Intelligent Augmented Reality Training for Motherboard Assembly

Giles Westerfield · Antonija Mitrovic · Mark Billinghurst

Abstract We investigate the combination of Augmented Reality (AR) with Intelligent Tutoring Systems (ITS) to assist with training for manual assembly tasks. Our approach combines AR graphics with adaptive guidance from the ITS to provide a more effective learning experience. We have developed a modular software framework for intelligent AR training systems, and a prototype based on this framework that teaches novice users how to assemble a computer motherboard. An evaluation found that our intelligent AR system improved test scores by 25 % and that task performance was 30 % faster compared to the same AR training system without intelligent support. We conclude that using an intelligent AR tutor can significantly improve learning compared to more traditional AR training.

Keywords Augmented reality · Intelligent tutoring · Constraint-based tutors · Training for assembly skills · ASPIRE

Introduction

Augmented Reality (AR) allows the user's view of reality to be combined with virtual content that appears to be spatially registered in the real world (Azuma 1997). AR technology has the potential to revolutionize education due to its unique ability to visually convey abstract concepts and present 3D information in context with real objects. For example, chemistry students could physically assemble virtual atoms

G. Westerfield · M. Billinghurst
The Human Interface Technology Laboratory NZ, University of Canterbury, Christchurch, New Zealand

G. Westerfield
e-mail: giles.westerfield@gmail.com

M. Billinghurst
e-mail: mark.billinghurst@canterbury.ac.nz

A. Mitrovic (✉)
Intelligent Computer Tutoring Group, Department of Computer Science and Software Engineering, University of Canterbury, Christchurch, New Zealand
e-mail: tanja.mitrovic@canterbury.ac.nz

(Singhal et al. 2012); students could rotate the earth and sun to explore the relationship of sunlight and day and night (Shelton and Hedley 2002; Kerawalla et al. 2006).

In addition to assisting with the teaching of abstract concepts, AR can also help with real-world tasks. For example, AR windshield displays can be incorporated into vehicles to highlight the edge of the road and identify hazards at night or in heavy fog (Tonnis et al. 2005). Navigation is another large application domain, where AR displays can direct users to their destinations or provide additional information about landmarks (Narzt et al. 2006; Dunser et al. 2012). Numerous other AR applications have been developed for areas as diverse as archaeology (Eve 2012), car design (Noelle 2002), gaming (Mueller et al. 2003), medicine (Navab et al. 2012), and tourism (Bartie and Mackaness 2006).

The ability to combine abstract concepts and 3D spatial information in the context of real-world objects makes AR an ideal tool for training in situations which require manipulation of objects, such as manual assembly and maintenance tasks. Whether a person is putting together furniture or repairing a car engine, these types of tasks are inherently spatial in nature, and can be difficult to teach without close instruction and supervision. Unfortunately, personalized human assistance is not always available or cost effective. Many systems include instruction manuals containing diagrams that detail the necessary steps to be performed, but these can be difficult and time consuming to interpret, and only show static information. Video tutorials can be more effective because they harness the power of animated visual instruction, but the user must repeatedly switch between the video and the real-world environment.

AR has the capacity to deliver hands-on training where users receive visual instructions in the context of the real-world objects. For example, instead of reading a paper manual, a person could look at a car engine while an AR application uses virtual cues to shows the parts that need to be adjusted and the sequence of steps required. In this way AR has the potential to provide a more intuitive, interactive and efficient training experience, and could provide new possibilities for rapid skill development.

While there has been much research into the use of AR to assist with assembly and maintenance, existing systems generally focus on improving user performance while using the AR interface as opposed to teaching the user how to perform the task without assistance. Most systems guide the user through a fixed series of steps and provide minimal feedback when the user makes a mistake, which is not conducive to learning. The learning experience is the same for every user, and there is little regard for whether learning is actually taking place.

In contrast, Intelligent Tutoring Systems (ITSs) provide customized instruction to each student (Psotka and Mutter 1988; Woolf 2008). ITSs have been applied successfully to a variety of instructional domains, such as physics, algebra, genetics and database design (VanLehn et al. 2005; Koedinger et al. 1997; Corbett et al. 2013; Mitrovic 2012). Typically, the interfaces employed are text-based or 2D graphical applets, which limit their ability to convey spatial or physical concepts. Notable exceptions include simulation-based ITSs and game-based learning environments, which provide 3D simulation interfaces such as Tactical Iraqi (Johnson 2007). There have been some studies on combining ITSs with Virtual Reality, but very few examining the combination of ITSs with AR. The integration of AR interfaces with ITSs creates new possibilities for both fields and could improve the way we acquire practical skills.

This paper presents MAT, a Motherboard Assembly Tutor. An earlier version of the paper (Westerfield et al. 2013) was presented at AIED 2013; in this paper we provide additional discussion on the technologies used, and the development process. We start by reviewing previous research on using Augmented Reality to provide training for assembly and maintenance tasks, and then present the architecture of MAT and discuss the development process employed. The main research question of our project is whether intelligent AR-based training enables users to learn and retain assembly skills more effectively than traditional AR training approaches. To address this question, we performed a small evaluation study. The results strongly support our conclusion that using an intelligent AR tutor can significantly improve learning outcomes over traditional AR training.

Related Work

Training for manual assembly and maintenance is one type of learning that can benefit significantly from the use of AR because these "hands-on" tasks are inherently spatial and lend themselves naturally to visual instruction. Earlier research in utilizing AR for training has largely involved procedural tasks where the user follows visual cues to perform a series of steps, with the focus on maximizing the user's efficiency while using the AR system. Caudell and Mizell (1992) developed one of the first industrial AR applications, which assisted with assembling aircraft wire bundles. Their goal was to improve worker efficiency and lower costs by reducing reliance on traditional templates, diagrams and masking devices normally employed in the assembly process. The AR display used simple wire-frame graphics to show the path of the cable to be added to the bundle, but the user evaluation showed that there were a number of practical concerts to be solved before deployment in a real aircraft factory (Curtis et al. 1999).

Another early investigation involved the use of AR to assist with car door lock assembly (Reiners et al. 1999). This system used 3D CAD models of the car door and the internal locking mechanism, and guided users through the linear assembly process in a step-by-step fashion, responding to voice commands to move between the steps. However, the prototype was not stable enough for novice users, so some introductory training was required to gain any tangible benefit from the AR system.

These early studies led to the formation of several research groups dedicated to exploring the use of AR for industrial applications. ARVIKA was a group based in Germany whose mission was to use AR to support working procedures in the development, production, and servicing of complex technical products and systems, including automobile and aircraft manufacturing and power plant servicing (Friedrich 2002). Their focus was on practicality and applicability, since most previous AR prototypes were too unwieldy to be integrated successfully into industrial workplaces. The researchers conducted usability tests to evaluate ergonomic aspects of AR hardware and software, the time–cost and quality effects of using AR in the work process, and the benefit of AR telepresence, which allows specialists to provide remote assistance to field technicians. The studies found that the use of AR in industrial contexts can be extremely beneficial, and that the expensive nature of AR systems is often offset by reduced development time and improved product quality. For example, design engineers were able to rapidly evaluate ergonomic aspects of different aircraft cockpit

prototypes by overlaying virtual layout elements over real cockpit mockups, significantly streamlining the design process.

Another research group, Services and Training through Augmented Reality (STAR), was formed between research institutes in the USA and Europe around the same time (Raczynski and Gussmann 2004). The primary focus of STAR was to develop new AR techniques for training, documentation and planning purposes. One of the resulting prototypes allows a technician to capture video of the work environment and transmit the images to an off-site specialist. The specialist then annotates the video with drawing and text, which appears in the worker's augmented view in a spatially registered fashion. The researchers found that this method of remote collaboration was an effective means of communicating physical procedures and that it allowed a person with expertise to share his/her knowledge efficiently with multiple trainees in different locations.

Henderson and Feiner (2009) developed an AR application to support military mechanics conducting routine maintenance tasks inside an armored vehicle turret. In their user study involving real military mechanics, they found that the use of AR allowed the users to locate components 56 % faster than when using traditional untracked head-up displays (HUDs) and 47 % faster than using standard computer monitors. They also discovered that in some cases the AR condition resulted in less overall head movement, which suggested that it was physically more efficient. The evaluation also included a qualitative survey, which demonstrated that the participants found the Augmented Reality condition to be intuitive and satisfying for the tested sequence of tasks.

In addition to large industrial applications, Augmented Reality has been used to assist with assembly on a smaller scale. A study conducted by Tang et al. (2003) prompted users to assemble toy blocks into specific configurations using several different forms of instruction: traditional printed media, instructions displayed on an LCD monitor, static instructions displayed via a see-through Head-Mounted Display (HMD), and spatially-registered AR instructions also using a HMD. The researchers found that AR instructions overlaid in 3D resulted in an 82 % reduction in the error rate for the assembly task. They also found that the AR approach was particularly useful for diminishing cumulative errors, i.e. errors resulting from previous assembly mistakes. Another study by Robertson et al. (2008) used a similar set of test conditions and found that users assembled toy blocks more quickly using 3D registered AR than with 2D nonregistered AR and graphics displayed on a HUD.

These toy block assembly studies provide valuable insight, but the tasks performed are somewhat abstract in nature. AR has also been applied to real-world assembly tasks in non-industrial settings. One such study conducted by Baird and Barfield (1999) involved the assembly of components on a computer motherboard. The participants were asked to perform the task using a number of different instructional media: printed material, slides on a computer monitor, and screen-fixed text on opaque and see-through HMDs. The researchers observed that the users completed the assembly task significantly faster and with fewer errors when using the HMD displays. This motherboard assembly task is similar to the one used in our project, but Baird and Barfield's system did not employ spatially-registered AR and did not utilize an ITS. In addition, their evaluation concerned itself only with the performance of users and did not test knowledge retention after the training was complete. Other similar studies have

demonstrated positive results for the integration of AR with real-world assembly and maintenance tasks in various domains, including furniture assembly (Zauner et al. 2003), medical assembly (Nilsson and Johansson 2007) and laser printer maintenance (Feiner et al. 1993).

There have been studies investigating the combination of ITSs with Virtual Reality, such as (Mendez et al. 2003; Evers and Nijholt 2000; Fournier-Viger et al. 2009), but very few examining the combination of ITSs with AR. The integration of AR interfaces with ITSs creates new possibilities for both fields and could improve the way we acquire practical skills. A few projects claim to have created intelligent AR applications, but in practice these systems are minimally intelligent and do not employ domain, student and pedagogical models to provide adaptive tutoring. For example, Qiao et al. (2008) developed an AR system that teaches users about the instruments in a cockpit. Their system detects which cockpit component the user is looking at and then displays relevant information describing the component's function. This context-based interface is very different from the kind of intelligence that is employed in ITSs.

Feiner et al. (1993) developed a prototype that employed what they call Knowledge-based Augmented Reality. Their system employed a rule-based intelligent back-end called IBIS (Intent-Based Illustration System), to dynamically generate graphics based on the communicative intent of the AR system at any particular moment. The communicative intent is represented by a series of goals, which specify what the resulting graphical output is supposed to accomplish. For example, a goal could be to show a property of an object, such as its location or shape, or to show a change in a property. Feiner and his colleagues demonstrated their system with a prototype that assists users with laser printer maintenance. While this system is intelligent in how it generates the graphics for the user, it is neither intelligent from a training/tutoring standpoint nor adaptive.

The Architecture and Development of MAT

The Motherboard Assembly Tutor is an intelligent AR system for training users how to assemble components on a computer motherboard, including identifying individual components, and installing memory, processors, and heat sinks. Figure 1 shows the system's architecture, which is designed to be as modular as possible so that it can be easily adapted for new assembly and maintenance tasks. The display elements and the domain model must be customized for each type of task, but the underlying software architecture, scaffolding algorithms and other back-end processing remains the same.
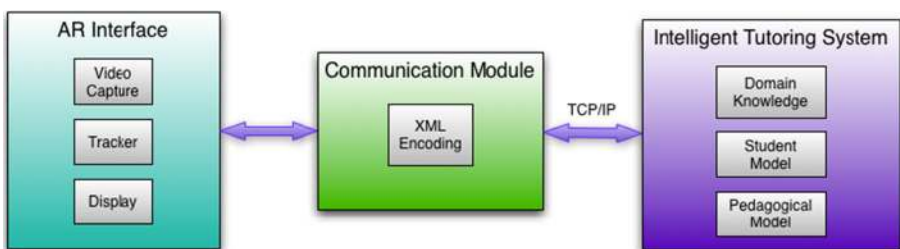

Fig. 1 The architecture of MAT

The communication module relays information between the AR interface and the ITS. The ITS we developed is a constraint-based tutor which controls what the user sees via the interface, and the AR interface tells the ITS what the user is doing. The AR interface encapsulates the video capture, tracking system, display and keyboard input. It uses 3D graphics, animations, audio and text, which are blended with the student's view of reality via a head-mounted display. The interface uses a camera to observe the student's behaviour, and the communication module sends the necessary data to the ITS via XML remote procedure calls over a TCP/IP network connection. The ITS analyzes the student's action by matching it to domain constraints, and provides feedback about student performance. In this section, we describe the ITS first, followed by the description of the AR interface.

Developing the Intelligent Support

The intelligent tutoring support was developed in ASPIRE,[1] an authoring system and deployment environment for constraint-based tutors (Mitrovic et al. 2009). We followed the standard procedure for developing constraint-based tutors in ASPIRE: no changes were done to the authoring system to accommodate MAT. As required by ASPIRE, the first stage of the authoring process involves describing characteristics of the task. In the case of MAT, the assembly task is procedural in nature and consists of 18 steps to be completed, such as opening the processor enclosure and inserting the processor in the correct orientation. The second stage of authoring consists of composing an ontology of the relevant domain concepts. We developed the domain ontology for MAT in ASPIRE's domain ontology editor, which is illustrated in Fig. 2. Each concept in the domain ontology has a number of properties and relationships to other domain concepts. For example, in the case of a memory slot, an important property is an indicator of whether the slot is open or not, since the slot must be opened before the memory can be installed. This property is represented as a Boolean value. There are 14 domain concepts in the ontology.

Next, we specified the solution structure by indicating which ontology concepts are involved with each problem-solving step. For example, installing computer memory involves four steps: (1) identifying and picking up the memory component, (2) opening the locking levers at the ends of the memory slot, (3) aligning the memory with the slot in the correct orientation, and (4) pushing the memory down into the slot until it locks. Each of these steps has at least one concept associated with it, and each concept has properties that are used to determine whether the student's solution is correct. In the case of the *open locking levers* step, the ITS uses the Boolean *isOpen* property of the *MemorySlot* concept to determine whether the slot has been successfully opened or not. The value of the Boolean property is set via the AR interface, which is described in the next section.

The following step was to create the interface that the students would use. ASPIRE generates text-based interfaces automatically (based on the domain information, domain ontology and the problem and solution structures), but it also allows for the default interface to be replaced with one or more Java applets (Mitrovic et al. 2007). ASPIRE also communicates over a network via a remote procedure call (RPC)
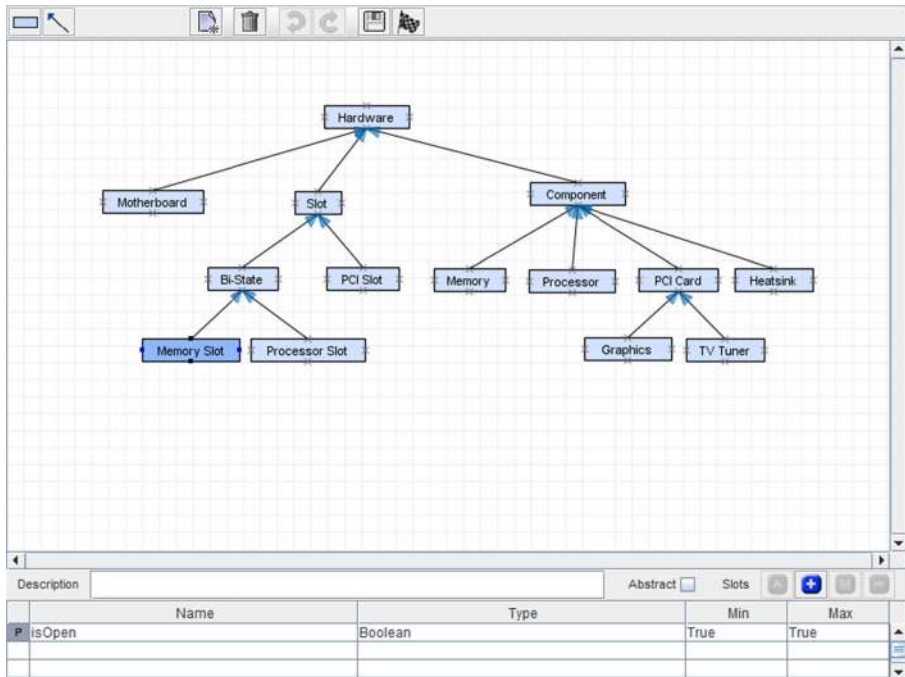
---

[1] http://aspire.cosc.canterbury.ac.nz/

Fig. 2 Composing the domain ontology of MAT in ASPIRE

protocol, which allows it to communicate with an external AR interface. In the case of MAT, the AR front-end communicates with ASPIRE over a network.

We then specified a set of problems with their solutions. The problem structure describes steps that apply to all motherboards, while a particular problem and associated solutions apply to a specific brand and model of motherboard. ASPIRE allows multiple solutions to be specified for each problem. In the case of motherboard assembly, there is often only one way to correctly install each component, but this is not always the case. For example, a memory module can be inserted into one of several slots, and a heat sink can sometimes be installed in more than one orientation. Accepting these different configurations as correct solutions gives the student more flexibility when solving the problem and enhances learning.

Using the information provided in the domain ontology, problem/solution structures, and the set of problems with solutions, ASPIRE generated the domain model consisting of 275 constraints. Each generated constraint contained feedback messages referring to the names of properties and concepts from the domain ontology; ASPIRE generates such feedback from templates. In order to make the automatically generated feedback messages more understandable, the author needs to re-write them in some cases. We tailored the constraints by changing those automatically generated feedback messages so that the feedback is more useful for the students.

When the domain model is completed, and the author is satisfied with the constraints, the tutor is ready to be deployed to ASPIRE-Tutor, the tutoring server of ASPIRE. ASPIRE-Tutor performs all standard ITS functions, such as student modelling, providing feedback and problem selection. Every action the student performs is matched to the constraints. If the action is correct, the student receives positive feedback. In the opposite case, if there is a violated constraint, the student receives feedback about the mistake made.

The author can create individual accounts for students and add them to groups, each of which can have customized settings. These include specifying the type of feedback to be supplied as well as the progression between the feedback levels as the student makes mistakes. Typically the system is set to provide minimal help initially by only indicating that that student's solution is incorrect, but not showing how it is wrong. The tutor then provides more and more detailed feedback hints as the student struggles with the problem until finally the full solution is revealed.

AR Interface Design

The AR interface presents problems and other information from the ITS to the student. The tracking module calculates the pose of the computer motherboard and its components relative to the camera affixed to the head-mounted display. This serves two fundamental purposes: (1) It allows the display module to render 3D graphics on top of the video frame in such a way that the virtual models appear to be part of the real world, and (2) the tracker sends information about the relative positions of the motherboard components to the ITS, which allows it to analyze the user's behaviour and provide feedback. The bulk of the work performed in the tracking module is handled by the underlying osgART software library (Looser et al. 2006), which is a C++ framework that combines computer graphics, video capture and tracking, and makes it easy to create AR applications. The osgART library uses the ARToolkit marker tracking approach (Kato and Billinghurst 1999) and black square visual tracking images.

The display module is responsible for everything the user sees through the head-mounted display. The HMD chosen for the project is a video-see-through device, meaning the user looks at a screen that displays a video reproduction of their first-person perspective via a camera attached to the front of the HMD. As the result of this choice, the first responsibility of the display module is to obtain video frames from the camera and draw them on the screen. After each frame is rendered, virtual graphics can be drawn on top of the video background in order to create the illusion that they exist within the real scene. All of the graphics are generated by the OpenSceneGraph[2] computer graphics library (OSG), which has been integrated into the osgART software package. OSG is based on the standard OpenGL[3] API, and provides a robust scene graph structure. In addition to built-in support for materials, textures, lighting and shaders, OSG has a set of plug-ins that allow it to handle a wide variety of file formats for images, 3D models and sound.

We used the 3D Studio Max[4] modelling software to generate accurate 3D models of the components to be installed on the computer motherboard, including memory, processor, graphics card, TV tuner card and heatsink. Models were also produced for relevant parts of the motherboard itself, such as the processor enclosure and memory securing mechanisms. Other 3D models of virtual cues, such as arrows, were created to guide the user through the tutoring process. Figure 3 shows a first-person view of the display for the TV tuner installation task, with a virtual model of the TV tuner card showing where the real card should be inserted. The virtual models were animated to

---

[2] www.openscenegraph.org
[3] www.opengl.org
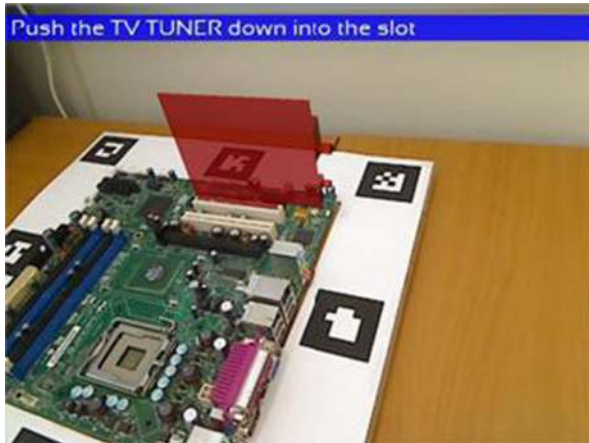[4] usa.autodesk.com/3ds-max

Fig. 3 First-person view of the AR display for part of the TV tuner installation task

illustrate the proper installation procedures. For example, the graphics card is visibly pushed downward into the PCI express slot, and the processor enclosure is opened before the processor is inserted. The animations were embedded into the exported 3D model files, which can be loaded directly into the display module by the appropriate plug-in in the OpenSceneGraph software library.

In addition to the spatially-registered 3D models, we developed a screen-aligned head-up display (HUD) for displaying text messages from the ITS. As the user looks around, the HUD components always stay in the same place on the screen. The ITS messages consist of instructions and positive/negative feedback. The text is displayed across the top of the screen and is highlighted with a semi-transparent background that changes color based on the message type. Instructions are blue (such as in Fig. 3), positive feedback is green and negative feedback is red (Fig. 4). The HUD also utilizes text-to-speech technology to read the messages to the user, using the Microsoft Speech API.

The hardware setup for the AR interface consists of a head-mounted display, a camera, a MS Windows computer and the ARToolkit fiducial markers used for tracking (Fig. 5). An Intel motherboard was selected for use with the computer assembly, as well as five generic hardware components to be installed: memory, processor, graphics card, TV tuner card and heatsink. At least one unique marker was attached to each component to enable the system to identify and track its position. The motherboard itself was mounted on a sturdy wooden surface and surrounded with a configuration of eight separate markers. This group of markers works together with the tracking system to limit the effects of marker occlusion as users look around and move their arms during the installation procedures. As long as the camera can see at least one of the eight markers, the tracking system is able to determine the relative position and orientation of the motherboard.

The HMD and camera combination chosen for the project is the Wrap 920AR model produced by Vuzix,[5] which has a resolution of 1,024×768 pixels with a 31-degree horizontal field of view (shown in Fig. 5). It supports stereoscopic viewing, and the

---

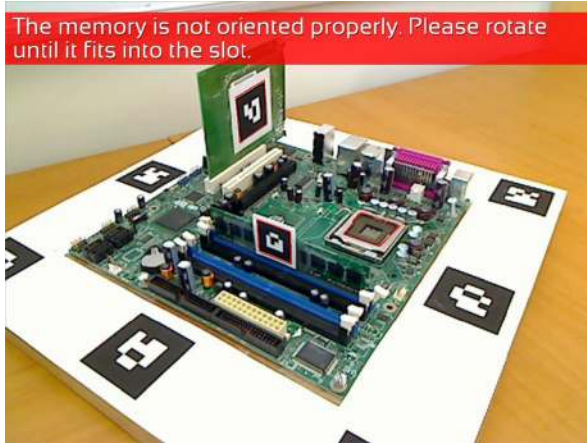[5] http://www.vuzix.com/consumer/produces\wrap920ar.html

Fig. 4 A negative feedback message for the memory installation task

front of the display is outfitted with two cameras for stereo video capture at 640×480 at 30 frames per second. The device connects to a computer via the standard VGA interface and also delivers audio via earbud headphones.

Experiment

We conducted a study in which we compared the intelligent AR system with a traditional AR tutor. The goal of the study was to determine the difference in knowledge retention between the two approaches. The evaluation was split into two phases: a
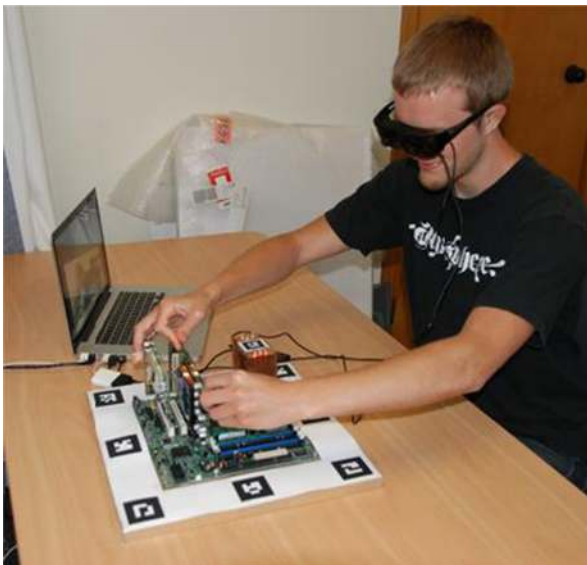

Fig. 5 A participant using the tutor, wearing the head-mounted display. The camera view is combined with the AR content, thus combining virtual models with the real world view

training phase and a testing phase (without the tutor) that measured the extent to which the participants retained the knowledge they acquired.

Both tutors presented the assembly steps to students in the same order. The student interface and the visual/oral instructions for each step were the same in both tutors, so the only differences lie in the features directly related to the ITS. In both tutors, the student indicates that he/she is finished with the current step by pressing a button. If the solution is incorrect, the intelligent tutor prevents the student from proceeding to the next step and provides a specific feedback message (coming from the violated constraint), while the traditional tutor always proceeds regardless.

There were 16 participants (11 males and 5 females) who were randomly allocated to one of the conditions. All of the participants were university students and aged between 18 and 45 years old. The experimental group used the intelligent AR tutor, while the control group used the traditional AR tutor. Great care was taken to select participants with minimal experience with computer hardware assembly. To measure this, all participants were given a written pre-test asking them to identify the five hardware components and their position on the motherboard. Following the pre-test, the participants were given an orientation to the AR tutor (intelligent or traditional) and its operation procedures. After they put on the head-mounted display, the tutor guided them through the process of identifying and installing five motherboard components: memory, processor, graphics card, TV tuner card and heatsink. After all of the components were assembled, the tutoring phase was complete and the participants were given a written post-test that was similar to the pre-test to measure how well they learned from the tutor. The two written tests covered the same material, but were not identical.

Immediately after the written post-test, the participants were asked to perform a physical post-test in which they attempted to assemble the motherboard components once more, this time without the help of the tutor. The aim of the physical post-test was to measure how well the participants retained the physical assembly knowledge gained from the tutoring process. Given only the name of each component, the participants had to correctly identify and install them one by one. In addition to qualitative observations, a number of quantitative measures were taken during this process, including task completion time and error counts.

Finally, the participants completed a questionnaire, which asked them to provide detailed feedback about their experience with the tutor. In addition to asking about prior hardware experience, the questionnaire contained a variety of questions with Likert-scale ratings. These asked the participants to indicate whether they thought the tutor was effective, whether they were satisfied with the 3D AR content, whether they thought the AR training system was more effective than other types of media such as videos or paper manuals, and whether they felt physically or mentally stressed during the tutoring process. Participants also had the opportunity to provide additional written feedback.

Results

Table 1 summarizes the written pre-test and post-test scores for the two groups. The maximum score on each test was 10 marks. There was no significant difference between the two groups on the pre-test performance. There was also no significant difference in the times both groups spent on working with the tutoring systems. The

Table 1 Mean and standard deviations for two groups

Group	Pre-test	Post-test	Normalized gain	Time (s)	Errors
Experimental	2.50 (2.27)	9.13 (1.13)	0.66 (0.26)	56.56 (11.31)	0.50 (0.93)
Control	2.63 (1.92)	6.63 (1.77)	0.40 (0.21)	81.13 (21.11)	1.00 (0.93)

performance of both groups increased significantly between the pre- and the post-test, yielding $t(7) = 7.165$, $p < .0002$ for the experimental group, and $t(7) = 5.291$, $p < .002$ for the control group. Both of these values are significantly less than the Bonferroni-corrected α value of .0083 (.05/6), which makes a very strong case for the effectiveness of both tutors.

Using the ITS significantly improved learning. The post-test performance of the experimental group is significantly higher than that of the control group ($t(14) = 3.374$, $p < .005$). This is less than the Bonferroni-corrected value of .0083 (.05/6), so the intelligent AR tutor produced a significantly better learning outcome than the non-intelligent AR tutor. There is also a significant difference between the normalized learning gains of the two groups ($t(14) = 2.198$, $p < .05$). The effect size (Cohen's d) is 0.981, which is a significant improvement.

Table 1 also reports the number of errors made and the total completion time to install all five motherboard components during the physical post-test. The errors generally fit into two categories: failing to match a name with the correct component, or incorrectly performing an installation procedure. There was no significant difference on the number of errors made, but the experimental group participants completed the physical task significantly faster than their peers ($t(14) = 2.9$, $p < .02$).

The questionnaire responses were positive for both tutors (Table 2). The responses were on the Likert scale from 1 (not very much) to 7 (very much). None of the ratings were significantly different between the two groups. The participants had no (44 %) or little experience (69 % rate themselves three or lower) with motherboard assembly prior to the study. None of the participants rated themselves very experienced. 81 % of the participants rated their level of agreement at six or higher with the statement that the

Table 2 Mean questionnaire responses

Question	Experimental	Control
Previous experience with motherboard assembly	3.4	2.1
AR tutor able to teach the procedure	6.0	6.0
Easy to perform the tasks	6.2	6.4
AR tutor was effective	6.1	6.1
Satisfied with feedback (graphics and text)	6.1	5.5
AR tutor more effective than a manual	5.8	6.5
AR tutor more effective than instructional video	5.6	5.5
AR tutor is interesting to use	6.5	6.8
Level of frustration and stress while using the tutor	1.4	1.9
Level of physical stress	1.6	1.8

tutor was able to teach them the procedure. Most participants felt that the visual step-by-step instructions were very helpful, allowing them to proceed at their own pace. When asked whether the AR tutor was more effective than the paper manual, 81 % of the participants rated their level of agreement at six or higher. When comparing the AR tutor with the video, 56 % had the same level of agreement. The immersive first-person experience provided by the head-mounted display was engaging, and the system as a whole was interesting and fun to use (94 % rated their interest level at six or higher). Some responses can be attributed to the novelty factor associated with AR, the participants generally found the tutors to be both effective and entertaining. Many of the experimental group participants found the ITS feedback very helpful.

One criticism stemmed from the fact that the textual instructions were screen-aligned in typical HUD fashion. Reading the text required the participants to shift their focus from looking into the scene to looking at the text displayed on the surface of the screen. It may have been more natural to use spatially-registered text that appeared within the scene to keep the students immersed in the AR environment. Other criticisms addressed the tracking performance. The virtual content would sometimes jiggle or disappear entirely when the tracking system was unable to obtain enough information about the markers. These issues could be addressed with a more robust tracking approach, perhaps one that utilizes multiple cameras and tracks the natural features of the motherboard components without markers.

While the participants found determining the correct position of the components to be relatively easy, determining the proper orientation was more difficult. This was partially due to a lack of orientation cues in some of the virtual content shown. The memory and processor are essentially symmetrical in shape, and it can be difficult to determine which direction the virtual rendering is facing when there are no distinguishing features. In these cases, it would be helpful to have some additional AR cues to help the student infer the correct orientation. One idea would be to attach virtual arrows to the motherboard slot as well as the actual component to be inserted, prompting the student to line up the arrows with each other. When this type of orientation mistake occurred, the intelligent AR tutor was able to detect the error and inform the student that the orientation was incorrect. The participant was required to correct the mistake before being allowed to proceed. The traditional tutor was unable to observe or correct errors, and they often went unnoticed by the student. In these cases, the student typically made similar mistakes during the post-test. This supports the claim that the ITS feedback improved the learning outcome over the traditional AR training approach, particularly where it was easy to make a mistake.

The results of the study confirm the hypothesis that the use of ITSs with AR training for assembly tasks significantly improves the learning outcome over traditional AR approaches.

Conclusions

Augmented Reality has been repeatedly shown to improve education and training through visualization and interactivity, but most AR training systems are not intelligent. In this paper we have shown how to combine an AR interface with an Intelligent Tutoring System to provide a robust and customized learning experience for each user. To demonstrate this, we created a prototype application that teaches users how to assemble hardware components on a computer motherboard. An evaluation found that

our intelligent AR system improved test scores by 25 % and that task performance was 30 % faster compared to the same AR training system without intelligent support. From these results, we conclude that adding intelligent tutoring can significantly improve the learning outcome over traditional AR training.

There are several limitations of our study. We had a small set of participants, and therefore it would be necessary to repeat the study with a bigger population of students. Furthermore, some of our participants have had limited experience with motherboard assembly, which might have affected the results. The tutors could be enhanced to teach declarative knowledge (i.e. educate participants about components) rather than focusing on the assembly procedure alone. It would also be interesting to conduct an experiment with additional conditions, in which participants will undergo training using video or written manuals, or an ITS without the AR interface.

There are many future research directions that could be explored. For example, the intelligent AR tutor could be extended by integrating a virtual character into the tutoring environment. Research has shown that virtual characters can be beneficial in tutoring situations as they increase student motivation (Johnson et al. 2000; Liu and Pan 2005; Gulz and Haake 2006). A 3D virtual character would allow the ITS to inhabit the world with the user, where it could give verbal instructions, make gestures and demonstrate installation procedures.

Tracking is another area in which the intelligent AR tutor can be improved. The current solution uses a fiducial marker-based approach, which has limited accuracy, poor resistance to occlusion and obtrusive markers. There are a number of better tracking approaches such as natural feature tracking or using multiple cameras to reduce the effect of occlusion. Stereoscopic cameras and depth mapping could be used to determine the three-dimensional shapes of objects. This would allow the system to generate a model of the environment on the fly, and adapt to new scenarios such as different brands of computer motherboards and components. It could also enable more complex training tasks that require more robust tracking.

Finally, more user studies need to be conducted in a wider range of training domains. Our results have shown the value of using an intelligent AR tutor in training for motherboard assembly, but similar studies need to be performed in other assembly or maintenance tasks to determine the educational benefits of intelligent AR training fully.

References

Azuma, R.T. (1997) A survey of augmented reality. *Presence: Teleoperators and Virtual Environments*, pp. 355–385.

Baird, K. M., & Barfield, W. (1999). Evaluating the effectiveness of augmented reality displays for a manual assembly task. *Virtual Reality, 4*(4), 250–259.

Bartie, P. J., & Mackaness, W. A. (2006). Development of a speech-based Augmented Reality system to support exploration of cityscape. *Transactions in GIS, 10*(1), 63–86.

Caudell, T., & Mizell, D. (1992). Augmented reality: an application of heads-up display technology to manual manufacturing processes. *Proceedings of the 25th Hawaii International Conference System Sciences, 2*, 659–669.

Corbett, A., MacLaren, B., Wagner, A., Kauffman, L., Mitchell, A., Baker, R. (2013) Differential impact of learning activities designed to support robust learning in the Genetics cognitive tutor. In *Proceedings of the. Artificial Intelligence in Education*, pp. 319–328.

Curtis, D., Mizell, D., Gruenbaum, P. & Janin, A. (1999) Several devils in the details: making an AR application work in the airplane factory. *Proceedings of the International workshop on Augmented reality: placing artificial objects in real scenes*, pp. 47–60.

Dunser, A., Billinghurst, M., Wen, J., Lehtinen, V., & Nurminen, A. (2012). Exploring the use of handheld AR for outdoor navigation. *Computers and Graphics, 36*(8), 1084–1095.

Eve, S. (2012). Augmenting phenomenology: using augmented reality to Aid archaeological phenomenology in the landscape. *Archaeological Method and Theory, 19*(4), 582–600.

Evers, M. & Nijholt, A. (2000) Jacob - An animated instruction agent in virtual reality, *Advances in Multimodal Interfaces*, pp. 526–533, Springer.

Feiner, S., Macintyre, B., & Seligmann, D. (1993). Knowledge-based augmented reality. *Communications of ACM, 36*, 53–62.

Fournier-Viger, P., Nkambou, R. & Nguifo, E. (2009) Exploiting partial problem spaces learned from users' interactions to provide key tutoring services in procedural and ill-defined domains. In Dimitrova, V., Mizoguchi, R., du Boulay, B., Graesser, A. (Eds.) *Proceedings of the 14th International Conference Artificial Intelligence in Education*, pp. 383–390.

Friedrich, W. (2002) ARVIKA - Augmented reality for development, production and service. *Proceedings of the International Symposium Mixed and Augmented Reality*, pp. 3–4.

Gulz, A., & Haake, M. (2006). Design of animated pedagogical agents—a look at their look. *Human-Computer Studies, 64*(4), 322–339.

Henderson, S. J. & Feiner, S. (2009) Evaluating the benefits of augmented reality for task localization in maintenance of an armored personnel carrier turret. *Proceedings of the 8th International Symposium Mixed and Augmented Reality*, 135–144.

Johnson, W. L. (2007). Serious use of a serious game for language learning. R. Luckin et al. (Eds.) *Proceedings of the Artificial Intelligence in Education, 158*, 67–74.

Johnson, W. L., Rickel, J. W., & Lester, J. C. (2000). Animated pedagogical agents: face-to-face interaction in interactive learning environments. *Artificial Intelligence in Education, 11*(1), 47–78.

Kato, H. & Billinghurst, M. (1999) Marker tracking and HMD calibration for a video-based augmented reality conferencing system. *Proceedings of the 2nd IEEE and ACM Int.Workshop on Augmented Reality*, pp. 85–94.

Kerawalla, L., Luckin, R., Seljeflot, S., & Woolard, A. (2006). Making it real: exploring the potential of augmented reality for teaching primary school science. *Virtual Reality, 10*(3–4), 163–174.

Koedinger, K. R., Anderson, J. R., Hadley, W. H., & Mark, M. A. (1997). Intelligent tutoring goes to the big city. *Artificial Intelligence in Education, 8*, 30–43.

Liu, Z. & Pan, Z. (2005) An emotion model of 3d virtual characters in intelligent virtual environment. In J. Tao, T. Tan, and R. Picard (eds.), *Affective Computing and Intelligent Interaction*, vol. 3784, Lecture Notes in Computer Science, pp. 629–636, Springer Berlin / Heidelberg.

Looser, J., Grasset, R., Seichter, H. & Billinghurst, M. (2006) OSGART-A pragmatic approach to MR. *Proceedings of the 5th IEEE and ACM International Symposium on Mixed and Augmented Reality*, pp. 22–25.

Mendez, G., Rickel, J. & de Antonio, A. (2003) Steve meets Jack: the integration of an intelligent tutor and a virtual environment with planning capabilities. *Proceedings of the Intelligent virtual agents*, pp. 325–332.

Mitrovic, A. (2012). Fifteen years of constraint-based tutors: what we have achieved and where we are going. *User Modeling and User-Adapted Interaction, 22*(1–2), 39–72.

Mitrovic, A., Martin, B. Suraweera, P., Milik, N., Holland, J. & Zakharov, K. (2007) *ASPIRE User Manual*, version 4.0, Intelligent Computer Tutoring Group.

Mitrovic, A., Martin, B., Suraweera, P., Zakharov, K., Milik, N., Holland, J., & McGuigan, N. (2009). ASPIRE: an authoring system and deployment environment for constraint-based tutors. *Artificial Intelligence in Education, 19*(2), 155–188.

Mueller, F., Agamanolis, S. & Picard, R. (2003). Exertion interfaces: sports over a distance for social bonding and fun. In *Proceedings of the SIGCHI Conference on Human factors in computing systems*, 561–568.

Narzt, W., Pomberger, G., Ferscha, A., Kolb, D., Müller, R., Wieghardt, J., Hörtner, H., & Lindinger, C. (2006). Augmented reality navigation systems. *Universal Access in the Information Society, 4*(3), 177–187.

Navab, N., Blum, T., Wang, L., Okur, A., & Wendler, T. (2012). First deployment of augmented reality in operating rooms. *IEEE Computer, 45*(7), 48–55.

Nilsson, S. & Johansson, B. (2007) Fun and usable: augmented reality instructions in a hospital setting. *Proceedings of the 19th Australasian Conference Computer-Human Interaction: Entertaining User Interfaces*, pp. 123–130.

Noelle, S. (2002) Stereo augmentation of simulation results on a projection wall by combining two basic ARVIKA systems. *International Symposium Mixed and Augmented Reality*, 271–322.

Psotka, J. & Mutter, S.A. (1988) *Intelligent Tutoring Systems: Lessons Learned*. Lawrence Erlbaum Associates.

Qiao, Y., Xie, X., & Sun, T. (2008). Design for the cockpit intelligent tutoring system based on augmented reality. *Proceedings of the International Symposium Computational Intelligence and Design, 2*, 224–227.

Raczynski, A. & Gussmann, P. (2004) Services and training through augmented reality. *Proceedings of the 1st European Conference Visual Media Production*, pp. 263–271.

Reiners, D., Stricker, D., Klinker, G. & Müller, S. (1999) Augmented reality for construction tasks: doorlock assembly. *Proceedings of the International workshop on Augmented Reality: Placing Artificial Objects in Real Scenes*, pp. 31–46.

Robertson, C. M.,MacIntyre, B. & Walker, B. N. (2008) An evaluation of graphical context when the graphics are outside of the task area. *Proceedings of the 7th IEEE/ACM International Symposium Mixed and Augmented Reality* (Washington, DC, USA), pp. 73–76, IEEE Computer Society.

Shelton, B. & Hedley, N. (2002) Using augmented reality for teaching earth-sun relationships to undergraduate geography students. Augmented Reality Toolkit. *Proceedings of the 1st IEEE International Augmented Reality Toolkit Workshop*. Darmstadt, Germany.

Singhal, S., Bagga, S., Goyal, P., & Saxena, V. (2012). Augmented chemistry: interactive education system. *International Journal of Computer Applications, 49*(15), 1–5.

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*, (New York, NY, USA), pp. 73–80.

Tonnis, M., Sandor, C., Lange, C. & Bubb, H. (2005) Experimental evaluation of an augmented reality visualization for directing a car driver's attention. *Proceedings of the 4th IEEE/ACM International Symposium on Mixed and Augmented Reality*, IEEE Computer Society, pp. 56–59.

VanLehn, K., Lynch, C., Schulze, K., Shapiro, J. A., Shelby, R., Taylor, L., Treacy, D., Weinstein, A., & Wintersgill, M. (2005). The Andes physics tutoring system: lessons learned. *Artificial Intelligence in Education, 15*, 147–204.

Westerfield, G., Mitrovic, A. & Billinghurst, M. (2013) Intelligent Augmented Reality Training for Assembly Tasks. *Proceedings of the 16th International Conference Artificial Intelligence in Education*, Springer LNAI 7926, pp. 542–551.

Woolf, B. (2008) *Building intelligent interactive tutors: student-centered strategies for revolutionizing e-learning*. Morgan Kaufmann.

Zauner, J., Haller, M., Brandl, A. & Hartmann, W. (2003) Authoring of a mixed reality assembly instructor for hierarchical structures. *Proceedings of the 2nd IEEE/ACM International Symposium Mixed and Augmented Reality*, p. 237.