Multimodal Interfaces for Educational Virtual Environments

Dimitrios Christopoulos, Athanasios Gaitatzes

Abstract— **Educational applications often are slow to leverage and use new interaction devices in order to bring new value and allow new forms of gameplay. Following decades of research on how to use 3D simulation and Virtual Environments in education, attention has recently turned to exploring Multi-User-Virtual-Environments for the educational community. In the following paper we present the results of a pilot simulation battle, created for educational purposes combining the positive aspects of multi-user virtual environments, edutainment VR applications and new Human Computer Interaction (HCI) interfaces. We present the technology used, as well as an evaluation case study of the human-computer interaction results.**

Index Terms—**Virtual Reality, Natural Interfaces, Educational Applications, Multi User Environments**

I. Introduction

The project presented is a recreation of an existing historical event for educational purposes; its concept is based on the familiarity of children and adolescents with computer games. Our focus was to study the feasibility of learning in a non-traditional way by creating an educational experience for class-sized audiences using computer games technology and modern human computer interfaces. The actual educational information passed, gives students the opportunity to learn about the ancient city of Syracuse, its prominent position in the ancient world and the historical context of the conflict between Rome and Carthage. Some details about Archimedes and his position in the global intellect are also explained. Initial evaluation results concerning the usability and the perceived knowledge of classes participating in the program are being presented.

II. Motivation

There are several well-known Multi-User-Virtual-Environments (MUVE) like Active Worlds, Barbie Girls, Gaia Online, Second Life, The Sims Online, but hardly any of them can be considered educational even-though they do provide the functionality of users adding educational features. However, recently attention has turned to exploring their use to support learning, and several research groups have been creating MUVEs and investigating their effectiveness. Designed for the educational community, they embed tasks or problems within a virtual environment or context. Users can explore the environment and examine digital objects. Typically, there is also a means to communicate with other users and online agents. Users may select an avatar to represent them in the environment, thus providing the opportunity to try out a persona in a non-threatening environment [1].

Research suggests that educational MUVEs should not solely focus on the virtual environment; support from the teacher and time for self-reflection are also important. So students also participate in teacher-led discussions that encourage reflection and sharing of ideas. Such teacher guidance is an important part of the learning process when using MUVEs [2].

Using high technology 3D simulations for educational purposes is not a new concept; in the past two decades immersive virtual reality (VR) has attracted the attention of many researchers and educators [3][4]. However, widespread uptake isn't yet apparent mainly due to the high dependency on specialized hardware and failure to build engaging high profile software for multiple users accessible to the public.

The amount of new innovative human computer interaction devices has increased significantly through the past years. From the keyboard and mouse we have moved to joystick devices, camera tracking devices, accelerometers, touch screens, dance pads, speech recognition and devices that can position the users head and hand in a virtual environment with six degrees of freedom. With the advent of the Wii controller the devices became more intuitive and open to the public in an easy to use context, but the games also helped since they were designed based on specific controller devices [5]. Existing educational applications often fail to leverage and use these new interaction devices to bring new value to educational applications by allowing new forms of gameplay. They are often treated as gadgets or gimmicks and therefore discarded.

Our aim was to combine the positive aspects of multi-user virtual environments, edutainment VR applications and new Human Computer Interaction (HCI) interfaces to study the feasibility of their usage in education. Augmenting

Manuscript received March 27, 2009. This work has been developed within the research project 'e-nterface", that has been co-funded by the General Secretariat of Research and Development of Greece, under the 3rd Community Framework of support.

D. Christopoulos is with the Foundation of the Hellenic World, Poulopoulou 38, 11851 Athens, Greece. Phone +30 212 254 8196, Fax +30 212 254 3838, E-mail: christop@fhw.gr.

A. Gaitatzes is with the Foundation of the Hellenic World, Poulopoulou 38, 11851 Athens, Greece. E-mail: gaitat@fhw.gr.

educational software using new technological advances is essential to its progress and accessibility [6][7].

III. Goal and Description of the Project

The project represents a virtual battlefield recreation of the siege by sea, of the ancient city of Syracuse by the Romans in 212 BC. It allows the collaboration of two users, which assume the roles of the historical generals Hippocrates from Syracuse and Marcellus from Rome. Both users have to control war machinery to either capture or defend the city walls. The user playing the roman general Marcelus controls 3D representations of roman siege ships in order to capture the city walls and conquer the city (Figure 1). Its opponent playing the Syracusian general Hippocrates has to prevent the capture of the city walls using the famous war machinery inventions of Archimedes (Figure 2). The educational purpose of this multiplayer set-up was the transmission of historical information, presentation of the battle through the great scientific inventions of Archimedes and information about its life. The whole experience is targeted towards users between 9-15 years of age and uses established video game techniques and metaphors like 3D graphics, cinematic story line, high action based interaction and dramatic closure. The design principle of the application allows the learning of historical information, through a pleasant experience, without having the conscious perception of participation in a formal learning procedure. Human computer interaction (HCI) models were of primary importance since by allowing natural forms of interaction, a high degree of immersion and suspension of disbelieve can be achieved. Gesture based, optical marker based, speech based and classic joystick control models were implemented and evaluated.

Albeit educational, such simulations often present a greater challenge in contrast to normal computer games. Not only has the rendered environment, the buildings, the weapons and siege ships to be realistic and appealing as in modern games but also historically correct. The buildings and city walls were created according to historical references, the roman siege ships were equipped and used according to the warfare [8] of that time period and the war machinery of Archimedes was recreated from historical blueprints. Four war machines of Archimedes [9] were reconstructed, the Catapult, which was used to toss large, rocks over great distances, the scorpion used to throw large arrows with precision, the famous mirrors which gathered the sun beams to one exact spot and were used to heat up and burn ships and the grab, a formidable gigantic claw mounted on the city walls, which was used to grab underlying ships heap them out of the water and sink them. Realistic rendering of people with lip-syncing and animation, physically based modeling of the environment, special effects, multiplayer support, presentation on a variety of VR setups in stereo are also a small list of the features implemented in order to realize this simulation.

Figure 1. Screenshot from the Battle Simulation. Roman Siege ships are attacking the wall. Defensive weapons are already mounted on the wall.

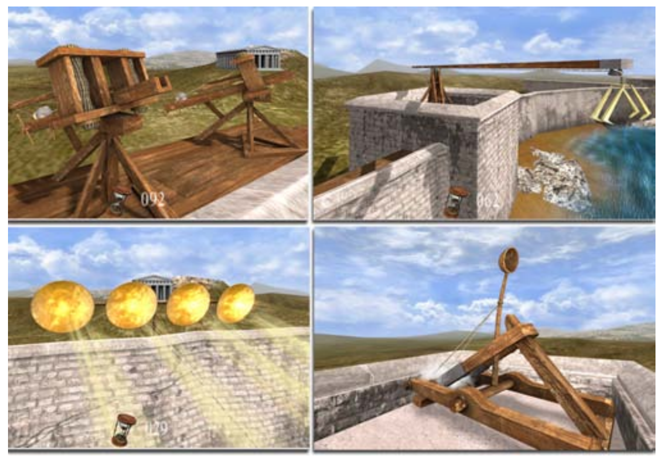

Figure 2. The inventions of Archimedes. Top left the Scorpions, top right the Claw, bottom left the Mirrors and bottom right the Catapult.

IV. Architecture of the Multi-User Environment

In any multiplayer game there are three problems that need to be addressed: timing, identification and determinism.

First you need to be able to synchronize time to be able to identify when things happen in the game and when messages arrive or are being sent. Second, identification is about being able to direct the correct data to the correct area in the game. There must be a way to uniquely identify each machine and have all other connected machines aware of that identification and "whom they are talking to". There is also a need for in-game identification to make sure data/messages are being sent to/from the right objects. So each object is issued a unique (but synchronized) identifier across all machines. Thirdly, determinism is the predictability of the game. Only necessary data is sent across the network, as to ensure matching game states across all instances of the game. So a very deterministic game will only require small amounts of information (ie: the exact choices the user made) whilst a completely non-deterministic game will require all the info that describes a state to be sent, sometimes as often as at every time step.

The virtual reality software used for implementing the

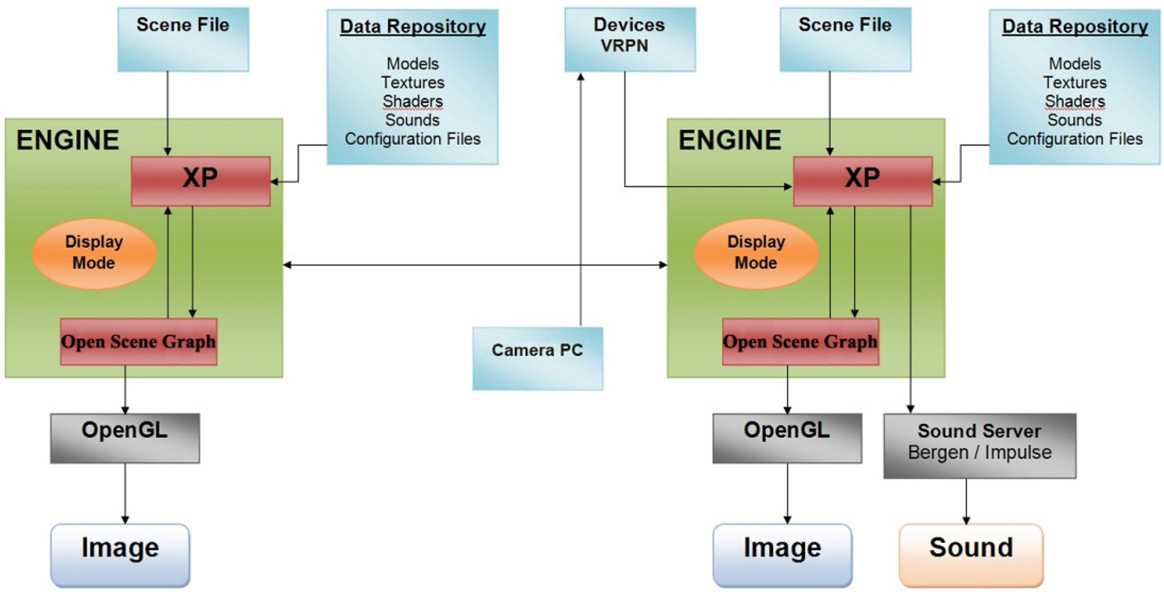

Figure 3. Diagram of the architecture of the multiuser environment.

multiplayer architecture was a high level framework that executed script files that described the virtual world and the user interactions. It used different libraries to control different aspects of the virtual environment like the user input, the synchronization, the audio reproduction and the visualization of the three-dimensional graphics. The framework was developed in individual elements on top of the OpenSceneGraph library [10], which was mainly used for its high performance graphics characteristics. Figure 3 shows the architecture of the whole environments that was developed using the Linux OS.

A custom protocol was developed and designed for the synchronization of multiple clusters / units consisting of a central unit (Master) which could synchronize multiple sub-units (slaves) in order to have consistent representation of the virtual world from multiple viewpoints on multiple screens. The communication between the units was via a 1Gbps LAN network, and the synchronization was achieved through special synchronization packets sent to all units by the central unit. These packets contained information on the position of the observer, information from various points of entry such as keyboard / mouse or a special control device (joystick), but also the local time of the central unit. Each secondary unit expected to receive a package from the central unit to move the simulated virtual world forward by one frame. When the secondary units received their package, they acknowledged it moving also one step forward in the simulation. Since all units had the same deterministic virtual world loaded and shared the position of the observer, the information from the various entry points and the local time of the central unit (ie everyone had the same core clock), it was guaranteed that all machines would run on the same frame.

For the visualization of the game in stereo and non-stereo modes, in display setups but also in large stereo screens we developed a library that could handle arbitrary display surfaces and viewing modes [11]. This way it became more versatile in its demonstration possibilities. Designed for multiple channel setups in monoscopic, passive stereo, active stereo or left/right individual eye operation. It was suitable for configuring an application to run on a CAVE-like environment, a Reality Center, a domed theatre, a Power wall, or any other display configuration, as well as on single screen desktop VR systems and HMDs. It was easily configurable through a simple yet effective XML script, which allowed multiple configurations to be present in a single file and share some common features if necessary.

V. Interfacing with External Capturing Devices

As described previously an open HCI interface framework for accepting input from various interaction devices was a primary design decision. The open source VRPN [12] protocol was used as a base layer upon which a specific server was implemented to provide the desired input abstraction. The Virtual-Reality Peripheral Network (VRPN) system provides a device-independent and network-transparent interface to peripherals. The idea is to have a PC or other host at each VR station that controls the peripherals (tracker, button device, haptic device, analog inputs, sound, etc). VRPN provides connections between the application and all of the devices using the appropriate class-of-service for each type of device sharing this link. The application remains unaware of the network topology. Note that it is possible to use devices that are directly connected to the machine that the application is running on, either using separate control programs or running all as a single program.

The protocol was enhanced with a server object class, which would act as an interface between the various input devices and the final application. The application was designed to be controlled by a game pad, which meant reading the x/y axis state of the joystick for controlling the game objects and button information for starting actions. The server could accept input from a vast array of input devices like stereo cameras for optical motion tracking, optical marker systems, voice recognition systems, traditional joystick interfaces or even mouse keyboard input and map that data to the values needed in order to emulate a joy pad. This architecture kept the implementation of the multiplayer game separate from of how to collect and interpret the input. To add a new device only the server object had to be changed to perform the mapping operation into joy pad data, this ability proved very important on solving various HCI issues in the game itself and supporting variety of interaction modes. Furthermore the ability to offload the input processing to a separate machine and communicate over TCP/IP was essential to achieve interactive frame rates when camera based input and image based processing was used [13].

VI. Human Computer Interaction

Emerging interface devices must be used intelligently; someone can't just "throw" devices at an existing application, but rather must change the original interaction mode or insert the new interface in a natural place where it enhances value. The primary concern is for the users to feel immersed in a virtual world and not as if they are using a device.

The application had to be easy and simple to learn and use and at the same time have a multitude of communication metaphors implemented for the interaction between player and machine. The users should easily understand the task to be performed and it should be easy and natural to execute those tasks using the specific input devices. The software had to handle all user inputs gracefully even the wrong ones and give feedback to the users about their mistake. The graphics visualization had to respond to those inputs with no delay or the user would loose his/hers immersion from the environment.

Since the target audience was between the ages of 9-15, it was very important to capture the attention on critical moments when historical or educational information was presented, thus avoiding the distraction, which usually happens when Virtual Reality experiences are shown to children or teenagers [14][15]. The excess energy and the anticipation in audiences of these ages categories often distracts them from paying attention and stay focused on a particular subject. A twofold approach was used to keep the attention of the audience. First the avatar of Archimedes appears and introduces the audience to what is going to happen and why he needs the audiences help, thus creating a first step to immersion by preparing the visitor and providing the basis an emotional bond [16]. Once this is achieved the avatar Archimedes is used to pass educational information during the actual game. Secondly with the pretension of traveling back in time to the city of Syracuse, the audience is taken on a dazzling roller coaster ride helping to release in a controlled manner the excitement before entering the stage

Figure 4. The animated avatar of Archimedes talking about the siege of Syracuse.

where Archimedes is met again presenting the historical information (Figure 4). The interactive part where the audience is experiencing the information just communicated by Archimedes (his inventions, the usage of war machines and siege ships during the Syracuse siege by the Romans) is introduced as a multiplayer game where two participants compete against each other controlling the siege machines and ships. After the end of the game sessions, which are considered the climax and most entertaining moments of the application, the ending sequence is presented which again communicates historical information about the actual outcome of the battle and the fate of Archimedes.

The game was tested with various input implementations. Initially using two wireless joypad devices, which required usage of its joystick and buttons to navigate and initiate actions. This traditional HCI interface mode, required the users to push the joystick into the direction the ship should travel in order to avoid the war machines and reach the city walls or change the direction and target of the war machines. The buttons initiate the war machine action attacking whichever ship was targeted or when commanding the siege ships the firing of arrows towards the city wall and the lowering of the tower bridge in order to board the wall. The second interaction mode implemented was using two optical motion tracking systems which allowed the users to control the objects by stepping left, right, forward and back and initiate actions by raising their hands or giving voice commands. This interface replaced the interaction device with the body of the user proving to be the easiest to learn and use. Finally along the lines of the last interaction mode a camera marker tracking system was implemented where each user was holding a marker pattern. Moving the marker in front of a web cam controlled the selected weapons. Actions were triggered either by voice commands or wireless joypads buttons. This interface replaced only the joystick part with natural body movements through a defined marker object, in combination with a speech recognition system it also provided often an even easier to user interface since the body movements were more limited.

VII. Evaluation Results

Our evaluation took place at the facilities of the Foundation of Hellenic World. Two passive stereo powerwalls for multiplayer play were setup side by side. The test groups were classes of age 9-15 from schools visiting the center (Figure 5).

It was observed that the avatar Archimedes could successful keep the attention of the audience taking over the role of the teacher or guide. The roller coaster ride was also exceptionally well taken, introducing the exciting and energy releasing element needed to stay focused.

During the game most of the users were already used to video games and had no difficulties playing with the joypad. After an initial adaptation period also the optical interactions systems successfully were used successfully, users with less or no gaming experience managed to adapt better to these interaction methods and play the game with more experienced games on par.

In the case study we conducted the users found it was easy to learn the tasks at hand and more than 60% said that the interaction methods involving optical devices were appropriate for the given task. The system responded very well, when wrong commands were given and the users were able to correct their mistakes. The virtual environment and the game-play were designed in such a way that they were engaging to the users. The users had a sense of control on the game-play and more than 60% responded positively when answered if this specific method was more natural in attaining the historical information. Finally most users were satisfied with the overall performance of the system and the amount of historical information presented.

VIII. Conclusion

Realism and the game play scenario in computer games certainly play an important role. Today's games have certainly mastered the visualization part of the games. The method though with which the users interact with the system is another important aspect of the experience. With so many input devices to choose from, aspects like natural devices or even intuitive devices come into play. But we have to overcome pre-attained knowledge. Lots will argue that a mouse is an intuitive device for user input but is it a natural device? Certainly users today, having played many computer games with mouses and joysticks will feel that these devises are familiar and therefore natural. But what is more natural than personal movements in front of a camera to play a game. Can you really control the bat of a baseball game with a joystick or is it more natural to swing a bat in front of a camera and having that action transferred in the game. We still have a long way to go in designing intuitive interfaces and that is what will distinguish the educational applications of tomorrow.

Acknowledgment

The external component and software for the optical motion tracking systems which has been used in our tests has been developed by the researchers of the Center for Research and Technology Hellas / Informatics and Telematics Institute (ITI) within the project "e-nterface".

References

Figure 5. Evaluation of the application

[1] E. Dieterle, J. Clarke. (in press). *Multi-user virtual environments for teaching and learning.* In M. Pagani (Ed.), Encyclopedia of multimedia technology and networking (2nd ed). Hershey, PA: Idea Group

[2] R. Francis, (2006). *Towards a theory of a games based pedagogy*. Paper presented at the JISC Innovating e-Learning 2006, online.

[3] C. D. Wickens, *Virtual reality and education. System, Man and Cybernetics*, (1) 842-847, Chicago, IL, USA, 1992

[4] Gaitatzes, A.G., Christopoulos D., Roussou M., "Virtual Reality Interfaces for the Broad Public," in the Proceedings of Human Computer Interaction 2001, Panhellenic Conference with International Participation, Patras, Greece, 7-9 December 2001.

[5] S. Sreedharan, E. S. Zurita, B. Plimmer. *3D input for 3D words.* Proceedings of the 19th Australasian conference on Computer-Human Interaction: Entertaining User Interfaces. Adelaide, Australia 2007. pp. 227-230.

[6] C. P. Lim, D. Nonis, J. Hedberg, (2006). *Gaming in a 3D multiuser virtual environment: Engaging students in Science lessons*. *British Journal of Educational Technology*, *37*(2), 211-231.

[7] D. J. Ketelhut, C. Dede, J. Clarke, B. Nelson (2006). *A multi-user virtual environment for building higher order inquiery skills in science.* Paper presented at the Annual Conference of the American Educational Research Association, San Francisco, CA

[8] Duncan B. Campbell, Adam Hook. *Siege Warfare in the Roman World*: 146 BC-AD 378. Osprey Publishing, 2005

[9] B. C. Hacker. Greek Catapults and Catapult Technology*: Science, Technology, and War in the Ancient World*, Technology and Culture, Vol. 9, No. 1 (Jan., 1968), pp. 34-50

[10] http://www.openscenegraph.org/projects/osg (last visited 25 March 2008)

[11] A. Gaitatzes, G. Papaioannou, D. Christopoulos, G. Zyba. *Media Productions for a Dome Display System*, Proc. ACM Symposium on Virtual Reality Software and Technology (VRST '06), pp. 261-264, 2006.

[12] R. M. Taylor II, T. C. Hudson, A. Seeger, H. Weber, J. Juliano, A.T. Helser, "*VRPN: A Device-Independent, Network-Transparent VR Peripheral System*," Proceedings of the ACM Symposium on Virtual Reality Software & Technology 2001, VRST 2001. Banff Centre, Canada, November 15-17, 2001.

[13] S. Pettifer, et al. DEVA3: *Architecture for a Large Scale Virtual Reality System*. in Proc. ACM Symposium in Virtual Reality Software and Technology 2000 (VRST'00). 2000. Seoul, Korea.

[14] M. Virvou, G. Katsionis : On the usability and likeability of virtual reality games for education: *The case of VR-ENGAGE in Computers & Education* Volume 50, Issue 1, January 2008, Pages 154-178

[15] S. Mumtaz, (2001). *Children's enjoyment and perception of computer use in the home and the school.* Computers and Education, 36, 347-362.

[16] K. Wojciechowski, B. Smolka, H. Palus, R.S. Kozera, W. Skarbek and L. Noakes : *EMPATHIC AVATARS IN VRML FOR CULTURAL HERITAGE* in Computer Vision and Graphics International Conference, ICCVG 2004, Warsaw, Poland, September 2004, Proceedings