Designing a Multimodal Human-Robot Interaction Interface for an Industrial Robot

Chapter · August 2016

DOI: 10.1007/978-3-319-21290-6_26

3 authors:

Bogdan Mocan
Universitatea Tehnica Cluj-Napoca

Mircea Fulea
Universitatea Tehnica Cluj-Napoca

Stelian Brad
Universitatea Tehnica Cluj-Napoca

Designing a Multimodal Human-Robot Interaction Interface for an Industrial Robot

Bogdan Mocan[1], Mircea Fulea[1] and Stelian Brad[1]

[1]Technical University of Cluj-Napoca, Dept. of Engineering Design and Robotics, Romania

bogdan.mocan@muri.utcluj.ro; mircea.fulea@staff.utcluj.ro; stelian.brad@staff.utcluj.ro

Abstract. This paper presents a framework for a multimodal human-robot interaction. The proposed framework is intended to bring important contributions to the development of human robot interaction in order to facilitate intuitive programming and to enable easily adapting to changes in robot task without the need of using skilled personnel. The key elements of this system are speech and hand gesture recognition, text programming, and interaction capabilities that allow the user to take over the control of the robot at any given time. Furthermore, our approach is focus on robot task. A user can express his/her preference for one or more modalities of interaction so that selected modalities fit user's personal needs.

Keywords: multimodal user interfaces · industrial robots · programming methods · multimodal interaction

1 Introduction

Over the last decades the industrial robots have become more powerful and intelligent. Thus, in many cases an investment in industrial robots is seen as a vital step that will strengthen a company's position in the market because it will increase the production rate and the process efficiency and it will reduce the operating costs [1]. Industrial robots based automation represents the best solution for both productivity and flexibility [2, 3]. However, in small and medium enterprises (SMEs) robots are not commonly found. Even though the hardware cost of industrial robots has decreased, the integration and programming costs for robots make them unaffordable for SMEs [4, 5]. It is, thus, quite difficult to motivate a SME, which is constantly under market pressure, to carry out a risky investment in robots [2], [6,7]. Typically for those SMEs, that have frequently changing applications, it is quite expensive to afford professional programmers or technicians, and therefore a human robot interaction solution is demanded [8,9]. Therefore, today's industrial robots do not offer rich human-robot interaction (multimodal interaction), are not simple to program for end-users, and the programming procedures are time consuming [10, 11].

The traditional online robot's programming can be done in three ways: (i) jogging an

industrial robot with 6 degrees of freedom with a joystick with two degrees of freedom is very time consuming and cumbersome; (ii) the operator doesn't get any visual feedback of the process result before the program has been generated and executed by the robot; (iii) many iterations are needed for even the simplest task [6], [11].
Offline programming environments like RobotStudio from ABB Company solve some disadvantages described above. But also the off-line programming software presents several problems in many industry applications, particularly, when the robot task or the robot trajectory needs frequent change [7].
On the other side, multimodal interfaces allow users to move effortless between different modes of interaction, from visual to speech and touch, according to changes in context or user preference [8]. These interfaces have the advantage of increased usability and accessibility [12, 13]. In multimodal interfaces, the weaknesses of one modality can be offset by the strengths of another [12]. Accessibility determines how easy it is for people to interact with the robot. Thus, multimodal interfaces can increase robot task efficiency, though perhaps not significantly, as pointed out by reference [14].
This paper propose an approach that enable the development of a multimodal interface that facilitate an intuitive programming of an industrial robot. Our goal is to give an industrial robot the ability to communicate with its human operators in a more intelligent way, thus making the programming of industrial robots more intuitive and easy.
This paper is structured on five sections. Following this introduction, section two describes the related works regarding multimodal interfaces for programming the industrial robots. Section three emphasis the design process and the interface architecture that lead to an intuitive programming of an industrial robot. In section four is described a software demonstrator for multimodal interaction and finally, the conclusions are discussed in fifth section.

2 Related work

Even though the literature on multimodal systems approach is still scarce, various studies have shown that multimodal interfaces may be preferred by users over unimodal alternatives, because they can offer better flexibility and reliability and interaction alternatives to better meet the needs of diverse users with a range of usage patterns and preferences [15, 16, 17]. Since the introduction of the "Media Room" concept by reference [18], many other systems have been developed based on multimodal interaction with the user. Researchers have established different methods for implementation of such systems [19, 20, 21]. All multimodal systems are in a way similar in the sense that they receive inputs from different devices/equipment and combine the information to build a common semantic meaning from these inputs [16].
Finite-state multimodal parsing has been studied by researchers within reference [22] and they present a method to apply finite-state transducers for parsing inputs. In reference [23] speech modality alone has been used to command an industrial robot through switching between preprogrammed tasks.
Reference [24] pointed out that coupling of speech recognition and dialog management (multimodal approach) can improve the performance of a system. A good study on

incremental natural language processing and its integration with a vision system can be found in reference [25] and also in reference [6], [26]. A multimodal interaction scheme is very convenient for robot programming because combining two or more interaction modalities can even provide improved robustness [27]. Reference [28] highlights the possibility of combining the speech and static gestures in order to program a robot for grasping and manipulate an object.
Augmented reality (AR) also provides great opportunities for Human Robot Interaction (HRI), and has been widely used in tele-robotics because AR allows the operator to work as if he is present at the remote working environment [9], [29, 30]. Through tablet PCs and head mounted displays it is possible to visualize and generate robot paths through a pointing device [30]. In their work related in reference [30] they visually tracked marker to define collision-free paths for the robot. Once the path is generated a virtual robot simulates the behavior of the robot on the screen.
Also references [6] and [30] combined both augmented and virtual reality environments together with higher level voice commands to remote operate an industrial robot.
Analyzing the available literature it can be seen that the majority of research efforts are focused on providing a suitable robotic programming method for SMEs, by improving the existing programming methods (e.g. online and offline programming methods) or combining different programming modalities in so called multimodal robot interaction using new concepts such as Augmented Reality.
Progress in online programming of an industrial robot is largely based around sensor and control technologies to assist the operator in creating complex robot motion more easily. Integrating the Augmented Reality with offline programming software is originated from the idea of making human-robot interaction more interactive and flexible. With the development of more powerful CAD/CAM/PLM software, robotic vision, sensor technology, etc., new programming methods, like multimodal interactions, suitable for SMEs are expected to be developed in years to come.

3 Competitive design of the multimodal interaction system

Designing multimodal systems for industrial robots is challenging. A set of multimodal myths have proven to be useful in designing multimodal systems. As it was pointed out in reference [32]. The classic design approaches and insights from personal computer environments do not necessarily translate well to industrial robots multimodal environments. Designing a multimodal interface for industrial robots, have to focus on robot tasks and operator needs. The set of tasks that an industrial robot can perform are directly affected by the robot pose, namely by the robot configuration at one moment [3]. Specifically, the followings are the challenges of the industrial robot programming: 1) *Obstacle avoidance*: in case that a robot approaches to an obstacle, the operator needs to modify a teaching position; 2) *Joint limit avoidance*: for example, when it reaches an axis limit of a wrist or when occur a singularity pose of the robot, the operator has to take out the robot from that position; 3) *Model task specific robot trajectories*: operator has to develop the logic of the task and to integrate adequately target points within the robot trajectory.

Having the challenges regarding classical robot programming by using competitive design principle and methods like AHP (Analytical Hierarchy Process), PUGH (Pugh Method), TRIZ and QFD (Quality Function Deployment) we developed a methodology (Figure 1) for identifying the right operators requirements, process needs, objective functions, and the best combination of the multimodal interface inputs.

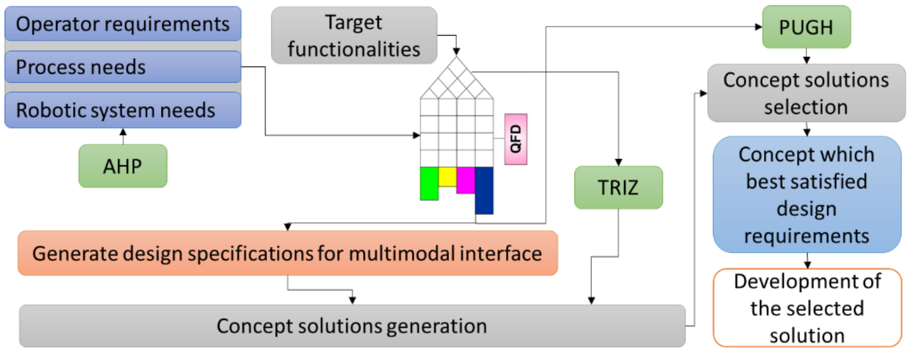

Fig. 1. Roadmap for selecting the inputs for the multimodal interface

The proposed methodology for selecting the inputs for the multimodal interface consists of the following steps (Figure 1):
Step 1: Define operator requirements, process needs and robotic system needs under the form of a set of requirements and rank these need-related requirements. For ranking, tool like AHP method might be used.
Step 2: Define a set of target functionalities in accordance to the intuitive programming needs and deploy them against the set of need related requirements. Thus, value weights of target functionalities in relation to the multimodal programming needs are determined. Correlations between target functionalities have to be also established. QFD-type relationship and correlation matrices could be used to fulfil this step.
Step 3: Formulate vectors of innovation for each negative correlation between target functionalities and for each challenging target. TRIZ method is a powerful tool to fulfil this process. The resulted vectors of innovation represent paths towards which creativity and skills must be directed when concept solutions (minimum three concept solutions) are elaborated.
Step 4: Formulate design specifications for multimodal interface having in mind the needed target functionalities.
Step 5: Generate minimum three concept solutions for the multimodal interaction. Inputs from TRIZ method are expected to be integrated within the generated solutions.
Step 6: Evaluate the solutions that were generated at step 5 and select the solution that best satisfies the planed performance for the multimodal interaction interface (see step 2). Pugh method could be used to fulfil this step.
Step 7: Result from step 6 is used for further development (detailed design and planning at component level). Use-cases, modelling languages (e.g. UML [1]) and other specific tools for software analysis and design could be used to support this step.
The identified operator requirements, process needs and robotic system needs for a

multimodal interface that will facilitate the intuitive programming of a robot, were: a) automatic presentation of contextually-appropriate information (18%); b) easy to use by an operator (25%); c) easy to set up (16%); d) graphical intuitive interface (24%); e) clear indication on the robot's display the logical next step (17%). The percentage written within the brackets represent the rank for each requirement, identified with the help of AHP method. The selected inputs for the multimodal interface were speech and hand gesture recognition, backed up by text programming.
Based on design guidelines and results highlighted above the proposed framework/ architecture of multimodal interface is composed of four functional modules, as illustrated in Figure 2. The first module (multimodal interaction) translates hand gestures and voice command into a structured symbolic data stream. The second module (actions interpretation) selects the appropriate set of primitives based on the user input, current state, and robot sensor data. The third module (prioritized execution) selects and executes primitives based on the current state, sensor inputs, and the task given by the previous step. Finally, the fourth module facilitate the translation of the actions, voice command and gesture into instructions into robot native programming language to be integrated within a robot programming task. The multimodal interaction between human operator and robot have to be supported by an intuitive graphical interface.

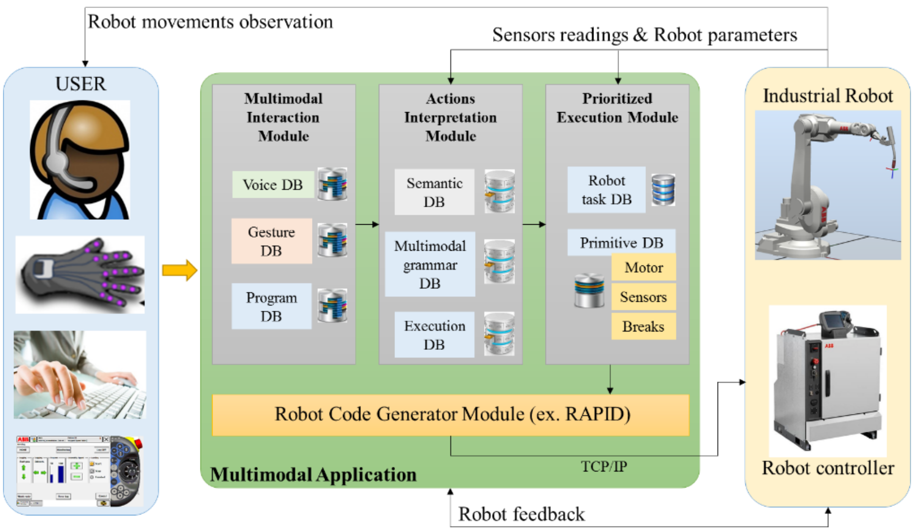

Fig. 2. The multimodal interface framework

4 Software demonstrator for multimodal interface

We have developed a software demonstrator (Figure 2) to exemplify the effectiveness of the multimodal interaction in robotic systems. The demonstrator is still under development, until this moment we have established the communication with the robot controller IRB ABB 1600 (S5) and we was able to control the opening and closing of a gripper mounted on the robot by the hand gesture. The demonstrator was also able to

control the robot by text programming directly from your computer. The demonstrator (multimodal interface) has the capability to interrupt commands preemptively. The demonstrator facilitated verification the operating mode through sequential programming and demonstrate effectiveness of client-server communication through ABB TCP/IP facility. The client server application allows one user at a time to access the server in order to transmit the coordinates of a point inside the robots workspace, which the robot is needs to reach. Also the client can specify what kind of motion type the robot will perform into reaching that point, having the options: linear, joint of circular.
Using the multimodal interface (software demonstrator) the operator instantly can program the robot, using graphical interface for text programming and sensor glove for gesture commands (Figure 3). This approach will enable production engineer to focus on the manufacturing requirements, rather than to the robot programming issues.

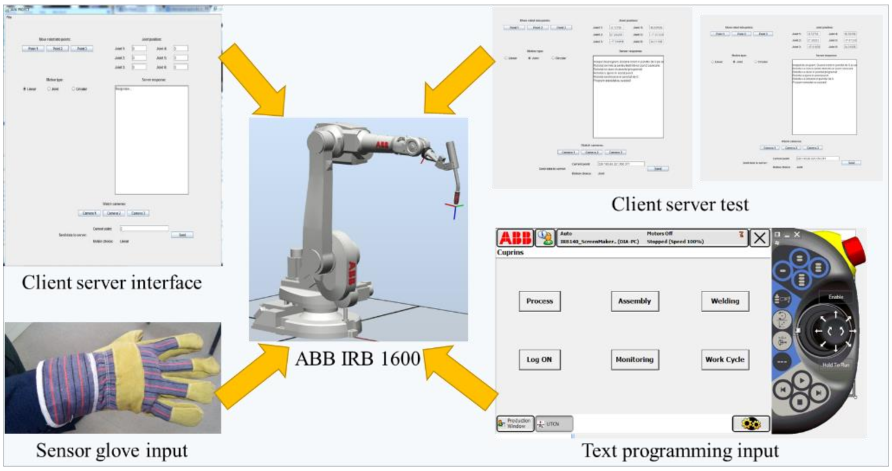

Fig. 3. Multimodal interaction demonstrator

5 Conclusions and further works

In this paper, we have described an overall architecture for interactive multimodal industrial robot programming interface and have illustrated the framework using one demonstrator. The programming approach offers, through an intuitive interface using intuitive text programming, hand gestures and speech recognition, the ability to provide interactive feedback to the industrial robot to coach it throughout the programming and execution phases. The user's intent is caught in the form of a sequential robot program, and the flexibility given to the user by the framework through real-time interaction and an intuitive interface allows the captured intent to be closer to the user's real intent.
The demonstrator verified the communication with the robot controller and the effectiveness of the text and partial hand gesture programming (till now we are able to open and close a gripper by the sensor gloves) within multimodal programming and execution approach, including the capability to interrupt commands preemptively.

To attain a comprehensive multimodal interface robot programming system, several elements still need to be further developed and others to be added in the future. We have to develop more the multimodal recognition module, to integrate speech interpretation module and to test more the actions interpretation module and the prioritized execution module. Although the programs generated by the current version can be re-executed, they are limited to simple robot task sequences. To expand the generality of the paradigm, we need to add the ability to define complex structures such as conditional logics and looping.

Acknowledgment. Support within the project POSDRU/159/1.5/S/137516 PARTING is acknowledged with gratitude.

References

1. Mocan, B., Fulea, M., Brad, S.: Framework for Developing a Multimodal Programming Interface Used on Industrial Robots, Robotica&Management, Vol. 19, No. 2/2014, pp. 29-33 (2014)
2. Fulea, M., Popescu, S., Brad, E., Mocan, B., Murar, M.: A Literature Survey On Reconfigurable Industrial Robotic Work Cells, Proceedings of the ROBOTICS 2014 - VIth International Conference on Robotics , Bucharest, (2014)
3. Mocan, B., Buna, D., Fulea, M., Brad, S.: Increasing the efficiency of robotic manufacturing systems by layout optimization, Proceedings of the ROBOTICS 2014 - VIth International Conference on Robotics , Bucharest, (2014)
4. Kongn, J., Zhang, W.Y., Yu, N., Xia, X.J.: Design of human-centric adaptive multimodal interfaces, Int. J. Human-Computer Studies 69, 854–869, (2011)
5. Brad, S., Mocan, B., Brad, E., Fulea, M.: Leading Innovation to Improve Complex Process Performances by Systematic Problem Analysis with TRIZ, TFC2014, Global Innovation Convention, EPFL, Lausanne, Procedia CIRP, Elsevier, (2014)
6. Akan, B., Ameri, A., Curuklu, B., Asplund, L.: Intuitive Industrial Robot Programming Through Incremental Multimodal Language and Augmented Reality, 2011 IEEE International Conference on Robotics and Automation, Shanghai, pp. 3934-3939 (2011)
7. Mocan, B., Fulea, M., Brad, E., Brad, S.: State-of-the-Art and Proposals on Reducing Energy Consumption in the Case of Industrial Robotic Systems, Proceedings of the 2014 International Conference on Production Research – Regional Conference Africa, Europe and the Middle East, Cluj-Napoca, Romania, 1-5 July, pp. 328-334, (2014)
8. Turk, M.: Multimodal interaction: A review, Pattern Recognition Letters 36, pp. 189-195 (2014)
9. Hirzinger, G., Bals, J., Otter, M., Stelter, J.: The DLR-KUKA success story: Robotics research improves industrial robots, IEEE Robotics and Automation Magazine 12 (3), 2005, pp. 16-23, (2005)
10. Mocan, B.: Performance planning of arc welding robotic systems using specific tools for quality planning and systematic introduction of innovation - part I, II, III, Calitatea - acces la succes Vol. 13, nr. 127 April 2012, pp. 80-85, (2012)
11. Mocan, B., Fulea M.: Offline Programming of Robotic Arc Welding Systems, Acta Technica Napocensis, S: Applied Mathematics and Mechanics, vol. 1, Issue 54, pp. 173-178, (2011)
12. Brick, T., Scheutz, M.: Incremental natural language processing for HRI. Proceeding of the ACM/IEEE Int. Conference on Human-Robot Interaction – HRI ’07, pp. 261-265, (2007)
13. Brad, S., Fulea, M., Brad, E., Mocan, B.: Smart deployment of demonstrators into successful

commercial solutions, Proceedings of the 24th CIRP Design Conference, Milan, Italy, March 14th -16th, pp. 165-175, (2014)
14. Horst, P., Schreck, G. Willnow, C.: SME-Service networks for cooperative operation of robot installations, Int. Conf. Emerging Solutions for Future Manufacturing Systems, February, 2006, pp. 339-346, (2006)
15. Marin, R., Sanz, P., Nebot, P., Wirz, R.: A Multimodal Interface to Control a Robot Arm via the Web: A Case Study on Remote Programming, IEEE Transactions on Industrial Electronics, 52(6):1506–1520, (2005)
16. Dumas, B., Lalanne, D., Oviatt, S.: Multimodal interfaces: a survey of principles, models and frameworks, Human Machine Interaction. Lect. Notes Comput. Sci. 5440, 3–26, Springer, (2009)
17. van Wassenhove, V., Grant, K.W., Poeppel, D.:Visual speech speeds up the neural processing of auditory speech. Proc. Nat. Acad. Sci. 102, pp. 1181–1186, (2005)
18. Bolt, R.A.: Put-that-there: Voice and gesture at the graphics interface. In International Conf. on Computer Graphics and Interactive Techniques, vol. 14, pp 261-264, (1980)
19. Xiao, B., Girand, C., Oviatt, S.L.: Multimodal integration patterns in children, in International Conference on Spoken Language Processing, pp. 629–632, (2002)
20. Oviatt, S., Lunsford, R., Coulston, R.: Individual differences in multimodal integration patterns: what are they and why do they exist? In ACM SIGCHI Conference on Human Factors in Computing Systems, vol. 2, No. 7, pp. 241–249, (2005)
21. Bohus, D., Horvitz, E.: Facilitating multiparty dialog with gaze, gesture, and speech. In: ACM International Conference on Multimodal Interfaces (ICMIMLMI), Beijing, (2010)
22. Johnston, M., Bangalore, S.: Finite-state multimodal parsing and understanding, Proceedings of the 18th conference on Computational linguistics, pp. 369-375, (2000)
23. Pires, J.N.: Robotics for small and medium enterprises: control and programming challenges. Industrial Robot, 2006
24. Fuegen, C., Holzapfel, H., Waibel, A.: Tight coupling of speech recognition and dialog management-dialog-context dependent grammar weighting for speech recognition. In Eighth International Conference on Spoken Language Processing, pp. 2-5, (2004)
25. Schlangen, D., Skantze, G.: A general, abstract model of incremental dialogue processing. In European Chapter Meeting of the ACL, pp. 710-718, (2009)
26. Hsiao, K.-Y., Vosoughi, S., Tellex, S., Kubat, R., Roy, D.: Object schemas for responsive robotic language use. Proceedings of the 3rd International Conference on Human Robot Interaction - HRI '08, pp. 231-235, (2008)
27. Iba, S., Paredis, C.J.J., Khosla, P.K.: Interactive multi-modal robot programming. Proceedings 2002 IEEE International Conference on Robotics and Automation, pp. 161-168, (2002)
28. McGuire, P.C., Fritsch, J., Steil, J.J., Roethling, F., Fink, G.A., Wachsmuth, S., Sagerer, G., Ritter, H.: Multi-Modal Human-Machine Communication for Instructing Robot Grasping Tasks, pp. 1082-1089. Int. Conference on Intelligent Robots and Systems (IROS), (2002)
29. Fang, H., Ong, S. K., Nee, A.Y.-C.: Robot Programming Using Augmented Reality. IEEE, Sept. (2009)
30. Jara, C.A., Candelas, F.A., Gil, P., Fernandez, M., Torres, F.: An augmented reality interface for training robotics through the web. Communication, pp. 189-194, (2005)
31. Marin, R., Sanz, P., Sanchez, J.: A very high level interface to teleoperate a robot via Web including augmented reality. Proceedings 2002 IEEE International Conference on Robotics and Automation, (May):pp. 2725-2730, (2002)
32. Oviatt, S.: Ten myths of multimodal interaction, Commun. ACM 42 (11), 74-81, (1999)
33. Quek, F.: Toward a Vision-Based Hand Gesture Interface, Virtual Reality System Technology Conference, Singapore, pp. 17-29, (1994).