Using Voice Assistants as HMI for Robots in Smart Production Systems

Javad Ghofrani[1] and Dirk Reichelt[1]

HTW Dresden
javad.ghofrani@gmail.com, dirk.reichelt@htw-dresden.de

Abstract. Smart voice assistant systems are widely used in the smart home and entertainment domain. Despite of the fact that touch screens are used in industry for giving commands to the robots and production machines, the use of voice or video assistants is still limited to smart home and technologies like Apple Siri or Amazon' Alexa. In this paper, we discuss the feasibility of applying existing digital assistant systems in industrial contexts and various aspects of their usage

Keywords: Robotic, industry 4.0, IIoT, digital assistant, voice recognition

1 Introduction

Using sound activated systems to control various devices such as smart home, BMW cars and also unmanned aerial vehicles [1, 5] is more and more intended. In comparison to mouse and keyboard, touch screen are more flexible since they occupy less space while making it possible to use different layouts in one place. Nevertheless, working with touch screens distracts the user from his main task. Sound based assistant systems compensate this shortage. In these systems, the natural language interfaces create a layer of abstraction and try to hide big amount of complexity to work with system interfaces. The aim of this paper is to propose the application of smart digital assistant for communicating with robots and to prepare a research project for the use of digital, voice-controlled assistance systems for communication with robots in smart production systems.

2 Related Work

Problem of processing natural languages is the first era in which many works have been done. Ortiz [7] investigates and discusses the challenges which the developers should consider when they developing a new interface. Furthermore, Milhorat et al. [6] address the existing challenges in the way of implementing the personal digital assistant which works with voice. Both of these papers mentions the understanding the language, handling the knowledge, and reasoning as important aspects of this field. In the work of Park et al. [8] a speech recognition system is used to send the movement directives to a robot in the unsafe area (fire). They

S. Kolb, C. Sturm (Eds.): 11th ZEUS Workshop, ZEUS 2019, Bayreuth, Germany, 14-15 February 2019, published at http://ceur-ws.org/Vol-2339

mention the problem of natural language processing, noise of the environment and network resilience as problems which can occur during their experiment. Solorio et al. [10] categorize the facing challenges in the implementation of IoT into standards, data management, security, privacy, and the lack of flawless technology in speech recognition. Polyakov et al. [9] investigates the implementation of a voice assistant system which is not equipped with cloud based services but locally exploiting the deep neural networks for analyzing the voice commands. They mention the lack of compatibility of middle-wares, security sphere of use cases, as well as unrealistic thinking about ubiquitous voice assistants as the problems in applying the voice assistant systems without cloud computing. However, cloud computing is not the only possible solution to solve the speech recognition problems. Other solutions are developed to overcome the latency and security issues of solutions based on cloud computing. Fog computing and edge with a lot of smart, connected endpoints are some of these solutions.

Replacing the Human Machine Interface (HMI) with speech commands are the next challenge which is considered in publications and scientific works. Kawai et. al [3] proposes a voice-activated system as a replacement of conventional inputs (mouse and keyboard) to get and put some objects in the 3D environment. It works like a calling a menu and selecting an item from it. While Green [2] investigates the various possible way to talk with the equipment, installed in smart home. Fernández et al. [1], similar to Meszaros et al. [5], investigates the Speech Command Interaction for giving command to the flying devices. They implemented a feedback mechanism which acknowledges the performing the commanded action. In the opinion of Keϊpuska and Gamal [4] a combination of different input channels such as voice, text, image, gesture and etc. are the way that the next generation of smart assistant systems are going to work.

3 Smart Digital Assistants as Interface for Controlling Robots

As mentioned in previous sections, using alternative channels, especially voice, to communicate with robots is not a new field of work. Various works have investigated the application of voice recognition techniques and converting them to the proper commands to communicate with various parts of smart environments, e.g. smart home or smart cars. Hence, the usage of such channels in industrial environments is an attention point by researchers and practitioners. Recent advances in the personal digital assistants and their usage in smart home and driving cars make them a proper candidates to be applied to the industrial environments for controlling robots, especially collaborative robots (cobots). Using smart digital assistant systems, the complexities of voice recognition and natural language processing are pushed to the side of the companies which provide these technologies. Furthermore, the developers are capable of implementing their own skills according to their needs. It bring more flexibility for the case of applying them in industrial environments and solutions. Nevertheless, applying voice recognition in the industry has additional challenge compared to smart

home. Since in production environments, especially involving robots, there is a lot of noise that disturbs the interaction using voice. Furthermore, in industrial environments, there is a need for distributing and orchestrating a larger number of assistants to cover a machine shop or production line. Considering these problems and the problems that we are still not aware of, we are going to analyze feasibility of applying smart digital assistants to the industrial robots. For this purpose, we are going to emphasize potential problems that could occur if we apply such assistant systems in smart production systems. We plan to collect data from various sources such as literature studies, surveys, conducting experiments and interviews in first step. We are going to propose the following research questions: (i) which factors prevent the application of smart assistant systems within the industrial context? (ii) is there any solution, best practice and work-around to overcome these factors? (iii) which points should be considered in implementation of such systems at most? Answering these research questions enables us to determine the specifications of our system.

We propose our notion of a smart assistant systems for controlling the cobots in an smart production system similar to industry 4.0 testbed on HTW Dresden[1]. In such scenarios, there is workplace in which a human works together with a cobot. The cobots are able to perform some tasks more precisely than human, e.g. gluing. In the scenario without a smart voice assistant system, the human should put a certain part in a certain place on the desk and then use flexpendant to select the task–here gluing which is located under tasks menu–and then cobot starts to perform the task by taking the part from specific area of the desk. In other scenario with a smart sound assistant, the user can talk to his micro and tell the robot "YuMi, take the piece from left side of desk and glue it". The keyword "YuMi" will activate the smart assistant system. It processes the sentences after the keyword and sends them to a program that is responsible for controling the cobot. This program should be developed within our project. This program could be similiar to skills of Amazon Alexa that will be activated by some phrases such as "glue". It generates the proper control signals for performing the task and send them to the cobot.

In this project, we will be able to measure the quality of communication in the industrial environments between human and robot. In addition, we are going to investigate the differences in these communication ways beside their advantages and disadvantages. After implementing our solutions, we are going to test it under various circumstances such as noisy place or with some accent of the user. The effect of using various equipment such as Amazon Alexa or IBM Watson will be considered as well. The results of these tests will be used to recognize the challenges and to provide a comparison between different configurations and solutions. In this regard, our work will cover the acceptance of such smart assistant systems in the industrial environments using empirical methods such as qualitative surveys.

[1] https://www.htw-dresden.de/industrie40

References

1. Fernández, R.A.S., Sanchez-Lopez, J.L., Sampedro, C., Bavle, H., Molina, M., Campoy, P.: Natural user interfaces for human-drone multi-modal interaction. In: Unmanned Aircraft Systems (ICUAS), 2016 International Conference on. pp. 1013–1022. IEEE (2016)
2. Green, A.: C-roids: Life-like characters for situated natural language user interfaces. In: Robot and Human Interactive Communication, 2001. Proceedings. 10th IEEE International Workshop on. pp. 140–145. IEEE (2001)
3. Kawai, Y., Higashiyama, Y., Koyama, K., Okada, M.: A fundamental study on a natural-language-based 3d cg modeling. In: Systems, Man, and Cybernetics, 1999. IEEE SMC'99 Conference Proceedings. 1999 IEEE International Conference on. vol. 5, pp. 714–719. IEEE (1999)
4. Këpuska, V., Bohouta, G.: Next-generation of virtual personal assistants (microsoft cortana, apple siri, amazon alexa and google home). In: Computing and Communication Workshop and Conference (CCWC), 2018 IEEE 8th Annual. pp. 99–103. IEEE (2018)
5. Meszaros, E.L., Chandarana, M., Trujillo, A., Allen, B.D.: Speech-based natural language interface for uav trajectory generation. In: Unmanned Aircraft Systems (ICUAS), 2017 International Conference on. pp. 46–55. IEEE (2017)
6. Milhorat, P., Schlogl, S., Chollet, G., Boudy, J., Esposito, A., Pelosi, G.: Building the next generation of personal digital assistants. In: Advanced Technologies for Signal and Image Processing (ATSIP), 2014 1st International Conference on. pp. 458–463. IEEE (2014)
7. Ortiz, C.L.: The road to natural conversational speech interfaces. IEEE Internet Computing (2), 74–78 (2014)
8. Park, S., Kim, Y., Matson, E.T., Jang, H., Lee, C., Park, W.: An intuitive interaction system for fire safety using a speech recognition technology. In: Automation, Robotics and Applications (ICARA), 2015 6th International Conference on. pp. 388–392. IEEE (2015)
9. Polyakov, E., Mazhanov, M., Rolich, A., Voskov, L., Kachalova, M., Polyakov, S.: Investigation and development of the intelligent voice assistant for the internet of things using machine learning. In: Electronic and Networking Technologies (MWENT), 2018 Moscow Workshop on. pp. 1–5. IEEE (2018)
10. Solorio, J.A., Garcia-Bravo, J.M., Newell, B.A.: Voice activated semi-autonomous vehicle using off the shelf home automation hardware. IEEE Internet of Things Journal (2018)