Agents as Tutors in Augmented Reality

Natalia Álvarez
Computer Science Department
EAFIT University
Medellín, Colombia
email: nalvare2@eafit.edu.co

Helmuth Trefftz
Computer Science Department
EAFIT University
Medellín, Colombia
email: htrefftz@eafit.edu.co

ABSTRACT
Interaction between virtual entities and human beings should be similar to the normal interaction between human beings. Verbal interaction is the usual way of interaction among humans, but the interfaces that allow verbal communication between a human being and an agent (virtual human) have a limited vocabulary, need a lot of training and are expensive. In this paper we propose an Augmented Reality System with an unidirectional interface between an agent and a person. By using gestures, animations and adequate texts the agent can show a user how to perform a simple task. We hope that the tests to be conducted will show how useful this unidirectional interface is.

KEY WORDS
Virtual Reality, Augmented Reality, Agents, Interactive Virtual Environments.

1 Introduction

There is a growing interest on virtual objects superposed over real ones. For instance: virtual humans or other kind of characters, who simulate physical presence to perform dangerous or impossible movie scenes replacing real people, or agents that can work as experts to lead people inside complex environments.

Mixing virtual and real objects in a real environment can also be used for training applications, using animated agents as virtual characters capable of training a person on how to perform simple tasks. Simple tasks are considered those consisting of sequential steps, where the decisions are taken based upon a set values of known variables. These agents had been successfully used in immersive virtual environments[1], in which people are completely isolated from reality (immersion).

One of the objectives of the work reported here is to use agents as tutors inside a real ambient supplemented with virtual objects. If agent is one of those objects, then we have an Augmented Reality System like the ones defined by Azuma and other authors[2].

2 Augmented Reality

Azumas definition of Augmented Reality[2] is not only restricted to the sense of sight, it is open for all senses and several ways of communication. Milgrams continuum[3] figure 1, show how wide the possibilities are for applications that can be developed in the spectrum that goes from a real environment to a completely virtual environment. Augmented Reality is a region in the continuum covered by Hybrid Reality.

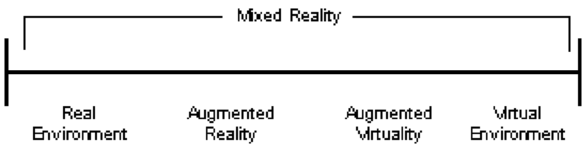

Figure 1. Milgrams reality-virtuality continuum

3 Agents

The presence of an agent in an educational environment makes it more substantial. Because it is not only the verbal communication that matters in the learning progress of an individual: movements and gestures get the users attention, in fact, visual self-feeding might indicate that an action is right or wrong without the system interfering in the process.

As mentioned in [4], there are two clear advantages in non-verbal communication. First, it increases the bandwidth between the individuals and the computer and also increases the availability for the computer to become a friendlier tool of learning. Nevertheless, another advantage associated with augmented reality focus on keeping in touch with the real world.

Interactive systems demand autonomous agents which are able to perform tasks and achieve goals regardless the environments complexity, work in collaborative tasks interacting with other agents. A learning environment is unpredictable, this is why a tutor agent must have a coherent behaviour inside the environment, such abilities allow the agent to solve questions, and to respond to stimulus generated by people, agents or other objects in the system.

4 Related Work

Adele and Steve are agents developed by The Center for Advance Research in Technology for Education (CARTE) at the university of Southern California (USC)

Three notable agents have been developed at the University of North Carolina: Herman the bug[5] used to teach botanical physiology and anatomy, talks and leads the students to build plants. Cosmo[6], is another agent used to teach about routing mechanisms and network topology, travelling through several sub-networks. WhizLow, developed to study a computer CPU, focused on the computer architecture.

PPP Persona[7] was developed in the German Research Center for Artificial Intelligence. Persona is an animated agent aimed at providing online support. It uses voice to comment suggestions and modify its schedule depending of the users actions.

5 System Description

All the agents mentioned above either inhabit a completely immersive virtual environment or are part of multimedia applications. The closest work to this work it that done by Thalman and Balcisory at the Swiss Federal Institute of Technology[8]. This system does not provide verbal interaction but uses an operator to indicate inputs and trigger actions in the agent.

The unidirectional interface proposed in this document, uses verbal communication or adequate text on the agent side and allows people to point at objects, make movements and select options. All these elements build a rich communication protocol, people use the same methods the children use when they are not able to speak. Gesture communication is not as effective as the verbal communication but it is very natural. The advances in artificial vision, and in several ways of gesture tracking can be interpreted by computers inside a specific context.

The first stage is done by using patterns as marks in the real world to align real and virtual objects. The system receives the users actions using a virtual cursor, a wand that seems to be in the users hand. The wand serves as a selector to pick objects avoiding the hand to interfere with the pattern recognition. Handling this cursor requires some practice to those users who have not used a mouse before. But the training does not take long. A virtual agent has restricted movements and expressions compared with the human possibilities, which make the use of verbal and written mediums necessary at agent side. Other didactical elements such as colors or symbols are helpful when emphasizing critical stages in the tasks. Mixing those ways of communication and taking advantage of peoples ability to get in touch using all them, it is possible to build less expensive systems that are nevertheless useful for training people in simple tasks, such as handling of electronic controls, machinery assembly and printer repair.

5.1 System Configuration

The central hardware platform consist of a Pentium IV at 2.3 GHz. In order to get the real world images, a Creative WebCam is used. A Polhemus electromagnetic tracker is used to track the users hand position. The images captured by the camera are combined in the CPU with the virtual objects, the augmented scene is drawn in the monitor or the glasses and the mixed result can be seen on both types of displays see figure 2.

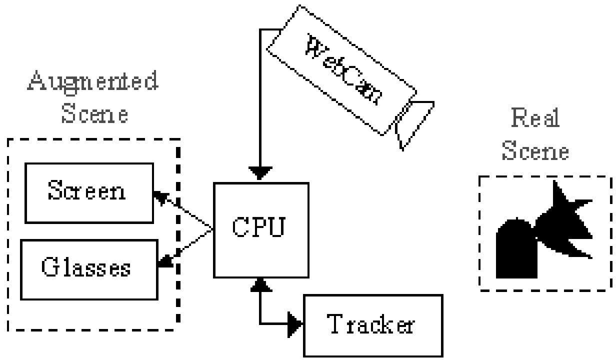

Figure 2. System Configuration

5.2 Software Architecture

The System architecture is shown in figure 3; ARToolkit[9] is used for optical tracking. ARToolkit is a freely distributed software from the HIT lab at the University of Washington. It searches the environment looking for patterns, those patterns become the origin of the coordinate system in the mixed image as result of the complete process. The interface with the magnetic tracker was modified code downloaded from internet[10]. The virtual objects are 3D models loaded to OpenGL, the loader of these objects is available at [11].

The agent is conformed by three modules:

1. Sensors: A set of functions to receive events from the mouse, the polhemus or the environment.
2. Interpreters: Those are the functions used to decide the action took by the agent, depending on the information available in the sensors and the agent states.
3. Executors: Functions used to perform the actions dictated by the interpreters.

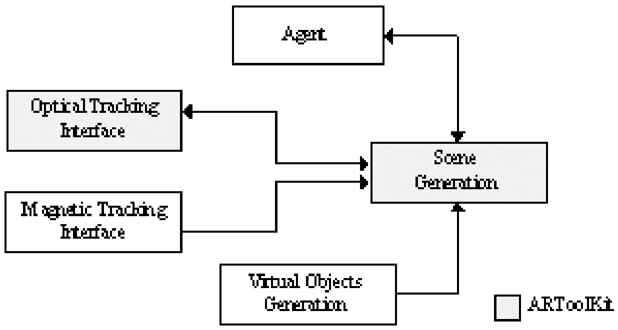

Figure 3. Software Architecture

6 Tests and Results

6.1 Tests

To test the interface we use a task which consists in matching couples of related objects. A cube is used as patterns container; it has a different pattern at each face. A virtual object is drawn over each side. An example of the task follows: a switch related with a bulb are shown together with other objects aside the cube, the bulb must be moved until it intersects the switch, thereby describing that they are a pair. When a matching couple is found, the corresponding objects disappear and the other objects must be matched.

The test consists in two phases: the initial face is a demonstration performed by the agent, as described in figure 4.

In the second phase the user performs the task, under the supervision of the agent.

6.2 Results

In order to evaluate the effectiveness of the system, an experiment will be performed. A control group will receive a written description of the task. The experimental group will be tutored by the agent. Times to Complete the Task will be collected and processed statistically in order to determine if there is a substantial difference. We plan to conduct the experiments in the next days, and the results

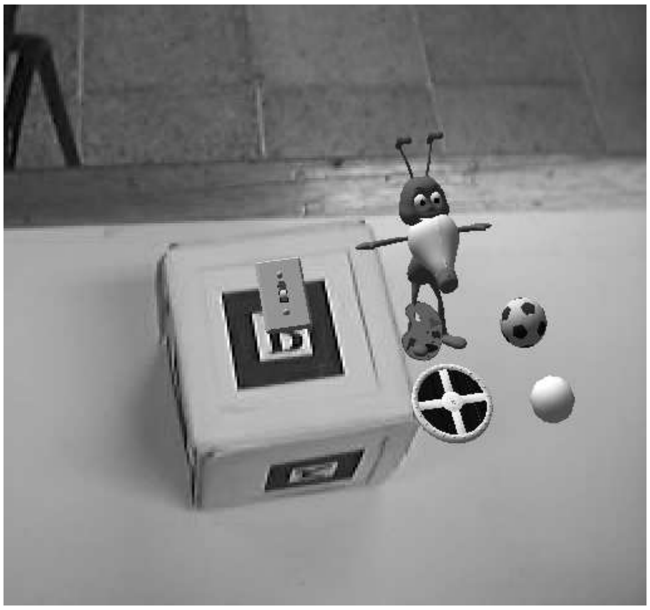

Figure 4. Agent Demonstration

will be ready for the conference, if the paper is accepted.

7 Conclusions and Future Work

If the results are as expected, and the Times to Complete the Task are actually smaller for the experimental group, we can conclude that the proposed means of interaction are an effective means of tutoring. If that is the case, we plan to use the system in an actual training environment in an assembly plant. We also plan to incorporate other modalities of interaction (voice synthesis/recognition, glove-gesture recognition, etc), and experiment on the effectiveness of several combinations of modalities.

References

[1] J. Rickel and W. Johnson. Task-oriented collaboration with embodied agents in virtual worlds. In *Embodied Conversational Agents, Cassell, J., J. Sullivan, S. Prevost, et al.* Eds. Boston: MIT Press, 2000.

[2] R. Azuma, Y. Baillot, R. Behringer, S. Feiner, J. Blair, S. Blair, and M. Blair. Recent advances in augmented reality. *IEEE Computer Graphics and Applications*, 25(6), Nov-Dec 2001.

[3] P. Milgram and F. Kishino. A taxonomy of mixed reality visual displays. *IEICE Transactions on Information Systems*, Vol E77-D(12):1321 – 1329, December 1994.

[4] W. L. Johnson and J. W. Rickel. Animated pedagogical agents: Face-to-face interaction in interactive learning environments. *Multimedia Tools and Applications*, 11:47 – 78, 2000.

[5] B.A. Stone and J.C. Lester. Dynamically sequencing an animated pedagogical agent. In *Proc. of AAAI-96*, volume 1, pages 424 – 431, August Portland, Oregon, 1996.

[6] J. Lester, J. Voerman, S. Towns, and C. Callaway. Cosmo: A life-like animated pedagogical agent with deictic believability. In *Proc. of the IJCAI97*, 1997.

[7] Hans J. W. Spoelder, Luc Renambot, Desmond Germans, Henri E. Bal, and Frans C. A. Groen. Man multi-agent interaction in VR: A case study with robocup. In *VR*, page 291, 2000.

[8] S. Balcisoy and D. Thalmann. Interaction between real and virtual humans in augmented reality. In *Tenth Computer Animation Conference*, pages 31 – 38. IEEE, June Geneva, Switzerland, 1997.

[9] H. Kato and M. Billinghurst. Marker tracking and hmd calibration for a video-based augmented reality conferencing system. In *Proc. of the IEEE ACM IWAR '99*, pages 85 – 94, October San Francisco, California, 1999.

[10] S. Konno. Polhemus driver package. http://www.cybergarage.org.

[11] R. A. Reierstad. 3ds loader. http://apron.morrowland.com.