
Elena Stoll¹*, Jan Reiher¹, Karishma Jagdish Gavali¹ and Dietrich Kammer¹

¹University of Applied Sciences Dresden, Faculty of Informatics/Mathematics, Friedrich-List-Platz 1, 01069, Dresden, Germany

Abstract

Smart factories still require manual work by humans to remedy defects in automated production lines. Given the shortage of highly skilled workers and the demand to include people with language barriers or cognitive challenges in the workforce, novel assistance systems must be introduced to the factories. We propose an Adaptive Visual Assistance systems using Spatial Augmented Reality (AVISAR), which not only adapts to different repair tasks and the layout of a manual workplace in the factory, but also to the individual human skills and needs. This should result in a more inclusive work environment and the well-being of a more diverse workforce. Spatial AR promises to be a more accessible approach than head-mounted displays. In this contribution, we present the conceptual framework and a first working prototype of the AVISAR system.

Keywords:
spatial augmented reality, mixed reality, adaptivity, human-centered computing

1. Introduction

Automation is a key driver for smart factories of the future. However, there still remain tasks for human workers when defects are detected in the automated production lines. Sustainability goals demand that defects are manually repaired in order to reduce scrap in production. This issue becomes even more severe when considering individual production processes with a batch size of 1. Given the demographic shift and the shortage of skilled workers, fewer expert workers remain for these manual tasks. It is also desirable to include people with cognitive challenges or language barriers in such processes. For quick on-boarding and ongoing guidance of such personnel, human-computer interaction can provide appropriate solutions [1]. The aim is to not only ensure operational efficacy, but also cultivate the sustainable well-being of the workforce by offering them tailored support in smart factories of the future.

We contribute conceptual considerations for an adaptive assistance system in industrial manufacturing based on Spatial Augmented Reality (SAR) as well as a partial implementation of the
system for future studies. AVISAR (Adaptive Visual Assistance System using Spatial Augmented Reality) is geared towards manual workstations in smart factories and features adaptivity on three distinct levels: human needs and skills, tasks, and the layout of the workstation. SAR is a branch of Augmented Reality (AR) that does not require users to wear specific glasses or handle additional devices such as tablet computers. The physical space is augmented with virtual information, e.g. using projectors, which makes it easily accessible for a diverse range of users. In contrast to wearable devices such as head-mounted displays, SAR is less obtrusive and tiresome to use over an extended period of time.

The next section reviews adaptive assistance systems and SAR solutions from the literature. Section 3 presents our concept and implementation of AVISAR. Consequently, we report an informal evaluation and discussion of the system in Section 4.

2. Related Work

Adaptive Assistance Systems. Assistance systems in manufacturing support workers in their assembly activities on the shop floor without replacing them, overruling them, or exposing them to risks [2]. Their main purpose is to compensate deficits (e.g. due to old age, lack of skills, or disabilities) and to expand existing skills. There are sensory, physical, and cognitive assistance systems [3], using visual tools, pictograms, or projection systems (such as AVISAR). Individual use of assistance systems gains importance in human-centered work of Industry 5.0. While current systems normally offer standardized views and functionalities, newer systems can be adapted to the experiences, cognitive abilities, and needs of groups or individuals [4].

Based on the conviction that “work should adapt to people and not people to work”, the field of human factors/ergonomics in human-computer interaction (HCI) has long investigated how adaptive user interfaces (AUI) can improve usability [5]. The field of intelligent user interfaces (IUI) leverages algorithms and machine learning to realize adaptations [6]. Personalized support and user experience of AUI are based on individual user profiles, analysis of the user style or state, work context, mental workload prediction [7], or integration of user traits from psychological theory [8]. Technological innovations such as artificial intelligence and improved sensing technologies spur this development at software and hardware level.

Adaptive interfaces are usually mentioned coincidentally with adaptable interfaces. While adaptive interfaces automatically adapt to changes in environmental conditions or user characteristics [9], adaptable interfaces require the initiative of the human user. The transition between the two concepts is fluid and intermediate stages are referred to as “levels of automation”, e.g. by [10, 11]. In the following, we present frameworks and implementations of adaptive assistance systems on the shop floor. Historically, they appear after machine-centered, flexible, and reconfigurable systems from the age of automation and mass production [12]. These new systems reflect a human-centered notion of adaptivity, where technology is used to adapt production processes to humans or consumer behavior and market dynamics on the macro level [13].

According to Peruzzini at al., adaptivity is the ability of the entire production system to adjust to new conditions, including machines, material flow, control systems, and personnel [14]. Applying human-centered and contextual adaptation rules positively affects system usability and process performance for elderly workers. Schlund and Kostolani [15] present
a framework for personalized work systems that integrates cognitive features for dynamic adaptivity across the dimensions operator (human), work equipment (machine), work process, workplace, and work environment. The authors cite implementations such as an adaptive projection of work instructions and other information to illustrate their concept. The INCLUSIVE System [16] is a general framework for industrial interaction systems that adapts to the skills and capacities of human operators. The framework relies on three modules: Measure (human capabilities), Adapt (the interaction), and Teach (unskilled workers) to tailor interaction with complex manufacturing systems to individual user requirements and particularly provide guidance and training. Objective measurements and user feedback obtained in industrial use cases showed the effectiveness of the approach. Similarly, Oestreich et al. present a conceptual approach for designing adaptive assistance systems in a human-centered way and create personalized user experiences, e.g. support learning, performance, motivation, or ergonomics [17].

Petzoldt et al. [18] examine human-centered assistance systems for manual assembly, including implementation approaches. Functionalities include both product- and process-related (assembly instruction, automatic configuration and calibration, progress recognition, quality control) as well as human-centered support dimensions (individualization, qualification, motivation, ergonomics). Bertram et al. provide an overview of assistance systems supporting manual work by visualizing and supervising tasks [19]. The authors propose mandatory aspects for future intelligent manual working stations such as assisted work instructions, recognition of products and working context, and autonomous learning ability, finding that adaptable integration in production is the least considered. Finally, Mark et al. identify nine potential user groups for assistance systems in production [20]. According to the authors, projection-based cognitive assistance systems with step-by-step picture or video-based task instructions and integrated quality control are especially suited for unskilled, inexperienced, mentally handicapped, migrant, or content-flexible users.

Research shows that human-centeredness in manufacturing includes both adaptations to operators and products as well as the working environment. Adaptations are achieved by means of an adaptation strategy that is guided by worker and process states, user information and interactions, and digital system information. Feedback loops may be integrated to continuously improve adaptations. A need for group-specific or personalized adaptations in order to compensate for cognitive deficits or missing skills is evident.

**Spatial Augmented Reality.** According to the reality-virtuality continuum [21], the area between the completely real and the completely virtual is referred to as Mixed Reality (MR). MR encompasses Augmented Virtuality (AV) and Augmented Reality (AR). SAR is a specific type of AR, which blends the real and virtual worlds in real time. The key characteristic of SAR is that no additional devices such as head-mounted-displays, data glasses, tablets, or wearables are required, since digital information is projected directly into the surroundings and onto objects [22]. This enables hands-free interaction coupled with digital augmentation, making it a promising technology for industrial assembly. Hence, all kinds of users can easily access the augmented workstation without any preparation or personal setup of additional devices. Next, we present SAR assistance systems designed for industrial use.

Zigart and Schlund tested the industrial readiness of a SAR system in the field of assembling electronic control panels with skilled workers and students [22]. Skilled workers suggested to
use the current system for new activities and training but not for series production. According to the authors, improvements could include automatic detection for switching to the next step and adapting the instructions to the qualification level. Hornacek et al. present a SAR system that enables and simplifies the effect of a keystone correction with an adjustable mirror and a downward-facing camera [23]. Their perspective correction has been tested for planar surfaces only, but with adjustments to warping and scene geometry recovery, it can also be applied to non-planar surfaces. The dynamic projection system by Rupprecht et al. detects human poses and gestures via the YOLOv3 algorithm and projects instructions relative to the worker using pictograms, simplified images, or elementary contours [24]. A comparison of static in-view (display information somewhere in the operator’s field of view) and guided in-situ (display of information at the place of action) instructions reveals that in-situ instructions have a shorter average total task time and perform slightly better in terms of usability and cognitive workload. Similarly, Uva et al. compared SAR to paper instructions, proving a reduction of completion time and error rates at manual workstations when using SAR [25].

Zhou et al. describe a SAR system which aims to improve the accuracy and efficiency of automotive spot welding inspection through dynamic in-situ visualization [26]. Specific geometric shapes indicate the inspection method. According to the authors, industrial applications of SAR include quality assurance, material handling, maintenance and repair, lowering of skill-level requirement, and training. The collaborative welding system by Tavares et al. uses projection to convey information to a human operator while orchestrating tasks between the operator and a welding robot [27]. SAR improves operator productivity and drastically reduces manufacturing errors. Mengoni et al. use SAR to support manual work in future smart factories and detect risks to human safety and the musculoskeletal system [28]. Usability, performance, and effectiveness proved to be higher than with pick by light system using LEDs, especially for difficult tasks.
3. Concept

Based on previous research on SAR and adaptivity, we describe an adaptivity framework for manual workstations on the shop floor in manufacturing. We then present a first prototypical implementation.

3.1. Adaptivity framework

Our framework addresses adaptivity in the dimensions of machines and factory layout, workers, and collaborative tasks, emphasizing human-centeredness and the workers’ well-being. The framework consists of four consecutive phases that form a closed loop (see Figure 1). Adaptation triggers from the operator, task, and workplace dimensions initiate an adaptation loop. The digitized information is analyzed software-side in an adaptation strategy and a decision is made about which adaptation to perform in the assistance system. Since our scenario is a SAR application, adaptations are limited to the projection and the projector. Finally, the adaptations are evaluated whether they are beneficial so that readjustments can be made accordingly. Each of the four phases is explained in detail below.

Adaptation Triggers. Potential triggers for adaptations in the operator, task, and workplace dimensions may be static or situational/dynamic. Static triggers relate to long-term characteristics such as individual prerequisites and skills of operators, general product dimensions, work instructions, or the layout of the workplace. Situational triggers change dynamically during work, and include the mental and physical condition, emotions or interactions of operators, faulty product assembly or changing environmental conditions in the factory such as lighting conditions or temperature. Depending on the type of trigger, different methods are suitable for monitoring and forwarding them to the adaptation strategy. Static triggers could be provided by user profiles or digital twins, while sensors are particularly suitable for capturing situational conditions.

Adaptation Strategy. An adaptation strategy decides on adaptations depending on the triggers. Various algorithmic implementations from simple, rule-based systems to machine learning algorithms are conceivable. However, producing human-centered decisions is of central importance. We propose a modular software architecture with components that analyze triggers and initiate adaptations. The calibration module adapts the projection to the workplace and surfaces. The interaction module detects movements and navigation gestures, whereupon the projection changes dynamically. The content module adapts the content projected, e.g. to the qualification of operators or incorrect assembly. The cognitive module plays a key role by adapting the system, e.g. to physical or mental limitations, experience, or situational conditions of workers. The motivation module ensures engaging adjustments such as short interruptions or the use of gameful design in case of fatigue or declining production speed, as proposed by [29, 30]. The health module processes health-threatening triggers such as extreme temperatures, poor lighting conditions, insufficient breaks, or bad poses. Finally, the notification module warns of critical situations.

Adaptation. Potential adaptations of the SAR system include the projection and the projector. Changes in the projection are possible for content (e.g. displayed information, language), style (e.g. shapes, colors, sizes, animation), or tone (e.g. level of detail, wording), and the projector
could be rotated, turned or height-adjusted.

**Evaluation.** The purpose of the evaluation phase is to ensure that the adaptations are beneficial for operator engagement and well-being, as well as for human-machine collaboration, in order to adjust the adaptation strategy if necessary. We propose a mixture of subjective and objective measures that can either be incorporated into the work process in real time or used for long-term evaluation. Both within and between subjects study designs are conceivable. Between subjects is suitable to compare user groups that have different barriers. Within subjects can be used to investigate how the interface performs in comparison to previous versions of the interface and if significant individual differences appear when adapting the novel interface. Objective measures include changes in stress, fatigue, accuracy, and error rates after preceding adaptations, and can be determined relatively well on a situational basis. Acute stress triggers changes in emotions, cognition, behavior, and physiological responses that can be continuously and unobtrusively recorded with sensors. Physiological measurements include heart rate, heart rate variability, skin conductivity, and thermal images. The coding of behavior may encompass location, gestures and interaction, gaze direction, facial expression, or speech. During acute stress, for example, the heart rate typically increases, whereby a greater variability of the heart rate signifies higher adaptability of the organism. These changes can be recorded using wearables such as a smartwatches or fitness trackers. In order to properly interpret measurements, stressful states should be compared with a no-stress baseline or restful states [31]. Alternatively, subjective self-report scales can be a viable option for long-term recording stress responses [32, 33]. Fatigue results from mental or physical exertion and is a multidimensional phenomenon with mental, physiological, and cognitive components lacking clear biological markers [34]. Measurement is usually based on uni- or multidimensional scales that capture the severity or type of fatigue. Specific performance aspects such as reduced reaction time, alertness, and short-term memory can indicate work fatigue. Measures of engagement and well-being, which are expressed, for example, in usability, user experience and on a psychological, physical, social and emotional level [29], need to be collected asynchronously to work. It is possible that these long-term surveys will result in the need to undertake general changes to the design and functions of the assistance system.

### 3.2. Use Case and Setup of the Prototype

AVISAR is located at the manual workstation of an Industry 4.0 model factory in which simplified smartphones consisting of a circuit board and two fuses are produced. Incorrectly assembled products are sent to manual repair. Our technical setup consists of a full HD ViewSonic M2e LED pico projector, a Leap Motion Controller for gesture recognition, and a 500mm x 100mm white, non-reflective projection surface mounted to the aluminium frame of the workstation (see Figure 2). The projector and the Leap Motion Controller are facing the product and the white projection surface from above, last of which is positioned at the operator’s waist height. There are two projection areas: On the product itself, in-situ projection highlights the faulty area, e.g. a missing fuse on the circuit board. Repair instructions are shown on the white surface. Several adaptation triggers are implemented in the prototype, not all of which work automatically yet. Gesture recognition can be used to confirm repair steps so that the next step is highlighted. We implemented static poses over the instruction area, which allows confirmation
Figure 2: Setup of the AVISAR system and sample repair instructions.

when holding a hand over one of the repair steps for a certain amount of this. This dwell time is visualized by a circular loading animation. Moreover, gesturing over the repair area mutes the projection in order to avoid misleading visual cues on the user’s hands. In addition, three user profiles have been integrated for operators with barriers in terms of qualifications, language, or digital skills. In the future, RFID tags will be integrated to automatically recognize the profiles. For simulating various situational tasks, six typical assembly errors were implemented, which should be automatically detected via image recognition in the future. Marker-based calibration of the projection to the layout of the workstation is also in preparation.

4. Discussion

AVISAR has already gone through several development stages. An earlier version, which did not yet include barrier profiles and gesture recognition, was evaluated in an informal user survey by visitors to an Industry 4.0 model factory at university. 14 people in total took part in our survey (with an average age of 39). The participants were professionals with backgrounds in engineering, chemistry, process technology, and digitalization. Students and department heads took part as well. AVISAR was shown as part of the complete tour of the factory and the functionality and ideas behind the concept of the system were demonstrated. Afterwards, the participants answered 8 questions on a 6 point Likert scale about the general recognisability and arrangement of the texts and projections, icons and animations, as well as confidence to work independently using the assistance. Overall, the feedback was positive. Projections were easily recognizable, instructions were clear and users felt able to carry out repairs independently using the projection. Based on the feedback, interactivity was integrated and icons were revised so that they would look more realistic. Suggestions for improvement included repositioning the projection surface to take the strain off the user’s neck. Overall, the results give us confidence that AVISAR is a valuable tool to support manual work in manufacturing.

Regarding the practical implementation and interoperability with existing systems in smart factories, there are some important considerations. AVISAR consists of a TypeScipt/JavaScript web application and integrates the Leap Motion Controller for gesture recognition. It thus requires a PC with Windows, MacOS, or Linux, but it is possible to realize the architecture with other software and hardware. A more recent version of the Leap Motion Controller supports deployment on a Raspberry Pi, which makes the setup more affordable, smaller and scalable.
For gesture recognition, there are various other approaches for camera-based recognition using computer vision and machine learning [35, 36] or sensing methods such as WiFi signal sensing or capacitive sensing exist [37]. To ensure that the adaptive assistance system blends seamlessly into existing industrial production systems, common industry standards such as the platform-independent machine-to-machine communication standard OPC UA should be used. Unless no other solution is already provided, real objects or humans, such as products delivered to a manual workstation or operators interacting with the assistance system, may be identified using RFID transponders. To ensure secure mounting and prevent mechanical damage to the projector and further components (e.g. gesture sensor and RGB camera), extra device protection may be required.

5. Conclusion and Outlook

In this paper, we proposed an Adaptive Visual Assistance system using Spatial Augmented Reality (AVISAR), which adapts to tasks, the layout of a manual workplace, and individual human skills and needs to enable a more inclusive work environment and the well-being of the workforce. We presented a conceptual adaptivity framework and a first working prototype of the AVISAR system.

Our next step will be to finalize the automatic adaptation of the system to user profiles and errors as well as the marker-based calibration of the projection to suitable projection surfaces. After that, further adaptation triggers will be incorporated into the user profiles for improved individualized support, training, and on-boarding for people with cognitive challenges or other special prerequisites. Moreover, sensors for recording the worker’s state will be integrated, starting with the measurement of electrodermal activity (EDA) to determine emotional arousal and stress levels. However, sensors are to be integrated with care, as it also touches on ethical issues and the fundamental rights of operators. The EU Artificial Intelligence Act (AI Act) [38] explicitly classifies AI-based emotion recognition in the workplace as an unacceptable risk and prohibits it unless for safety or medical reasons. Finally, we will draw attention to the evaluation phase of the AVISAR application by assessing performance, usability, and user experience. We are currently planning to conduct a between subjects laboratory study in which 60 participants with illiteracy, language barriers, and lack of familiarity with digital technologies carry out six different manual repair tasks with the help of AVISAR. Repair instructions in AVISAR are displayed in a different language or only with icons. By surveying stress levels, accuracy, task completion times, behavior patterns, and usability, we hope to gain insights into how adaptations foster an inclusive working environment for user groups with different needs. In follow-up studies, we intend to explore further and personalized adaptations and recruit real workers from industry as participants. It is essential that our assumptions about user groups and their individual skills and needs undergo a thorough investigation to ensure that AVISAR offers true benefits for workers in industrial production. Our aim is to create an adaptive assistance system that can also be used in similar contexts in industrial production in order to promote the well-being of workers in the long term.
Acknowledgments

On behalf of Elena Stoll and Karishma Jagdish Gavali, this research was funded as part of the project “Produktionssysteme mit Menschen und Technik als Team” (ProMenTaT, application number: 100649455) with funds from the European Social Fund Plus (ESF Plus) and from tax revenues based on the budget passed by the Saxon State Parliament.

References

[10] P. Burggräf, M. Dannapfel, T. Adlon, M. Kasalo, Adaptivity and adaptability as design


