USING SPEECH RECOGNITION, HAPTIC CONTROL AND AUGMENTED REALITY TO ENABLE HUMAN-ROBOT COLLABORATION IN ASSEMBLY MANUFACTURING

RESEARCH PROPOSAL

Patrik Gustavsson
patrik.gustavsson@his.se

2016-05-15

University of Skövde

Volvo Cars Corporation

SUPERVISORS

Anna Syberfeldt, anna.syberfeldt@his.se
Lihui Wang, lihuiw@kth.se
Rodney Lindgren Brewster, rodney.lindgren.brewster@volvocars.com

ABSTRACT

In recent years robots have become more adaptive and aware of the surroundings which enables them for use in human-robot collaboration. By introducing robots into the same working cell as the human, then the two can collaborate by letting the robot deal with heavy lifting, repetitive and high accuracy tasks while the human focuses on tasks that needs the flexibility of the human. Collaborative robots already exists today in the market but the usage of these robots are mainly to work in close proximity.

Usually a teaching pendant is used to program a robot by moving it using a joystick or buttons. Using this teaching pendant for programming is usually quite slow and requires training which means that few can operate it. However, recent research shows that there exist several application using multi-modal communication systems to improve the programming of a robot. This kind of programming will be necessary to collaborate with a robot in the industry since the human in a collaborative task might have to teach the robot how to execute its task.

This project aims to introduce a programming-by-guidance system into assembly manufacturing where the human can assist the robot by teaching the robot how to execute its task. Three technologies will be combined, speech recognition, haptic control, and augmented reality. The hypothesis is that with these three technologies an effective and intuitive programming-by-guidance system can be used within the assembly manufacturing industry. This project have three main motivators: Allowing workers, with no robot programming expertise, to teach the robot how to execute its task in an assembly manufacturing system; Reducing the development time of the robot by introducing advanced programming-by-guidance technology; Showing that augmented reality can add additional information that is useful when programming the robot.

CONTENTS

1 INTRODUCTION

In recent years robots have become more adaptive and aware of the surroundings which enables them for use in human-robot collaboration. By introducing robots into the same working cell as the human, then the two can collaborate by letting the robot deal with heavy lifting, repetitive and high accuracy tasks while the human focuses on tasks that needs the flexibility of the human. In order to quickly adapt to changes in the production, be it the introduction of new products or the personification between different operators, the operator should be able to teach the robot how the process should be done, similar to when a coach instructs other workers how to do a specific task. However, there are some limitations to how the operator can instruct a robot, usually this is done using a teaching pendant to instruct the robot where to find the different parts, work objects, etc. The main problem with the teaching pendant is that it requires training and only a few can operate it.

Programming-by-guidance is a concept that has been introduced in several papers, that shows that programming of an industrial robot can be made more efficient by introducing multi-modal communication systems, that is more natural for human beings. So far these papers have mainly introduced these technologies as another programming language, instead of using regular teaching pendants. This project will first explore the possibilities of programming-by-guidance using the three technologies speech recognition, haptic control, and augmented reality, then apply it into an assembly manufacturing system where this programming-by-guidance of the robot can be utilized in the normal production.

This project is funded by University of Skövde, Knowledge Foundation and Volvo Cars Corporation located in Skövde.

1.1 AIM AND OBJECTIVES

This project aims to reduce the gap of collaborating with industrial robots by introducing a new type of programming-by-guidance into the assembly manufacturing. Current collaborative robots are mostly used to work in close proximity with the human in an industrial setting, but in order to fully be able to collaborate with the robot they need to assist each other. Therefore, the programming-by-guidance will be one step on the way where the human can assist the robot by teaching the robot how to execute its task. This project will combine three technologies, speech recognition, haptic control, and augmented reality to achieve an effective and intuitive programming-by-guidance system that can be used within the assembly manufacturing industry. The project have three main motivators:

- Allowing workers, with no robot programming expertise, to teach the robot how to execute its task in an assembly manufacturing system.
- Reducing the development time of the robot by introducing advanced programming-by-guidance technology.
- Showing that augmented reality can add additional information that is useful when programming the robot such as:
    - Coordinate systems of work objects and tools
    - Path of the robot for collision detection
    - Robots possible movements when moving the robot by hand

In order to achieve the aim for this project five objectives have been defined that step by step add up to the programming-by-guidance system. First the three technologies needs to be researched to find state-of-the-art solutions that will function in an assembly manufacturing setting for human-robot collaboration. This will add up to the complete artifact of the three technologies which then can be applied on a use-case within assembly manufacturing. From these criteria the following five objectives have been defined:

1. Identify and implement appropriate speech recognition systems that can function in industrially noisy environments such as assembly manufacturing.
2. Identify and implement haptic control systems in order to jog/move the robot in different ways, covering all the movement types of a regular teaching pendant, and possibly adding more movement types and functionalities.
3. Identify and implement appropriate augmented reality interfaces which can show additional information onto the real world, in order to improve the usability of the speech recognition and the haptic control.
4. Define human-robot collaboration use-cases for assembly manufacturing, based on the functionalities of the three technologies from objectives 1-3, and select one specific use-case.
5. Develop a system that includes all three technologies identified in objectives 1-3, ensuring the necessary functionalities for the selected use-case in objective 4. Then evaluate the system's functionalities, performance, usability and reliability.

1.2 SCOPE AND LIMITATIONS

This project will focus its research on rotary joint controlled robots, specifically six axis robots because they have six degrees of freedom with redundant joints. Robots based on prismatic joints such as Cartesian arms have less complexity. Therefore, when the technology for robots with rotary joint have been established, it will be easier to apply it for robots with prismatic joints.

The robots used within this project will be limited to lightweight robots that has been certified as collaborative robots. This is because the project aims to implement this programming-by-guidance system in an industrial environment which would require to certify the safety of non-collaborative robots, which is not part of this project's scope.

1.3 STRUCTURE OF REPORT

This report follows the structure defined by "Writing a project proposal" in "Real world research" [1] and is adapted to include the chapters necessary as defined by the regulations of research in the University of Skövde [2].

Section 2 describes the theoretical background of the research with focus on the gaps that this research intend to fill. Section 3 describes the overall plan of the project including methodology, time plan, publication plan and expected results.

2 BACKGROUND AND PURPOSE

This section describes the current technology for human-robot collaboration, the different techniques intended for this research project and the gaps that this research intend to fill.

2.1 THEORY

Human-robot collaboration is a research area that has been ongoing for several years and has major interests in the academic community and the industry [3, 4]. The idea of humans and robot working together will open up many possibilities, especially within the assembly industry increasing the flexibility, adaptability and reusability of assembly systems [5]. In today's market there are already several industrial robotic manufacturers that provide collaborative robots such as the UR3, UR5, UR10 series from universal robot, YuMi robot from ABB, LBR IIWA from KUKA and baxter from rethink robotics. The use of collaborative robots are still limited in the industry and mostly used to work in close proximity, but recent research have shown promising results by using these robots for more collaborative work such as [6, 7]. There are, however, other robotic system in the industry [5] used to improve the ergonomics for the humans. These kind of robotic systems originated from the work of [8], who introduced the concept of cobots, assisting robotic systems used for the automotive assembly line.

Speech recognition is a heavily researched topic that has several usages for smartphones, home automation, medical documentation, car interactions, games, etc. Speech recognition can be used both in close proximity, typically head mounted microphones, and on distance, microphone arrays such as the microphones used in Kinect for gaming. When the raw signal enters a speech recognition system the signal first goes through a front end processing unit which aims at reducing undesired noise and outputs feature vectors which describes the input speech [9]. Then the feature vectors are decoded using acoustic models, dictionary and language model which produces the best word sequence. The best word sequence can then be used for different features such as speech-to-text, executing commands, etc.

Haptic control is the use of haptic feedback technology in order to control a device, in this research that device is the collaborative robot. Haptic feedback is generally divided into two areas tactile and kinesthetic. Tactile refers to the feeling of touch, vibration, pressure e.g. the tissue of the hand feels the surface of an object. Kinesthetic refers to sensing the location, velocity and weight based on force, angles and positions of joints (proprioceptors) e.g. feeling the weight and approximate geometry in the hand based on the muscles exertion and angles of joints.

Augmented reality is a topic that has received more and more attention the last two decades. Augmented reality is used to add computer generated information onto the real world [10, 11]. For the user it appears as the virtual objects and real objects coexists in the same space. Augmented reality usually consist of two main technologies, display and tracking. The display can be head mounted (video or optical see through), hand held (e.g. smartphone, tablet), or spatial display (computer screen). The tracking technology can be vision based (live video, depth camera), or sensor based (optical sensors). With the tracking technology the computer can apply information on the right location in the display.

2.2 RELATED WORK

The regular teaching method is tedious and time consuming while requiring considerable technical expertise which becomes a problem for the unskilled worker [3]. Multiple research projects have been done to combine different sensor technology in order to efficiently program an industrial robot, [12] uses a force sensor that can be used to track the force put onto the tool, [13] created a system for aluminum wheel deburring, with vision, force and position sensors which allowed the robot to follow a marked line. Another research [6] shows a programming-by-guidance approach by using the robot manipulator as a kinesthetic haptic control, based on the proprioceptive measures (joint positions, joint velocities and motor torques). Augmented reality has been used in a robot programming system where it was combined with gesture control in [14] this paper showed that programming with augmented reality and gesture based control can potentially reduce the programming time of an industrial robot.

2.3 PURPOSE

All recent projects within this area have partly researched the usage of different sensory techniques for robot programming and some for augmented reality. But the missing part is the combination of all three technologies, speech recognition, haptic control and augmented reality for programming-by-guidance. This project focuses on whether this programming-by-guidance can be used in the context of human-robot collaboration for an assembly manufacturing situation.

3 THE PLAN OF WORK

This section describes the work that is planned for this research project including, methodology with research strategy, overall time plan, publication plan and what results that are expected after completing this project.

3.1 METHODOLOGY

This research follows the design and creation strategy defined by Oates, B.J. [15], which is the research focusing on developing artifacts within information technology. The types of artifacts within information technology are constructs, models, methods and instantiations [16]. This project aims to develop an artifact that demonstrates the feasibility and performance of using the three technologies for programming-by-guidance, which is an instantiation artifact. It is argued that novel instantiations are simple extensions of novel constructs, models and methods. However, it is realized within computer science that instantiations are the real proof and evidence that the underlying constructs, models and methods work.

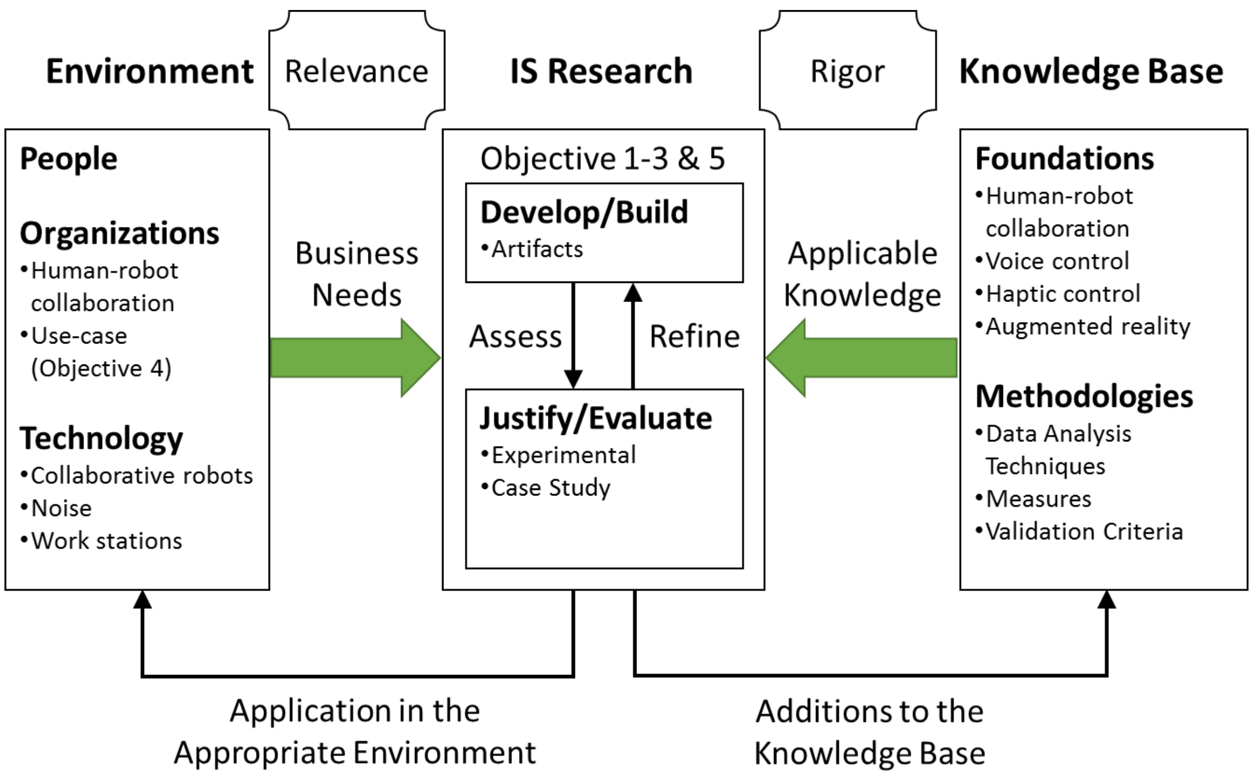


FIGURE 1: THE INFORMATION SYSTEM RESEARCH FRAMEWORK AS DEFINED BY HEVNER ET AL. [17], ADAPTED TO THIS RESEARCH BY SHOWING WHERE THE OBJECTIVES OF THIS RESEARCH FITS INTO THE FRAMEWORK.

The project follows the information system research framework, as defined by Hevner et al. [17], as the model for understanding, executing and evaluating this research. This framework combines behavioral-science and design-science paradigms in order to compare and position them in the research. Behavioral-science focuses more on the development and justification of theories to meet business needs while design-science focuses on building and evaluating artifacts designed to meet the business needs. Hevner et al. [17] argues that these paradigms are inseparable i.e. that building and evaluation of artifacts informs theory while development and justification of theories informs design for artifacts.

Figure 1 shows the overall definition of the Information System Research Framework defined by Hevner et al. [17], with specific information related to this research and its objectives. This framework clarifies how different parts of a research relates to each other. The environment specifies the need for an artifact, and it is necessary to consider the needs to attain a relevance of the research. The knowledge base constitutes the foundation and methodologies to execute research, and by appropriately applying existing knowledge the research becomes rigor. Information System Research is where new artifacts and theories are developed and evaluated in order to contribute to the knowledge base while at the same time satisfy the environment.

The process of executing a design and creation strategy is not straightforward but more like an iterative cycle [15, 17]. Oates, B.J. [15] explains this iterative process: development of an idea leads to increased understanding, thinking about a tentative solution increases the awareness of the problem, failure leads to new insights, etc. This iterative cycle is learning by making which is the incorporated process of this research illustrated in figure 1, which starts by exploring the different technologies in objectives 1-3. When each technology and their potential in human-robot collaboration is fully understood, then the final programming-by-guidance system can be finalized in objective 5.

The environment defines the goals, tasks, problems and opportunities which defines the business needs. These needs are formed based on the perception of the people within the organization. They are assessed and evaluated based on the organizational strategies, structure, culture and processes. The needs are defined in relation to existing technologies, limitations and possibilities. From the start of this project the organizational strategy or the trend of the industry points to human-robot collaboration as illustrated in figure 1. When the three technologies have been fully explored then a use-case can be defined which is based on the business needs.

The knowledge base defines the foundations and methodologies which is the building blocks used to execute information system research. The foundations consists of theories, frameworks, instruments, constructs, models, methods and instantiations used in the developing and building phase. Methodologies on the other hand provides guidelines of how to justify and evaluate the artifact or theory. From the start the current state-of-the art technology is known within the different areas, but as the project goes on creating artifacts and possibly new theories, then this project will iteratively add to the knowledge base.

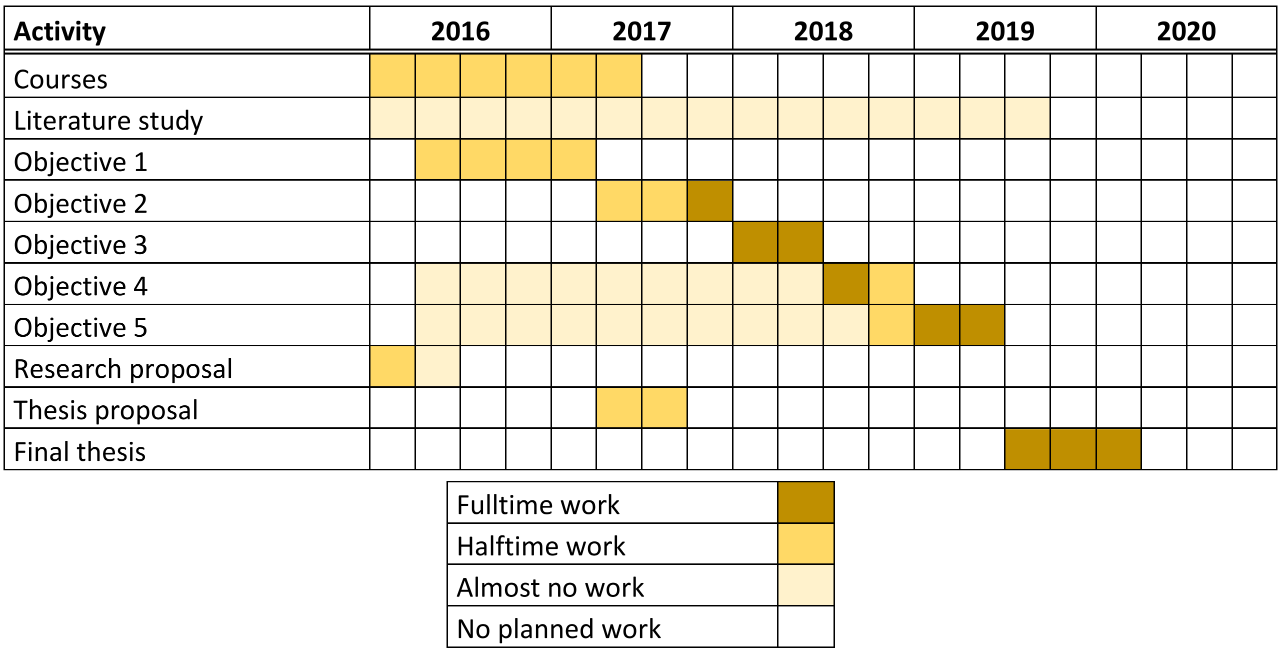
3.2 RESEARCH PLAN

The research plan is defined in table 1 and covers the overall plan of the PhD Project, with the colors that defines the workload explained at the bottom. The years are split up into four quarters in this plan excluding the summer vacation. The plan includes the overall parts of the project such as when the courses are planned, doing literature review, when the different objectives are planned, and the milestones in the project.

The courses are planned to be finished in the mid of 2017, that way more focus can be put into the research for two and a half years. The research will therefore have medium workload until mid of 2017, since the courses are run in parallel. However, if possible the courses will be adapted as much as possible to fit the research project. After the mid of 2017 more workload will be put into the research project.

During the research project it is important to keep track of the state-of-the art technology, and new findings in order to not fall behind. Therefore, literature study will be parallel with the whole research project, but will only be done in the background.

TABLE 1: OVERALL RESEARCH PLAN OF THE PHD PROJECT, THE COLORS ARE EXPLAIN AT THE BOTTOM



There are five objectives that needs to be fulfilled in order to achieve the overall aim of the project:

- Objective 1, speech recognition, is the first step, the courses are done in parallel and therefore a year will be necessary due to halftime work, and are planned to be finished at quarter 1 of 2017.
- Objective 2, haptic control, is the second step, during this time courses and thesis proposal will be partly done in parallel, therefore only halftime work will be done between quarter 2-3 of 2017, but more at quarter 4 of 2017 this will increase to fulltime work.
- Objective 3, augmented reality, is the third step, here there are no other activities planned at the same time and therefore fulltime work is allocated during quarter 1-2 in 2018.
- Objective 4, usage of the system, is ongoing during objective 1-3 because it is important to have the understanding of what the different technologies can be used for. Therefore most of the work might already be done and therefore most can be finished during quarter 3 of 2018.
- Objective 5, design of the programming-by-guidance system, the design of the system will be part of objectives 1-4 since these requires some kind of experimental system, but the final work to finish the design of the system will be done during the end of 2018 to mid 2019.

There are three main deadlines of the PhD project which follows the general study plan of the University of Skövde, these are Research proposal, Thesis proposal and Final thesis.

- This report finalizes the research proposal which is done in May 2016.
- Thesis proposal is planned to be done at quarter 2-3 of 2017, and will be done in parallel with other work, therefore two quarters will be required.
- Final thesis is planned to be done at quarter 3 in 2019 to quarter 1 2020. This part will require quite a lot of work and will summarize the results from all the other objectives, therefore three quarters of a year are planned to finalize the thesis.

The project is planned for five years in total, it started in April 2015 and is planned to be finished in the first quarter of 2020.

3.3 COURSE INTEGRATION AND PUBLICATION PLAN

For the different parts of the project, the plan is to adapt courses for the PhD project. For each activity of the project, the plan is to write papers for conferences and or journals. Table 2 shows how each activity of the project is connected to courses and publications. The demonstrator course will be useful not only for objective 1 as shown in Table 2, but will create the experimental framework that can be used during the whole project.

TABLE 2: COURSE INTEGRATION AND PUBLICATION PLAN.

Activity	**Courses**	**Publications**
Research proposal	Research methodology course	DIVA
Objective 1 – Speech recognition	Demonstrator course to create experimental framework Course on signal processing for speech recognition	Swedish Production Symposium International journal of automation and computing
Objective 2 – Haptic control	Course on haptic control	ACM/IEEE International Conference on Human Robot Interaction IEEE Transactions on Haptics
Thesis proposal		DIVA
Objective 3 – Augmented reality		ACM/IEEE International Conference on Human Robot Interaction Journal of Intelligent & Robotic Systems
Objective 4 – Use cases		International journal of automation and computing
Objective 5 – Design of the system		IEEE/RSJ International Conference on Intelligent Robots and Systems Robotics and Computer-Integrated Manufacturing
Final thesis		DIVA

The journals in this publication plan are selected based on impact factor, topic relevance and whether they exist in the NSDs publication channels [18], which is a database used to rank different journal papers. This database decides whether publications in journals may be funded, and the weights of these funds. The journals have then been selected based on the impact factor to decide the level difficulty to publish in the journals, and this correlates to the competence level in the PhD studies and the interest of the topic. Finally only journals with relevant content have been selected.

3.4 EXPECTED RESULTS

After this project, the expectation is to have a fully functional programming-by-guidance system of an industrial robot that can be used by a worker, with no expertise in robot programming, to easily teach the robot how to execute its task. These specific tasks can be, to create a simple program that moves around with different movement types and activates a tool, a calibration that tells the robot the new position of a work object or product, the introduction of a new product in an already existing program.

The programming-by-guidance system is also expected to reduce the general development time of lightweight collaborative industrial robots by replacing the regular teaching pendant to quicker jog/move the robot to a position without losing the accuracy.

4 REFERENCES

1. Robson, C., *Real world research : a resource for users of social research methods in applied settings*. 2011, Chichester :: Wiley. xxi, 586 p. :.
2. Skövde, U.o. *Examination inom utbildning på forskarnivå vid Högskolan i Skövde* 2013 [cited 2016 2016-05-13]; Available from: http://www.his.se/Policies/Forskning_utbildning_forskningsniva/foreskrifter-examination-forskarniva.pdf.
3. Pires, J.N., *New challenges for industrial robotic cell programming.* Industrial Robot-an International Journal, 2009. **36**(1): p. 5-5.
4. Haddadin, S., et al., *Towards the Robotic Co-Worker*, in *Robotics Research: The 14th International Symposium ISRR*, C. Pradalier, R. Siegwart, and G. Hirzinger, Editors. 2011, Springer Berlin Heidelberg: Berlin, Heidelberg. p. 261-282.
5. Krüger, J., T.K. Lien, and A. Verl, *Cooperation of human and machines in assembly lines.* CIRP Annals - Manufacturing Technology, 2009. **58**(2): p. 628-646.
6. Ragaglia, M., et al., *Accurate sensorless lead-through programming for lightweight robots in structured environments.* Robotics and Computer-Integrated Manufacturing, 2016. **39**: p. 9-21.
7. Cherubini, A., et al., *Collaborative manufacturing with physical human–robot interaction.* Robotics and Computer-Integrated Manufacturing, 2016. **40**: p. 1-13.
8. Akella, P., et al. *Cobots for the automobile assembly line*. in *Robotics and Automation, 1999. Proceedings. 1999 IEEE International Conference on*. 1999.
9. Padmanabhan, J. and M.J.J. Premkumar, *Machine Learning in Automatic Speech Recognition: A Survey.* Iete Technical Review, 2015. **32**(4): p. 240-251.
10. Wang, X., S.K. Ong, and A.Y.C. Nee, *A comprehensive survey of augmented reality assembly research.* Advances in Manufacturing, 2016. **4**(1): p. 1-22.
11. Azuma, R.T., *A survey of augmented reality.* Presence-Teleoperators and Virtual Environments, 1997. **6**(4): p. 355-385.
12. Zengxi, P. and Z. Hui. *Robotic machining from programming to process control*. in *Intelligent Control and Automation, 2008. WCICA 2008. 7th World Congress on*. 2008.
13. Zhang, H., et al. *On-Line Path Generation for Robotic Deburring of Cast Aluminum Wheels*. in *2006 IEEE/RSJ International Conference on Intelligent Robots and Systems*. 2006.
14. Lambrecht, J., et al., *Spatial programming for industrial robots through task demonstration.* Int J Adv Robotic Sy, 2013. **10**(254).
15. Oates, B.J., *Researching information systems and computing*. 2006, London ; Thousand Oaks, Calif.: SAGE Publications. xvii, 341 p.
16. March, S.T. and G.F. Smith, *Design and natural science research on information technology.* Decis. Support Syst., 1995. **15**(4): p. 251-266.
17. Hevner, A.R., et al., *Design science in information systems research.* MIS Q., 2004. **28**(1): p. 75-105.
18. AS, N. *Scientific journals, series and publishers*. 2016 [cited 2016 2016-05-13]; Available from: https://dbh.nsd.uib.no/publiseringskanaler/Forside.