

Head-Mounted Displays in Industrial AR-Applications: Ready for Prime Time?

Hitesh Dhiman¹, Sascha Martinez¹, Volker Paelke³, and Carsten Röcker^{1,2}

¹Fraunhofer IOSB-INA, Lemgo, Germany

{hitesh.dhiman, sascha.martinez, carsten.roecker}@iosb-ina.fraunhofer.de

²Ostwestfalen-Lippe University of Applied Sciences, Lemgo, Germany
carsten.roecker@hs-owl.de

³Bremen University of Applied Sciences, Bremen, Germany
volker.paelke@hs-bremen.de

Abstract. The latest generation of head-mounted displays such as HoloLens provide mixed reality capabilities that claim to better integrate the real and virtual worlds. In this paper, we would like to share our experiences in implementing a user interface for an assembly assistance system using the HoloLens. We carried out a preliminary evaluation of the applicability of mixed reality using the perspective of developers and expert users in an assembly scenario that allows us to operate and compare two interfaces - a state-of-the-art projector display system and the HoloLens. We believe our findings may contribute towards a better understanding of the effects of new display technologies such as the HoloLens in developing and using assistance systems in other fields as well. Areas that may be of future research are also highlighted.

Keywords: Human Machine Interaction · Assembly Assistance System · Qualitative Study · HoloLens

1 Introduction

It is anticipated that trends like shorter innovation and product life cycles as well as an increase in mass customization will result in highly complex products that have to be produced in relatively small batches [4,5]. The complexity of assembly typically leads to a complexity in automation, which might be too expensive to implement given the short life cycle [6]. Consequently, manual assembly is another viable option for many companies.

The assembly operation is designed to be a part of a larger product manufacturing environment (Fig. 1). Custom parts are loaded onto a moving conveyor prior to the assembly station, while regular parts are available in the bins directly in front of the user. Once the assembly station receives an order, the user is guided through the assembly process. Each assembly tutorial consists of various steps which are divided into two categories, pick and assemble. In the pick step the user is shown the right component to pick, and in the assembly step the components are assembled followed by a confirmation. Some parts of the assembly station are designed to be user configurable, for example the height of the workplace can be adjusted at any time. Several such systems are available commercially (see, e.g., [7] or [8] for an overview).

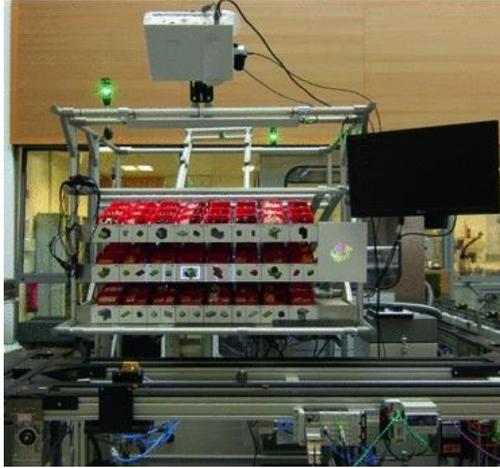


Fig. 1. Use case of an assembly assistance system, image taken from [12].

The proliferation of Augmented Reality (AR) in industrial applications [17] has enlarged the choice of display and interaction technologies that can be used on these systems. Early approaches in this field made use of head-mounted displays (HMDs) in the form of data glasses. Although the use of HMDs showed improvement in speed for picking tasks in comparison with paper instructions [4], they were shown to have various usability issues, such as the inaccuracy of stereoscopic projection, limited field of view, focusing issues and physical discomfort [1]. As a result, most state-of-the-art assembly assistance systems in the last years have utilized projection-based displays [11, 12, 18]. For instance, the Ulixes ‘Der Assistent’ system uses a projector in conjunction with a touch screen display.

Projection-based systems, however, also come with a few disadvantages. Firstly, the area of the workplace in such a system is limited to the area of projection. Any objects, parts and boxes that lie outside this zone cannot be augmented. Secondly, the placement of the assembly bins needs to be designed such that the projection is not blocked. Although the height of the worktable is adjustable, the placement and configuration of the boxes is fixed. In short, projection-based systems suffer from a lack of flexibility that may limit their use scenarios other than assembly assistance.

In the context of Industry 4.0, when looked at from the employees’ point of view, a flexible workplace is considered as one which is networked, personalized, and mobile [3], as it may be necessary to retool and reconfigure both the entire production line as well as the assembly stations in order to respond to variations in customer orders. In addition, every employee may prefer a particular placement and arrangement of tools and parts. In this case it would be desirable if the display technology is aware of the user’s environment, such that it is able to respond to the changes in configuration. Not many AR devices were able to achieve this ability until recently, when Microsoft released the HoloLens display which uses depth sensors to map the user’s surroundings and make them a part of the digital augmentation. Our focus in this paper is to find out if the HoloLens offers a more flexible display technology than projection, and if it reduces the usability issues of HMDs as reported above.

2 State of the Art

Most of the work in the field of assembly assistance has applied various display technologies, which we broadly categorize into two groups: dynamic systems and static systems.

Static systems use some form of AR technology that is installed onto the system itself. For instance, some systems make use of a light that is affixed in front of, or above, every bin on the rack; this light indicates the part to be picked and feedback to the system is provided via integrated touch or proximity sensors [23]. A second class of systems [2, 18] uses a projector located above the assembly station that projects images onto the workplace and bins. In this case, hand gestures are detected via an additional 3D sensor placed close to the projector. The projector-based system was demonstrated to be faster and more accurate in comparison to a smart glass-based system [1].

Nonetheless, an obvious limitation of these systems is that physical installation limits their flexibility and scalability, both of which cannot be achieved without considerable economic and manual effort. Further, the placement of bins has to take into account the nature of projection – bins can only be placed within the area of projection, which itself is limited per workstation. For instance, an area of 100cm x 60cm while sufficient for a single major assembly may not be suitable for multiple minor assembly stations spanning across a few meters. Using multiple projectors and 3D sensors to achieve the desired result becomes considerably expensive.

Dynamic systems make use of a head mounted device [10,12,13,14,15]. Ideally, this type of a system allows itself to be adjusted and scaled since the display and the workplace are decoupled. Most of the research work in this area has been conducted using smart glasses which overlay 2D or 3D images onto a display. Although smart glasses have been shown to improve productivity as compared to paper-based instructions [24], in comparison to projection-based systems, they have considerable drawbacks, namely, reduction of wearing comfort due to high weight, need for optical adaptability, and a narrow field of view [1,16].

The studies mentioned previously have compared in-situ projection with HMDs that vary in their technical capabilities. For instance, Zheng et al. [13] used a Google Glass device which is a form of peripheral display, while Khuong et al. [14] used a Vuzix Wrap 920AR which is capable of displaying “stereoscopic 3D video content” but lacks the capability to render content on its own [9]. Although commercial projection technology has been available for a couple of decades now, the market for commercial HMDs is still evolving. The latest generation of HMDs such as HoloLens classify themselves as ‘Mixed Reality’ glasses, where Mixed Reality is defined as a combination of Augmented Reality (AR) and Augmented Virtuality (VR) [25]. In our literature review, no study has thus far compared a MR HMD to a projection-based display.

3 Overall Research Goal and Procedure

The aim of this paper is to explore the advantages and disadvantages of using an MR capable HMD as compared to a projection-based display on an assembly assistance system. We would like to firstly understand the effect of MR on information representation, and secondly, if MR brings an obvious improvement or feature add to the

existing state of the art assembly assistance system such that it can possibly replace a projector as the primary display. Further, we also wish to define research questions that were until now not possible to be studied in our application scenario due to the technical limitations in HMD display technology.

For the purpose of this study we developed two instances of one and the same assembly assistance system, the only difference being the way in which information is presented. One system consists of a projection-based display, the other system uses a HoloLens (Fig. 3). A preliminary evaluation is carried out by expert users who have developed and used these assembly assistance systems.

4 Design and Implementation of Evaluation Prototype

4.1 General Setup of Evaluation Prototype

The prototype systems consist of two height adjustable workbenches, each with two rows of workbins whose height, tilt and depth can be adjusted (Fig 2).

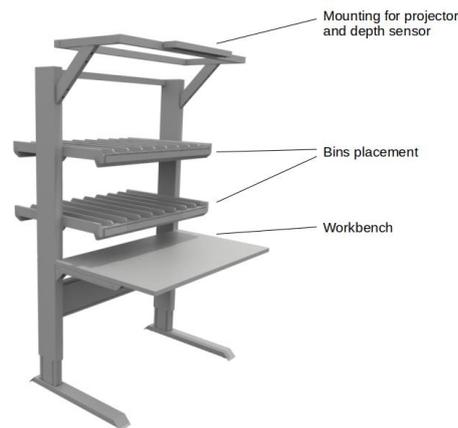


Fig. 2. Assistance System Setup.

The first system contains a projector and a depth sensor attached vertically above the workbench, whereas on the second system only the depth sensor is used (Fig. 3). Data from the depth sensor is processed via a hand recognition algorithm as developed by Büttner et al. [18]. The algorithm can detect if the correct part has been picked, or it can be used to make specific areas on the table touch sensitive. On both the systems, an LCD touchscreen is mounted above the bins in order to choose specific assembly instruction tutorials. The assembly system software runs on a PC that also acts as a server and is able to connect to multiple clients.



Fig. 3. System setup with projection-based display (left) and HoloLens (right).

4.2 Prototype using the Projection-based Display

The projector displays content directly onto the workbench (Fig 3, left). Along with the depth enabled gesture recognition, this transforms the entire workbench into a touchscreen display, where specific, predefined areas are configured as touch inputs. In the pick step the correct bin is highlighted, and in the assembly step the assembly instruction and a confirmation icon are projected. Once the physical setup is configured, the projection is stable in repeated usage with change in table height or lighting conditions.

4.3 Prototype using the HoloLens

The hardware of the HoloLens consists of an inertial measurement unit (IMU), 4 environment understanding cameras, 1 depth camera, 1 ambient light sensor, a 12MP photo/video camera and 4 microphones. The display contains two see through holographic lenses utilizing waveguide technology. The entire unit weighs 579 grams. On the software side, libraries that provide gaze tracking, gesture recognition, speech recognition and spatial understanding are available for developer use. Applications are programmed via the Unity game engine which provides ray casting, graphics, physics and interaction support.



Fig. 4. HoloLens: sensor assembly (left) and display assembly (right), images taken from [19].

The main feature that sets the HoloLens apart from other VR and Immersive headsets is “Spatial Mapping”. According to Microsoft, “Spatial mapping provides a detailed representation of real-world surfaces in the environment around the HoloLens, allowing developers to create a convincing mixed reality experience. By merging the real world with the virtual world, an application can make holograms seem real. Applications can also more naturally align with user expectations by providing familiar real-world behaviors and interactions.” [20]. Spatial mapping allows holograms to exhibit physical behavior, for example occlusion and physics.

A second related, but important feature is that of a “Spatial Anchor”. As stated on the Microsoft website, “A spatial anchor represents an important point in the world that the system should keep track of over time. Each anchor has a coordinate system that adjusts as needed, relative to other anchors or frames of reference, in order to ensure that anchored holograms stay precisely in place.” [21] One anchor point can serve as a frame of reference for more than one hologram in a local proximity. A typical use case is that of a board game, in which holographic objects need to be placed on a flat surface. In our use we observed that anchor placement is primarily vision based, achieved through the environment understanding cameras.

4.4 System Operation with HoloLens

Two ‘virtual screens’ placed in front of the user display a picture or video of the assembled part, and an additional, larger screen in the middle shows text instructions (Fig 5, left). A green holographic arrow indicates where the correct component is located (Fig. 5, right), which the user picks to advance to the next step. The assembly step displays the intended assembly, which the user needs to confirm by tapping on a holographic buzzer button placed on the table (Fig. 6, left).



Fig. 5. Instruction screens (left) and picking instruction (right).

The user can navigate through the tutorial steps, or cancel a tutorial by tapping on the touchscreen placed in front left (Fig. 4). The position of the arrow and the virtual button is user configurable prior to starting the assembly tutorial.



Fig. 6. Confirmation button (left) and calibration setup (right).

One of the primary requirements of the assembly application is that the visual instructions need to be stable in all conditions of use. This implies that the placement of the visual objects should not be affected by factory conditions or physical adjustments by the user. With a top mounted projector this requirement is easily met; with the HoloLens this is achieved by making use of a QR code which serves as a physical marker. Prior to using the assembly station, the user needs to perform a ‘calibration’ where a digital image of the QR code is superimposed onto the physical code [Fig. 6]. The calibration makes use of feature points and OpenCV to perform the alignment, after which the position of the QR code is saved as a spatial anchor. All the holograms displayed in the workplace use this anchor as a reference point.

5 Preliminary System Investigation

To evaluate this setup, we constructed a scenario in which the HoloLens replaces the projector in everyday use for our application. In order to achieve that, the HoloLens would have to surpass both the technical and functional capabilities of the projection-based display. From a technical perspective, we compared the stability of the MR display of the HoloLens with the existing 2D projection implementation under different everyday scenarios. For a functional evaluation, expert users were asked to use the system, complete an assembly task, and their responses were recorded.

Display Stability. By stability is meant the ability of the system to maintain a stable projection. A test was conducted by simulating typical workplace use cases. Each situation was repeated 10 times and the results recorded.

a) Start/stop Scenarios and Pauses

This test simulates beginning a shift, taking a break and resuming work. The HoloLens is taken off and placed on the desk without quitting the assistance application. Work is resumed after 5 minutes of inactivity. In 8 out of 10 cases, the projections were stable.

In the remaining 2 cases a drift was observed and a recalibration was needed to correct this behavior.

b) Table Height Adjustment

The table's height can be changed anytime during the assembly. This is where we found the HoloLens had most trouble. The holograms would stick to the workplace in only 5 out of 10 cases, that is, only 50% of the time, in the remaining 5 cases they did not follow the desk, i.e. they were left either floating in air or intersecting with the physical workplace. In order to bring the holograms back into place a re-calibration had to be performed.

Subjective Evaluation. In order to compare the HoloLens with previous smart glass implementations, we use the same metrics as in the prior study conducted by Büttner et al. [1]. The results here were obtained via observation and discussion with expert users following the completion of the assembly task. The following points were considered:

a) Robustness under Bright Lighting Conditions

The evaluation was conducted in a factory like environment with a mix of both natural and artificial light. In the previous study it was reported that users had difficulty when the background lighting was set to above 500 lux, with the HoloLens, users did not report any issues due to ambient lighting.

b) Restrictions of Field of View

The Vuzix Star 1200 glasses used in the previous study had a field of view of 35 degrees, while the HoloLens offers a field of view of 70 degrees. Compared to the projection display this amounts to about 70% coverage of the workstation at a distance of 0.5 meters from the assembly bins. Still, most users noticed the restriction in the field of view immediately.

c) Fixation of the Overlay

In the previous study the users reported discomfort due to a misalignment of the interpupillary distance, and also because the position of the glasses on the nose bridge affected the digital overlays. In the case of the HoloLens no such issue was observed.

d) (Un-)Natural Overlays

It was noted in the previous study that "HMD are beneficial for showing virtual objects in a free space, but when augmenting existing physical objects, the overlay may still be perceived as a separate entity and therefore be perceived as unnatural" [1]. When using the HoloLens, however, the integration between the real and the virtual world was unnoticed. In fact, the level of realism created other UI quirks - users expected the virtual objects to behave as real-world objects, for instance most users expected that the buzzer button (Fig. 6) needs only be tapped and not pressed or clicked, which led us to re-adjust the 3D camera's sensitivity to detect a tap at a certain height above the desk.

e) Stereoscopic View and Focusing Issues

No problems due to focal distance were reported. The holograms appeared to be in focus at all virtual distances.

f) Smart Glasses and Optical Aids

The HoloLens' placement on the head can be adjusted such that the display sits in front of optical glasses, all users in our evaluation were able to use the HoloLens without any discomfort.

g) Drawbacks of Wearable Devices in Industrial Settings

In the prior study the HMD was wired, which was noted to restrict mobility. Since the HoloLens has an on-board battery, it can be used wirelessly. Users could perform tasks such as picking up a custom piece from the assembly line situated a few meters away, however, at 579g the weight of the HoloLens is high compared to other smart glasses. The on-board battery life is 2-3 hours of active use which makes it infeasible for use in an 8-hour shift. Although it is possible to use the HoloLens while charging, it comes at the cost of reduced mobility.

6 Discussion

In our view, the HoloLens is a significant step up from existing smart glasses. It is still too heavy and has a battery life unsuitable for regular shift use in manufacturing environments, and the limited field of view can hinder the workflow in cases when UI elements go out of view. In our case, when the green holographic arrow pointing to the correct bin was outside the field of view, some users thought that the UI was inactive.

A related issue was regarding the distance of hologram placement. Microsoft recommends that the holograms be placed “in the optimal zone – between 1.25m and 5m” from the user [20], but this may not be possible if the bins are to be kept within arm's reach.

We also realized that the design of the workplace itself was also proving to be a hindrance in using the HoloLens effectively. For instance, the coexistence of 2D and 3D objects in the same space puzzled users as to which objects were interactable. In addition, users expected holograms to follow the rules of object placement in the real world, and drifts in hologram positions were particularly confusing.

Another problem was noticed when users wanted to discuss or ask questions about the UI. With the projection-based display it is possible to have more than one user interact with the UI. On the HoloLens this ability needs to be programmed into the application, and more than one HoloLens is required, which means more development effort and costs.

On the positive side, in the course of development and use, we found several features of the HoloLens had an obvious advantage, predefined areas are configured as touch inputs over a projection display. While a projector is mounted on the workstation, the HoloLens has no such limitation as long as an anchor can be located in the visible environment. A projector needs a 2D plane to display information, therefore a 3D space needs to be constructed as a staggered collection of non-overlapping 2D planes. For instance, each new row of bins needs to be placed a few cm behind the row below, which restricts the number of rows of bins that can be used before the 2D area available extends beyond the projector's reach. In case of the HoloLens, the 3D space around a physical object is available for augmentation, therefore any kind of bin arrangements

can be used as long as there is enough space for virtual objects to exist along with real objects. With a HMD such as the HoloLens the workplace can be reconfigured in real time.

A benefit of dissociating the display from the physical space is the possibility to allow more than one device to exist in the same space. In the course of our development we were able to use 2 HoloLens devices in the same workspace. This opens up the opportunity to segment the UI to cater to specific roles. For instance, the assembly supervisor may wish to see a different view of the assembly process than the assembly operator.

Lastly, the HoloLens extends the interaction modalities of the workstation as it supports speech, gaze and gesture-based interaction on its own. Along with the 3D sensor and the touch screen display, this amounts to more options from a usability perspective.

7 Topics for Future Work

This preliminary study was carried out by expert users to compare two display technologies, however, it remains to be seen if the use of MR displays such as HoloLens actually results in usability and productivity improvements in assembly tasks of various levels of difficulty. In future we would like to conduct a more detailed study with larger user groups.

The technological advancements which the HoloLens offers also brought to light the lack of usability research in using MR HMDs. Design guidelines in this regard are few and far between, the only resource available so far is Microsoft's own set of guidelines which are based on the experience of HoloLens' development [22].

The effect of MR technology on industrial workplace design may also be an area of future research. As this technology matures, it may be worthwhile to investigate how the workplaces of the future may be designed to accommodate both real and MR environments.

Studies on using HMDs such as the HoloLens in 8-hours shifts in industrial settings have not been carried out, therefore the occupational health effects of these displays are also unknown.

8 Conclusion

In this paper, we presented an early expert evaluation of two different display types for an assembly assistance system – a HoloLens mixed reality head mounted display and an existing state of the art projection-based display. Preliminary feedback indicates that the HoloLens has greatly improved on the limitations of HMD displays, but due to its weight and low battery life, it cannot replacement projection-based displays in assembly assistance at present. However, assistance systems are also used in other areas of factories, for instance training systems, where the HoloLens may fulfill the technical and functional requirements. Nonetheless, a deeper investigation is needed to understand the effects of mixed reality-based displays in industrial settings.

References

1. Büttner, S., Röcker, C., Sand, O. (2016). Using Head-Mounted Displays and In-Situ Projection for Assistive Systems – A Comparison. In: *Proceedings of the 9th ACM International Conference on Pervasive Technologies Related to Assistive Environments* (PETRA'16), ACM New York, USA, Article No. 44.
2. Funk, M., Mayer, S., Schmidt, A. (2015). Using In-Situ Projection to Support Cognitively Impaired Workers at the Workplace. In: *Proceedings of the 17th International ACM SIGACCESS Conference on Computers and Accessibility*, pp. 185-192.
3. Gold, S. (2017). *Tomorrow's Factory Jobs*, available online: <https://www.siemens.com/innovation/es/home/pictures-of-the-future/industry-and-automation/digital-factories-personalized-workstations.html>.
4. Guo, A., Raghu, S., Xie, X., Ismail, S., Luo, X., Simoneau, J., Gilliland, S., Baumann, H., Southern, C., Starner, T. (2014). A Comparison of Order Picking Assisted by Head-Up Display (HUD), Cart-Mounted Display (CMD), Light, and Paper Pick List. In: *Proceedings of the 2014 ACM International Symposium on Wearable Computers*, pp. 71-78.
5. Lasi, H., Fettke, P., Feld, T., Hoffmann, M., (2014). Industry 4.0. In: *Business & Information Systems Engineering*: Vol. 6, No. 4, pp. 239-242.
6. Orfi, N., Terpenney, J., Sahin-Sariisik, A. (2011). Harnessing Product Complexity: Step 1—Establishing Product Complexity Dimensions and Indicators. In: *The Engineering Economist*, Vol. 56, No. 1, pp. 59-79.
7. https://www.assemblysolutions.de/index_en.html
8. <https://www.boschrexroth.com/de/de/produkte/produktneuheiten/montagetechnik/active-assist>
9. <https://www.vuzix.com/Products/LegacyProduct/6>
10. Paelke, V., Röcker, C., Koch, N., Flatt, H. (2016). User Interfaces for Cyber-Physical Systems: Expanding the Designer's Toolbox. In: *at – Automatisierungstechnik*, Vol. 63, No. 10. De Gruyter, Oldenbourg, pp. 833-843.
11. Paelke, V., Büttner, B., Mucha, H., Röcker, C. (2017). A Checklist-Based Approach for Evaluating Augmented Reality Displays in Industrial Applications. In: S. Trzcielinski (Ed.): *Advances in Ergonomics of Manufacturing: Managing the Enterprise of the Future*. Springer, Heidelberg, Germany, pp. 225-234.
12. Sand, O., Büttner, S., Paelke, V., Röcker, C. (2016). smARt.Assembly – Projection-Based Augmented Reality for Supporting Assembly Workers. In: Lackey S., Shumaker R. (Eds.): *Virtual, Augmented and Mixed Reality*, Lecture Notes in Computer Science, Vol. 9740, Springer International Publishing, Switzerland, pp. 643-652.
13. Zheng, X. S., Foucault, C., Silva, P. M., Dasari, S., Yang, T., Goose, S. (2015). Eye-Wearable Technology for Machine Maintenance. In: *Proceedings of the 33rd Annual ACM Conference on Human Factors in Computing Systems (CHI'15)*, ACM Press, NY, USA, pp. 2125-2134.
14. Khuong, B. M., Kiyokawa, K., Miller, A., Viola, J. J., Mashita, T., Takemura, H. (2014). The Effectiveness of an AR-Based Context-Aware Assembly Support System in Object Assembly. In: *Proceedings of the IEEE Conference on Virtual Reality (VR'14)*.
15. Biocca, F., Tang, A., Owen, C., & Xiao, F. (2006). Attention Funnel: Omnidirectional 3D Cursor for Mobile Augmented Reality Platforms. In: *Proceedings of the Conference on Human Factors in Computing Systems (CHI'06)*, ACM Press, NY, USA, pp. 1115-1122.
16. Grubert, J., Hamacher, D., Mecke, R., Bockelmann, I., Schega, L., Huckauf, A., Tumler, J. (2010). Extended Investigations of User-Related Issues in Mobile Industrial AR. In: *Proceedings of the IEEE International Symposium on Mixed and Augmented Reality (ISMAR'10)*. IEEE Computer Society Press, pp. 229-230.
17. Fellmann, M., Robert, S., Büttner, S., Mucha, H., Röcker, C. (2017). Towards a Framework for Assistance Systems to Support Work Processes in Smart Factories. In: A. Holzinger P. Kieseberg, A M. Tjoa, E. Weippl (Eds.): *Machine Learning and Knowledge Extraction*.

Proceedings of the International Cross Domain Conference for Machine Learning & Knowledge Extraction (CD-MAKE'17), LNCS 10410. Springer, Heidelberg, Germany, pp. 59 - 68.

18. Büttner, S., Sand, O., Röcker, C. (2017). Exploring Design Opportunities for Intelligent Worker Assistance: A new Approach Using Projection-Based AR and a Novel Hand-Tracking Algorithm. In: A. Braun, R. Wichert, A. Maña (Eds.): *Ambient Intelligence. Proceedings of the European Conference on Ambient Intelligence and the Internet of Things (AmI'17)*, LNCS 10217. Springer, Heidelberg, Germany, pp. 33 - 45.
19. https://developer.microsoft.com/en-us/windows/mixed-reality/HoloLens_hardware_details
20. <https://developer.microsoft.com/en-us/windows/mixed-reality/hologram>
21. https://developer.microsoft.com/en-us/windows/mixed-reality/spatial_anchors
22. https://developer.microsoft.com/en-us/windows/mixed-reality/about_this_design_guidance
23. <http://flexmation.com/components/pick-to-light-systems/>
24. Tang, A., Owen, C., Biocca, F., Mou, W. (2003). Comparative Effectiveness of Augmented Reality in Object Assembly. In: *Proceedings of the Conference on Human Factors in Computing Systems (CHI'03)*. ACM Press, NY, USA, pp. 73-80.
25. Milgram, P., Kishino, F. (1994). A Taxonomy of Mixed Reality Visual Displays. In: *IEICE Transactions on Information and Systems*, Special issue on Networked Reality, Vol. 77, No. 12, pp. 1321-1329.