Design Criteria for AR-Based Training of Maintenance and Assembly Tasks

Sabine Webel, Ulrich Bockholt, and Jens Keil

Fraunhofer Institute for Computer Graphics Research IGD
Fraunhoferstr. 5, 64283 Darmstadt, Germany
{sabine.webel,ulrich.bockholt,jens.keil}@igd.fraunhofer.de

Abstract. As the complexity of maintenance tasks can be enormous, the efficient training of technicians in performing those tasks becomes increalingly important. Maintenance training is a classical application field of Augmented Reality explored by different research groups. Mostly technical aspects (e.g tracking, 3D augmentations) have been in focus of this research field. In our paper we present results of interdisciplinary research based on the fusion of cognitive science, psychology and computer science. We focus on analyzing the improvement of AR-based training of maintenance skills by addressing also the necessary cognitive skills. Our aim is to find criteria for the design of AR-based maintenance training systems. A preliminary evaluation of the proposed design strategies has been conducted by expert trainers from industry.

Keywords: Augmented Reality, training, skill acquisition, training system, industrial applications.

1 Introduction

As the complexity of maintenance and assembly tasks can be enormous, the training of the technician to acquire the necessary skills to perform those tasks efficiently is a challenging point. A good guidance of the user through the training task is one of the key features to improve the efficiency of training. Traditional training programs are often expensive in terms of effort and costs, and rather inefficient, since the training is highly theoretical. Due to the complexity of maintenance tasks, it is not enough to teach the execution of these tasks, but rather to train the underlying skills. Speed, efficiency, and transferability of training are three major demands which skill training systems should meet. In order to train the maintenance skills, the trainee's practical performance of the training tasks is vitally important.

From previous research it can be derived, that Augmented Reality (AR) is a powerful technology to support training in particular in the context of industrial service procedures. Instructions on how to assemble/disassemble a machine can directly be linked to the machines to be operated. Various approaches exist, in which the trainee is guided step-by-step through the maintenance task. Mostly technical aspects (tracking, visualization etc.) have been in focus of this research field. Furthermore, those systems function rather as guiding systems than training systems. A potential danger of Augmented Reality applications is that users become dependent

R. Shumaker (Ed.): Virtual and Mixed Reality, Part I, HCII 2011, LNCS 6773, pp. 123–132, 2011.

on Augmented Reality features, and as a result they might not be able to perform the task, when those features are not available or when the technology fails. That is to say, an AR-based training system must clearly differ from an AR-based guiding system; it must really train the user instead of only guiding him through the task. This can be only achieved by involving cognitive aspects in the training.

Industrial maintenance and assembly can be considered as a collection of complex tasks. In most cases, these tasks involve the knowledge of specific procedures and techniques for each machine. Each technique and procedure requires cognitive memory and knowledge of the way the task should be performed as well as fine motor "knowledge" about the precise movements and forces that should be applied. Hence, the skill, which is responsible for a fast and robust acquisition of maintenance procedures, is a complex skill. In this context, *procedural skills* can be considered as the most important skill in industrial maintenance tasks. Procedural skills are the ability to follow repeated a set of actions step-by-step in order to achieve a specified goal. It is based on getting a good representation of a task organization: What appropriate actions should be done, when to do them and how to do them.

Within a cooperation of engineering and perceptual scientists we explored the training of industrial maintenance. Here we focused on training of procedural skills. By analyzing the use of Augmented Reality technologies for enhancing the training of procedural skills, we aim for finding design criteria for developing efficient AR-based maintenance training systems. Therefore, a sample training application has been developed. We present preliminary results of the evaluation conducted by maintenance trainers from industry.

2 Related Work

As the complexity of maintenance and assembly procedures can be enormous, the training of operators to perform those tasks efficiently has been in focus of many research groups. Numerous studies presented the potential of Augmented Reality based training systems and its use in guidance applications for maintenance tasks. One of the first approaches is using Augmented Reality for a photocopier maintenance task [1]. The visualization is realized using wireframe graphics and a monochrome monoscopic HMD. The tracking of objects and the user's head is provided by ultrasonic trackers. The main objective is to extend an existing two dimensional automated instruction generation system to an augmented environment. Hence, only simple graphics are superimposed instead of complicated 3D models and animations.

Reiners et al. [2] introduce an Augmented Reality demonstrator for training a doorlock assembly task. The system uses CAD data directly taken from the construction/production database as well as 3D-animation and instruction data prepared within a Virtual Prototyping planning session, to facilitate the integration of the system into existing infrastructures. For the tracking they designed an optical tracking system using low cost passive markers. A Head Mounted Display functions as display device.

Schwald et al. describe an AR system for training and assistance in the industrial maintenance context [3], which guides the user step-by-step through training and maintenance tasks. Magnetic and infrared optical tracking techniques are combined to obtain a fast evaluation of the user's position in the whole set-up and a correct projection for the overlays of virtual information in the user's view. The user is equipped with a lightweight helmet, which integrates an optical see-through HMD, a microphone, headphones, and a 3D-positioning sensor. The headphones offer the user the possibility to get audio information on the procedures to achieve. Via the microphone the user can easily interact with the system by using speech recognition. The 3D-positioning sensor is used to determine the position of the objects of interest in 3D space in relation to the user’s position. That way, 3D augmentations are directly superimposed with their real counterparts, whereby the parts of interest are highlighted. Besides, also information about how to interact with the counterparts can be visualized. The paper discusses the usage of the system, the user equipment, the tracking and the display of virtual information.

In [4] a spatial AR system for industrial CNC-machines, that provides real-time 3D visual feedback by using a transparent holographic element instead of using user worn equipment (like e.g. HMD). Thus, the system can simultaneously provide bright imagery and clear visibility of the tool and work piece. To improve the user's understanding of the machine operations, visualizations from process data are overlaid over the tools and work pieces, while the user can still see the real machinery in the workspace, and also information on occluded tools is provided. The system, consisting of software and hardware, requires minimal modifications to the existing machine. The projectors need only to be calibrated once in a manual calibration process.

An Augmented Reality application for training and assisting in maintaining equipment is presented in [5]. Overlaid textual annotations, frames and pointing arrows provide information about machine parts of interest. That way, the user's understanding of the basic structure of the maintenance task and object is improved. A key component of the system is a binocular video see-through HMD, that the user is wearing. The tracking of the position and orientation of equipment is implemented using ARToolKit [6].

The work of Franklin [7] focuses on the application of Augmented Reality in the training domain. The test-bed is realized in the context of Forward Air Controller training. Using the system, the Forward Air Controller (trainee) can hear and visualize a synthetic aircraft and he can communicate with the simulated pilot via voice. Thus, the trainee can guide the pilot onto the correct target. The system can provide synthetic air asset stimulus and can support the generation of synthetic ground based entities. Positions and behavior of these entities can be adapted to the needs of the scenario. The author concluded that the impact of Augmented Reality for training depends on the specific requirements of the end user and in particular on the realism of the stimulation required. According to the author, this is influenced by the means of the required stimulation, the criticality on how the synthetic stimulation is used, the dynamism and complexity of the training environment and the availability of a common synthetic environment.

3 Training of Procedural Skills

As mentioned before, *procedural skills* are the ability to follow repeated set of actions step-by-step in order to achieve a specified goal and reflect the operator's ability to obtain a good representation of task organization. This skill is needed in the performance of complex tasks as well as simple tasks. Procedural skills are based on two main components: procedural knowledge and procedural memory. Procedural knowledge enables a person to reproduce trained behavior. It is defined as the knowledge about how and when (i.e. in which order) to execute a sequence of procedures required to accomplish a particular task [8]. Procedural knowledge is stored in the procedural memory, which enables persons to preserve the learned connection between stimuli and responses and to response adaptively to the environment [8]. Generally speaking, procedural skills develop gradually over several sessions of practice (e.g. [9]) and are based on getting a good internal representation of a task organization. Therefore, the training of procedural skills should address the development of a good internal representation of the task and the execution of the single steps in the right order in early training phases.

3.1 Enhancement of Mental Model Building

It has been explored that the performance of a learner of a procedural skill becomes more accurate, faster, and more flexible when he is provided with elaborated knowledge (e.g. [10],[11]). This means that the learner's performance increases when how-it-works knowledge ("context procedures") is provided in addition to the how-to-do-it knowledge ("list procedures") (e.g. [10]). According to Taatgen et al., when elaborated knowledge is given, the learner is able to extract representations of the system and the task, which are closer to his internal representation, and as a result performance improved [10]. This internal, psychological representation of the device to interact with can be defined as *mental model* [12]. In order to support the trainee's mental model building process, the features of the task which are most important for developing a good internal representation must be presented to the trainee. It has been suggested, that "the mental model of a device is formed largely by interpreting its perceived actions and its visible structure" [13]. The mental model building is mainly influenced by two factors: the actions of the system (i.e. the task and the involved device) its visible structure.

Transferring this into the context of procedural skill training, two aspects seem to be important for supporting the building of a good mental model: One is providing an abstract representation of the system, what constructs a better understanding of how it works. The other one is providing the visual representation of the system, which will strengthen the internal visual image. It has been found, that people think of assemblies as a hierarchy of parts, where parts are grouped by different functions (e.g. the legs of a chair) [11]. Hence, the hypothesis is that the displayed sub-part of the assembly task should include both the condition of the device before the current step (or rather the logical group of steps to which the current step belongs) and the condition after. This hypothesis is based on the work of Taatgen et al. [10], in which it is shown that instructions which state pre- and post-conditions yield better performance than instructions which do not. Reviewing this it can be concluded, that the user's mental model building process can be improved by using visualization elements providing context information.

4 Design Strategies

It has been shown, that guided experience is good for learning, but an active exploration of the task has to be assured as well (e.g. [14],[15]). A too strong guidance of the trainee during training impedes an active task exploration and harms the learning process. Active exploration naturally occurs when transferring the information about the task during training is accompanied with some difficulties, forcing the trainee to independently explore the task. If such difficulties are reduced (e.g. by showing the user in detail how to solve the problem), active exploration may not take place. Strong visual guidance tools impede active exploration, because they guide the trainee in specific actions and thus inhibit the trainee's active exploratory responses [16]. This can be illustrated using the example of a car driver guided by a route guidance system: this driver typically has less orientation than driver who is exploring the way with the help of maps and street signs. Also reproducing the way, when he has to drive it again, is more difficult for the driver who used the route guidance system. From all this it can be concluded, that the training system should include visual elements that allow for reducing the level of provided information. Furthermore, it should contain elements, which guide the trainee through the training task by improving the trainee's comprehension and internalization of the task, while active exploration is not inhibited.

4.1 Adaptive Visual Aids

An important issue when designing AR-based training systems is how much information should be visualized in the different training phases. A basic understanding of how much information the trainee needs during learning can be obtained by observing studying people. Examining the learning behavior of a student studying procedural processes using textbooks or written notations, the following characteristics can be observed: First of all, for each step the student marks a couple of words, a sentence or an excerpt in the running text and writes annotations at the side margin. He studies the process by going repeatedly through this learning material. In the first cycles, the student reads the marked text and the accordant annotations to catch information about the single steps and to put them in order. With the increasing number of performed studying cycles the information that he needs to decide and reproduce the single steps of the procedure decreases. When he starts studying he needs more detailed information about the single steps, because the learning of the single steps is in focus. With the growing development of an understanding of the single steps, the learning of how the steps fit together (i.e. of the procedure) comes increasingly to the fore.

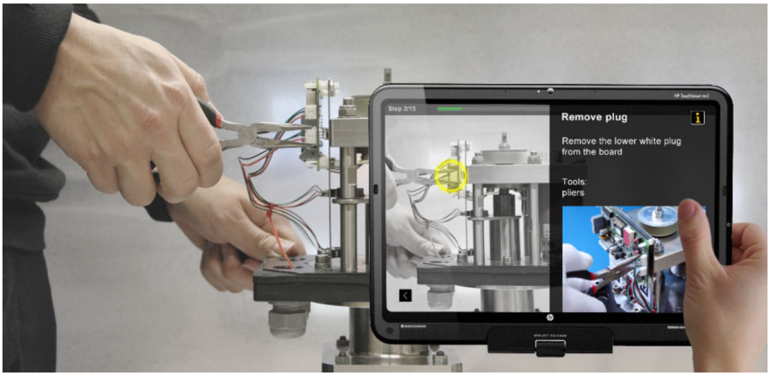

Fig. 1. Adaptive Visual Aid in the training application: a pulsing yellow circle (pointer) highlights the area of interest; the detailed instruction is given on the plane (content object)

Transferring this observation into the context of training, the mapping of the visualized information level to the different training phases can be hypothesized as follows: In early phases, a clear and detailed instruction about the current step should be provided in order to train the trainee in understanding and performing the single steps. This can be realized by using *adaptive visual aids (AVA)* consisting of overlaid 3D objects (pointer) and/or multimedia instructions (content) that is displayed on user demand (see Fig. 1). Alternatively, the pointer can act as object/area highlight while the content provides the detailed multimedia instruction. During the training the level of presented information should be gradually reduced (e.g. only 3D animation, then only area highlight with some buzzwords or a picture, then only area highlight, etc.). Hence, both *AVA pointer* and *AVA content* object can provide a variable amount of information. The pointer consists of at least one virtual object overlaid on the camera image (like traditional Augmented Reality overlays). Hence, it presents also the spatial component of the information. The pointer object can contain for example complete 3D animations, 3D models, or highlighting geometries (e.g. pulsing circle).

The AVA content object consists of a view-aligned virtual 2D plane and different multimedia data visualized on that plane. Thus, it can provide multimedia information that is clearly recognizable for the user. The data displayed on the plane can contain text, images, videos and 3D scenes rendered in a 2D image on the plane, or any combination of those elements. That is, it can contain detailed instructions (e.g. a text description and a video showing an expert performing the task) or just a hint (e.g. a picture of the tool needed to perform the task).

4.2 Structure and Progress Information

Since providing abstract, structural information about the task can improve the trainee's mental model building process, and hence the acquisition of procedural skills (see chapter 3.1), visual elements displaying information about the structure of the training task should be included in the training system. Not only the structure of the task, but also the relation between the current status and the structure is important. That is, the position of the current state in the whole structure should be visualized as well. Thus, the trainee gets an overview of the training task and can arrange the current step in the structure of the task and use this information to refine his internal representation of the task. One possibility to visualize structural information is the use of progress-bars. Progress-bars provide an abstract overview of the trainee's current status in relation to the whole task (see Fig. 2).

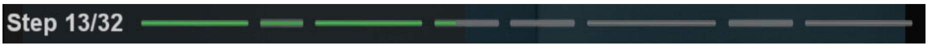

Fig. 2. An extended progressbar showing the user's progress inside the task and inside the mental groups (each part of the bar corresponds to a mental group of steps)

4.3 Device Display

As mentioned in chapter 3.1, the presentation of context information, such as logical units of sub-tasks, and the display of the device to maintain can support the trainee's mental model building. Moreover, the presentation of only relevant sub-parts of the device and the visualization of the pre- and post-conditions can further enhance the development of a good internal representation.

Based on these findings, the use of a *Device Display* is suggested. The Device Display is a visual element that provides information about successive steps, or rather sub-tasks, belonging to a logical group. That is, it provides information about a good mental model of the task. This can support the user in developing his internal representation of the task. The provided information includes also the condition of the device before the current step and afterwards. Thus, using the Device Display, the user can recognize a sub-goal of the task he has to perform. This can help him to understand "what" he has to do, and hence to deduce the next step to perform. In fact, the presentation of sub-goals actually forces the trainee to deduce the next step without using a more direct visual guidance.

The visualization of the Device Display is similar to the visualization of the AVA content (see Fig. 3, left). It consists of a view-aligned 2D plane and multimedia objects rendered on the top of this plane, which can be faded in/out on user demand. The objects displayed on the plane provide information about the grouped sub-tasks (i.e. mental group) and the condition of the device before and after the mental group. A text describes the objective of the mental group in a few words. For example, if the mental group comprises all steps for removing a machine cover, the text "Remove the cover of the valve" is displayed. Additionally, either a video of an expert's performance of the grouped sub-tasks, or a 3D animation presenting the sub-tasks including the device conditions is shown in the Device Display. Thus, for each mental group the best representation can be chosen. Also a progress-bar is displayed, that shows the user's progress inside the mental group. That way, supplemental information about the structure of the task, or rather of the mental model, is presented, what can further support the user's mental model building process.

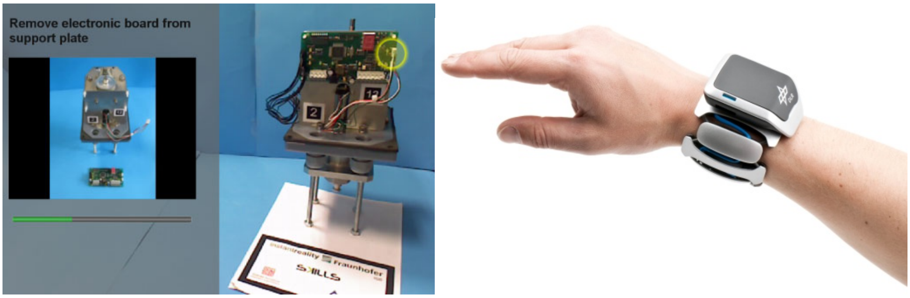

Fig. 3. Left: a “Device Display” at the left side of the window shows a video about a mental group of steps; Right: a vibrotactile bracelet developed by DLR (German Aerospace Center)

4.4 Haptic (Vibrotactile) Hints

The potential of vibrotactile feedback for spatial guidance and attention direction has been demonstrated in various works (e.g. [17]). Usually a lot of visual information has to be processed in complex working scenarios. In contrast, the tactile channel is less overloaded. Furthermore, vibrotactile feedback is a quite intuitive feedback, as the stimuli are directly mapped to body coordinates. Since it provides a soft guidance that "channels" the user to the designated target instead of directly manipulating his movements, it does not prevent the active exploration of the task. Thus, the mental model building process can be supported. Vibrotactile hints can be given by using simple devices like the vibrotactile bracelet shown in Fig. 3 (right). The bracelet developed by DLR is equipped with six vibration actuators which are placed at equal distance from each other inside the bracelet and hence also around the user‘s arm. The intensity of each actuator can be controlled individually. That way, various sensations can be generated, such as sensations indicating rotational or translational movements.

Such vibrotactile feedback should be used to give the trainee additional motion hints during the task training, such as rotational or translational movement cues, and to guide the trainee to specific targets. For example, if the trainee needs to rotate his arm for performing a sub-task, the rotational direction (cw or ccw) may be difficult to recognize in a video showing an expert performing the sub-task. Receiving the same information using a vibrotactile bracelet, the trainee can easier identify the rotational direction. Also translational movements can be conveyed. Apart from that, vibrotactile feedback can also be used for presenting error feedback, such as communicating whether the right action is performed (e.g. the right tool is grasped). This can prevent the user from performing errors at an early stage. In addition, vibrotactile hints can be used to provide slight instructions by directing the trainee’s attention to a body part.

5 Preliminary Tests and Conclusion

A preliminary evaluation has been conducted by four expert trainers from the food packaging industry (Sidel[1]). The training task is the assembly of a valve. The implemented training application consists of 32 steps, showing the sub-tasks which are necessary to assemble the valve. Haptic hints indicating rotational and translational movements of the user's right wrist have been implemented and provided using the vibrotactile bracelet described above. The trainers performed the training task using the realized AR training platform. Afterwards they filled out a questionnaire about the usability and functionality of the training system and the design strategies. Table 1 shows an extract of this questionnaire.

Table 1. Extract of the preliminary evaluation questionnaire

	SCALE		T1	T2	T3	T4	AVG
The information provided by the platform via displayed information was enough to understand the task.	1	7	4	5	6	6	5,25
The visualization of the different operations was enough for leaning the task?	1	7	5	5	6	6	5,5
Is there any critical information of the task missing?	N/A		no	no	no	no	
Please rate the general visualization utilities: spatial information, step information, captions, etc.?	1	10	6	7	8	8	7,25
Please rate the overview strategy?	1	10	10	7	8	8	8,25
Please rate the spatial pointer strategy? (AVA pointer)	1	10	6	4	8	7	6,25
Please rate the content aids display strategy? (AVA content, Device Display)	1	10	7	8	10	7	8
Please rate the context aids strategy? (progress bars)	1	10	6	7	6	8	6,75
Please rate the haptic hints strategy?	1	10	6	3	2	8	4,75
Please rate the playback/trainer-trainee based strategy?	1	10	6	8	10	8	8
From the functionality point of view, how do you rate the platform in overall?	1	10	6	6	9	8	7,25
What percentage of the task do you consider that you have learnt?	%		90%	70%	80%	10%	62,50%
What grade would you give to the AR platform as learning system?	1	10	8	7	8	5	7

We conclude from this, that the proposed design strategies, namely the use of Adaptive Visual Aids (AVAs), the provision of structure and progress information, the visualization of a Device Display and the integration of haptic hints, have a great potential for improving training of maintenance and assembly skills. The perception of the implemented haptic hints indicating movements turned out to be potentially valuable, but we have to refine the realization of the hints (i.e. the controlling of the vibration stimuli) in order to produce clear indications of the movements the trainee

[1] Sidel is one of the world's leaders of solutions for packaging liquid foods (http://www.sidel.com/).

has to perform. In our future work the training platform will be optimized according to the results of the preliminary tests (i.e. improvement of haptic hints, provision of error feedback) and evaluated by technicians working at Sidel.

References

1. Feiner, S., Macintyre, B., Seligmann, D.: Knowledge-based Augmented Reality. Commun. ACM 36(7), 53–62 (1993)
2. Reiners, D., Stricker, D., Klinker, G., Müller, S.: Augmented Reality for construction tasks: Doorlock assembly. In: Proc. IEEE and ACM IWAR 1998: 1st Int. Workshop on Augmented Reality, pp. 31–46 (1998)
3. Schwald, B., Laval, B.D., Sa, T.O., Guynemer, R.: An Augmented Reality system for training and assistance to maintenance in the industrial context. In: 11th Int. Conf. in Central Europe on Computer Graphics, Visualization and Computer Vision, Plzen, Czech Republic, pp. 425–432 (2003)
4. Olwal, A., Gustafsson, J., Lindfors, C.: Spatial Augmented Reality on industrial CNC-machines. In: SPIE Conference Series (2008)
5. Ke, C., Kang, B., Chen, D., Li, X.: An Augmented Reality-Based Application for Equipment Maintenance. In: Tao, J., Tan, T., Picard, R.W. (eds.) ACII 2005. LNCS, vol. 3784, pp. 836–841. Springer, Heidelberg (2005)
6. Kato, H., Billinghurst, M.: Marker tracking and HMD calibration for a video-based Augmented Reality conferencing system. In: Proc. IEEE and ACM IWAR 1999, pp. 85–94. IEEE Computer Society, Los Alamitos (1999)
7. Franklin, M.: The lessons learned in the application of Augmented Reality. In: Virtual Media for Military Applications, pp. 30/1–30/8 (2006)
8. Tulving, E.: How many memory systems are there? American Psychologist 40(4), 385–398 (1985)
9. Gupta, P., Cohen, N.J.: Theoretical and computational analysis of skill learning, repetition priming, and procedural memory. Psychological Review 109(2), 401–448 (2002)
10. Taatgen, N.A., Huss, D., Dickinson, D., Anderson, J.R.: The acquisition of robust and flexible cognitive skills. Journal of Experimental Psychology: General 137(3), 548–565 (2008)
11. Agrawala, M., Phan, D., Heiser, J., Haymaker, J., Klingner, J., Hanrahan, P., Tversky, B.: Designing effective step-by-step assembly instructions. ACM Trans. Graph 22(3), 828–837 (2003)
12. Cañas, J.J., Antolí, A., Quesada, J.F.: The role of working memory on measuring mental models of physical systems. Psicológica 22, 25–42 (2001)
13. Norman, D.: Some observations on mental models. In: Mental Models, pp. 7–14. Lawrence Erlbaum Associates, Mahwah (1983)
14. Mayer, R.E.: Multimedia Learning. Cambridge University Press, Cambridge (2001)
15. Wickens, C.D.: Multiple Resources and Mental Workload. Human Factors 50(3), 449–455 (2008)
16. Gavish, N., Yechiam, E.: The disadvantageous but appealing use of visual guidance in procedural skills training. In: Proc. AHFE (2010)
17. Weber, B., Schätzle, S., Hulin, T., Preusche, C., Deml, B.: Evaluation of a vibrotactile feedback device for spatial guidance. In: IEEE- World Haptics Conference (2011)