Available online at www.sciencedirect.com

ScienceDirect

Procedia CIRP 72 (2018) 123–128

www.elsevier.com/locate/procedia

51st CIRP Conference on Manufacturing Systems

Human-robot collaboration – towards new metrics for selection of communication technologies

Patrik Gustavsson[a,*], Magnus Holm[a], Anna Syberfeldt[a], Lihui Wang[b]

[a]*University of Skövde, PO Box 408, Skövde 541 28, Sweden*
[b]*KTH Royal Institute of Technology, Kungliga Tekniska Högskolan, Stockholm 10044, Sweden*

* Corresponding author. Tel.: +46-500-448-507; fax: +46-500-416-325. *E-mail address:* patrik.gustavsson@his.se

Abstract

Industrial robot manufacturers have in recent years developed collaborative robots and these gains more and more interest within the manufacturing industry. Collaborative robots ensure that humans and robots can work together without the robot being dangerous for the human. However, collaborative robots themselves are not enough to achieve collaboration between a human and a robot; collaboration is only possible if a proper communication between the human and the robot can be achieved. The aim of this paper is to identify and categorize technologies that can be used to enable such communication between a human and an industrial robot.

Peer-review under responsibility of the scientific committee of the 51st CIRP Conference on Manufacturing Systems.

Keywords: human-robot collaboration; human-robot interaction

Nomenclature

AR	Augmented Reality
ASR	Automatic Speech Recognition
HRC	Human-Robot Collaboration
HRI	Human-Robot Interaction
TTS	Text-To-Speech

1. Introduction

Interaction with industrial robots have historically been limited to simple control panels with displays. The robots were either controlled by human guidance or operated almost independently from the user. Human-Robot Collaboration (HRC) tries to close the gap between robots and humans by introducing a shared workspace that enables a human and a robot to execute a specific task together [1, 2]. This combination utilizes the strengths of both the human and the robot, where the human has flexibility, adaptability and intelligence, while the robot has physical strength, repeatability and accuracy [3]. There are currently several industrial robot manufacturers that offer collaborative robots, e.g., [4-6], which have greatly advanced the research in HRC the last couple of years. However, to fully utilize the potential of HRC there are several issues that remains to be considered. One such issue is to achieve a proper communication between the robot and the human, which is a necessity to truly realize HRC. Today, robot manufacturers use the term "collaborative" mainly in the sense of force limitation required by the safety standard, which allows humans to work in the same area as the robot. Force limitation does, however, not enable collaboration but there must also be a way of communicating between the robot and the human. The collaborative robots generally support programming-by-guidance, which is without doubt an important feature for HRC, but not enough for enabling full two-way communication.

The research area of HRC belongs to the field of human-robot interaction (HRI), which covers all types of interaction between a human and a robot. HRI can be divided into two general categories: remote and proximate interaction [7]. In remote interactions the human and the robot are spatially separated from each other, while in proximate interaction the human and the robot are co-located sharing the same area. Since industrial HRC is focused on the collaboration between a

2212-8271 © 2018 The Authors. Published by Elsevier B.V.
Peer-review under responsibility of the scientific committee of the 51st CIRP Conference on Manufacturing Systems.
10.1016/j.procir.2018.03.156

human and a robot in the same working cell, only proximate interaction is of interest for this paper. Specifically, the paper focuses on communication technologies for enabling proximate interaction and how to successfully select the proper technologies for a specific scenario. For aiding the selection, the paper suggests a number of metrics to be used for identifying the best technologies. The paper targets technologies used for communication between a human and a robot, and excludes technologies for safety, social interaction, and trust factors.

A considerable number of papers presents HRI applications combining different communication technologies [8-10] and several papers also summarizes various communication technologies used in HRI. However, these papers either focuses on technologies that have been tested together [1, 7, 8], or how metrics can be used when evaluating a combination of technologies [11, 12]. These papers uses metrics based on characteristics such as:

- **Reliability**, that is, how well the technology functions in nominal condition
- **Robustness**, that is, how the technology functions in adverse conditions
- **Cognitive load**, that is, amount of mental effort when using the technology
- **Delay**, that is, processing time that is necessary before the action is interpreted

These characteristics consider performance of specific technologies, however, they do not consider how technologies matches different tasks in HRC applications. Therefore, this paper aims to improve and extend the current use of metrics by considering also the type of task to be carried out. As far as the authors are aware, there are no previous metrics or classification scheme that aid the selection of communication technologies for specific tasks within HRC applications, which make this paper unique. With proper selection metrics, the idea is to enable end-users to efficiently identify the most optimal technologies for a specific scenario.

The next chapter continues by listing state-of-the-art communication technologies that have been tested in HRI and HRC applications. Chapter 3 then proposes a metric set needed to reliably select communication technologies for HRC applications and categorizes the technologies found in chapter 2. Chapter 4 finally concludes the paper and discusses future work.

2. HRI communication technologies

HRI is not only limited to communication from human to robot, but an essential part of interaction is the feedback loop to the human, to facilitate the human's understanding of the decisions made by the robot [13]. In addition, the human may need information from the system to know what he or she needs to do. Therefore, communication technologies can be separated into human-to-robot and robot-to-human communication.

The papers [9, 14-16] discusses how multimodality improves flexibility and robustness of HRI. The flexibility is improved using complementary communication technologies where different modalities recognizes different type of messages. The robustness is improved by using redundant communication technologies where different modalities improve the recognition of the same message. This work categorizes technologies and does not consider the robustness, therefore, the separate technologies are considered. There could, however, be a situation where the combination of technologies generates a unique message, not possible by the individual technologies. In that case those technologies are considered as one entity. As an example the soft-buttons mentioned in [9], is such an entity.

In the next two subchapters, the human-to-robot and robot-to-human communication technologies are described in further detail.

2.1. Robot-to-human communication technologies

Augmented reality (AR) is a technology that overlay digital information onto the real world and demonstrates promising results in HRI [8, 10, 17]. The technology provides several advantages such as displaying information where it is needed, highlighting different objects, showing how a motion can be executed, etc. To enable the technology some sort of hardware device is used, these devices can be categorized into: spatial, hand-held and head-mounted devices [18]. Different types of optics can be used to visualize information on the devices: video, optical and retinal affects the view of the user, while hologram and projection affects the visualization of the real world. AR technologies using spatial devices can be separated into spatial monitor (affects the view of the user) and spatial projection (affects the visualization of the real world), because these two categorization affects the type of task that they can be used for. AR using hand-held and head-mounted devices only uses optics affecting the view of the user and does not require additional categorization.

Text-To-Speech (TTS) technologies provides an artificial way of providing understandable audible output for the human [19]. This technology is used today in smartphones, cars, laptops, etc. TTS has also been suggested for HRI [8], to allow the robot to express itself using speech. Devices for TTS can be categorized into head-mounted or freestanding. The audio signal can be delivered in a non-spatial and spatial way. Spatial sound allows the user to locate it in a three dimensional space, which has been useful when searching and navigating through AR environments [20]. Both head-mounted and freestanding technologies can be used for spatial and non-spatial sound and do not need additional classification.

Pick-by-light and pick-by-voice are communication technologies common in modern warehouses [21]. Pick-by-light uses small lamps installed on each storage compartment. This aids by lighting up the compartment that the human should pick from. However, this system is not flexible because lamps or displays needs to be installed on every compartment. Therefore, a pick-by-vision system is suggested to overcome these problems, using AR glasses to highlight the different compartments. Pick-by-voice supports the worker using TTS instructions. The reliability of this technology degrades in noisy environment, and it is questionable whether the human would appreciate being told what to do with a monotone voice.

However, one objective with TTS synthesis is to make the speech indistinguishable from that produced by human [19], in which case the monotone voice will not be a problem.

2.2. Human-to-robot communication technologies

Haptic controls such as controls using force-torque sensors, joint-torque sensors, impedance or admittance, have the ability to physically control a robot by guiding it with the hand [3, 22, 23]. In comparison to traditional methods such as joystick or buttons, the efficiency can be increased by a multitude, and require less training to work with. There are two main approaches of controlling a robot, in Cartesian space and in joint space. Controlling a robot in Cartesian space may produce singularities if a redundant robot arm is used. However, controlling a robot in joint space will not produce such errors. Force-torque sensors mounted on end effector can be used to control a robot in Cartesian space but not in joint space, making them less flexible. Torque sensors, or compliance can be incorporated into each joint enabling control both in joint and Cartesian space, making them more flexible. Haptic control is therefore divided into two categories, end effector based and joint based.

A virtual impedance control has been tested in [24] for collision avoidance to ensure the safety of the operator. This was implemented with Kinect sensor using the detected skeleton to change the robot path to avoid collision. Although virtual impedance is used in this case for collision avoidance, other instances of impedance has been used to control the robot accurately such as [23]. This suggests that virtual impedance could be used for guiding the robot, but this has not been tested so far.

Automatic speech recognition (ASR) is the process of converting an audio signal into recognizable sentences for the system. ASR has been used in several instances in HRI to tell the robot what to do [8, 9, 14, 16, 25]. It shows good promise in HRC, because the human can interact in a way that is natural in human-to-human communication. This technology provides a way to communicate without removing hand or focus from current activity. Devices used for ASR can be divided into two categories, head-mounted and distant. Distant devices can use technologies such as omni- and unidirectional microphones, microphone arrays, etc. Microphone arrays can provide additional information such as direction of the speech, to filter out other voices. However, such filtering information is mainly used to improve robustness, which is not the focus of this paper. Therefore, ASR is divided into distant and head-mounted devices.

Gesture recognition provides an interface allowing the human to use gestures to interact with a system [26]. Such interaction includes pointing at an object to highlight it, giving thumbs up to indicate good quality, grasping the hand to demonstrate a gripping command, nodding the head to indicate affirmative decision, etc. Gesture recognition has been used in HRI using, vision based technologies [10, 14, 16, 25, 27], and glove based technologies [28, 29]. Several of the vision-based gesture recognition papers uses the inexpensive Microsoft Kinect as vision system. Vision based technologies may have better flexibility in comparison to glove based systems, but they face difficulties in covering gestures from all directions.

A multimodal HRI system has been tested in [9] that consists of a robot, a projector, and three input modalities. The input modalities are gaze recognition, ASR, and so called soft-buttons. Human gaze is realized with eye-tracking glasses, the ASR uses a head-mounted microphone, and the soft buttons are a combination of tracking the hand using vision sensors, i.e., hand gestures, with a projector that displays buttons onto a workbench. The projector can also be used for displaying other information, such as assembly instructions at the gaze of the human using eye-tracking technology. The authors also mention another application where the gaze can be used to detect which button the human wants to activate.

Gesture and ASR have been combined in [25] to control an artificial robot with nine navigational commands, such as forward, back, stop, northeast, etc. The paper demonstrates that these technologies can be used for proximate interactions, making them possible in a HRC setting. In this case a Kinect camera is used for both gesture recognition and distant ASR. Using this setup the robustness is greatly improved when combining the two modalities.

Screens have been used to display facial expressions (emotions) [30], to improve the feedback loop to the human. The emotional states of the face can help the operator prioritize which task to execute, guiding the attention of the human. This technology improves the interaction between the human and the robot. However, by itself the technology cannot be utilized and is therefore excluded from the paper.

3. HRC task-based metrics

A new, more sophisticated set of metrics is suggested in this paper for selecting communication technologies in different HRC applications. This metric set is based on how a technology conforms to specific HRC tasks based on the following categories:

- **Extent of usage**, that is, how many HRC tasks that the technology can be used for
- **Flexibility**, that is, how the technology can be extended with more features
- **Duration**, that is, from the time an action starts until it ends
- **Additional classification**, that is, classification of the technology based on how it affects HRC applications

In subchapter 3.1-3.4 these categories are described further. In subchapter 3.5, different communication technologies are summarized based on the four categories.

3.1. Extent of usage

Extent of usage is defined by how many basic tasks that a technology can communicate, the more tasks the higher extent of usage that technology has. Depending on the task, one or several communication messages are needed for the human and robot to collaborate. These messages are categorized into several types based on the information they contain. The message types were derived from the usage of communication technologies in HRI, described in chapter 2, with the mindset to cover all possible HRC tasks. The message types are categorized as follows:

1. **Command messages** communicates what the robot or human should do, e.g., next, reject and stop commands. These messages do not require any real-world information or additional data.
2. **Data messages** communicates data to the human or robot, such as quantity, dimension, strings, etc.. The data could contain, quantity of products to produce, article number, instruction, etc. These messages contain data without real-world information.
3. **Highlighting messages** communicates where in the physical world the robot or human should execute its work. For example, to point out an object to work with, or to visualize from where a component should be collected. These messages require real-world positional information.
4. **Demonstration messages** communicates a continuous work flow of how to execute a specific task, e.g., showing the human or robot how an object needs to be assembled. These messages require real-world positional information with recording of motion.
5. **Guidance messages** communicates how the robot should move to execute its task by physically moving the robot, e.g., teach a motion, move to safe location, calibrate robot, flexible fixture. These messages require a continuous flow of robot and real-world positional/force information
6. **Option messages** communicates to the human what alternative options are available depending on the scenario, e.g., alternative motion constraint, alternative processes. These messages require context information from the current state of the system.

With these message types, the authors believe most of the tasks within HRC applications can be communicated. Therefore, it should be enough to measure the performance based on the tasks instead of a specific application.

Message types 1-4 are suitable for both robots and humans, but the type of communication technology may differ. For example: command messages using audio as communication media may use ASR for robots and TTS for humans. Guiding messages are, however, only suitable for robots, because humans have enough sensory-motor skills and intelligence to know how to move based on highlighting and/or demonstration messages. Therefore, guiding messages for humans are excluded from table with robot-to-human communication technologies. Similarly, option messages are only suitable for humans because the robot already has full knowledge of what can be done in a specific scenario, but the human can be presented with different options to know what he or she can do. Therefore, option messages are excluded from table with human-to-robot communication technologies.

Communicating the identity of the operator is a special case that is important in the industry for traceability. However, technologies developed for identification, e.g., voice recognition, face recognition, RFID tags, can generally not be used for the previously mentioned communication tasks. They may use the same hardware as another technology, but the purpose of the technologies differs so they cannot be used in each other's context.

Depending on the application multiple message types may be necessary to complete a task. In [9] it is demonstrated how positioning of information at humans gaze can be used, which is the combination of human-to-robot highlighting message (gaze of human), and robot-to-human data message (projecting info on workbench).

3.2. Flexibility

Flexibility is defined by whether the physical interface can be used for multiple features within a specific task. The flexibility is classified in four levels based on the findings in chapter 2:

- **Not-applicable –** for technologies that cannot be used in that specific task, e.g. ASR cannot be used for demonstration messages because it cannot contain a recording of motion.
- **Special use-cases –** for technologies that can only be used in few instances, e.g., joint force control can be used to push robot and therefore implying that the robot should continue.
- **Poor flexibility –** for technologies that can be used in a general purpose, but cannot easily be extended to support most features, e.g. gesture recognition can be used for a smaller set of commands, because the human has limited ability to produce gestures.
- **Good flexibility –** for technologies that can easily be extended for most features, e.g. head-mounted AR can be used for most highlighting messages, because it can produce any visual artifact for the human.

3.3. Duration

Duration requires empirical studies to be quantified for a specific task. However, this information can still be estimated based on the findings in chapter 2 using the following classification scheme:

- **Not-applicable –** for technologies that cannot be used in that specific task.
- **Poor duration –** for technologies that can execute the specific task but requires considerable more time in respect to low-duration technologies. E.g. buttons and joystick for guidance messages require considerable more time than using haptic control, as mentioned in chapter 2.
- **Good duration –** for technologies that can execute the specific task, in approximately the same time in respect to low-duration technologies. E.g. Gesture recognition, gaze recognition, and soft-buttons all have equal duration for highlighting messages, because the recognition processing for all these technologies have similar performance.

3.4. Additional classification

In some cases the metrics will produce the same results, even if the hardware changes. To further improve the selection process of technologies, three additional classifications are defined, based on how technologies affect HRC applications:

- **Wearable**, that is, whether the technology requires the user to wear the hardware, which affects requirement of protecting gear
- **Limited coverage**, that is, whether the position of the user or the shape of the workspace/workpiece affects the readability of the message
- **Hand usage**, that is, whether the users hand(s) are necessary to use the technology, which removes hand(s) from work task

These categories do not require quantification measures and are simply stated yes (symbol ✓) or no (without symbol).

3.5. Suggested metrics to be used for selecting communication technologies

To guide the end-user in the selection of communication technologies, the various technologies are classified for each message type based on the scheme presented in Table 1.

Table 1. Scheme used for estimating a technology measurement values based on flexibility and duration.

Meaning	Symbol
No or Not applicable	
Good flexibility and good duration	●
Good flexibility and poor duration	◐
Poor flexibility and good duration	◑
Poor flexibility and poor duration	○
Special use cases	-
Yes	✓

Table 2 and 3 presents the communication technologies discussed in chapter 2 with the classification suggested in the paper, that is, the new metrics. Technologies for human-to-robot are presented in Table 2, while Table 3 presents technologies for robot-to-human communication. Using these two tables, the idea is that an end-user can easily select the proper communication technologies for a specific scenario.

Table 2. Categorization of human-to-robot communication technologies.

	Wearable	Limited coverage	Hand usage	Command messages	Data messages	Highlighting messages	Demonstration messages	Guidance messages
Gesture recognition								
Vision based		✓	✓	◑	-	●	●	○
Glove based	✓		✓	◑	-	●	●	○
Automatic speech recognition								
Head-mounted	✓			●	●			-
Distant		✓		●	●			-
Haptic control								
Joint based			✓	-				●
End effector based			✓	-				◑
Virtual impedance control [1]			✓	-				◑
Gaze recognition								
Head-mounted	✓			-		◑		
Stationary		✓		-		◑		
Buttons/Joystick								
Stationary			✓	◑	◑			◐
Soft-buttons		✓	✓	◑	○	◑		○

[1] Has not been tested

Table 3. Categorization of robot-to-human communication technologies.

	Wearable	Limited coverage	Hand usage	Command messages	Data messages	Highlighting messages	Demonstration messages	Option messages
Augmented reality								
Spatial monitor		✓		●	●	●	◑	◑
Spatial projection		✓		●	●	●	○	◑
Hand-held			✓	●	●	●	●	●
Head-mounted	✓			●	●	●	●	●
Text-To-Speech								
Head-mounted	✓			●	●	○		◐
Freestanding				●	●	○		◐
Pick-by-light								
Lamp based				◑		◑		◑

4. Conclusions

This paper presents state-of-the-art communication technologies for HRC. Shortcomings of current metrics for the selection of communication technologies in HRC have been identified. This paper, therefore, suggests new metrics to classify different communication technologies for use in HRC applications. The new metrics focuses on three characteristics; extent of usage, flexibility, and duration. Extent of usage is measured by how many communication message types a technology can be used for. The message types are divided into six categories; command, data, highlighting, demonstrating, guidance, and option messages. The performance of the technologies when used in each message type is then classified based on flexibility and duration. The communication technologies are additionally classified into wearable, limited coverage, and hand usage to further improve the selection process.

Using the two tables defined in the paper, that cover various technologies for human-to-robot and robot-to-human communication and their various strengths and weakness, the work task of selecting the proper communication technologies for a specific HRC scenario is simplified. The long-term ambition is to extend the results further in the future and eventually provide a comprehensive document that future researchers, developers and integrators can utilize for selecting communication technologies in HRC applications.

This paper uses a classification scheme for determining measurement values, but further work should focus on how to quantify the measurements values. Empirical studies should then be used to evaluate the real-world effectiveness of the metrics.

References

[1] B. Chandrasekaran and J. M. Conrad, "Human-Robot Collaboration: A Survey," *Ieee Southeastcon 2015,* 2015.

[2] C. Lenz and A. Knoll, "Mechanisms and capabilities for human robot collaboration," in *The 23rd IEEE International Symposium on Robot and Human Interactive Communication*, 2014, pp. 666-671.

[3] J. Krüger, T. K. Lien, and A. Verl, "Cooperation of human and machines in assembly lines," *CIRP Annals - Manufacturing Technology,* vol. 58, pp. 628-646, // 2009.

[4] KUKA. (2017, 21st of December). LBR IIWA. Available: https://www.kuka.com/en-se/products/robotics-systems/industrial-robots/lbr-iiwa

[5] ABB. (2017, 21st of December). *YuMi - Creating an automated future together*. Available: http://new.abb.com/products/robotics/yumi

[6] U. Robots. (2017, 21st of December). Collaborative robots from Universal Robots. Available: https://www.universal-robots.com/products/

[7] M. A. Goodrich and A. C. Schultz, "Human-robot interaction: a survey," *Found. Trends Hum.-Comput. Interact.,* vol. 1, pp. 203-275, 2007.

[8] S. A. Green, M. Billinghurst, X. Chen, and J. G. Chase, "Human-Robot Collaboration: A Literature Review and Augmented Reality Approach In Design," *International Journal of Advanced Robotic Systems,* vol. 5, pp. 1-18, Mar 2008.

[9] A. Bannat, J. Gast, T. Rehrl, W. Rosel, G. Rigoll, and F. Wallhoff, "A Multimodal Human-Robot-Interaction Scenario: Working Together with an Industrial Robot," *Human-Computer Interaction, Pt Ii,* vol. 5611, pp. 303-+, 2009.

[10] J. Lambrecht, J. Kr, x00Fc, and ger, "Spatial programming for industrial robots based on gestures and Augmented Reality," in *2012 IEEE/RSJ International Conference on Intelligent Robots and Systems*, 2012, pp. 466-472.

[11] A. Steinfeld, T. Fong, D. Kaber, M. Lewis, J. Scholtz, A. Schultz, *et al.*, "Common metrics for human-robot interaction," presented at the Proceedings of the 1st ACM SIGCHI/SIGART conference on Human-robot interaction, Salt Lake City, Utah, USA, 2006.

[12] R. R. Murphy and D. Schreckenghost, "Survey of Metrics for Human-Robot Interaction," *Proceedings of the 8th Acm/Ieee International Conference on Human-Robot Interaction (Hri 2013),* pp. 197-+, 2013.

[13] J. Scholtz, *Human-Robot Interactions: Creating Synergistic Cyber Forces*, 2002.

[14] S. Rossi, E. Leone, M. Fiore, A. Finzi, and F. Cutugno, "An Extensible Architecture for Robust Multimodal Human-Robot Communication," *2013 Ieee/Rsj International Conference on Intelligent Robots and Systems (Iros),* pp. 2208-2213, 2013.

[15] J. de Gea Fernández, D. Mronga, M. Günther, T. Knobloch, M. Wirkus, M. Schröer, *et al.*, "Multimodal sensor-based whole-body control for human–robot collaboration in industrial settings," *Robotics and Autonomous Systems,* vol. 94, pp. 102-119, 2017/08/01/ 2017.

[16] I. Maurtua, I. Fernandez, A. Tellaeche, J. Kildal, L. Susperregi, A. Ibarguren, *et al.*, "Natural multimodal communication for human-robot collaboration," *International Journal of Advanced Robotic Systems,* vol. 14, Jul 7 2017.

[17] J. Guhl, S. Tung, and J. Kruger, "Concept and architecture for programming industrial robots using augmented reality with mobile devices like microsoft HoloLens," in *2017 22nd IEEE International Conference on Emerging Technologies and Factory Automation (ETFA)*, 2017, pp. 1-4.

[18] A. Syberfeldt, O. Danielsson, and P. Gustavsson, "Augmented Reality Smart Glasses in the Smart Factory: Product Evaluation Guidelines and Review of Available Products," *IEEE Access,* vol. 5, pp. 9118-9130, 2017.

[19] Y. Tabet and M. Boughazi, "Speech synthesis techniques. A survey," in *International Workshop on Systems, Signal Processing and their Applications, WOSSPA*, 2011, pp. 67-70.

[20] D. Rumiński, "An experimental study of spatial sound usefulness in searching and navigating through AR environments," *Virtual Reality,* vol. 19, pp. 223-233, November 01 2015.

[21] R. Reif and W. A. Gunthner, "Pick-by-vision: augmented reality supported order picking," *Visual Computer,* vol. 25, pp. 461-467, May 2009.

[22] A. Cherubini, R. Passama, A. Crosnier, A. Lasnier, and P. Fraisse, "Collaborative manufacturing with physical human–robot interaction," *Robotics and Computer-Integrated Manufacturing,* vol. 40, pp. 1-13, 8// 2016.

[23] L. Roveda, F. Vicentini, N. Pedrocchi, and L. M. Tosatti, "Impedance Control based Force-tracking Algorithm for Interaction Robotics Tasks: An Analytically Force Overshoots-free Approach," *Icimco 2015 Proceedings of the 12th International Conference on Informatics in Control, Automation and Robotics, Vol. 2,* pp. 386-391, 2015.

[24] S. Y. Lo, C. A. Cheng, and H. P. Huang, "Virtual Impedance Control for Safe Human-Robot Interaction," *Journal of Intelligent & Robotic Systems,* vol. 82, pp. 3-19, Apr 2016.

[25] Z. Lei, Z. H. Gan, M. Jiang, K. Dong, and Ieee, "Artificial Robot Navigation based on Gesture and Speech Recognition," *2014 International Conference on Security, Pattern Analysis, and Cybernetics (Spac),* pp. 323-327, 2014.

[26] S. Mitra and T. Acharya, "Gesture Recognition: A Survey," *IEEE Transactions on Systems, Man, and Cybernetics, Part C (Applications and Reviews),* vol. 37, pp. 311-324, 2007.

[27] M. V. d. Bergh, D. Carton, R. D. Nijs, N. Mitsou, C. Landsiedel, K. Kuehnlenz, *et al.*, "Real-time 3D hand gesture interaction with a robot for understanding directions from humans," in *2011 RO-MAN*, 2011, pp. 357-362.

[28] D. Lu, Y. Yu, and H. Liu, "Gesture recognition using data glove: An extreme learning machine method," in *2016 IEEE International Conference on Robotics and Biomimetics (ROBIO)*, 2016, pp. 1349-1354.

[29] M. Simão, P. Neto, and O. Gibaru, "Natural control of an industrial robot using hand gesture recognition with neural networks," in *IECON 2016 - 42nd Annual Conference of the IEEE Industrial Electronics Society*, 2016, pp. 5322-5327.

[30] B. Sadrfaridpour and Y. Wang, "Collaborative Assembly in Hybrid Manufacturing Cells: An Integrated Framework for Human-Robot Interaction," *IEEE Transactions on Automation Science and Engineering,* vol. PP, pp. 1-15, 2017.