Virtual Reality and Augmented Reality as a Training Tool for Assembly Tasks

A.C. Boud, D.J. Haniff, C. Baber, and S.J. Steiner

The University of Birmingham
The School of Manufacturing and Mechanical Engineering
a.c.boud@bham.ac.uk

Abstract

In this paper we investigate whether Virtual Reality (VR) and Augmented Reality (AR) offer potential for the training of manual skills, such as for assembly tasks, in comparison to conventional media.

We present results from experiments that compare assembly completion times for a number of different conditions. We firstly investigate completion times for a task where participants can study an engineering drawing and an assembly plan and then conduct the task. We then investigate the task under various VR conditions and context-free AR.

We discuss the relative advantages and limitations of using VR and AR as training media for investigating assembly operations, and we present the results of our experimental work.

1. Introduction

In recent years, the terms virtual reality (VR) and augmented reality (AR) have each received a growing amount of interest and support that has seen the development of a number of different fields of investigation. However, although VR and AR rely on different technologies providing very different solutions, both technologies are often 'branded' in the same category. For example, Wilson [1] categorises AR as a form of VR, while Drascic and Milgram [2] describe VR and AR in terms of a 'reality-virtuality continuum', where AR is towards the real world end of the continuum and VR is at the opposite extreme.

VR can be defined as a three dimensional computer generated environment, updating in real time, and allowing human interaction through various input/output devices. By providing a variety of representations, e.g. 2D or 3D, desktop or immersive, VR can offer users the opportunity to explore virtual objects at levels of detail appropriate to the work activity.

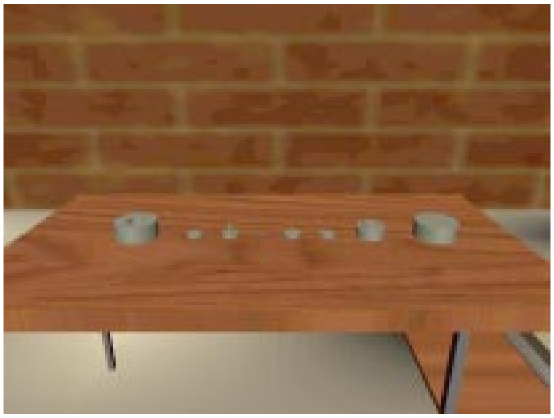
AR can be defined as the *enhancement* of the real world by a virtual world, which subsequently provides additional information [3]. AR itself can then be classified into two types: computationally *context-free* and computationally *context-aware*. The context-free systems use a see-through head mounted display with the computer image projected in front of the user by a half-silvered mirror. The context-aware systems overlay graphics or other media onto a real image by sensing the context in which it finds itself. Context-aware AR can only be achieved through the use of computer-vision or electronic sensors. In comparison, VR simulates rather than supplements the real world.

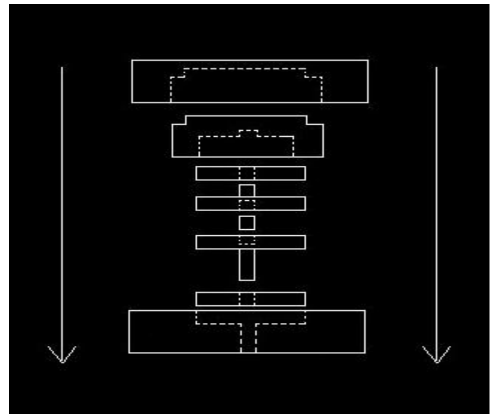
There has been little discussion comparing these technologies, mainly due to the differing benefits offered by both of these disciplines. In this paper we do not seek to suggest one as the optimal method at the expense of the other, but rather wish to present the relative advantages of each technology. As a basis of evaluation, we have chosen to investigate the use of both technologies as a training tool for assembly tasks.

2. Assembly

Assembly requires the manipulation and joining of parts to form a whole [4]. In order to achieve the apparently simple goal of placing a peg in a hole, a number of factors need to be considered, such as reaching for and grasping the peg, determining the relative positions of peg and hole, transporting the peg towards the hole (aiming), and inserting the peg accurately [5]. Each of these actions requires differing levels of haptic and visual guidance. If multiple pegs are to be inserted into multiple holes, then further consideration of a higher order of cognitive activity, such as planning, is required.

If assembly can be considered to have such cognitive components, then we ought to be able to demonstrate a





(a) **(b)**

Figure 1. Pictorial representation of the water pump assembly (a) VR conditions (b) AR condition

learning period (during which assembly performance improves), and to observe the effect of the instruction format on the learning of the assembly task. Work on assembly in the 1960s as reviewed by Seymour [6], clearly showed learning periods in assembly work.

Baggett and Ehrenfeucht [7] demonstrated that the presentation of instructions via a video could improve performance of assembly operations. One group of participants viewed a video of each subassembly being put together (with no information regarding final build), while another group viewed a video of subassemblies being put together in the correct sequence. The hypothesis was that the latter would facilitate learning of the assembly sequence better than the former. However, there was a clear interaction between knowledge of the assembly and activity, and impact of the video; for participants with little prior experience, the subassemblies plus sequence video led to better performance, while for participants with assembly knowledge there was no difference in the effect of video. Thus, simply seeing the assemblies being built was sufficient for experienced participants to be able to develop assembly plans.

While the work of Baggett and Ehrenfeucht [7,8] suggest that dynamic representations, i.e., through video, present superior training media to static, i.e., through paper, their work does not allow any interaction between the participant and the assembled item. This study investigates whether the application of VR and AR to permit direct manipulation of assembly components would facilitate training in this activity.

3. Experimental Work

Information for an operator to conduct an assembly task is conventionally obtained from a combination of engineering drawings and written assembly procedures. The investigation at the University of Birmingham has sought to develop a VR system where an operator could visualise the construction of a product, and then perform the assembly operations in the virtual environment (VE) under a number of different conditions in order to evaluate the application of VR as a tool to improve assembly completion times, in this case for a water pump assembly.

We also present a context-free AR system that displays a static pictorial representation of the water pump's assembly sequence. For the purpose of this study, the issue of attaching and overlaying representations to objects is unimportant due to the measurement of performance in terms of carrying out the task without using the AR system. The AR system is only used for training. However, work is being carried out at the University of Birmingham to develop a computer vision-based, context-aware AR system so that further studies can be carried out; for example, comparing the performance between context-free and context-aware systems.

The aim of the experimentation is to investigate the immediate impact of a given format of media on task performance to operator training and post-training.

3.1 Subjects

Five groups of five participants were selected from the student body of the School. All participants were from an engineering background and were familiar with 2D engineering drawings and assembly plans. None of the subjects had any previous experience of VR or AR.

3.2. Equipment

The VR conditions were conducted on a Silicon Graphics Indigo[2] maximum impact workstation, using either a conventional '2D' mouse and monitor (Desktop), a '2D' mouse, monitor and CrystalEyes stereoscopic glasses (Stereoscopic), and a Virtual Research 'Vr4' HMD, a Polhemus Fastrak tracking system and a '3D' mouse (Immersive). The model was generated using 3D Studio Max. with Divisions' dVISE VR software, and was constructed from approximately 7k polygons, and maintained at approximately ≥ 20 fps (see figure 1(a)).

For the context-free AR condition a see-through, monocular, monochrome, head mounted display (HMD) developed by Seattle Sight was used. The HMD was connected to a Pentium PC running at a speed of 200 MHz. A 2D static engineering diagram was used to provide the augmentation for the user. The diagram had a white outline and a black background; the black background allowed the real world to be seen more clearly (see figure 1(b)).

3.3. Method

Participants were asked to assemble a water pump (consisting of 8 separate components) in the real world, after receiving information from either:

- Conventional 2D engineering drawings (Conventional)
- Desktop VR using a monitor and 2D mouse (Desktop)
- Desktop VR using stereoscopic glasses to provide a 3D images, but still using a 2D mouse (Stereoscopic)
- Immersive VR using a HMD and '3D' mouse (Immersive), and
- Context-free AR (AR)

To complete the assembly using conventional 2D drawings, each participant was shown an assembly plan and an engineering drawing of a water pump assembly, and then given 10 minutes to study the construction. The participant then assembled the water pump. Each task completion time was measured.

The VR participants were firstly given a brief introduction on how to use the VR software and opportunity experience the interaction techniques for each variation. The 2D engineering drawing was then given to the participant to study for 2 minutes in order to assess the correct sequences of assembly operations, whereby the participant could then investigate the assembly operation in the VE for the remaining 8 minutes. The participants then completed the assembly operations in the VE using one of the specific interaction techniques, and were then asked to complete the assembly task on the water pump in the real world. The time taken to complete the assembly of the water pump was recorded in each case.

In the AR condition, participants were asked to don the headset and adjust the display to enable them to view the image clearly. Once the headset was correctly fitted, participants viewed the image on the screen to construct the pump. A maximum of 8 minutes was allotted to this task. The subjects were then asked to construct the pump without the AR system and the completion time was monitored.

3.4. Results

Figure 2 indicates that task completion times are longer when the participants train to assemble the water pump using the 2D drawing before assembling the real product, in comparison with the other conditions. There was a significant difference between the best VR condition and AR system ($t=2.132$, $df=4$, $p < 0.01$). However, there are no significant findings between the different forms of VR interaction, although participants suggested that immersive VR was more 'intuitive', as they were able to manipulate 3D objects in 3D space.

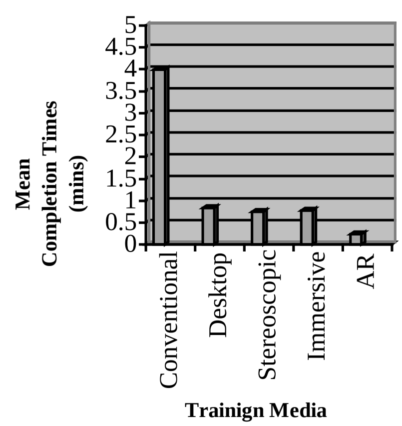


Figure 2. Comparison of total assembly completion time in the real world using different interaction techniques

The participants were asked for their opinions of both the VR and AR systems at the end of the experiments. From this, several issues were highlighted which may have contributed to the slower performance times in the VE in comparison to the real world. 8 out of the 20 users indicated that their control of the system was slow; 7 users indicated that the VR system limited the participant to conduct the task using only one hand. In the real world however, both hands would be used. Finally, 4 users indicated difficulty in selecting objects and ‘reaching through’ the computer generated components, as no haptic feedback was provided. In addition, for the AR condition the subjects were generally positive about the usefulness of the technology.

3.5. Discussion

One of the most important theories of skill acquisition stems from the work of Fitts [9], who suggested a three-stage development process. Initially, people learn the basic procedures and properties of the object. This is the cognitive stage. Secondly, the procedures and knowledge of the objects become ‘chunked’ into sequences of action. This is the associative stage. Finally, the sequences of actions are combined into a smooth pattern of activity. Each stage requires a decreasing level of overt conscious control, with the final stage representing ‘skilled’ activity.

It is proposed that assembly tasks completed using 2D drawings and assembly plans are based on the cognitive phase, where participants require a longer period of time to calculate the correct sequence of actions. Using VR to conduct the assembly tasks uses the cognitive and associative stages, where plans are developed to complete the tasks. VR and AR therefore ought to be superior for learning sequences.

However, assembly not only requires a learned sequence, but also a learned motor behaviour. VR participants were able to investigate assembly sequences through a number of VR conditions, before physically assembling the water pump. By enabling participants to investigate tasks at the level of detail appropriate to the work activity, VR offers improvements over conventional media.

The subjects in the AR system could construct the real life pump and simultaneously access the virtual information for guidance. The mapping between the training and the actual task is therefore more tightly coupled than for the other conditions. AR can therefore facilitate fast learning for simple assembly tasks, as the operator is able to conduct the actual task whilst referencing additional training material. In terms of Fitt’s theory of skill acquisition, the AR system allowed the subject to cognize, chunk information and to become partially skilled by the use of augmentation while performing the training for the task. By the time the subject had to perform the real task, their skill level was starting to develop.

There are, however, problems that need to be addressed with both VR and AR. one of the main problems of VR technology is the lack of provision of haptic feedback, and this plays an important role when manipulating objects in VEs, especially for immersive applications. Although there are a number of commercially available devices, several limitations are still encountered. (See Burdea [10] for a review of force feedback technology for VR). Previous work by Boud et al., [11] has addressed this problem with the use of instrumented objects (IOs).

AR on the other hand, allows the user to have tactile feedback through the manipulation of the real objects. Additional virtual information supplements the user’s knowledge by providing instructions in the same manner as a piece of paper. AR allows the user to attend to the set of instructions and the real world without having to refer to separate media, such as an instruction manual. Context-free and context-aware AR systems have been used in a variety of applications to provide support for the user. For example, Webster et al., [12] used AR to aid in the construction of architectural structures, and Feiner et al., [3] used an AR system to provide information on printer maintenance. However, AR systems have not been examined as a training tool where the user is expected to use the AR system to build up their knowledge of the task and then perform the task without the aid of the AR system.

With context-free AR, the user still needs to attend to an image at a specific focal length and then refocus on the real world [13]. The images can be merged using context-aware AR systems; however, problems such as registration between the virtual objects and the real objects needs to be overcome [14].

VR does however, have the advantage of separating the real world from the virtual world in terms of the accessibility of the real world objects. Hence, in a manufacturing environment, operators could be trained for an assembly operation during a product’s design cycle, before a physical prototype has been manufactured. With AR, the physical objects need to be present in order for the subject to be trained; this is not the case for VR.

4. Conclusion

The VR and AR conditions were found to out-perform the 2D engineering drawing condition. The two types of realities therefore offer the advantage of improved performance for assembly training over the conventional approach of studying a 2D drawing and then being asked to perform the task. The VR conditions demonstrated that there were no significant differences between the VR technologies for this particular application, however, there was a significant difference between the best VR condition and the AR system. VR allows the user to manipulate objects without the use of the real objects and hence offers benefits in applications such as manufacturing, where operators can be trained to assemble a product before the product has been physically manufactured. In terms of training, VR is more flexible than AR in that the environment in which it can be used can be separated from the real environment. For example, in some cases it may not be practical to interact with the real objects due to their non-availability and their associated costs. However, AR does improve performance times as the training is conducted on the real objects.

Both AR and VR have relative merits for training purposes, but their use relies upon the particular application. Therefore before employing these technologies it is important to investigate the task to ensure the benefits offered by both technologies are maximised.

Acknowledgements

The work undertaken in this project would not have been possible without the invaluable assistance of Simon Batterbee, an undergraduate student in the School of Manufacturing and Mechanical Engineering (1994-1997).

References

[1] Wilson, J.R., "Virtual environments and ergonomics: needs and opportunities", *Ergonomics, 40(10)*, 1998, pp. 1057-1077.

[2] Drascic and Milgram "Perceptual Issues in Augmented Reality", *SPIE Volume 2653: Stereoscopic Displays and Virtual Reality Systems III*, Ed. M.T. Bolas, S.S.Fisher, J.O. Merritt, Jan Jose, California, USA, January-February 1996, pp. 123-134.

[3] S. Feiner, B. MacIntyre, D. Seligmann, "Knowledge-based Augmented Reality", *Communications of the ACM*, 36(7), 1993, pp. 53-62.

[4] J.L. Nevins, and D.E. Whitney, "Assembly research", *Automatic, 16*, 1980, pp. 595-613.

[5] R.F. O'Connor, C. Baber, M. Musri, and H. Ekerol, "Identification, classification and management of errors in automated component assembly tasks", *International Journal of Production Research 31(8)* 1993, pp. 1853-1863.

[6] W.D. Seymour, *Industrial Skills*, London:Pitman. 1967.

[7] P. Baggett, and A. Ehrenfeucht, "Building physical and mental models in assembly tasks", *International Journal of Industrial Ergonomics, 7 (3),* 1991, pp. 217-227.

[8] P. Baggett, and A. Ehrenfeucht, "Conceptualizing in assembly tasks", *Human Factors 30(3),* 1988, pp. 269-284.

[9] P.M. Fitts, "The information capacity of the human motor system in controlling the amplitude of movement", *Journal of Experimental Psychology, 47*, 1954, pp. 381-391.

[10] G. Burdea, *Force and Touch Feedback for Virtual Reality*, Wiley-Interscience Publication, 1996.

[11] A.C. Boud, C. Baber, S.J. Steiner, "Virtual Reality: A Tool for Assembly", *Presence: Teleoperators and Virtual Environments.* In Press.

[12] A. Webster. S. Feiner, B. McIntyre, W. Mussie, and T. Krueger, "Augmented Reality in Architectural Construction, Inspection and Renovation", *Proceedings of ASCR Third Congress on Computing in Civil Engineering*, Anaheim, CA, June 17-19 1996, pp. 913-919.

[13] C. Baber, D. Haniff, L. Cooper, J. Knight, and B.A. Mellor, "Preliminary Investigations into the Human Factors of Wearable Computers", In Ed. H. Johnson, L. Nigay, & C. Roast, *People and Computers XIII*, Berlin: Springer-Verlag, 1998, pp.313-326.

[14] R.T. Azuma, "A Survey of Augmented Reality", *Presence: Teleoperators and Virtual Environments, MIT Press, 6*,(4), 1997, pp.355-385.