Towards Optimal Worker Assistance: Investigating Cognitive Processes in Manual Assembly

Christian Stoessel[1], Mathey Wiesbeck[2], Sonja Stork[1], Michael F. Zaeh[2] and Anna Schuboe[1]

[1] Experimental Psychology, Ludwig-Maximilians-Universität München, Munich, Germany

[2] Institute for Machine Tools and Industrial Management, Technische Universität München, Munich, Germany

Abstract

The integration of cognitive systems in currently humanly dominated, manual assembly environments is the core issue of the project ACIPE (Adaptive Cognitive Interaction in Production Environments). This paper presents a novel concept of workbench augmentation for adaptive human worker observation and guidance. An increase in worker performance is expected from a cognitive assistance system integrated into the workbench, which can support the worker via context sensitive, adaptively generated instructions at the right time, location and with appropriate content. In order to assist the worker adequately, a deeper understanding of mental workload and related cognitive processes and limitations during manual assembly is needed. The augmented workbench serves both as a research tool for detecting cognitive bottlenecks in manual assembly, as well as an implementation platform for the worker assistance system, leading to a more efficient manual assembly performance.

Keywords:

Manual assembly, worker assistance, human cognition, mental workload, augmented reality

1 INTRODUCTION

The integration of cognitive systems in currently humanly dominated, manual assembly environments is the core issue of the project ACIPE (Adaptive Cognitive Interaction in Production Environments). It is tightly integrated in the excellence cluster 'Cognition for Technical Systems' (CoTeSys), a large scale, long-term research initiative funded by the Deutsche Forschungsgemeinschaft (German Research Foundation, DFG). The cluster CoTeSys brings together Munich-based researchers from engineering, computer science, psychology and neurosciences to strive for the realization of cognitive technical systems. The cluster's research findings will be implemented in different demonstration platforms. Within this context, a crucial goal of ACIPE is to develop concepts and prototypical technical implementations, which will allow the realization of a truly context-sensitive guidance and assistance system for manual workplaces.

2 COGNITION IN PRODUCTION ENVIRONMENTS

Skilled human workers turn mechanical workshops into today's most flexible and widely applicable forms of production. However, this flexibility and generality comes at very high production costs, which restrict the use of mechanical workshops to building prototypes and a limited range of highly specialized and valuable products. The source of flexibility in this case can be easily identified: the cognitive capabilities of the humans that operate it. Humans can perceive their environment with multiple sensory organs, plan actions, learn and adapt behaviors and they can interact in multiple ways with their surroundings. Moreover, humans can do this robustly despite changing contexts and situations. The realization of comparable cognitive capabilities in technical systems bears an immense potential for the creation of industrial automation systems that are able to overcome today's boundaries. Conventional automation systems fail to account for the demand on flexible manufacturing of highly variant and customer individual products [1]. In order to reach a high degree of flexibility and efficiency, automated processes, such as guidance systems, have to be integrated with human workers [2]. Such a tight interaction between humans and machines becomes only possible by introducing intelligent interfaces that enable adaptive support during assembly. Adaptive in this context means the real-time integration of sensory information on the assembly progress, the production environment and the current state of the worker in order to generate a context-sensitive set of instructions and appropriately accommodate the information output to the worker. This knowledge will allow for interactive multimodal guidance and support of the worker in manual and semi-automatic assembly.

2.1 Understanding cognitive processes in manual assembly

When striving for an assistance system that can support the worker adequately, it is of crucial importance that the system delivers the right kind and amount of information at the right time and place so that the worker can receive and process it effortlessly. It is therefore not sufficient for a system to be aware of the environment and the current state of the product in the manufacturing process, but additionally it has to incorporate information on the cognitive processes involved during manual assembly. By cognitive processes we understand the entirety of human mental functions dedicated to information processing, such as perception, attention, memory, problem solving and action. Only on the basis of such knowledge it can be ensured that the assistance system chooses the right kind of support which does not overstrain the human worker. This is particularly important in those situations where high demands on the worker's cognitive capacities threaten an efficient and faultless completion of the workpiece. The research presented in this article focuses on the investigation of cognitive processes and respective cognitive bottlenecks during a manual assembly task. It provides a framework how these insights can be incorporated in the technical realization of an assistive workbench.

The information processing framework

In order to disentangle the cognitive processes involved in manual assembly, it is useful to adopt an information processing framework. Such a framework (e.g., Figure 1, left) is based on the notion that environmental stimuli, which are taken up by the human through the sensory organs, are processed in a roughly serial manner through a couple of defined processing stages, until finally a response is generated to act upon the environment. This basic taxonomy is useful in breaking up the cognitive processes involved in manual assembly into different substeps. Since in manual assembly usually some instruction detail has to be detected, identified, translated into an appropriate motor response and this response has to be executed correctly, this framework seems especially appropriate to structure the involved cognitive processes. An example of how this framework can be applied to a manual assembly task is provided in Figure 1 (right). Such a framework, although oversimplified, is nevertheless useful to describe specific sources of human error: As human processing resources are limited, they have to be distributed or allocated to specific mental processes, and delays or errors in performance can arise if processing resources are insufficient or not allocated appropriately.

Selective Visual Attention

The authority deciding which items (features, objects…) will be processed at all, and which information is passed over to the next processing stage, is referred to as attention. Even though we can deploy attention covertly without moving the eyes to a certain region in the environment, we cannot attend to an unlimited number of objects at different spatial locations simultaneously. During manual assembly for example, we cannot attend to the workpiece and the instruction manual at the same time if they are spatially separated. There are competing theories in psychological literature on the exact role, locus and timing of attention, but they all agree on its selective character, its ultimately limited resources and the measurable time costs it takes to switch attention from one location or object to another. Attention can be directed reflexively, by external, exogenous cues, like a salient object in the environment, or voluntarily (endogenously), by instruction or intention. It is evident that an assistance system should guide the worker to attend to the right locations and objects, and avoid unnecessary shifts of attention. The worker can work more efficiently, if a strong exogenous cue guides her/his attention (for example a red piece among grey ones, or a highlighted box surrounded by nonhighlighted boxes) and if she/he knows where to look next for the next relevant piece of information.

Multiple Task Performance

For most tasks, demands on our attention become first obvious when we have to perform several tasks simultaneously: While driving and conducting a conversation at the same time is usually a simple task, conversation might come to a halt if the driving situation becomes difficult, and all available attention has to be allocated on the driving scene. Similarly, in an assembly context, the worker has to divide his attention between the instruction manual, the work piece, and

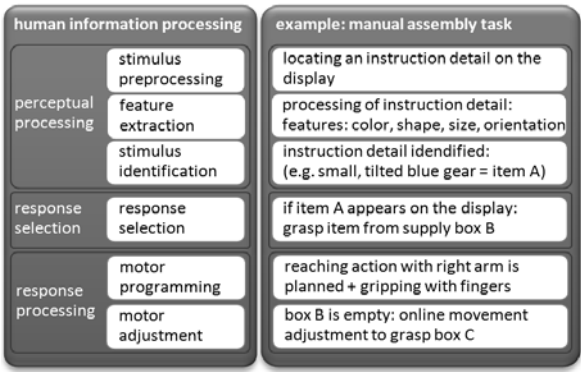

Figure 1: Left: Human information processing stages (adapted from Sanders, 1990, [3], p.37) Right: Mapping to corresponding processing stages within a manual assembly task

the factory environment. Humans have to allocate attention appropriately to aspects of the tasks, and in order to optimize performance, they have to find the right combination of attentional strategies (attending to different sources of information) and perceptual strategies (only extracting the necessary amount of information from a source) [3]. Efficient, coordinated performance depends on the ability to switch from one task (component) to another. Whenever the worker has to switch tasks, (e.g., when the next assembly step requires a new set of parts/tools) some cognitive processing time is needed to adapt. In order to establish an efficient assembly workflow, a switch of tasks should be reduced to the necessary minimum. In order to avoid task switching, information to the worker should be organized coherently, for example the sequence of assembly instructions and the spatial organization of the workplace should be in a logical and comprehensible manner.

Mental Workload

The previous section described the important role of attention in selecting information, and how it needs to be carefully distributed to objects, spatial location and tasks on the background of limited cognitive resources. Equally important from an applied perspective is to measure and evaluate the mental demands imposed by tasks and to describe how people cope with multiple demands on their attention in order to avoid performance break down and errors. The term **mental workload** has been used in applied psychology and human factors literature to describe information processing demands imposed by the performance of cognitive tasks [3]. Wickens suggested that different tasks draw on different mental resources with their specific capacity limitations, and only if two tasks need the same resource, decreased performance is observed [4]. He proposed a four-dimensional resource model described along the dimensions **input modalities** (visual - auditory), **processing codes** (spatial-verbal), **processing stages** (perception – central processing - responding) and **response modality** (manual-verbal). This model has become popular to describe mental workload, as it has been proven useful to represent tasks in a relatively simple framework and allows making predictions on concurrent performance of multiple tasks. In the context of a manual assembly operation, this framework could be applied, for example, to the concurrent tasks of mounting some parts and skipping through assembly instructions. If both operations have to be carried out simultaneously and involve the same response modality (e.g. both require manual intervention) the workload is higher due to increased demands in hand coordination, compared to a situation where the task of flipping pages in the instruction manual can be carried out concurrently through a verbal response. Therefore, task design should avoid cognitive bottlenecks within the same dimension.

3 TOWARDS A COGNITIVE WORKER ASSISTANCE SYSTEM FOR MANUAL ASSEMBLY

In order to reach the goal of providing an adequate support for the worker, progress has to be made on several levels.

On a theoretical level, the cognitive processes involved during a manual assembly task and processing bottlenecks need to be understood. The previous paragraph provided a framework of information processing stages during manual assembly. It stressed the importance of directing the worker's attention appropriately, to present the task coherently and with a minimum of attention switching, and, more generally, to design tasks based on these findings. Only if sufficient progress is made in understanding the cognitive demands during manual assembly, an assistance system can be effectively tailored to increase worker performance.

Secondly, on an implementation level, we investigate by which technology we can enable an assistance system to appropriately adjust to the worker. If the system should be able to adjust the worker support online, immediately, and in a context-sensitive way, then we are in need of a tracking and sensor technology that provides constant information on the environment, the state of the assembly process, and the actions or even intentions of the worker. A further technological challenge is the mode of information presentation to the worker. Especially if we deal with untrained workers or highly customized products, the worker has to rely strongly on the assembly instructions. The work instructions must describe concisely what needs to be done, in which succession, and with which tools and materials. Consequently, the instructions need to be carefully planned in order to minimize operator learning time, while at the same time being economical to prepare, reproduce, distribute and change [5]. While in many manual assembly workplaces today paper instructions manuals are still abundant and have clear advantages on the side of their production and reproduction, their disadvantages and limitations are also obvious: They are static in respect to the content that is displayed (e.g. no animation, context highlighting possible), their content cannot change dynamically with changing products and individual worker needs, physical size is fixed and takes up workspace, and manual assembly workflow has to be interrupted if pages need to be flipped manually. These limitations are contrary to the demands on a flexible assistance system, which should provide intuitively understandable assembly instructions and should be able to issue context-sensitive warnings or direct the worker's attention by exogenous cues, while at the same time minimally interfere with the assembly process. Regular display technology (e.g. mounting a TFT monitor to the workbench) can surmount already some of these limitations: e.g. several product instructions can be provided on the same physical instruction medium, animations or 3D inspections of a virtual workpiece are possible. However, it is still limited to a fixed location in space and needs to be attended by the worker, who consequently has still to perform cognitive mapping and search tasks to match the instructions with the real assembly scenario. **Augmented Reality (AR)** technology seems to be an appropriate candidate to deliver instruction information interactively at the right location in space and at the right time when needed.

3.1 Potential of Augmented Reality for manual assembly

Augmented Reality technology, such as head mounted displays (HMD) and tracking systems, has been employed in an impressive variety of applications (e.g., in architecture, medicine, entertainment, production or maintenance). AR has the potential to be successfully applied within an industrial working context (e.g. the assembly line), and to deliver updated, accurate, and useful information at the place where it is needed (e.g. a representation of the next product to be assembled), which will 'eventually lead to shorter production times, less training efforts, the reduction of errors, and finally to lower production costs' ([6], p. 284). Tang and colleagues have pointed out that AR technology has the potential of reducing head and eye movements, reducing the costs of attention switching and supporting spatial cognition and mental transformation [7]. However, successful working AR scenarios in manual assembly environments are rarely found. This discrepancy between being highly adequate from a theoretic point of view and the lack of successful implementation is explainable by the many technical problems that these systems are still facing. In the manual assembly domain, AR scenarios have been developed for example in car door assembly, cable wiring, mainboard assembly or toy settings. Common to these approaches was the use of a HMD, which was used to overlay the real world with additional, spatially coregistered pieces of information. To summarize the findings, AR technology has shown beneficial mostly only in cases when the tasks were sufficiently complex (e.g. [8]). It was most effective during the selection phase of assembly, where workers have to find the right piece and position where to put it, but not during manual execution of the task. In many cases, the technical limitations – strain during continuous wearing of HMDs, their limited resolution, contrast and field of view, problems with registration, unsatisfying speech based interaction – outnumbered the benefits [8,9].

The approach we took tried to combine the benefits of AR technology – providing additional information on top of the real world environment at the right time and location where it is needed – without having to face the drawbacks that the use of HMDs – co-registering techniques and speech based interaction – usually brings about.

4 WORKPLACE SETUP

In order to develop an adaptive interaction model to support the worker more efficiently, two lines of research were pursued in parallel: The first challenge was to develop a working environment that satisfies the ergonomic constraints for effortless manual assembly, and includes a well-adjusted combination of display and tracking technology. The worker can thereby be presented with individually tailored instructions at the time and place when and where she/he needs it. The second challenge was to develop a realistic and at the same time adequate scenario and measurement

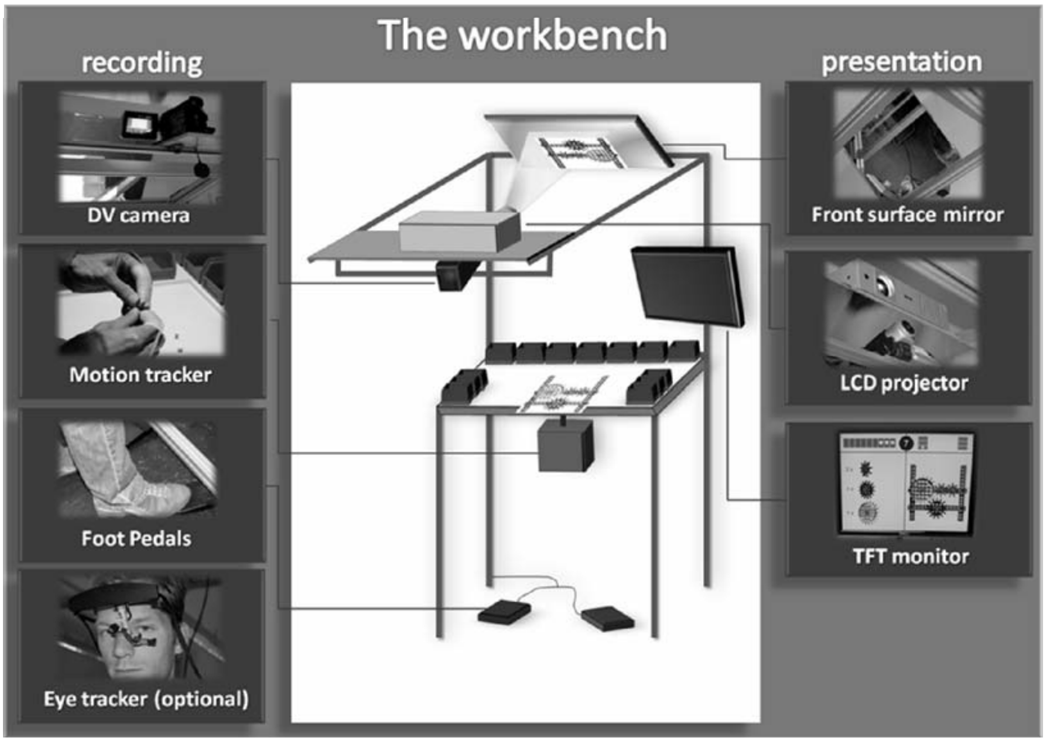

Figure 2: Schematic depiction of the workbench.

methods which allow investigating the cognitive processes involved in assembly tasks in great detail, in order to achieve a better understanding of where workers are facing problems and why they arise. Cognitive bottlenecks need to be identified and specific processing stages need to be differentiated (see section 2). We met these challenges by designing a special workbench, which was equipped with a customized information presentation technique, a foot pedal based interaction concept and a tracker system in order to achieve an online feedback of human behavior to the machine (Figure 2).

4.1 Information Presentation

One of the crucial points in assisting the worker is the manner in which instructions are presented. Following the ideas of worker-adaptive information display [2] and contact analog information presentation in accordance with attentional and visual processing mechanisms in humans, we opted for an augmented reality solution where the instructions become an integrated part of the workbench. As most difficulties in creating a usable AR system were related to HMDs and tracking technologies we decided against such a system. Instead, the relatively static environment of a manual assembly workspace offered ideal conditions to make use of a fixed projection augmented reality setup. By projecting assembly information directly onto the workbench, it is possible to combine the advantages of an LCD projector over a HMD (increased contrast, luminance, resolution and colorspace, unhampered field of view, no strain on nose or head) with the possibilities to provide the information at the spatial location where the worker would need it, and to provide **contact analog** information on static objects within the work space. Furthermore, this approach has the potential of adapting the instruction display in real time according to the worker's needs. This kind of AR support to a manual assembly task also bears the benefit of being cheaper to acquire, easier to install and operate and more likely to be tolerated by the worker for a daily use than HMD based approaches. In order to project the information onto the working plane, a standard workbench for manual assembly was extended overhead with a horizontally mounted LCD projector. As this setup should be operable not only in a laboratory environment but also within a real factory setting with variable lighting conditions, care was taken that the projector provided enough luminous power (4000 lumen), while at the same time providing a good projection ratio (projection distance : projected image diameter) and low operating noise. A front-surface mirror was installed at an angle of 45° at the front end of the workbench (Figure 2) so that the projected image got reflected to cover the whole working area. With this setup, we were able to provide contact analog information, for example a highlighting of only those boxes where parts for the current assembly step had to be taken from. In addition to the projection directly onto the working surface by means of the LCD projector and the front-surface mirror, this setup includes a 20" TFT monitor for alternative or complimentary information presentation that is located at an ergonomically favorable position, can be adjusted in angle and is easily (de)mountable.

4.2 Interaction Concept

While in many interaction concepts for AR applications in production environments a speech recognition system is employed, these systems were often error-prone due to insufficient speech recognition for commands that were altered by slang or background noise, and consequently annoying to the user (e.g. [9]). Particularly a noisy factory environment seems inappropriate to use speech as input modality for worker commands. In order to provide an interaction mode which is not based on speech or manual

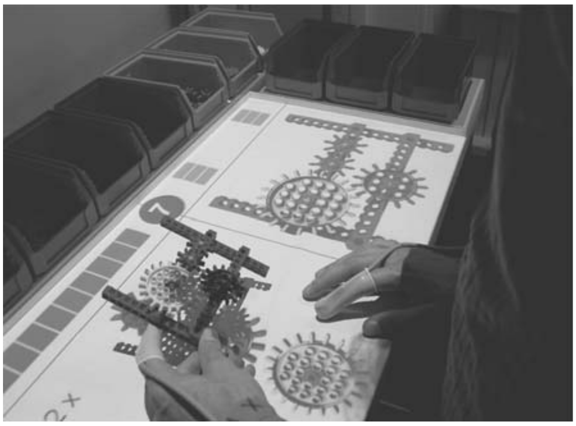

Figure 3: The augmented workbench including projected instruction and contact analog highlighting (boxes).

operation of some input device (which would interrupt the manual assembly workflow), we devised two custom built foot pedals which can be connected over USB to a control computer. The foot pedals are placed at easily reachable and adjustable positions, one operated by the right foot and one by the left foot. By operating the right foot pedal the worker can advance one step in the assembly guide, by operating the left foot pedal the worker can go back one step. Even though the ultimate goal is to develop a system that automatically adjusts to the worker's skills, experience and current mental workload and that way for example automatically detects mistakes and offers instant solutions, it is yet important to provide some means of user interaction so that the worker does not perceive himself as being controlled by the system. A foot-pedal input system realizes such a user control through a simple forward-backward flipping through the assembly steps, while, which is important, the focus of view, attention and bimanual handling can rest on the work piece.

4.3 Performance Measurement

The workbench offers the possibility to mount tracking technology by which the worker's performance can be registered online, interpreted in terms of skill level and current workload, and instructions can be modified accordingly. Currently, the workbench is equipped with a Polhemus Liberty motion tracking system (Figure 3) and a DV camera which can record worker behavior over the mirror. Behavioral variables, such as time-to completion for a certain assembly substeps, error rates and error categorization give a first measure of worker performance. Under controlled experimental conditions, these measures are helpful to compare e.g., different instruction modes or the effect of instruction complexity on assembly performance. As an example, the time it takes from instruction presentation until the worker initiates a movement can be regarded as an indirect measure of how easily the worker could decode the presented information. The scenario also incorporates an eye tracking system, which is especially valuable to identify the worker's fixation at a point in time and is an excellent estimation for attention allocation. Fixation patterns can reflect search patterns on which the workbench and instruction layout can be optimized.

5 EXPERIMENTAL STUDY

An experimental study was designed to test and validate the proposed workbench setup and gather empiric data of how worker performance is modulated by task setup and complexity. In order to quantify the benefits of projected and contact analog information, three presentation conditions were chosen: In condition 1, the instructions were presented on a 20” TFT monitor mounted to the workbench. In condition 2, instructions were projected directly onto the working area, and in condition 3 the worker received additional contact analog information on which parts were needed in the current step (Figure 3). Performance data was gathered with continuous motion tracking, foot pedal registering and video recordings of the assembly and complemented with a short questionnaire focusing on usability issues.

5.1 Study Design

Thirty subjects (20m/10f) participated in this pilot study. Most of the participants were students (age 20-30, average 23.9 years) with dominantly technical engineering background. The task was to assemble three models with LEGO Technic bricks according to a detailed instruction (Figure 3). In order to account for individual differences in past experience with LEGO bricks, a short training phase was conducted in which the participants had to build a simple LEGO model. Each participant had to perform three assembly tasks in counterbalanced order. Participants received the instructions in one of the three presentation modes (monitor, projection or contact analog). The assembly tasks were self-paced. The use of LEGO bricks to simulate a real assembly tasks offers a couple of advantages over similar construction systems: Firstly, they can be combined purely manually without the use of additional tools, second they are available in an unrivaled variety of colors and shapes, third, they are consistent in the way the parts can be mounted, and no particular motor skills or specialized training is necessary. The three models that had to be built were all similar in basic structure – they resembled a windmill – but differed in detail. Each model had to be constructed from a sequence of defined assembly steps that differed systematically in complexity with regard to the amount of pieces, the diversity of pieces and the class of assembly (e.g. ‘gear’ or ‘frame’). Performance was measured by assembly time per construction step, the number and type of errors made and the efficiency of grasping movement trajectories.

5.2 Results

Results showed that complexity of the assembly step had a huge impact on assembly performance: certain assembly steps (e.g. gear assembly) could be performed quite fast (on average 6.8 seconds per part), while others (e.g. gear fitting) took much longer (Figure 4). It could also be shown that the proposed projection method of instruction presentation was in general beneficial to the worker, especially in those situations where assembly instructions were highly complex (Figure 4). In relatively simple assembly steps, no difference of instruction type was observed.

6 CONCLUSIONS

In order to achieve more efficient and less error prone manual assembly, workers can be supported by a context-sensitive guidance and assistance system. Especially in the case of complex and highly variable assembly, productivity could be

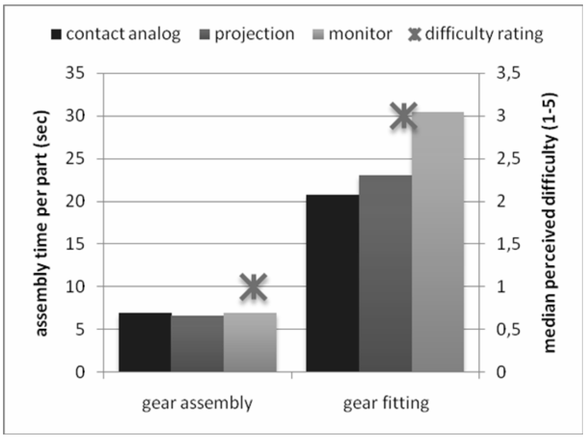

Figure 4: Average mounting times per part for two types of assembly steps. Contact analog and projection presentation mode benefit only for the more complex task (gear fitting). Step complexity reflected by user difficulty rating ('x', ranging from 1: not difficult at all, to 5: very difficult).

raised if the human is supported by an assistance system which provides instruction presentation tailored to the capabilities of the worker, the current situation, and helps to avoid and overcome errors.

The setup as it was proposed presents a first step towards such an assembly assistance system. We have shown that with limited technical effort it is possible to augment a standard workbench with instruction information directly on the workbench, reducing the need to switch attention back and forth between instructions and assembly area. Furthermore, the foot-pedal interaction concept allows continuous bimanual assembly. This scenario can serve as a research platform to investigate cognitive processes during manual assembly in greater detail. The experimental pilot study showed that the proposed approach including instruction projection directly onto the working plane and context-sensitive highlighting of objects led to fewer errors and faster assembly on more complex assembly steps. Contact analog information was especially helpful in drawing the worker's attention to relevant work-pieces and their location in cases where assembly parts were very similar to each other and hard to disambiguate.

7 FUTURE DIRECTIONS

So far, we included mainly behavioral assessments of worker performance. In order to investigate the underlying cognitive processes more detailed, this setup will be complemented with an eye tracking system (Figure 2). The combined movement and vision data will provide valuable insights of how efficiently the instruction information is taken up by the worker. Insights from controlled experiments will feed the design process of adaptive, worker tailored instruction presentation. Wireless tracking technologies need refinement to track the state of the environment, the work piece and the manual movements of the worker. The information gathered by the system may be fed into an environmental and a human model, which in turn dynamically influence the instruction presentation. Online analysis of movement patterns could be used to classify workers as belonging to different skill classes, and adopt instruction display accordingly: a novice on this task would need a different instruction presentation than an expert. Furthermore, assembly errors could be detected and even prevented by immediate feedback to the worker.

Further research on cognitive processes during manual assembly and further development of the assistive workbench will help to support the worker better in challenging assembly situations. By this adaptive accommodation of machines to human needs, human skill can be integrated tightly into the production process to achieve high degrees of efficiency and flexibility, where conventional automation fails.

8 ACKNOWLEDGEMENTS

Special thanks go to Frank Wallhoff, Alexander Bannat, Florian Friesdorf, Rainer Ullman and Cordula Vesper for their supportive collaboration in the ACIPE project. Furthermore, we would like to acknowledge the funding support of the Deutsche Forschungsgemeinschaft (DFG).

9 REFERENCES

[1] Putzer, H., Onken, R., 2003, COSA - A generic cognitive system architecture based on a cognitive model of human behavior, Cognition, Technology & Work , 5:140-151.

[2] Zaeh, M.F., Lau, C., Wiesbeck, M., Ostgathe, M., Vogl, W., 2007, Towards the Cognitive Factory, Proceedings of the 2nd International Conference on Changeable, Agile, Reconfigurable and Virtual Production, Toronto, Ontario, Canada.

[3] Johnson, A., Proctor, R.W., 2004, Attention: Theory and practice, Sage Publications Inc, Thousand Oaks, CA, USA.

[4] Wickens, C.D., Hollands, J.G., 2000, Engineering Psychology and Human Performance (3rd. ed.), Prentice Hall Inc, Upper Saddle River, NJ, USA.

[5] Crowson, R., 2006, Assembly Processes: finishing, packaging, and automation, Taylor & Francis Group, New York, NY, USA.

[6] Regenbrecht, H., 2006, Industrial Augmented Reality Applications, in: Haller, Thomas and Billinghurst, Emerging technologies of Augmented Reality: Interfaces and Design, 283-304, Idea Group Publishers, Hershey, PA, USA.

[7] Tang, A., Owen, C., Biocca, F., Mou, W., 2003, Comparative Effectiveness of Augmented Reality in Object Assembly, Proceedings of the SIGCHI conference on Human Factors in Computing Systems, 73-80, Ft. Lauderdale, FL, USA.

[8] Wiedenmaier, S., Oehme, O., Schmidt, L., Luczak, H., 2003, Augmented Reality (AR) for Assembly Processes: Design and Experimental Evaluation, International Journal of Human-Computer Interaction, 16(3): 497-514.

[9] Schmidt, L., Wiedenmaier, S., Oehme, O., Luczak, H., 2005, Benutzerzentrierte Gestaltung von Augmented Reality in der Produktion, in: Stary, Mensch & Computer 2005: Kunst und Wissenschaft - Grenzüberschreitungen der interaktiven ART, 51-60, Oldenbourg Verlag, München.