system uses C/S (Client-Server) architecture to solve the problem of user instructions' real-time transmission. HoloLens, as the server, can complete the functions of target information collection, model loading, spatial mapping and human-computer interaction, which enhances the mixing effect. The client, namely robot, completes the corresponding task according to the change of the system target information.

TCP/IP (Transmission Control Protocol/ Internet Protocol) is the data transmission protocol of the Internet in the whole system, which provides reliable service and plays an important role in guaranteeing network performances. When Hololens sends a connection request, the robot accepts data and returns data of successful acceptance to Hololens, which indicates successful communication. During transmission, the server and client are under the same Local Area Network (LAN).
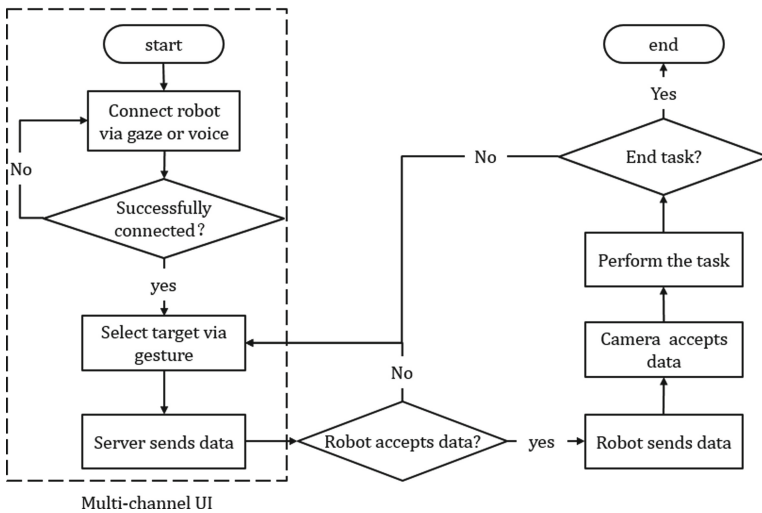

Fig. 4. The communication module of the system.

Robot Vision. After receiving the comments, the robot distinguishes and grabs the designated component by recognizing the size, shape or other characteristics through the visual sensor.

An image-based visual recognition control system was developed for the robot to recognize and grab the designated component. Through hardware design of the system, the image processing, and recognition technology, the recognition and location of the target are achieved.

The basic process is divided into two steps (see Fig. 5). The first step is to train and test the image samples collected as algorithm training models, so as to debug and verify the reliability and stability of the algorithm and continuously improve the fitting degree of the

model. The operations of involved label image samples include distortion correction, target features extraction, build feature set and hand-eye calibration. Building target feature sets is usually a model training process. The other step is to classify and recognize the target image with the optimal model after training, which includes repetitive distortion correction, target features extraction, target recognition and calibration conversion [21]. Finally, the target coordinates of the base coordinate system are obtained.
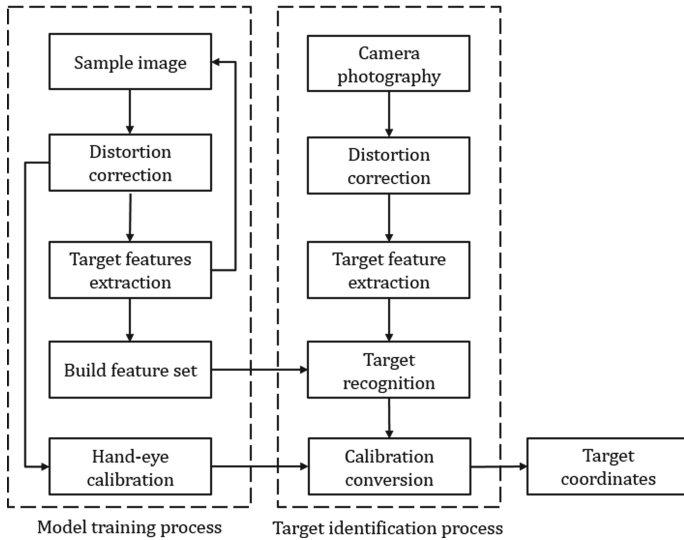

Fig. 5. The basic flowchart of the image-based visual recognition system

The architecture of the robotic computer vision module in this project is shown in Fig. 6. The camera matches the objective according to users' instructions and get its coordinates. The information is transmitted to the robot to grasp the target and accomplish the task.

3.3   Case Study

The AR-based human-robot instruction system is applied to the case study for a simple assembly process. As mentioned above, the HRI system consists of 3 parts: the human operator, an AR device (Hololens), a 6-DOF robot with a camera. Wearing Hololens, the human operator can catch sight of the virtual scene and control the robot anywhere as shown in Fig. 7.

During the task, for example, obtain a ball through teleoperation, which was developed to verify the application of the developed method. First, gaze at the cube with "connect with Robot". Then select the component magic representing the ball with a gesture. At last, the robot moved and grabbed the ball, as shown in Fig. 8.

The results provide evidence that AR is one of the efficient methods to realize multi-channel human-robot interaction. This application will help pave the way for future applications of augmented reality to practical, industrial, and other applications.
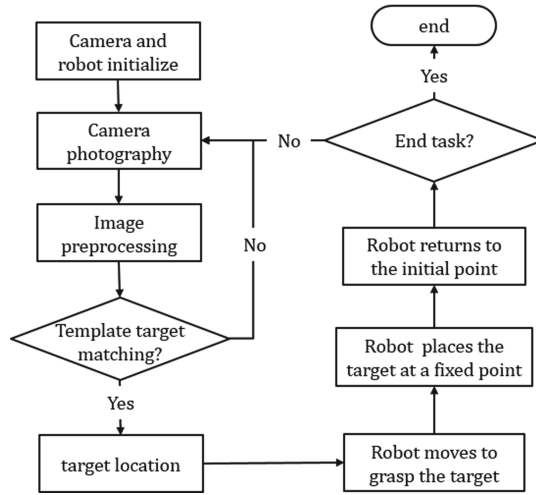
Fig. 6. The robotic computer vision module architecture

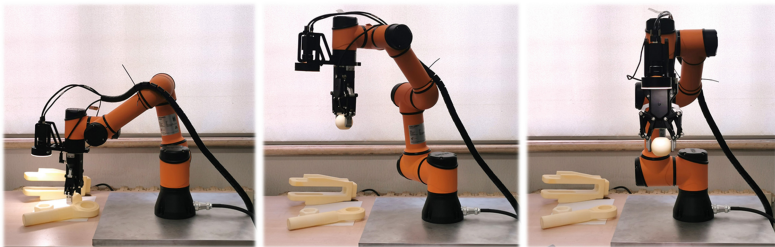**Fig. 7.** The user operate virtual UI with Hololens

Fig. 8. The robot action process of ball grabbing

4   Conclusion and Future Work

In this paper, we introduced and presented the implementation of an AR-based human-robot interaction system for industry application. Users can intuitively teleoperate a 6-DOFs industry robot to recognize and locate entity as long as the entity are within the robotic visual range. The multi-channel interaction of UI can be defined according to the scene and tasks, in order to maximize efficiency and convenience.

However, there are some limitations, which could be addressed in the future work. For example, the contents in the UI are determined and need to be uploaded again once modified, which couldn't suit abrupt environmental events and individual habits.

In the future, the vast possibility of the project will be revealed, and certain limitations from current abilities can be improved.

The ability to interact with robot is theoretically possible to be more natural. The virtual UI could be substituted by an invisible interface. Wearing Hololens, operators select directly real components via their visual ray, instead of clicking pictures on the visual interface. This ray can intersect with virtual collider superimposed on the real object. Within the above improvement, selected parts will not stick in the virtual interface and the sight line will not be blocked in this way as well, which improves the safety and flexibility of assembling.

Additionally, according to personas of operators, the spatial auxiliary information in virtual world can present intelligently. The system recognizes the human activities and forecasts user intention. After that, it will present auxiliary information and robotic motions as outputs. Meanwhile, the system is worth verifying to combine with JCS (Joint Cognitive Systems) and to be used in remote collaboration with robot. Offer more value relevance information for users in every case, which will enhance the individuation and productivity of tasks.

We believe that AR technology is an ongoing pursuit in the field of HRI. The finding of this paper has provided a new paradigm for AR-based applications and imparted novel insight to HRI for industry.

References

1. Wang, X.V., et al.: Closed-loop augmented reality towards accurate human-robot collaboration. CIRP Ann. **69**(1), 425–428 (2020)
2. Guo, J., et al.: Real-time Object Detection with Deep Learning for Robot Vision on Mixed Reality Device. IEEE (2021)
3. Kästner, L., Lambrecht, J.: Augmented-Reality-Based Visualization of Navigation Data of Mobile Robots on the Microsoft Hololens - Possibilities and Limitations (2019)
4. Huang, J.: A non-contact measurement method based on HoloLens. Int. J. Performabil. Eng. **14**, 144 (2018)
5. Krupke, D., et al.: Comparison of Multimodal Heading and Pointing Gestures for Co-Located Mixed Reality Human-Robot Interaction. IEEE (2018)
6. Guhl, J., Nguyen, S.T., Kruger, J.: concept and architecture for programming industrial robots using augmented reality with mobile devices like Microsoft HoloLens. Institute of Electrical and Electronics Engineers Inc., Limassol, Cyprus (2017)
7. Perzanowski, D., et al.: Building a multimodal human-robot interface. IEEE Intell. Syst. App. **16**(1), 16–21 (2001)

8. Vahrenkamp, N., et al.: Workspace analysis for planning human-robot interaction tasks. IEEE Computer Society, Cancun, Mexico (2016)

9. Muhammad, F., et al.: Creating a Shared Reality with Robots. IEEE Computer Society, Daegu, Republic of Korea (2019)

10. Hietanen, A., et al.: AR-based interaction for human-robot collaborative manufacturing. Robot. Comput.-Integr. Manuf. **63**, 101891 (2022)

11. Stone, K.A., et al.: Augmented Reality Interface for Industrial Robot Controllers. Institute of Electrical and Electronics Engineers Inc., Orem, UT, United States (2020)

12. Guhl, J., Hügle, J., Krüger, J.: Enabling human-robot-interaction via virtual and augmented reality in distributed control systems. Procedia CIRP. **76**, 167–170 (2018)

13. Xue, C., Qiao, Y., Murray, N.: Enabling Human-Robot-Interaction for Remote Robotic Operation via Augmented Reality. Virtual Institute of Electrical and Electronics Engineers Inc., Cork, Ireland (2020)

14. Liang, C., et al.: Robot Teleoperation System Based on Mixed Reality. Osaka Institute of Electrical and Electronics Engineers Inc., Japan (2019)

15. Azuma, R.T.M.: Augmented Reality A Reality. San Francisco OSA - The Optical Society CA, United States (2017)

16. Microsoft HoloLens | Mixed Reality Technology for Business. https://www.microsoft.com/en-us/hololens. Accessed 28 Jan 2022

17. AUBO Collaborative Robot. https://www.aubo-cobot.com/public/. Accessed 28 Jan 2022

18. Meng, Y., Zhuang, H.Q.: Autonomous robot calibration using vision technology. Robot. Comput. Integr. Manuf. **23**(4), 436–446 (2007)

19. Fan, L., et al.: Using AR and Unity 3D to Support Geographical Phenomena Simu-lations. Institute of Electrical and Electronics Engineers Inc. Virtual, Nanjing, China (2020)

20. Furlan, R.: The future of augmented reality: Hololens - Microsoft's AR headset shines despite rough edges [resources-tools and toys]. IEEE Spectr. **53**(6), 21 (2016)

21. Wang, J., et al.: Research on automatic target detection and recognition based on deep learning. J. Visual Commun. Image Represent. **60**, 44–50 (2019)