2019 IEEE International Conference on Artificial Intelligence and Virtual Reality (AIVR)

# Augmented Reality for Human-Robot Cooperation in Aircraft Assembly

Andreas Luxenburger*, Jonas Mohr*, Torsten Spieldenner§, Dieter Merkel*, Fabio Espinosa*, and Tim Schwartz*
*Cognitive Assistants, §Agents and Simulated Reality
German Research Center for AI (DFKI GmbH)
Saarbrücken, Germany
{firstname.lastname}@dfki.de

Florian Reinicke, Julian Ahlers and Markus Stoyke
Technology / Development
Broetje Automation GmbH
Rastede, Germany
{firstname.lastname}@broetje-automation.de

***Abstract*—Augmented Reality (AR) is often discussed as one of the enabling technologies in Industrie 4.0. In this paper, we describe a practical application, where Augmented Reality glasses are used not only for assembly assistance, but also as a means of communication to enable the orchestration of a hybrid team consisting of a human worker and two mobile robotic systems. The task of the hybrid team is to rivet so-called stringers onto an aircraft hull. While the two robots do the physically demanding, unergonomic and possibly hazardous tasks (squeezing and sealing rivets), the human takes over those responsibilities that need experience, multi-sensory sensitiveness and specialist knowledge. We describe the working scenario, the overall architecture and give design and implementation details on the AR application.**

***Index Terms*—Augmented Reality, aircraft manufacture, mobile robots, teamwork, hybrid teams, Industrie 4.0, worker assistance**

## I. Introduction

Augmented Reality (AR) aims at extending physical, real-world contents by virtual, interactive objects which are usually registered to the real world in 3D. Since both the presentation of overlays as well as the interaction with these typically involve different input and output modalities, applying concepts of AR in Industrie 4.0 can create immersive and flexible applications for assisting workers in their daily tasks (e.g. [1]–[3]). An example would be the visual highlighting of a set of rivet positions, which have to be checked regarding quality.

While AR glasses like the Microsoft HoloLens provide real-time tracking of a worker's head pose from which an approximated direction of gaze can be inferred, one can go one step further and adapt the style of presentation to the current situation by incorporating the worker's position and focus. In turn, this contextual information can be used to guide a worker's focus of attention to specific regions of interest belonging to a specific task that has to be fulfilled. In particular, when using strategies of contextually adapted presentation, one typically refrains from rigorous, static insertions in the worker's field of view, which would potentially impede ongoing work. In contrast, we present a situation-adaptive, multimodal notification system for workers, which uses a level of detail (LOD) approach similar to [4] for first guiding the worker unintrusively to the desired working location and finally exposes an intuitive feedback mechanism for reporting the quality state of rivets and task acknowledgment. Both aspects, the worker assistance as well as the feedback mechanisms, are also used to coordinate a so-called hybrid team, i.e. a team consisting of humans and robots.

Funded by the German Federal Ministry of Education and Research under grant numbers 01IS16026A (Hybr-iT) and 01IW17004 (TRACTAT)

## II. System Architecture

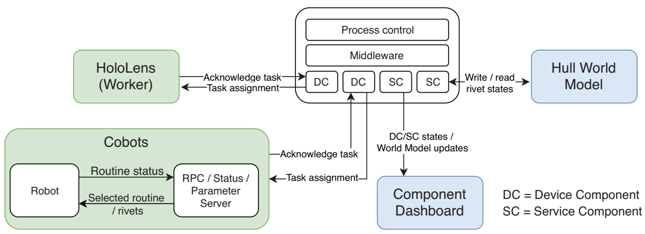


Fig. 1. Overview of the system, including a central controlling instance, devices and services.

Besides the two robots (see also Section VI), the demonstration system consists of several sub modules: a central controlling instance (based on BaSys 4.0; see https://www.basys40.de) that holds information about the current process, monitors the current working progress and assigns tasks to team members–which are modelled as device components–via a communication middleware. A distributed world model, which is based on the linked data paradigm (see Section V-A) is modelled as service components and provides information about the real world (i.e. the state of the aircraft hull and all rivet positions). A component dashboard allows to monitor the current state of the system. Finally, a Microsoft HoloLens is used as a means of communication with the human worker. Figure 1 shows a high level block diagram of the overall architecture.

## III. Multimodal, AR-based Worker Guidance System

### A. Adaptive Level of Detail (LOD) Guidance

The presentation of an upcoming task by our AR-system can basically be seen as a three-step procedure: First, the worker's

978-1-7281-5604-0/19/$31.00 ©2019 IEEE
DOI 10.1109/AIVR46125.2019.00061

attention should be shifted to the direction of the task. Second, the level of detail of the presented information of the task should be adapted according to the current position of the worker relative to the position of the task while providing a guiding mechanism. Finally, a fully detailed task description should be provided to the worker with the possibility of acknowledging the task by associated holographic widgets.

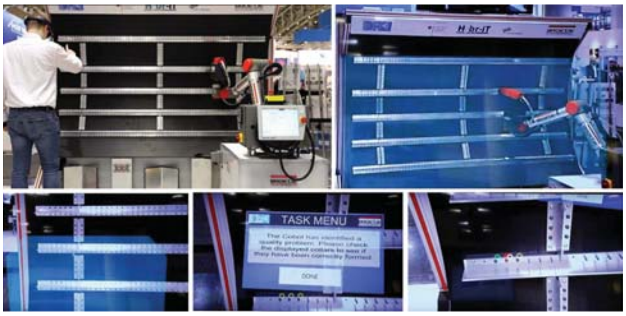


Fig. 2. Multimodal, AR-based worker guidance system for assisting quality assurance processes in aircraft production. *First row:* **(a)** Worker performs quality check. **(b)** Visual highlighting of associated working region from a distance. *Second row:* **(c)** More specified region as seen from decreased distance. **(d)** High detail information about worker task. **(e)** Interactive rivet markers with color codes for different quality states.

For shifting the worker's attention to the direction of the task, we used the spatial sound feature of the Microsoft HoloLens, which allows to provide a feeling from where in 3D space a certain sound signal comes from. This was implemented by attaching a virtual sound source emitting a pulsing notification sound to the task location. We scaled the volume of the sound signal in accordance with the current distance of the worker to the location and the angle between the worker's direction of gaze and the normal vector of the sound source.

Once the worker is facing the direction of the task, the impact of the guiding sound signal is reduced by our system and a visual component takes over. In addition to the actual task description and the set of rivet positions which has to be checked, the central system sends related meta-information to the AR component like associated spatial hierarchies as e.g. 3D regions of interest (ROI) in which the task itself is situated. In our demo scenario, as can be seen in Figure 3, we divided the aircraft hull into different sectors, which are then associated to corresponding levels of detail. Each sector corresponds to a ROI ranging from the whole hull (LOD-0) over left, middle and right sectors (LOD-1) to corresponding upper and lower regions (LOD-2), where the target stringer unit is located. The presentation of the rivet checking task with highest level of detail (LOD-3) finally consists of a set of colored, ring-shaped markers, which indicate the target rivet positions and are accompanied by a task menu, which describes the quality assurance task textually and provides a button for acknowledging a finished task.

After the worker has checked the desired positions, feedback about the quality state of a rivet can be reported by looking at the corresponding rivet marker and switching to quality-indicating colors (yellow: unchecked, green: `IO`, red: `NIO` where `IO` and `NIO` stand for *In Order* and *Not In Order*, respectively) either by tap gestures or with the help of a Bluetooth controller, which has been integrated into the testing mandrel. One should mention here, that the worker has the freedom to check the squeezed rivets in any order they choose and to report an intermediate result to the world model. The latter point is important, e.g. if the worker needs to take a break. In that case, the sealing robot can already work on the checked rivets and the worker can resume their work on the remaining rivets at a later time.

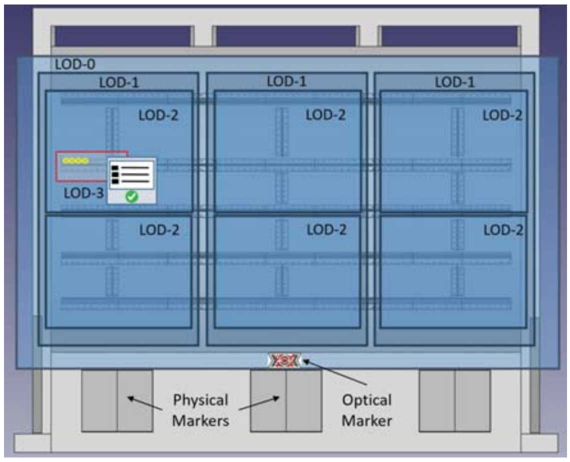


Fig. 3. LOD-based task presentation: depending on the distance of the worker to the task location, the level of information shown in AR is adapted from highlighting related sectors of the hull (LOD-0 to LOD-2) to marking the desired set of rivets (yellow rings) and showing an associated task description. Associated 3D models are registered with the help of an optical marker.

### *B. Implementations and Alignment of 3D Models*

Our implementations of the presented AR-system are based on the Unity game engine (https://unity.com) and Vuforia (https://engine.vuforia.com), an AR application development platform, which offers robust object and marker detection and that is integrated into Unity. In order to define a global reference point for holographic contents in the real-world production environment, we use the concept of spatial anchors. For the HoloLens a spatial anchor represents an important point in the real world, which is visually tracked by the system over time. All virtual contents are then positioned in the coordinate system of this anchor. The quality of the resulting alignment of real-world objects, such as the aircraft hull, sectors within the hull, stringer units in the sectors and rivet positions within one stringer unit and their virtual counterparts depends on two main aspects: first, the precision of the visual tracking of the spatial anchor and second the quality of the virtual representation of the real-world objects in terms of 3D CAD models.

While spatial anchors are often user-created by intersecting the user's gaze with the estimated mesh of the surroundings (which is a result of the spatial mapping of the HoloLens), we opt for a more precise definition of the spatial anchor with the help of a fixed optical marker. Vuforia offers a set of customised markers called VuMarks, which are able to encode object specific data, such as a URI, while simultaneously acting as AR-target. We placed one of these markers at a centered ground position of the aircraft hull (see Figure 3). Once the marker has been detected by Vuforia, we convert its 2D center to a 3D world position constituting the spatial anchor. Ideally, this procedure of registering virtual to real-world content has to be done only one time, as spatial anchors can be stored in the HoloLens.

However, as can also be seen in Fig. 2(e), the alignment suffers from drifts over time, which strongly depend on the worker's position relative to the spatial anchor. In order to provide a correction mechanism, we integrated interactive elements to the system's main menu for triggering a registration process where the worker has to look at the marker and perform a simple tap gesture in order to re-define the spatial anchor. The positional error of the virtual content increases with the distance of both the worker and the virtual content from the spatial anchor due to inaccuracies of the visual tracking and the well-known lever effect. In general, we found that the positional accuracy of the virtual overlays was good enough in order to provide a visual guidance for the worker and a helpful highlighting of the rivet positions, but there is still room for optimisations.

## IV. Description of the Production Scenario

The use-case for this demonstration is the fitting of an aircraft hull with so-called stringers and ties (see Figure 4). These are load-bearing components that usually are riveted to the hull, using so-called lockbolt-rivets. These rivets are inserted from one side into pre-drilled holes through the parts to be connected, and are then pulled off from the other side, using a special tool. The act of pulling-off or squeezing the rivets can be very strenuous for a worker, as the tool is rather heavy and produces a strong mechanical impact with each actuation. After squeezing, each rivet has to be checked for any inaccuracies or quality issues. This check is done by hand using a testing mandrel, but also through visual inspection as well as feeling it with bare fingers. If a rivet was checked to be okay, it has to be sealed, again using a tool to apply a special sealing compound. Although not highly dangerous, the sealant is slightly toxic and should therefore only be handled while wearing protective eye-wear, gloves and a respirator. It is therefore desirable to assign these tasks to robots.

Although it would surely be possible to program a robot to check the squeezed rivets, such a robotic system would only check for those parameters that were programmed into such a system. In contrast to this, a human checking the rivets can detect all sorts of issues, including those that are not directly related to the rivets themselves (e.g. a hairline crack in the hull). Accordingly, the hybrid team in the scenario consists of two mobile robots, one capable of squeezing rivets, the other one capable of sealing them, and a human worker, who is responsible for the quality control.

## V. World Model

### A. Linked Data Model

Providing understandable World data to several heterogeneous consuming clients requires a well designed integration layer that achieves both structural and (semantic) data interoperability between applications. For this, the Semantic Web community emphasizes the importance of a sufficiently semantically described server API (see [5], [6]). Following this notion, we model relevant robotic data, as well as data of material and workers, as Semantic Web resources. Our resulting model follows the Linked Robotic Things specification [7].

In a nutshell, each relevant datum is published as a `HTTP/1.1` endpoint with individual resolvable URI. Relations and dependencies between resources are expressed explicitly by a direct link to the URI of the relevant HTTP resources (*Linked Data*).

A `HTTP GET` request to a resource returns a semantic self-description in RDF (https://www.w3.org/TR/rdf-concepts/) [8]. Likewise, a `HTTP OPTIONS` request provides further information about interaction methods, such as real-time subscription updates, in a machine-interpretable representation. This description, along with links between resources explicitly modelled, allows clients to self-drivenly explore provided data, and infer further knowledge about individual resources.

In our application, clients use the SPARQL (https://www.w3.org/TR/sparql11-query/) query language to direct queries like, *"Return all rivets with certain status in a given region of the hull"*, directly to the server API. The return of this call would be a set of URIs pointing to the respective resources for the relevant rivets. The client can then use these URIs to read or modify the content of the resources, and by this the state of the world, using HTTP operations on the resources as specified by *Linked Robotic Things*. The World server is implemented as standalone application. It uses the ECA2LD framework [9] to lift the local data model of the server run-time to its Linked Data representation.

### B. Visualization of World Model and Basic Architecture

To give the worker a better overview of the current state of each rivet, a real-time 3D visualization of the hull was built (see Fig. 4). The application was implemented as a web-based application using Babylon.js and has been integrated into a component dashboard of BaSys 4.0, a virtual middleware for Industrie 4.0 applications.

The current state of each rivet (`EMPTY`, `INSERTED`, `SQUEEZED`, `CHECKED_IO`, `CHECKED_NIO` or `SEALED`) is visualized with the help of different 3D models and color codes (see Fig. 4) and gets updated in real-time, based on the world model.

## VI. Robotic System

The processes described before were carried out by two mobile cobot systems with specialized end-effectors

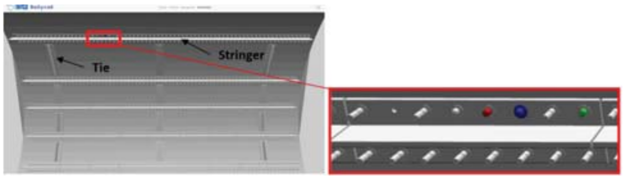


Fig. 4. *Left*: Real-time visualization of the aircraft hull. Black arrows and texts are not part of the visualization and were added to clarify terminology. *Right*: Different visualizations of the rivets and color codes show the current state of each rivet (f.l.t.r. INSERTED, EMPTY, INSERTED, SQUEEZED , CHECKED_NIO, SEALED, INSERTED, CHECKED_IO).

designed and assembled by Broetje-Automation. Both robotic platforms, i.e the squeezing robot and the sealing robot, each consist of a mobile MiR100 platform (https://www.mobile-industrial-robots.com), a Universal Robot (https://www.universal-robots.com) robotic arm and all the required components to operate the pneumatic tools.

### A. Robot Hardware

The challenge in developing the end effectors was to meet the maximum allowable load of the cobot, as well as generating the necessary process forces for the squeeze process. In addition to this, the process has to be as stable as possible while attaining fast cycle times. Additional sensors were integrated into the cobot system to achieve the required accuracy for the processes. First, physical markers (metal brackets with a specific angle) were integrated into the aircraft hull carrier, which can be recognized by the laser sensors of the mobile platform and enable an improved positioning in front of the workpiece. Furthermore, cameras were integrated into the end effectors of both cobots. Combining preknown position data derived from the CAD model with the improved platform positioning and camera-based referencing allowed for a precise targeting of individual rivet positions without the need to teach these positions manually.

### B. Robot Software Architecture

In order to establish a bi-directional communication with the cobots, a generic, RPC-based architecture was implemented (see Fig. 1), which can handle both robotic tasks of the described scenario. The function of the RPC server are in charge of translating and storing high-level commands received from the central controlling instance and to execute the corresponding low-level robot routines. The server offers functions for setting the needed robot routine along with the associated set of parameters, e.g. squeezing a specific set of rivets. Additionally, the overall execution state as well as detailed information about the selected routine can be retrieved, e.g. whether a rivet was properly processed or not. The robot's device component in the BaSys 4.0 middleware propagates this information to the world model, such that it can be accessed later. Regarding the concrete implementation, both RPC servers were running directly on the UR controllers, while the low-level robot routines were programmed and executed by using the controller's native operating system and programming environment (PolyScope).

## VII. Conclusions and Future Work

We described a practical scenario, in which AR glasses are used to guide a worker through their working steps as well as to provide the basic communication means to coordinate a hybrid team. The demonstration system described was shown live during the German Industry Fair in Hannover (April 1st – 5th 2019), i.e. 8 hours per day during 5 consecutive days. During that time, we could already identify some practical issues. Wearing AR glasses for longer periods of time is strenuous, where comfort is one issue that hardware companies are currently working on. Another aspect here is the frequent (cognitive) switching of the user between virtual and real objects. A simple countermeasure could be to show or hide the virtual objects on demand (triggered by gesture, hardware buttons, speech etc.). A user study examining these aspects are planned as future work. Currently, we are extending the overall system with an optimizing component that tries to allocate the working tasks in such a way that down-times of the robots are minimized while ensuring optimal working conditions for the human worker.

## Acknowledgment

We would like to thank our project partners at Airbus, BAuA, The Captury, DFKI Robotic Innovation Center, EngRoTec, Kuka and Volkswagen for valuable input during the discussions about this use-case as well as during the implementation phase. Special thanks to Daniel Porta for his BaSys 4.0 implementation and support.

## References

[1] V. Paelke, "Augmented reality in the smart factory: Supporting workers in an industry 4.0. environment," in *Proceedings of the IEEE Emerging Technology and Factory Automation (ETFA'14)*. IEEE, 2014, pp. 1–4.
[2] R. Masoni, F. Ferrise, M. Bordegoni, M. Gattullo, A. E. Uva, M. Fiorentino, E. Carrabba, and M. Di Donato, "Supporting remote maintenance in Industry 4.0 through Augmented Reality," *Procedia Manufacturing*, vol. 11, pp. 1296–1302, 2017.
[3] D. Mourtzis, V. Zogopoulos, and E. Vlachou, "Augmented Reality application to support remote maintenance as a service in the robotics industry," *Procedia CIRP*, vol. 63, pp. 46–51, 2017.
[4] S. J. Henderson and S. Feiner, "Evaluating the benefits of augmented reality for task localization in maintenance of an armored personnel carrier turret," in *8th IEEE International Symposium on Mixed and Augmented Reality*, 2009, pp. 135–144.
[5] T. Berners-Lee, J. Hendler, and O. Lassila, "The semantic web," *Scientific American*, vol. 284, no. 5, pp. 34–43, 2001.
[6] S. Mayer, R. Verborgh, M. Kovatsch, and F. Mattern, "Smart configuration of smart environments," *IEEE Transactions on Automation Science and Engineering*, vol. 13, no. 3, pp. 1247–1255, 2016.
[7] R. Schubotz, C. Vogelgesang, A. Antakli, D. Rubinstein, and T. Spieldenner, "Requirements and Specifications for Robots, Linked Data and all the REST," in *Proceedings of the Workshop on Linked Data in Robotics and Industry 4.0 (LIDARI'17) located at Semantics'17*. CEUR, 2017.
[8] E. Miller, "An introduction to the resource description framework," *Bulletin of the American Society for Information Science and Technology*, vol. 25, no. 1, pp. 15–19, 1998.
[9] T. Spieldenner, R. Schubotz, and M. Guldner, "ECA2LD: From entity-component-attribute runtimes to linked data applications," in *Proceedings of the International Workshop on Semantic Web of Things for Industry 4.0. Extended Semantic Web Conference (ESWC'18)*. CEUR, 2018, 12 pages.