Intuitive Robot Programming Using Augmented Reality

Sebastian Blankemeyer\textsuperscript{a}, Rolf Wiemann\textsuperscript{b}, Lukas Posniak\textsuperscript{c}, Christoph Pregizer\textsuperscript{c}, Annika Raatz\textsuperscript{a}

\textsuperscript{a}Leibniz Universität Hannover, Institute of Assembly Technology, An der Universität 2, 30823 Garbsen, Germany
\textsuperscript{b}Leibniz Universität Hannover, student
\textsuperscript{c}Viewlicity GmbH, Schäferkampsallee 42, 20357 Hamburg, Germany

\textsuperscript{*} Corresponding author. Tel.: +49-511-762-18249; fax: +49-511-762-18251. E-mail address: blankemeyer@match.uni-hannover.de

Abstract

The demands on companies caused by the markets are becoming more and more fast-moving and complex. Reasons are the increasing number of variants of products as well as reduced product life cycle times. This is particularly relevant for assembly tasks, since a very high flexibility and adaptability to varying ambient conditions must be ensured in this area. Therefore, a lot of steps are currently being carried out manually. However, if an automation is attempted in the field of assembly, companies are often facing a trade-off between a high degree of automation and flexibility of the production system. A concept to increase the automation rate in assembly tasks can be seen in the use of collaborative robots, so that the benefits of both, humans and robots, can be combined to accomplish the task. In doing so, one key issue is to simplify and thus to accelerate the programming process so that the necessary programming skills of employees can be reduced. Although some of today’s collaborative robots already offer good programming approaches like kinesthetic teaching, this article introduces a new and more intuitive programming method which is based on Augmented Reality. For this purpose, components of an assembly group are virtually linked with CAD models by using optical markers. The operator can then virtually assemble the components according to the assembly sequence. Results of first tests indicate, that once the assembly process is recorded, the robot can accomplish the assembly in reality. Finally, possibilities for future developments are presented.

© 2018 The Authors. Published by Elsevier B.V.
Peer-review under responsibility of the scientific committee of the 7th CIRP Conference on Assembly Technologies and Systems.

Keywords: augmented reality, intuitive robot programming, human-machine-interaction

1. Introduction

Short product life cycles and a high variety of designs demand flexible and intuitive programming methods for industrial robots so that companies can efficiently react to rapidly changing market situations. Conventional programming methods, such as teach-in methods or simulation programs, require specialized knowledge so that fast and flexible changes are difficult to realize [1,2]. For this reason, different approaches can be found in literature which aims at abstracting the programming procedure so that the robot’s program code is automatically generated [3]. Intuitive programming methods are developed especially in the field of human-robot collaboration. Collaborative robots enable humans and robots to work together within the same workspace and are already used in industry. These robots can be guided by hand, whereby points or trajectories can be stored (kinesthetic teaching). Thus, collaborative robots offer a suitable hardware design for flexible and simple programming. On the other hand, the software is often simplified by using program blocks to abstract the textual programming [4,5].

A further simplification of programming can be achieved by translating human actions directly into a robot program. Programming-by-demonstration (PbD) procedures are based on the imitation of human actions by the robot. By deriving motion paradigms, the demonstrated actions can also be transferred to other or similar tasks [6].

Another approach of programming robots is the use of gestures. In order to generate waypoints for determining the path of the robot, camera systems and special software tools are used to recognize such gestures and interpret them, respectively [7]. If such a camera system is combined with an augmented reality system, additional information and objects can be visualized spatially. Previous research projects on the use of augmented reality (AR) in the field of robot programming have used AR systems, which are, however, complicated and usually not available [8,9]. The development of powerful head-mounted displays (HMD) such as the Microsoft Hololens enables a highly simplified system structure. The Microsoft Hololens combines the AR and the camera system in a compact device. A permanent tracking of the environment makes it possible to place 3D-holograms into it. Furthermore, the integrated gesture and speech recognition permit an intuitive control and interaction with holograms. Since these functionalities are beneficial for the design of intuitive robot programming methods, the Hololens is used as an interface in this work.

First, the real
components are recognized by markers and overlaid with a virtual model. This can be displayed by the hololens directly at the workstation. The worker can then virtually assemble the product using gestures. The gestures are recognized by the integrated cameras and the assembly steps are thus recorded by the Hololens. Finally, the robot can carry out the process in reality through the acquired data.

At the beginning of this paper, an overview on already existing AR systems is given, which have been used and developed in the field of human-robot-collaboration. The concept and the implementation of a new programming method are then described, which will also be examined with regard to its accuracy. Based on a summary, an outlook on future work is given.

2. Augmented Reality in Robotics

Augmented reality can be described as a computer-aided transfer of virtual aspects to the real world. The Hololens is a device that can be used for AR applications. A characteristic of an AR system is the combination of the reality with a virtual world, an interaction in real time and a three-dimensional registration [10]. Based on these characteristics, an AR system usually consists of a display for visualizing virtual contents, a camera system for tracking the environment and communication interfaces. Basically, a distinction can be made between head-mounted displays (HMDs), handheld devices and projection systems, which have often intuitive control interfaces such as gesture and voice control.

Despite the development of more powerful hardware and software, the usage of AR systems is mainly limited to research and gaming applications. Main applications comprise installation and maintenance tasks [11,12]. In industrial practice, for example, Boeing successfully uses Google Glass to visualize assembly instructions [13]. M. Funk [14] developed various systems, with which employees are supported in assembling and commissioning tasks by virtual information and instructions, that are visualized via a projector. In addition, the user’s actions are detected with a camera. O. Sand et al. [15] developed a similar system to provide assembly and picking instructions, which is based on a projector as well.

The possibility to display information about the assembly process is also used in the field of human-robot collaboration. S. Makris et al. [16] developed a system, with which assembly steps can be visualized and information about the sequence of events can be displayed. In addition, the workspace and the planned trajectory as well as the current status of the robot can be visualized. An AR glass is used to display the virtual content and the system can be controlled via a smartwatch. O. Danieleson et al. [17] developed a similar system, which supports the worker with virtual instructions during the collaborative assembly. In doing so, the user receives information about the assembly steps performed by or in cooperation with the robot or about assembly steps of the user himself. The AR contents are visualized using a display and the components are detected by means of a marker tracking procedure.

Furthermore, there are systems that not only display information and instructions, but also support the user while programming a robot. In this case, the use of AR systems is focused on path planning and the simulation of the created robot program in order to detect and avoid collisions. H. Fang et al. [18] developed a method to show points in space by a hand-guided marker. Based on this, a path is calculated which can be simulated with a virtual robot model and transferred to the real robot. A laptop is used to display the virtual content. J. Chong et al. [19] developed a similar system, where besides saving points, it is also possible to show a trajectory. The procedure presented by Y. Pai et al. [20] is based on the same functional principle but it has been extended to include the possibility of programming pick-and-place tasks. Virtual information and content is visualized via a conventional computer monitor. J. Lambrecht [8] uses gestures that are detected by the Microsoft Kinect instead of markers. With the created system, real objects can be overlaid with virtual objects. In addition to path planning and simulations, simple pick-and-place tasks can also be programmed by manipulating the virtual parts and translating this action into a robot program. A tablet computer is used to control the system and to show the AR content. G. Reinhart [9] developed a system for programming trajectories using a pen. This is particularly suitable for programming in the field of welding and laser hardening. In comparison to the conventional teach-in method, a considerably time saving approach could be shown.

When considering the presented systems in the field of robot programming, it is noticeable that they are characterized by an intricate structure, which generally consists of a camera system and an AR display. In addition, tablets or laptops are often used to display AR contents, however, such systems are impractical in application and do not allow for free-handed operations. In contrast, the Microsoft Hololens connects a camera display with an AR display inside an HMD, which yields a simple system setup in total. In addition, unlike many other HMDs, holograms can be displayed in 3D and a permanently scanning process of the environment allows the user to place and fix holograms in the workspace. Furthermore, the gesture and voice control provides an intuitive control functionality.

3. Concept of programming

The goal of this work is to develop an intuitive programming tool for robots. As shown in the previous section, AR is a method to allow human-machines interactions. In this context, the Hololens offers new possibilities to realize communication and information processing, which are used to develop an application that converts virtual assembly steps into a robot program. In this way, the first step will be the generation of pick-and-place motion profiles for the robot to assemble components (see Fig. 1).

![Fig. 1. Flowchart of the developed robot programming system](image_url)

The process design is structured in such a way that the worker wears the Hololens which is linked with the robot at his
workplace. The components to be assembled can be reached by the robot and lie on optical markers that are needed for localization. Furthermore, the markers allow the connection between real and virtual components. They are necessary because the internal tracking possibilities of the hololens only allow the creation of a spatial mapping, but not the recognition of objects. Therefore, the Software Development Kit SDK developed by Vuforia is used for the hololens to determine the position and orientation of the components.

In order to program an assembly step, the user first of all carries out the assembly step virtually by moving the virtual component to the desired position. The virtual component is displayed to the user by the Hololens and is at the same position as the real component, whose position is determined by marker tracking. The start and end coordinates of the assembly step are determined and saved with respect to the internal coordinate system of the Hololens (CSH). To enable the robot to manipulate the real components, these coordinates must be transformed from CSH into the robot’s base coordinate system (CSR). The CSR is already shifted to the mounting aid as it simplifies the transformation calculation (see Fig. 2). The transformed coordinates are subsequently sent to the robot, which then performs the manipulation in reality.

In order to work with the obtained data, a program logic is implemented on the robot controller that uses the start and end coordinates of the moved component to automatically generate the assembly process. The path planning task between the individual poses is thus carried out directly by the robot controller.

The interaction with the virtual objects and the control of the application is realized via the Hololens’ standardized communication interfaces. By focusing on individual objects, the user can select components and interact with them using the so-called AirTap, which is similar to a mouse click. Based on this, the virtual components can be controlled and positioned with the help of manipulation gestures.

4. Implementation

Windows 10 is used as the operating system for Microsoft Hololens. The Universal Windows Platform (UWP) serves as a uniform development platform for Windows 10 applications. For the development of Hololens applications, Visual Studio is used in combination with the game development software Unity3D, for which Microsoft already provides basic functions such as gesture and speech recognition.

The basis of the application is the product, which is to be assembled virtually by the user as well as in reality by the robot. To display the product holographically, the CAD model of the product has to be converted to a mesh model and then imported into Unity3D. During this process, the assembly structure defined in the CAD model is retained. However, it can be changed in Unity3D as required. The program logic is created by scripts which are designed in such a way that they can be used on any model and thus facilitates the exchange of different models.

A toolbar with buttons is implemented as user interface. So, all functions of the program can be called up and settings can be made. In addition, a voice control allows an intuitive control of the application and the user is given feedback and instructions via a voice output.

4.1. Use case

To demonstrate the functionality by means of a sample application, this paper is concerned with the implementation of the assembly process of a helicopter model. The helicopter serves as a use case in the learning factory of the Institute of Production Systems and Logistics. For simplification, only two components are considered, but the application can be extended by additional components without any problems. Two skirting boards of the helicopter model are used as components, which are assembled on a mounting aid (see Fig. 3). In this work a Kuka KR6 R900 sixx robot is equipped with a suction gripper and connected to the Hololens via the Kuka technology package Robot-Sensor-Interface (RSI). The RSI enables real-time data exchange between the Hololens and the robot, based on the User Datagram Protocol (UDP).

4.2. Marker-Tracking

In order to be able to overlay and link the real components with the respective virtual components, the real components must be detected in the environment. Furthermore, a fixed spatial reference point has to be defined in order to transform poses from the Hololens’ coordinate system (CSH) into that of the robot (CSR) and vice versa. The Vuforia SDK available for the Hololens is used for the marker tracking.

The markers themselves are generated using the tool of AR-software developer Brosvision [21]. The tool takes all specifications into account, which are recommended by Vuforia to
of the assembly process between the Hololens and the robot,

4.4. Coordinate transformation

end coordinates of this assembly step are transformed and sent
the manipulation gesture is terminated. Afterwads, the start and
placed. The workpiece is then set to the target position when
position that can be determined from the assembly model, it
remains the
handle an object virtually, the user focuses a component and
by Microsoft and is therefore straightforward to implement. To
aid but not visualized.

fined by the assembly model, which is placed on the mounting
mounting aid. The target position of each subassembly is de-
uation of each component must be possible. As shown in Fig.
3, the complete product is assembled on the marker with the
mounting aid. The target position of each subassembly is de-
defined by the assembly model, which is placed on the mounting
aid but not visualized.

The components (here: skirting boards of the helicopter) are
moved with a manipulative gesture. This gesture is predefined
by Microsoft and is therefore straightforward to implement. To
handle an object virtually, the user focuses a component and
remains the AirTap as long as the movement is in progress. The
object can be moved simulatanously to the hand movement and
placed in the environment. If the component is near the target
position that can be determined from the assembly model, it
changes its color to signal to the user that the object can be
placed. The workpiece is then set to the target position when
the manipulation gesture is terminated. Afterwads, the start and
end coordinates of this assembly step are transformed and sent
to the robot, which manipulates the real component.

4.4. Coordinate transformation

In order to be able to exchange the start and end coordinates
of the assembly process between the Hololens and the robot,
they must be transformed (see Fig. 5). As soon as the marker
of the mounting aid has been detected, the local coordinate sys-
tem of the Hololens (CSL) is superimposed on the robot’s base co-
ordinate system. Basically, poses of objects are specified by the
Hololens in a global coordinate system. Therefore, the start and
end coordinates are first transformed into the local coordinate
system defined on the mounting aid marker. This is followed by
a transformation into the robot’s base coordinate system. Since
the Hololens application is programmed with Unity3D, which
uses a left-handed coordinate system, the coordinates are finally
transformed into the robot’s right-handed coordinate system.
With the described transformation, only the position coordi-
ates and not the rotation coordinates are transformed. With
the developed application, components could only be assem-
bled in z-direction so that only the rotation around the z-axis is
calculated.

4.5. Robot program and data exchange

After the coordinates have been transformed, they are sent
to the robot by the Hololens. The Kuka technology package
KUKA.RSI is used for data exchange. With this, a real-time
connection between the robot and an external system (here:
Hololens) can be established. The Hololens is connected di-
rectly to the robot via a local network and so that a simple
system setup is achieved. For data exchange and the program
logic of the assembly process, a robot program is implemented
in KukaRobotLanguage KRL, which is a special programming
language for Kuka robots. The coordinates sent by the Hololens
are put in the defined variables in this program. If the coor-
dinates and a start signal are received from the Hololens, the
programmed assembly sequence is started (see Fig. 6). Starting

![Fig. 4. Marker with mounting aid and Hololens local coordinate system (left) and virtual components (right)](image)

![Fig. 5. Coordinate transformation between the Hololens and the robot](image)

![Fig. 6. Robot pick-and-place motion](image)
direction. This is followed by a linear movement and the component is picked up. The target pose is also approached with a PTP movement and an offset. The component is then placed with a linear movement and the robot moves back to the waiting position. The user can then virtually demonstrate the next assembly step and start the robot again.

5. Accuracy and precision of marker tracking

An important feature when considering assembly processes is the required accuracy. If the production system does not provide the required accuracy, an assembly is not useful. The marker tracking is crucial for the accuracy of the robot assembly. The more precise the tracking, the more accurately the robot can move and place the components. In order to examine the accuracy of the tracking, two markers are placed at a distance of 0.5 m within the workspace (see Fig. 7). Both markers are scanned from different distances and angles with the Hololens. In doing so, the user is located in the middle of the markers and scans the markers by looking to the right and to the left. After both markers have been scanned, the distance of the reference marker to the local coordinate system defined on the marker (see Fig. 7) is compared to the real distance of the markers. This setup and procedure corresponds to the method developed for robot programming, in which the coordinates of the assemblies are also specified in the local coordinate system of the Hololens (CSL) and are then transformed into the robot’s base coordinate system.

To examine the accuracy and precision of the tracking process, the distance and angle between the Hololens and the robot are varied. At each of the 25 examined measuring points (5 different angles and distances) 20 measurements are made. Moreover, ISO 9283 [22] is applied for the validation. ISO 9283 is mainly used in the field of industrial robots, but is also used for examinations of tracking systems [23]. The accuracy \( A_s \) is defined as follows:

\[
A_s = \sqrt{(\bar{x} - x_{\text{ref}})^2 + (\bar{y} - y_{\text{ref}})^2 + (\bar{z} - z_{\text{ref}})^2}
\]

\[
\bar{x} = \frac{1}{n} \sum_{i=1}^{n} x_i; \quad \bar{y} = \frac{1}{n} \sum_{i=1}^{n} y_i; \quad \bar{z} = \frac{1}{n} \sum_{i=1}^{n} z_i
\]

(1)

The accuracy \( (A_s) \) indicates the center of gravity of the point cloud, i.e. the distance between the mean value of the measured values \((\bar{x}, \bar{y}, \bar{z})\) and the reference value \((x_{\text{ref}}, y_{\text{ref}}, z_{\text{ref}})\). According to ISO 9283, the precision \( P_s \) is defined by the distance of the measured values to the mean value \((\bar{l})\) and the standard deviation \( (\sigma_l) \) and is calculated as follows:

\[
P_s = \bar{l} + 3\sigma_l
\]

\[
\bar{l} = \frac{1}{n} \sum_{i=1}^{n} l_i
\]

\[
l_i = \sqrt{(x_i - \bar{x})^2 + (y_i - \bar{y})^2 + (z_i - \bar{z})^2}
\]

\[
\sigma_l = \sqrt{\frac{\sum_{i=1}^{n} (l_i - \bar{l})^2}{n - 1}}
\]

(2)

In Fig. 8 and Fig. 9 the accuracy and precision are shown with respect to the distance and angle to the Hololens. A cubic interpolation is made between the measuring points. It can be seen that the best results are achieved between 40° - 60° and 0.7 mm and 0.9 mm. Here, the accuracy is 1 mm - 2 mm and the precision is 3 mm - 5 mm. If the angle decreases and the distance increases, it is not possible to detect the markers (white area). In addition, when angles become large, the accuracy and precision decreases strongly to max. values of 9 mm resp. 45 mm. The determined accuracy is considerably worse than the pre-
cision stated by the robotic manufacturers (here: repeatability of Kuka KR 6: \( P_\text{R}=+/- 0.03 \text{ mm} [24])\). This accuracy was sufficient for this use case. Several tests have shown that the component could be gripped and set down correctly. This is due to the fact that the skirting boards center themselves on the mounting aid when they are placed on the mounting aid and thus bring themselves into the correct position.

6. Conclusion and Outlook

This paper is concerned with the development of a new Hololens application, so that a various components can be virtually assembled. By linking virtual and real components, an industrial robot can perform pick-and-place tasks based on start and end coordinates. This application has been implemented and successfully tested for two components as a sample application, but can be extended by using additional markers.

In order to determine the accuracy and precision of the marker recognition process, extensive measurements were carried out in this work. In the recommended scanning area an accuracy of 1 mm - 2 mm and a precision of 3 mm - 5 mm is reachable. During the testing of the system the components were gripped in almost all cases. However, this is a high degree of inaccuracy for robot applications.

In the following works, the accuracy of another marker tracking system, for example the AR-ToolKit, could be tested. On the other hand, the marker tracking could be replaced with an object recognition. This would allow an even more flexible and simple system structure. However, one marker for referencing the coordinate systems of the Hololens and the robot would still be necessary. Therefore it has to be evaluated if a higher degree of accuracy can be achieved by object recognition and the use of only one marker.

In order to further expand the robot’s programming options, possibilities for describing a flow logic and path planning should be implemented. With the current application only a predefined pick-and-place sequence with a free adjustable offset is possible. A more flexible programming of the path can be realized by tracking path points from the virtual movement of the component. Based on this data, a robot path could be interpolated. In addition, assembly paradigms could be developed for various components from which gripping, approaching and placing modalities could be generated automatically. During this, the application should be extended to include the possibility of mounting components in all spatial directions. To do this, the coordinate transformation have to be supplemented by a transformation of the rotation coordinates.

The goal should be a dynamic generation of the robot program by the Hololens, which is sent to the robot. This would require the implementation of a different type of connection between the robot and the Hololens, since the current RSI only allows data exchange.

References