Student Conference Proceedings
Vol 1, No 1, 2025, Article ID 2040, 4 Pages

Student Conference Lübeck

Proceedings Article

Projector-Based Augmented Reality Interface for Intuitive Robot Programming

Lugh Martensen [a,*] · Christoph Willibald [b,*]

[a] Student of Robotics and Autonomous Systems, Universität zu Lübeck, Lübeck, Germany
[b] Institute of Robotics and Mechatronics, German Aerospace Center, Oberpfaffenhofen-Weßling, Germany
[*] Corresponding author, email: lugh.martensen@student.uni-luebeck.de; christoph.willibald@dlr.de

Received 04 February 2025; Accepted 23 June 2025; Published online 15 July 2025

Abstract

Robotic sanding and surface finishing has the potential to help the craft sector to become more productive and more efficient. For small businesses, intuitive programming interfaces are required in order to adopt robotic systems without the need for additional training or personnel. For such an application, we propose an Augmented Reality (AR) system that uses a projector for visualization and a touchscreen-based web interface for input detection. In a user study, the proposed interface is compared to an AR application running on a head-mounted display (HMD). The results indicate that the new approach significantly reduces mental load and increases performance. The combination of projector and touchscreen is generally preferred by the users due to a high intuitiveness and ease of use when compared to the HMD application.

I. Introduction

Many smaller businesses face increasing economic pressure due to a shortage of skilled workers and competition from larger companies that are able to automate production [1]. Robotic assistance could help them to become more efficient by allowing workers to offload repetitive tasks to a machine. To make this profitable for small-scale production, a flexible, intuitive and user-friendly interface for robot programming is needed that can be used by workers without additional training.

In previous work [2], a *learning-from-demonstration* interface was developed for programming a robot to sand the surface of an unknown workpiece. It solves the aforementioned problem by incorporating Augmented Reality (AR) with a Microsoft HoloLens head-mounted display (HMD), to give the user feedback during the demonstration process. A study conducted in [2] confirms that the visualization reduces the knowledge gap between the human and the robot and helps users to understand how their demonstration affects the result. However, the interface also increases mental load, and users reported technical issues such as delays and inaccurate hand tracking.

This work explores a new AR interface to address these problems. Instead of using an HMD, it augments the real world with a projector by highlighting points on existing surfaces. Thanks to a calibration procedure which is performed beforehand, the projection is accurate even on complex surfaces that are not flat. To reduce the mental load by removing interaction elements from the user's field of view, and to increase familiarity with the device, we decided for a touchscreen-based web interface to obtain user input. During development, feedback from [2] was incorporated to address the mentioned points of criticism. We conducted a follow-up user study, which confirms that the new system does indeed improve the user experience. The results extend the state of the art by showing that a projector-based AR interface can increase intuitiveness and reduce mental load for robot programming tasks when compared to HMD-based AR. In the following, the developed system is explained and the results of the user study are presented and discussed.

II. Methods and Tools

In the projector-based setup, a depth camera is attached to the workbench to capture a point cloud of the robot's workspace (see Figure 1 and Figure 2). A projector is mounted on the ceiling. The projector and the robot are connected to a computer that communicates with an Android tablet over a wireless connection.

II.I. Calibration and Projection

The calculation of the projection is based on the idea that a projector works similar to a camera. Instead of focusing light rays through a lens and detecting them on a sensor plane, the projector does the opposite. It sends light through a lens that projects an image into 3D space. By applying the pinhole camera model to the projector, it is possible to compute a mapping from 3D points to the 2D image plane. This requires the intrinsic and extrinsic calibration parameters of the projector. They can be determined with a camera calibration software, as long as a sufficient number of correspondences between world and image space is known. We used DLR CalLab, a software developed by the German Aerospace Center (Deutsches Zentrum für Luft- und Raumfahrt e.V., DLR). The correspondences are obtained as follows: the camera takes a picture of a checkerboard pattern that is mounted to the robot (see Figure 1). The pattern corners at the transition of white and black squares are detected in the camera's pixel space, and their 3D position is calculated using the robot's forward kinematics. To get the corresponding projector pixels, each pixel coordinate is converted to binary. The pixel positions are then encoded with a sequence of binary images. In the first image, each pixel is colored according to the most significant bit of its binary coordinate representation, black for zero and white for one. In the second image, the pixels are colored according to the second most significant bit of their binary coordinate representation, and so on. This way, all pixel coordinates in a full HD image can be encoded with 22 binary images. These images are called *temporal binary structured light patterns* and they are explained in great detail in [3] and [4]. As shown in Figure 1, the projector successively displays the different patterns on the checkerboard. The camera captures the patterns and for every image, the brightness levels are determined at every checkerboard corner and assigned to represent either a one or a zero. With all 22 brightness levels at every checkerboard corner, the pixel coordinates are decoded. The resulting 3D positions and their corresponding projector pixel coordinates are passed to the camera calibration software, which returns intrinsic and extrinsic calibration parameters for the projector. After the calibration, points in the real world can now be highlighted using the pinhole camera model, as long as they are on a surface that is not obstructed by shadows.

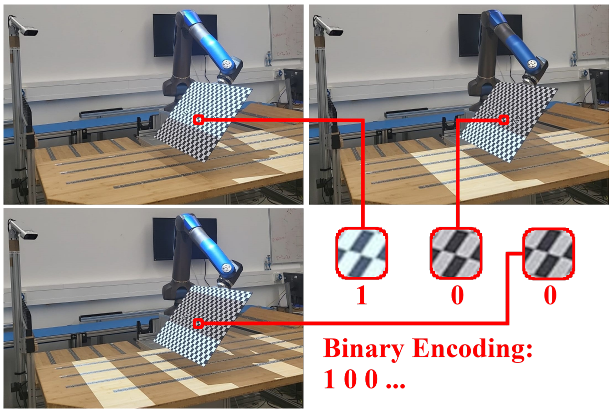

Figure 1: The calibration procedure for the projector and the camera. The projector displays a time-shifted binary encoding of the pixel coordinates. For every pattern, the camera checks the brightness at each corner point of the checkerboard pattern. Converting the brightness levels into ones and zeros, the binary encoding of the pixel's coordinate can be reconstructed.

II.II. Web Interface

The web interface is based on the Human Factory Interface (HFI), which was developed by DLR for the purpose of robot programming. The HFI provides a 3D environment that shows a digital representation of the robot and the workbench (see Figure 2). To enable a user to provide task instructions for surface finishing with the HFI, we developed and implemented the following five-step process:

Scan: A point cloud is captured and displayed. The user can confirm the point cloud or capture a new one.

Select: The robot, which has a grinding tool attached to it, switches into gravity compensation mode and can now be dragged around. The user demonstrates the sanding with the tool in the real world. The robot records the points the tool comes into contact with. Alternatively, the user can select points on the point cloud by drawing on the tablet. Either way, the application generates a polynomial that approximates the surface on which the selected points lie, and continuously updates this shape. It determines all the points of the point cloud that lie inside this polynomial, which we term inliers. As shown in Figure 2, both the selected points and the inliers can be shown in the HFI and through the projector.

Edit: In the HFI, the polynomial is visualized as a mesh. The user can crop the shape by drawing with a finger.

Generate: A trajectory is generated that allows the robot to sand the entire surface of the modified shape. It is shown in the HFI and with the projector.

Run: Executes the generated trajectory on the robot or in a simulation to verify the planned strategy for surface finishing. The progress is displayed in the HFI and in the real world.

10.18416/SCP.2025.2040

The HoloLens application from [2] was modified to include the same steps, which makes it comparable to the projector-based interface. The development will not be explained in detail here.

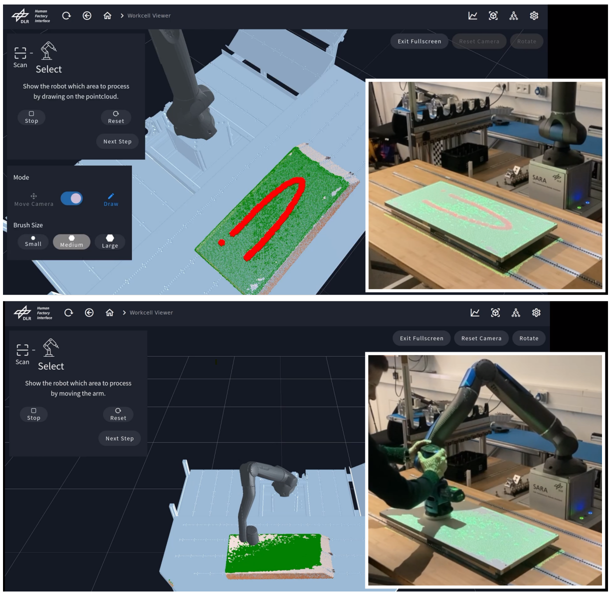

Figure 2: The second step (*select*) in the programming process. The selected points are shown in red, the inliers in green. Top image: Selection by drawing in the HFI. Bottom image: Selection by robot demonstration. On the right, the visualization of the projector in the real world is shown for both methods.

II.III. User Study

20 students and employees of DLR participated in a user study to systematically compare different user interaction concepts with respect to factors such as mental and physical demand, usability, or effectiveness in completing the task. Each participant had to program the robot to sand the surface of a wooden chair, using the following five interface configurations:

RO: No visualization, point selection with the robot's grinding tool.

HFI_RO: Visualization with the HFI running on the tablet, point selection with the robot's grinding tool.

HFI_HA: Visualization with the HFI running on the tablet, point selection by hand in the HFI.

HL_HA: Visualization with the HoloLens application, point selection by hand in the HoloLens application.

HL_RO: Visualization with the HoloLens application, point selection with the robot's grinding tool.

Different methods of point selection were tested because the study from [2] could not identify a superior method. Participants tested the configurations in different orders to obtain independent results. After each test, the user had to fill out the following questionnaires:

- NASA Task Load Index (TLX): measures subjective workload for human-machine interfaces [5].
- Questionnaire for the subjective consequences of intuitive use (QUESI): measures user satisfaction and intuitiveness of product interaction [6].
- User Experience Questionnaire (UEQ): Used to obtain six individual scores for attractiveness, perspicuity, efficiency, dependability, stimulation and novelty [7].
- General questions about model understanding, to evaluate if the user understood how the model was generated and how to influence this process.

At the end, participants were also asked to provide general feedback on which interfaces they liked the most and least.

III. Results and Discussion

For the evaluation, a repeated measures ANOVA was calculated. In case of violation of sphericity (Mauchly's sphericity test), Huynh-Feldt ($> .75$) or Greenhouse-Geisser ($< .75$) corrections were made. Post hoc tests with Bonferroni correction were performed to identify which methods differ significantly. Due to the large amount of data collected, only the most relevant significant differences are discussed in the following chapter. For quantitative results, please refer to Figure 3.

The overall TLX score for *HFI_HA* is significantly better than the score for any other configuration ($p < .05$ when compared to *RO* and $p < .001$ for *HFI_RO*, *HL_HA* and *HL_RO*). The same distribution can be found for the overall QUESI scores, including the respective confidence values. This indicates that *HFI_HA* reduces the subjective workload and makes the procedure more intuitive. *HFI_RO* does not show such significant benefits.

In the TLX category *mental demand*, *HFI_HA* performs significantly better than *HFI_RO* ($p < .05$), *HL_RO* ($p < .05$) and *HL_HA* ($p < .001$). *RO* is also significantly better than *HFI_RO* ($p < .05$), *HL_RO* ($p < .05$) and *HL_HA* ($p < .001$). For the QUESI category *subjective mental workload*, *HFI_HA* has a significantly higher score than *HFI_RO*, *HL_RO* and *HL_HA* ($p < .001$), while *RO* performs better than *HL_RO* and *HL_HA* ($p < .05$). *HFI_HA* and *RO* clearly reduce mental load. The good performance of *RO* is not surprising here, as this configuration does not provide an interface that the user has to get used to. However, it also receives a significantly lower score than every other configuration for the general questions about model understanding, with $p < .05$ when compared to *HL_RO* and $p < .001$ for any other configuration.

To measure programming performance, the QUESI categories *perceived error rate* and *achievement of goals* were analyzed. The *perceived error rate* is significantly higher for both HoloLens configurations when compared

10.18416/SCP.2025.2040

to *RO* ($p < .05$) or *HFI_HA* ($p < .001$). In the category *achievement of goals*, *HFI_HA* performs significantly better than *RO* ($p < .05$) and the two HoloLens configurations ($p < .05$ for *HL_RO* and $p < .001$ for *HL_HA*). This again highlights the good reception of *HFI_HA*. Compared to *RO*, the HoloLens configurations do not lead to a better performance. This can be explained by the trajectory generator, which gives good results even with a sparse selection of points and no modification of the generated mesh. This works well for relatively simple surfaces, but for more complex workpiece geometries, a better model understanding might be required.

The answers from the general feedback form clearly show that most users prefer the projector-based configurations. When asked about their favorite interface, *HFI_HA* was mentioned 13 times, followed by *HFI_RO* with three times. Additionally, both HFI-based configurations were the only ones that were not chosen as the least favorite by any participant.

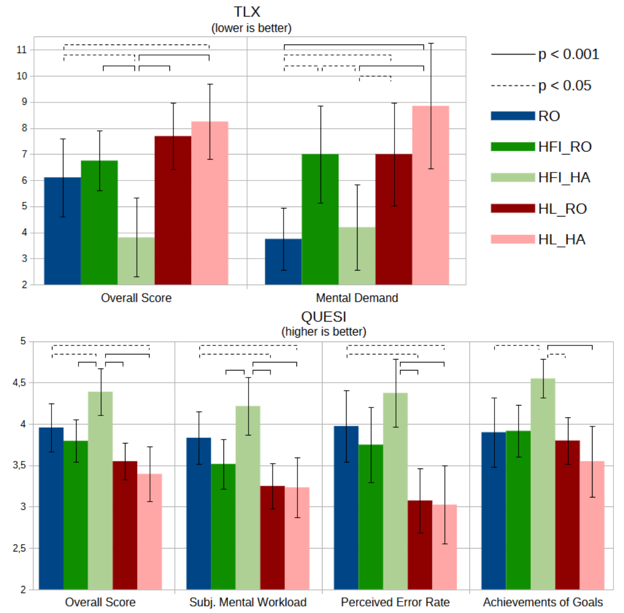

Figure 3: Different TLX and QUESI scores together with their 95% confidence intervals and different significance levels *p*.

IV. Conclusion

The results of the questionnaire show that the goal of creating a more intuitive and user-friendly application for the programming of robotic sanding was indeed achieved. The configuration with the projector and the point selection via touchscreen received better or at least similar scores in almost every category when compared to any HMD configuration. In particular, a lower mental load, better subjective performance, and higher user preference could be observed. The projector-based AR visualization could not be tested independently of the touchscreen interface, since it requires an additional device for capturing user input. For future works, a new input method should be developed to replace the HFI. This way, projector-based AR can be compared with HMD-based AR and with the touchscreen interface.

Acknowledgments

The work has been carried out at the Institute of Robotics and Mechatronics in Oberpfaffenhofen-Weßling, German Aerospace Center (Deutsches Zentrum für Luft- und Raumfahrt e.V., DLR) and supervised by Prof. Dr. Floris Ernst, Institute of Robotics and Cognitive Systems, Universität zu Lübeck, Lübeck.
Research funding: This work was supported in part by the German Federal Ministry of Education and Research (BMBF) through the “The Future of Value Creation — Research on Production, Services and Work” Program under Grant 02K20D032.

Author’s statement

Conflict of interest: Authors state no conflict of interest. Informed consent: Informed consent has been obtained from all individuals included in this study. Ethical approval: The research related to human use complies with all the relevant national regulations, institutional policies and was performed in accordance with the tenets of the Helsinki Declaration, and has been approved by the ethics committee of the German Aerospace Center. DeepL was used for linguistic fine-tuning of this paper.

References

[1] Lerosh project, Werk5 GmbH, 2022. URL: https://lerosh.de/en/project/ (visited on 01/19/2025).

[2] J. G. Eberle, Development of an augmented reality interface for intuitive robot programming, Technische Universität München, 2013, URL: https://elib.dlr.de/200286/ (visited on 01/19/2025).

[3] F. Leutert, Flexible augmented reality systems for robot-based production environments, Universität Würzburg, 2021. doi:10.25972/OPUS-24972.

[4] G. Taubin, D. Moreno, and D. Lanman. 3d scanning for personal 3d printing: Build your own desktop 3d scanner. *SIGGRAPH 2014*, 2014, doi:10.1145/2619195.2656314.

[5] S. G. Hart and L. E. Staveland, Development of NASA-TLX (Task Load Index): Results of empirical and theoretical research, in *Human Mental Workload*, ser. Advances in Psychology, 52, 139–183, North-Holland, 1988. doi:https://doi.org/10.1016/S0166-4115(08)62386-9.

[6] A. Naumann and J. Hurtienne, Benchmarks for intuitive interaction with mobile devices, in *Proceedings of the 12th International Conference on Human Computer Interaction with Mobile Devices and Services*, ser. MobileHCI ’10, 401–402, Association for Computing Machinery, 2010. doi:10.1145/1851600.1851685.

[7] M. Schrepp, A. Hinderks, and J. Thomaschewski. Design and evaluation of a short version of the user experience questionnaire (ueq-s). *International Journal of Interactive Multimedia and Artificial Intelligence, 4 (6), 103-108.*, 2017, doi:10.9781/ijimai.2017.09.001.