



Faculty of Informatics  
Masaryk University  
Czech Republic

---

# **Habilitation Thesis**

**Interactive Virtual and Augmented Reality  
Environments**

**Fotis Liarokapis, Ph.D**

**March 2015**

# Preface

Interactive computer graphics applications have gained a lot of attention over the past decade. In this broad field the two major technologies, virtual and augmented reality are interfering with consumer's life in a number of ways. Virtual reality has already become dominant in certain applications such as movies and video games where augmented reality has now started developing more robust applications.

The main goal of this thesis is to provide an overview of my most significant achievements in the areas of interactive virtual and augmented reality environments. My achievements are subdivided into four areas, which are: (a) procedural modelling, (b) virtual and augmented reality interfaces, (c) interactive environments, and (d) application domains. This work covers a complete set of methods and techniques from content generation to visualisation and interaction, and finally to the application into different domains.

This thesis is written as a commentary to a collection of 10 peer-reviewed journal papers and 6 peer-reviewed conference papers. My percentage contribution for each paper is estimated and included in the thesis as well as a brief description of my work. My personal contribution to the papers ranges between 10% to 100% with an average of approximately 40%.

# Acknowledgements

Firstly, I would like to thank all of my colleagues and collaborators who contributed to the papers that are provided in this thesis. I would also like to thank all of my colleagues at the Human Computer Interaction laboratory for their support. Special thanks to Petr Matula and Michal Kozubek for their inspiration. Finally, my greatest thanks go to my family and my girlfriend for their support and patience throughout the whole process of this thesis.

Parts of the work presented in this thesis have been supported by the EU IST Framework V programme, Key Action III-Multimedia Content and Tools, Augmented Representation of Cultural Objects (ARCO) project IST-2000-28336 and by the EPSRC Pinpoint Faraday project GR/T04212/01, called LOCUS.

# Table of Contents

Chapter 1 Introduction .....	1
1.1 Introduction.....	1
1.2 Motivation .....	1
1.3 Background .....	2
1.3.1 Virtual Reality.....	2
1.3.2 Augmented Reality .....	2
1.3.3 Procedural Modelling.....	3
1.3.4 Crowd Modelling.....	3
1.3.5 Serious Games.....	3
1.3.6 Human Computer Interaction.....	4
1.4 Goal and Overview .....	4
Chapter 2 Procedural Modelling .....	5
2.1 Introduction.....	5
2.2 Terrain Environments.....	5
2.3 Buildings and Cities .....	7
2.4 Behaviour of Crowd Simulation .....	8
Chapter 3 Virtual and Augmented Reality Interfaces.....	10
3.1 Introduction.....	10
3.2 Virtual Reality Interfaces .....	10
3.2.1 Indoor VR Interfaces.....	10
3.2.2 Mobile VR Interfaces .....	12
3.3 Augmented Reality Interfaces .....	13
3.3.1 Indoor AR Interfaces.....	13
3.3.2 Mobile AR Interfaces .....	15
Chapter 4 Interactive Environments .....	17
4.1 Introduction.....	17
4.2 Multimodal Interaction .....	17
4.3 Wireless Sensor Network Based Interaction .....	18
4.4 Brain-Computer Interaction .....	20
Chapter 5 Application Domains .....	22
5.1 Introduction.....	22
5.2 Virtual Archaeology .....	22
5.3 Urban Navigation .....	24
5.4 Higher Education.....	26
5.4.1 VR and AR in Education.....	26
5.4.2 Activity-Led Introduction to First Year Creative Computing .....	27
5.5 Serious Games and Virtual Environments .....	29
5.5.1 Serious Games Technologies.....	29
5.5.2 Learning as Immersive Experiences within Serious Games.....	29
Chapter 6 Conclusions and Future Work .....	31
6.1 Conclusions .....	31
6.2 Future Work .....	31
Chapter 7 References.....	32
Chapter 8 Appendix – Paper Reprints .....	38
8.1 Paper #1.....	39
8.2 Paper #2.....	48
8.3 Paper #3.....	53
8.4 Paper #4.....	62
8.5 Paper #5.....	67
8.6 Paper #6.....	72
8.7 Paper #7.....	86

8.8	Paper #8.....	108
8.9	Paper #9.....	125
8.10	Paper #10.....	138
8.11	Paper #11.....	145
8.12	Paper #12.....	155
8.13	Paper #13.....	165
8.14	Paper #14.....	174
8.15	Paper #15.....	190
8.16	Paper #16.....	212

# List of Figures

Figure 2-1 Procedural terrain [40].....	6
Figure 2-2 (a) Roman Settlement, (b) Vitruvian Temple Comparions .....	7
Figure 2-3 The urban crowd simulation displaying crowds of agents (a) No graphical complexity [44] (b) High realistic scenes.....	9
Figure 3-1 Indoor online VR Interfaces .....	11
Figure 3-2 Mobile VR Interfaces (a) Manual mode (b) GPS mode (c) VR view .....	12
Figure 3-3 Indoor AR Interfaces [54].....	14
Figure 3-4 Mobile AR Interfaces .....	15
Figure 4-1 Multimodal augmented reality interface [48] .....	18
Figure 4-2 Wireless Sensor Network Based Interaction.....	19
Figure 4-3 Brain-computer interaction [47].....	20
Figure 5-1 Archaeology (a) Complete solution and (b) Artefact visualisation in AR.....	23
Figure 5-2 Interfaces for presenting information retrieved from a mobile information system.....	25
Figure 5-3 Operation of the AR application (a) AR environment (b) Visualisation of educational content [59].....	26
Figure 5-4 3D etch-a-sketch. (a) Student-based drawing application [60], (b) student group's hardware interface [60].....	28
Figure 5-5 Learning as immersive experiences. (a) Four Dimensional Framework [63], (b) Meeting in-world in Second Life for virtual tour [62].....	30

# Abbreviations

2D	Two Dimensional
3D	Three Dimensional
ALL	Activity Lead Learning
AR	Augmented Reality
BCI	Brain Computer Interface
DOF	Degrees of Freedom
EEG	Electroencephalography
GPS	Global Positioning System
GUI	Graphical User Interface
HCI	Human Computer Interaction
HMD	Head Mounted Display
UMPC	Ultra Mobile Personal Computer
VE	Virtual Environment
VR	Virtual Reality
WSN	Wireless Sensor Network

# Chapter 1

---

## Introduction

### **1.1 Introduction**

The presented habilitation thesis consists of a collection of 16 publications, 10 peer-reviewed journal papers and 6 peer-reviewed conference papers. The introduction chapter presents the motivation of this thesis, followed by a brief background of the research areas covered, and then the goal and overview of this work. The next four chapters provide a summary of the main research contributions. Finally, the last chapter presents conclusions and future work.

### **1.2 Motivation**

Interactive virtual and augmented reality environments are becoming more and more appealing to a wider audience. The creation of realistic virtual and augmented reality environments is an important issue in computer animation, computer games, digital film effects, and simulation industries. In recent years, the computer and video games industry has overtaken both the film and music industries. The top revenue producers and the cost for developing commercial interactive applications nowadays usually require investments of several million dollars. This typically involves large teams of developers recruiting hundreds of workers, many of whom are artists and designers providing content for the decoration of rich virtual and augmented reality environments. While many creative companies have the necessary budget to develop these expensive interactive environments (i.e. movies, games, etc), which employ state of the art computer graphics and their applications, not all companies have the same resources. In addition, hardware improvements allow for better and faster tracking and visualisation devices that can be used for creating novel applications.



### **1.3 Background**

This section provides a brief overview of the main technologies included in the thesis such as: *virtual reality*, *augmented reality*, *procedural modelling*, *crowd modelling*, *serious games* and *human-computer interaction*.

#### **1.3.1 Virtual Reality**

The first virtual reality (VR) environment was originally introduced in the 1960s by Ivan Sutherland [1]. Since then there have been published many studies [2], [3], [4], [5]. The main characteristic of a VR system is that the user's natural sensory information is completely replaced with digital information. The user's experience of a computer-simulated environment is called immersion. VR systems can completely immerse a user inside a synthetic environment by blocking all the signals of the real world. The most common problems of VR systems are of emotional and psychological nature including motion sickness, nausea, and other symptoms which are created by the users' high degree of immersiveness [6]. VR systems are also sometimes called virtual environments (VEs); however, typically the term is referred to online virtual world applications. Nowadays more than 100 VEs exist and they provide excellent capabilities for creating effective distance and online learning opportunities through the provision of unique support for distributed groups (online chat, the use of avatars, document sharing etc.) [2].

#### **1.3.2 Augmented Reality**

The basic concept of augmented reality (AR) is to superimpose digital information directly upon a user's sensory perception [7], rather than replacing it with a synthetic environment as VR systems do. Both technologies process and display the same digital information and often make use of the same dedicated hardware but AR systems use more complex software approaches compared to VR systems [8]. In technical terms, it is not a single technology but a collection of different technologies that operate in conjunction, with the aim of enhancing the user's perception of the real world through computer-generated information [9]. This kind of information is usually referred to as virtual, digital, or synthetic information. The real world must be matched with the virtual in position and context in order to provide an understandable and meaningful view [10]. Users can work individually or collectively, experiment with computer-generated information and interact with a mixed environment in a natural way [11]. In the coming years, AR systems will be able to include a complete set of augmentation applied exploiting all people's senses [12]. Finally, a recent survey of AR describes some known limitations regarding human factors that developers need to overcome with [13].

### **1.3.3 Procedural Modelling**

A number of survey papers have been recently published in the areas of terrains [14], cities [15], and virtual worlds [16]. Procedural modelling can be considered as a set of formal production rules that specify how geometric shapes are created and transformed. Procedural modelling is mainly used to generate content for a number of aspects of the real environment including: terrains, buildings, cities, road structures, trees and vegetation. Muller and Parish [17] proposed a city generation approach that made use of self-sensitive L-systems to automatically lay out a set of streets and generate virtual architecture. Greuter et al. described a set of methods that allowed for the procedural generation of a ‘pseudo-infinite’ digital environment [18]. Wonka et al. [19] devised a variation on shape grammars for use in the construction of building facades, which they named split grammars. More recently, shape grammars have been extended through the use of context-sensitive shape rules [20].

### **1.3.4 Crowd Modelling**

The process of simulating huge crowds of intelligent agents in real-time is still a challenging task due to numerous different issues [21], [22]. The real-time simulation of crowds can be conducted using a variety of approaches. The most common methods involve employing a series of models and algorithms working in tandem with animate each agent. These include decision-making [23], pathfinding navigation [24], local steering mechanics [25] and agent perception systems [26]. Social forces models [27] can also be utilised to enhance crowd believability under certain situations. However, some form of quantification is required to assess the behaviour of agents within crowd simulations and past research has utilised perception as a tool for evaluating crowds [28], [29]. Realism is the degree of plausibility of the crowd behaviour whereas perceived realism is centred on the perception of humans.

### **1.3.5 Serious Games**

Serious games are part of a new emerging field that focuses on computer games that are designed for non-leisure but often, educational purposes. They have important applications in several distinct areas such as: military, health, government, and education [30]. Serious games have the capability of enabling learners to undertake tasks and experience situations that would have otherwise been impossible. The success of serious computer games in educational scenarios is based on the combination of audiovisual media that is prevalent in these games, which enhances the absorption of information in the learner's memory [31], [32]. Although the state-of-the-art in serious games technology is identical to the state-of-the-art in computer games, both types share the same technical infrastructure [2]. Moreover, there are two diverse

views on how serious games should be designed. One argues that while pedagogy is an implicit component of a serious game, it should be secondary to entertainment, meaning that a serious game that is not ‘fun’ to play with would be useless, regardless of its pedagogical content or value [33]. On the other hand, design methodologies exist for the development of games incorporating pedagogic elements, such as the four dimensional framework [34], which outlines the centrality of four elements that can be used as design and evaluation criteria for the creation of serious games. As a result, this approach focuses mainly on educational and pedagogy theories.

### **1.3.6 Human Computer Interaction**

Human-computer interaction (HCI) is the study of the interaction between humans and computer systems [35]. As a result, it is one of the most important issues when designing interactive environments [36][37]. The design and implementation of software user interfaces that will produce robust interfaces is interrelated with the use of HCI techniques. The integration of such interfaces into AR/VR systems can reduce the complexity of the HCI by using implicit contextual input information [38]. Nevertheless, the design and implementation of effective VR and AR environments is a difficult task and an area of continuous research. Nowadays, most common HCI rely on different types of sensors providing user-friendly applications. Typical techniques include: acoustic, mechanical, optical, electromagnetic, inertial, global positioning system (GPS), and electroencephalography (EEG). Multimodal systems combine natural input modes (i.e. speech, pen, touch, hand gestures, eye gaze, head and body movements) in a coordinated manner with multimedia system output [39].

## **1.4 Goal and Overview**

The goal of this thesis is to illustrate the most significant results in the area of interactive virtual and augmented reality environments. The main results are summarised in the next four chapters. Each chapter provides a brief overview of the contributions incorporating my contribution as well.

## Chapter 2

---

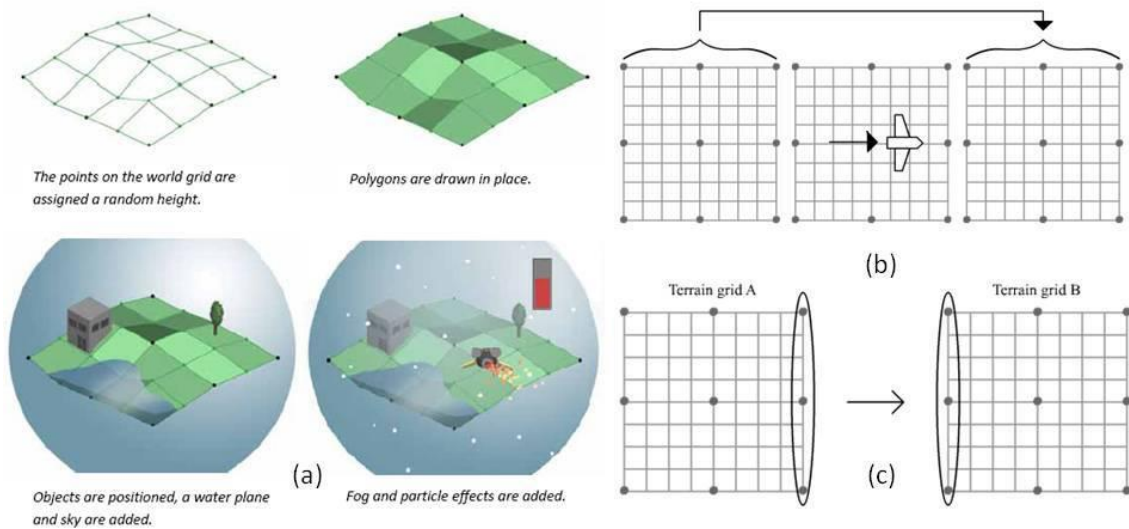
# Procedural Modelling

### **2.1 Introduction**

This chapter presents a set of techniques used for creating content for the computer graphics community as well as for research purposes. This includes the creative industry, VR and AR interactive applications. The focus of this chapter is on procedural modelling techniques for: (a) terrains, (b) buildings and cities and (c) behaviour of crowd simulation.

### **2.2 Terrain Environments**

A variety of methods for automatically creating detailed but also randomised terrain environments have been developed. The use of these procedural methods saves time and reduces the budget for creating effective computer graphics applications (i.e. games, VR, AR etc). This work explained some of the problems that can arise from this situation and described a variety of methods that can be used to overcome them [40]. These methods have been applied to a basic flight simulator, so that the results could be observed and evaluated. Figure 2-1 (a) illustrates an overview of the randomly generated environment for the proposed flight simulator which can be used for developing games and serious games. Heightmaps were generated using the diamond-square algorithm to provide surface detail.



**Figure 2-1 Procedural terrain [40]**

Based on a recursive algorithm, the level of detail can be adjusted as necessary, which can be an advantage when dealing with different methods that require different levels of processing power. To smooth the terrain, a two-dimensional (2D) Lorentz distribution as well as a Gaussian filter allowing for bell shapes were used. Next, the terrain was generated in a way so that it gives the illusion that it is infinite. Upon reaching the right side of the landscape, a tile of terrain is moved in front of the player to provide the illusion of endless terrain (Figure 2-1 b). Upon reaching the right side of terrain grid A, the values of the far-right points are copied to the far left points of terrain grid B (Figure 2-1 c). The other points of terrain grid B are then calculated via randomisation and mid-point displacement, as done for grid A. Moreover, a simplified method was implemented that made use of randomised positioning of vegetation models, rather than procedurally creating vegetation. Evaluation with two different types of user groups (remote and hallway) showed that overall the flight simulator is enjoyable, looks realistic for a gaming scenario and thus has the potential to be used for the development of serious games [40].

**Paper:** Noghani, J., Liarokapis, F., Anderson, E.F. Randomly Generated 3D Environments for Serious Games, Proc. of the 2<sup>nd</sup> IEEE International Conference in Games and Virtual Worlds for Serious Applications, IEEE Computer Society, Braga, Portugal, 25-26 March, 3-10, 2010.

**Contribution (40%):** Design of the architecture, implementation of smoothing techniques and advice on evaluation. Write-up of most of the paper (full text on section 8.1).

## 2.3 Buildings and Cities

This work proposed the development of a novel shape grammar [19] inspired by ‘CGA Shape’ [20] for describing Roman settlements derived from the writings of Vitruvius (Figure 2-2), initially with a focus on the description of classical Roman temples, meaning the main building of a religious site, excluding its courtyard [41]. Moreover, the technique was extended in generating complete Roman settlements. The construction of Roman temples included a large number of common elements found in Roman architecture, e.g. palaces shared many of these and often also incorporated temples themselves [42]. Structures generated from these Vitruvian rules can provide an exemplar of archetypal Roman architecture in a similar manner as the “Virtual Egyptian Temple” by Jacobson and Holden [43] which depicts architecture in ancient Egypt. Different approaches were taken for the various elements of the generated city. A weighted formula was designed for the purpose of citing a city location upon a heightmap, incorporating factors like the distance to the nearest body of water and the gradient of the land. Three methods of situating generic structures within a city were considered, including a probability distribution method that assigned buildings to districted allotments with a flexible degree of randomness.



**Figure 2-2 (a) Roman Settlement, (b) Vitruvian Temple Comparisons**

For the generation of the rest of the city, a novel formal grammar syntax was devised, capable of describing shapes in a deterministic and technical fashion. The grammar made use of superscripts preceding symbols for notating conditional rules, and superscripts and subscripts

following symbols for the purpose of adding attributes to existing symbols. In this way, architecture was described using grammar rules in a way that would be impractical or outright impossible through the use of traditional grammar syntax. The dominant feature of a temple is its main building, which in most Roman settlements would be built within the courtyard and the enclosure wall. While the overall makeup of this usually followed the same pattern – the ‘cella’ (the temple building’s enclosed room), fronted or surrounded by a portico and raised on top of a podium – there is considerable architectural variation possible in Roman temple construction. Temple buildings were built on a podium with steps only at the front or with steps on all four sides, with the number of steps in both cases being an odd number. The temples’ proportions would be such that the length of the main temple buildings would be twice the width of the temple with the length of the temple’s cella being 25% larger than the overall width of the temple [41].

**Paper:** Noghani, J., Anderson, E., Liarokapis, F. Towards a Vitruvian Shape Grammar for Procedurally Generating Classical Roman Architecture, Proc. of the 13<sup>th</sup> International Symposium on Virtual Reality, Archaeology and Cultural Heritage VAST 2012, Short and Project Papers, Eurographics, Brighton, UK, 19-21 November, 41-44, 2012.

**Contribution (30%):** Design of the architecture and advice on the implementation. Collaboration on the writing of the paper (full text on section 8.2).

## 2.4 Behaviour of Crowd Simulation

This work examined (a) the development of intelligent crowd simulation in virtual environments, and (b) a perceptual experiment to identify features of behaviour, which can be linked to perceived realism [44]. The urban crowd simulation developed as part of this research implements a range of real-time simulation techniques [23], [24], [25], [26], [27]. To carry out the psychophysical experimentation a platform was developed in the form of the urban crowd simulation. The results of this research can feedback into the development processes of simulating inhabited locations, by identifying the key features, in order to achieve more perceptually realistic crowd behaviour. Perceptual experimentation methodologies can be adapted and potentially utilised to test other types of crowd simulation, for application within computer games, or more specific simulations such as for urban planning or health and safety purposes. In the initial stage of the research, the perceived realism of agent crowd behaviour is evaluated through the features that shape behaviour traits. For example, for the velocity type, the behavioural annotation and so on, the graphical complexity is not essential to the core of the research (Figure 2-3, a).



**Figure 2-3 The urban crowd simulation displaying crowds of agents (a) No graphical complexity [44] (b) High realistic scenes**

Data is collected from experiments in the form of a perceived realism value between ‘0’ (completely unrealistic) and ‘1’ (completely realistic) [44]. Initial results with 32 participants completed the social forces experiment, using an online survey platform. The experiment consists of two key variables, one for each of the agent-based social forces. These variables were tested at specific trials, such as: (a) agent avoidance and (b) agent attraction. The majority of participants (94%) found that when the agent avoidance social force is present the behaviour of the agents is more realistic, and (95%) selected the videos with the agent attraction social force present to be more realistic. Results showed that the majority of the participants found the simulation with social forces to be more realistic than a simulation without. In the next stage of this research, the realism of the environment is also included (Figure 2-3, b) and another study will examine if there is a correlation with the previous results.

**Paper:** O'Connor, S., Liarokapis, F., Peters, C. An Initial Study to Assess the Perceived Realism of Agent Crowd Behaviour in a Virtual City, Proc. of the 5<sup>th</sup> International Conference on Games and Virtual Worlds for Serious Applications (VS-Games 2013), IEEE Computer Society, Bournemouth, UK, 11-13 September, 85-92, 2013.

**Contribution (30%):** Collaboration on the design of the architecture and advice on the experimental part. Collaboration on the writing of the paper (full text on section 8.3).



## Chapter 3

---

# Virtual and Augmented Reality Interfaces

### **3.1 Introduction**

This chapter demonstrates novel solutions developed in the area of virtual and augmented reality for both indoor and outdoor environments. In particular, a number of novel virtual and augmented reality interfaces are presented illustrating how these technologies can be used effectively for both types of environments.

### **3.2 Virtual Reality Interfaces**

A number of novel VR interfaces have been developed for both indoor and outdoor environments that can be categorised as: (a) indoor interfaces and (b) mobile interfaces.

#### **3.2.1 Indoor VR Interfaces**

VR systems vary from laboratory custom-made systems [45], to modern gaming environments (which rely on the functionality of commercial game engines [46], [47]) as well as on online virtual environments. The focus of this work was on the presentation of realistic graphics in an interactive online VR environment [49]. Online VR interfaces allow multiple users to access the content in an easy and convenient manner from remote locations. The most significant work performed here, was the integration of multimodal visualisation VR environments (which are connected to a database) to allow users to switch between web and VR views in real-time performance.

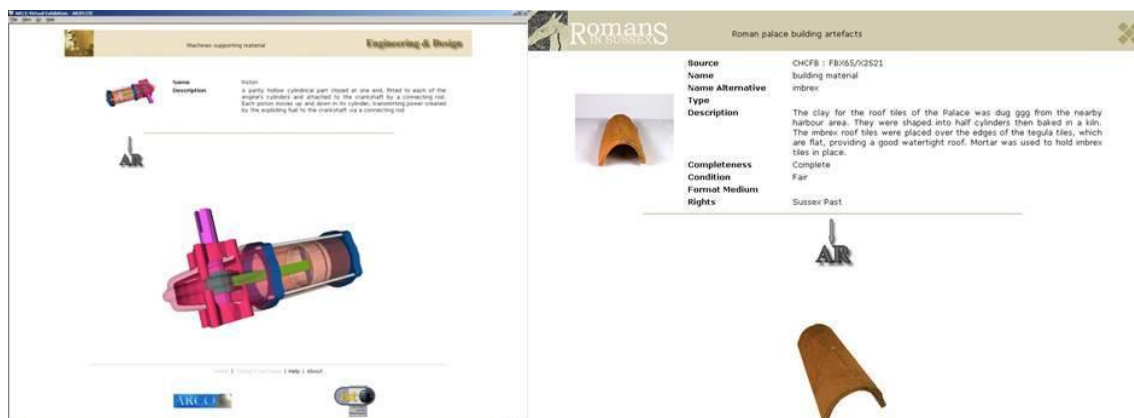


Figure 3-1 Indoor online VR Interfaces

Metadata is also associated with the virtual information and presented appropriately (Figure 3-1). In the Web-based interface a user can browse information presented in a form of 3D VRML virtual galleries or 2D Web pages with embedded multimedia objects. Virtual exhibitions can also be visualized in the Web browser in a form of 3D galleries [55]. In this visualization, users can browse objects simply by walking along the 3D environment (i.e. a reconstruction of a real gallery). Different interaction devices were also integrated to the system allowing users to manipulate 3D content in a more appealing manner.

**Paper:** Liarokapis, F., Mourkoussis, N., White, M., Darcy, J., Sifniotis, M., Petridis, P., Basu, A., Lister, P.F. Web3D and Augmented Reality to support Engineering Education, World Transactions on Engineering and Technology Education, UICEE, 3(1): 11-14, 2004.

**Contribution (80%):** Collaboration on the design of the architecture. Implementation of the most of the VR interface. Write-up of most of the paper (full text on section 8.4).

**Paper:** White, M., Mourkoussis, N., Darcy, J., Petridis, P., Liarokapis, F., Lister, P.F., Walczak, K., Wojciechowski, R., Cellary, W., Chmielewski, J., Stawniak, M., Wiza, W., Patel, M., Stevenson, J., Manley, J., Giorgini, F., Sayd, P., Gaspard, F. ARCO-An Architecture for Digitization, Management and Presentation of Virtual Exhibitions, Proc. of the 22<sup>nd</sup> International Conference on Computer Graphics (CGI'2004), IEEE Computer Society, Hersonissos, Crete, June 16-19, 622-625, 2004.

**Contribution (15%):** Collaboration on the design of the VR and AR architecture. Implementation of the most of the VR and AR interface. Write-up of parts of the paper (full text on section 8.5).

### 3.2.2 Mobile VR Interfaces

In this work, visualisation within the mobile virtual environment (the spatial 3D map) can take place in two modes: automatic and manual. In the automatic mode, a GPS automatically feeds and updates the spatial 3D map with respect to the user's position in the real space (Figure 3-2, b). This mode is designed for intuitive navigation. In the manual mode, the control is fully with the user, and it was designed to provide alternative ways of navigating into areas where we cannot obtain a GPS signal (Figure 3-2, a). Users might also want to stop and observe parts of the environment in which case control is left in their hands (Figure 3-2, c).



**Figure 3-2 Mobile VR Interfaces (a) Manual mode (b) GPS mode (c) VR view**

The immersion provided by GPS navigation is considered as pseudo-egocentric because fundamentally the camera is positioned at a height, which does not represent a realistic scenario. If, however, the user switches to manual navigation, any perspective can be obtained,

which is very helpful for decision-making purposes. While in a manual mode any model can be explored and analysed, therefore additional enhancements of the graphical representation are of vital importance [50].

**Paper:** Liarokapis, F., Brujic-Okretic, V., Papakonstantinou, S. Exploring Urban Environments using Virtual and Augmented Reality, *Journal of Virtual Reality and Broadcasting*, GRAPP 2006 Special Issue, Digital Peer Publishing, 3(5): 1-13, 2006.

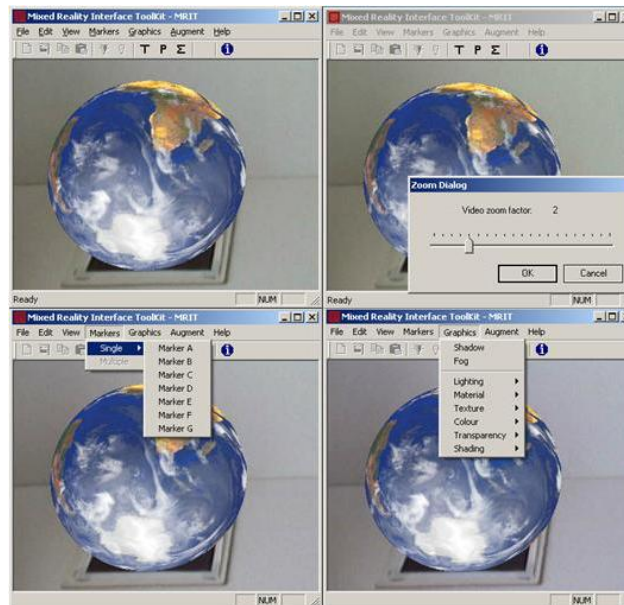
**Contribution (70%):** Collaboration on the design of the architecture. Implementation of the majority of the VR interface. Write-up of most of the paper (full text on section 8.6).

### 3.3 Augmented Reality Interfaces

Results from the previous sub-sections were used as an input for the implementation of the interactive AR interfaces and again can be categorised as: (a) indoor interfaces and (b) mobile interfaces.

#### 3.3.1 Indoor AR Interfaces

Human computer interaction techniques can offer greater autonomy when compared with traditional windows style interfaces. Although some research has been performed into the integration of such interfaces into AR systems [51], [52], [53] the design and implementation of an effective AR system that can deliver realistically audio-visual information in a user-friendly manner is a difficult task and an area of continuous research. However, it is very difficult to create experiences to eliminate these barriers [53] preventing even nowadays the creation of new AR applications. To address the above issues, a number of prototype AR interfaces were proposed and implemented.



**Figure 3-3 Indoor AR Interfaces [54]**

A prototype high-level AR architecture using a selection of cost effective software and hardware components to realise robust visualisation and interaction of virtual information for indoor environments was developed. The software libraries are based on the integration of computer vision, computer graphics and auditory techniques resulting in three prototype video-see through AR architectures and eventually to a general purpose AR interface. The main novelty of the interface is that they are capable of superimposing simultaneously digital information such as metadata, 2D images, 3D models, spatial sound and videos [54]. Spatial sound was simulated based on a linear approximation of distance to give the impression of 3D space. The greatest advantage of the proposed AR interface is that it allows participants to perform complex operations very accurately (Figure 3-3). Specifically, sometimes it is of crucial importance to superimpose objects in specific locations in the real environment. Using other methods it could take a great amount of time and effort (depending on the experience of the user) to achieve this and it will definitely not be very accurate. The graphical user interface (GUI) interaction techniques offer the solution to this issue using double point precision accuracy. Finally, it allows users to transfer data from the internet into a tabletop AR environment [55].

**Paper:** Liarokapis, F. An Augmented Reality Interface for Visualizing and Interacting with Virtual Content, *Virtual Reality*, Springer, 11(1): 23-43, 2007.

**Contribution (100%):** Design of the architecture and implementation of the AR interface. Write-up of the paper (full text on section 8.7).

**Paper:** White, M., Mourkoussis, N., Darcy, J., Petridis, P., Liarokapis, F., Lister, P.F., Walczak, K., Wojciechowski, R., Cellary, W., Chmielewski, J., Stawniak, M., Wiza, W., Patel, M., Stevenson, J., Manley, J., Giorgini, F., Sayd, P., Gaspard, F. ARCO-An Architecture for Digitization, Management and Presentation of Virtual Exhibitions, Proc. of the 22<sup>nd</sup> International Conference on Computer Graphics (CGI'2004), IEEE Computer Society, Hersonissos, Crete, June 16-19, 622-625, 2004.

**Contribution (15%):** Collaboration on the design of the VR and AR architecture. Implementation of the most of the VR and AR interface. Write-up of parts of the paper (full text on section 8.5).

### 3.3.2 Mobile AR Interfaces

The two most common tracking techniques used in AR applications include computer vision and external sensor systems. In this work both approaches were investigated but since the requirement was to have an AR system that could be operational anywhere and everywhere, the sensor approach was preferred. A GPS receiver and digital compass can provide sufficient accuracy for displaying points of interest in the approximate location relative to the user's position. At present, however, these sensor solutions lack the accuracy required for more advanced AR functionality, such as aligning an alternative facade on the front of a building in the real world scene. There is no need for a head-mounted display (HMD), since the screen on the device can be aligned with the real world scene. On the screen of the device, information can either be overlaid on imagery captured from the device's internal camera, or the screen can display just the virtual information with the user viewing the real world scene directly.



Figure 3-4 Mobile AR Interfaces

For the computer vision approach, road signs which are most of the time represented in black color on a white background, were used as an initial approach. Later on, road signs were replaced and distinctive natural features like door entrances, windows etc, have been experimentally tested to see whether they can be used as 'natural markers' (Figure 3-4). For the sensor approach a similar approach to section 3.2.2 was adopted. The main challenge however was to reduce the latency produced by the sensors (GPS and digital compass) as well as provide a textual based augmentation. The AR interface can then provide navigational information, in the form of distance and direction annotations, to guide the user to the location associated with those results [56].

**Paper:** Liarokapis, F., Brujic-Okretic, V., Papakonstantinou, S. Exploring Urban Environments using Virtual and Augmented Reality, *Journal of Virtual Reality and Broadcasting*, GRAPP 2006 Special Issue, Digital Peer Publishing, 3(5): 1-13, 2006.

**Contribution (70%):** Collaboration on the design of the architecture. Implementation of the majority of the VR interface. Write-up of most of the paper (full text on section 8.6).

**Paper:** Mountain, D., Liarokapis, F. Mixed reality (MR) interfaces for mobile information systems, *Aslib Proceedings*, Special issue: UK library & information schools, Emerald Press, 59(4/5): 422-436, 2007.

**Contribution (50%):** Collaboration on the design of the architecture and implementation of the VR interface. Write-up of half of the paper (full text on section 8.8).

## Chapter 4

---

# Interactive Environments

### **4.1 Introduction**

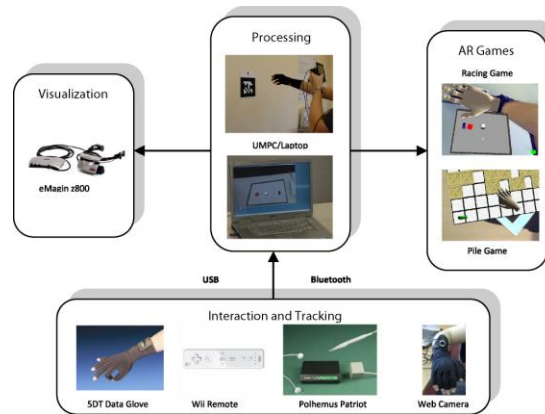
HCI is an important aspect of any computer system and this section is focused on illustrating different novel paradigms that were developed. Both VR and AR users can make use of more sophisticated hardware devices to perceive and interact with the environment. These can be categorized in three different areas including: multimodal, wireless sensor networks, and brain-computer interactions.

### **4.2 Multimodal Interaction**

In this section, tangible AR gaming environments that can be used to enhance entertainment using a multimodal interface were explored [48]. The main objective of the research was to design and implement generic tangible interfaces that are user-friendly in terms of interaction and can be used by a wide range of players, including the elderly or people with disabilities.

To allow for seamless interaction between the users and the superimposed environmental information, a number of custom interaction devices have been researched. In particular, six different types of interaction were implemented including: hand position and orientation, pinch glove interaction, head orientation, Wii interaction, and ultra mobile personal computer (UMPC) I/O manipulation. However, since usability and mobility were crucial, only a few interaction devices were finally integrated to the final architecture. In the final configuration players can interact using different combinations between a pinch glove, a Wiimote, a six degrees-of-freedom (DOF) tracker, through tangible ways as well as through I/O controls. An overview of the system is shown in Figure 4-1.





**Figure 4-1 Multimodal augmented reality interface [48]**

Two tabletop AR games have been designed and implemented including a racing game and a pile game. The goal of the AR racing game was to start the car and move around the track without colliding with either the wall or the objects that exist in the gaming arena. Initial evaluation results showed that multimodal-based interaction games can be beneficial in gaming. Based on these results, an AR pile game was implemented with the goal of completing a circuit of pipes (from a starting point to an end point on a grid). Initial evaluation showed that tangible interaction is preferred to keyboard interaction and that tangible games are much more enjoyable. From the research proposed many potential gaming applications could be produced such as strategy, puzzles and action games.

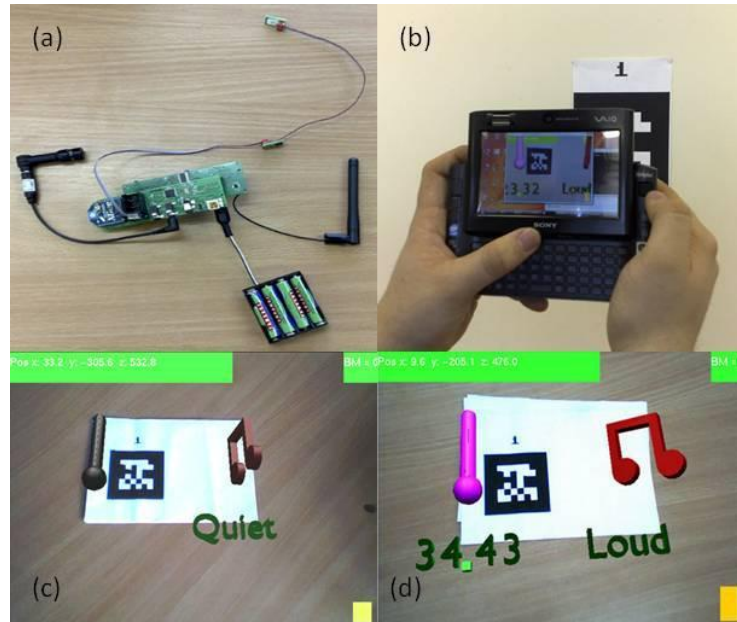
**Paper:** Liarokapis, F., Macan, L., Malone, G., Rebolledo-Mendez, G., de Freitas, S. Multimodal Augmented Reality Tangible Gaming, *Journal of Visual Computer*, Springer, 25(12): 1109-1120, 2009.

**Contribution (30%):** Contribution on the design of the architecture. Implementation of parts of the AR interface. Write-up of most of the paper (full text on section 0).

### 4.3 Wireless Sensor Network Based Interaction

Wireless Sensor Network (WSN) technology uses networks of sense enabled miniature computing devices to gather information about the world around them. While the gathering of data within a sensor network is one challenge, another of equal importance is presenting the data in a useful way to the user. A prototype mobile AR system for visualising environmental information including temperature and sound data was proposed [57]. Sound and temperature data are transmitted wirelessly to the client (which is a handheld device). Environmental

information is represented graphically, as 3D objects and textual information, in real-time performance (Figure 4-2). Participants visualise and interact with the augmented environmental information using a small but powerful handheld computer. The main contribution of this work is the visual representation of wireless sensor data in a meaningful and tangible way.



**Figure 4-2 Wireless Sensor Network Based Interaction**

In terms of operation, as soon as the temperature and sound sensors are ready to transmit data, visual representations including a 3D thermometer and a 3D music note as well as textual annotations are superimposed onto the appropriate marker. When environmental data is transferred to the AR interface, the colour of the 3D thermometer and the 3D music note change according to the temperature level and sound volume accordingly. Textual annotations indicate the sensor readings. For the temperature data, the readings from the sensors are superimposed as text next to the 3D thermometer. For the sound data, a different measure was employed based on a scale 0-4, where '0' corresponds to 'quiet', '1' corresponds to 'low', '2' corresponds to 'medium', '3' corresponds to 'loud' and '4' corresponds to 'very loud'.

**Paper:** Goldsmith, D., Liarakapis, F., Malone, G., Kemp, J. Augmented Reality Environmental Monitoring Using Wireless Sensor Networks, Proc. of the 12<sup>th</sup> International Conference on Information Visualisation (IV08), IEEE Computer Society, 8-11 July, 539-544, 2008.

**Contribution (30%):** Collaboration on the design of the architecture. Advice on the implementation of the majority of the VR interface. Write-up of most of the paper (full text on section 8.10).

#### 4.4 Brain-Computer Interaction

Non-invasive BCIs operate by recording the brain activity from the scalp with EEG sensors attached to the head on an electrode cap or headset without being surgically implanted. However, they still have a number of problems since they cannot function as accurately as other natural user interfaces and traditional input devices such as the standard keyboard and mouse. The current research done examined the application of commercial and non-invasive EEG-based brain-computer (BCIs) interfaces with serious games [47].



**Figure 4-3 Brain-computer interaction [47]**

Two different EEG-based BCI devices were used to fully control the same serious game (Figure 4-3). The first device (NeuroSky MindSet) uses only a single dry electrode and requires no calibration. The second device (Emotiv EPOC) uses 14 wet sensors requiring an additional training of a classifier. User testing was performed on both devices with sixty-two participants measuring the player experience as well as key aspects of serious games, primarily learnability, satisfaction, performance and effort. Recorded feedback indicates that the current state of BCIs can be used in the future as alternative game interfaces following familiarisation and in some cases calibration. Comparative analysis showed significant differences between the two

devices. The first device provides more satisfaction to the players whereas the second device is more effective in terms of adaptation and interaction with the serious game.

**Paper:** Liarokapis, F., Debattista, K., Vourvopoulos, A., Ene, A., Petridis, P. Comparing interaction techniques for serious games through brain-computer interfaces: A user perception evaluation study, *Entertainment Computing, Elsevier*, 5(4): 391-399, 2014.

**Contribution (40%):** Collaboration on the design of the architecture. Advice on the implementation of the serious game as well as in the BCI interface. Write-up of most of the paper (full text on section 8.11).

## Chapter 5

---

# Application Domains

### **5.1 Introduction**

This section presents how the above mentioned research can be applied for creating different applications such as archaeology, navigation, education, and serious games.

### **5.2 Virtual Archaeology**

A number of museums hold large archives or collections of artefacts, which they cannot exhibit in a low cost and efficient way. Another underlying issue is that museums simply do not have the space to exhibit all the artefacts in an educational and learning manner. Museums are interested in the digitising of their collections not only for the sake of preserving the cultural heritage, but also to make the information content accessible to the wider public in a manner that is attractive. Emerging technologies, such as VR, AR and Web3D are widely used to create virtual museum exhibitions both in a museum environment through informative kiosks and on the World Wide Web. This work surveyed the field, and while it explored the various kinds of virtual museums in existence, it discusses the advantages and limitations involved with a presentation of old and new methods and of the tools used for their creation.

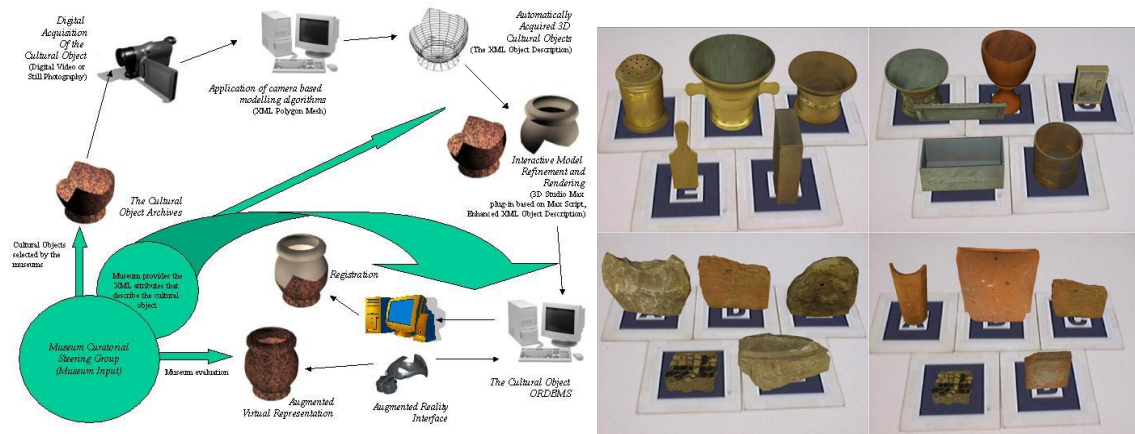


Figure 5-1 Archaeology (a) Complete solution and (b) Artefact visualisation in AR

The work also provided a complete tool chain starting with the stereo photogrammetry based digitization of artefacts, their refinement, collection and management with other multimedia data, and visualization using virtual and augmented reality (Figure 5-1, a). The generated system is a one-stop-solution for museums to create, manage and present both content and context for virtual exhibitions (Figure 5-1, b). Interoperability and standards are also key features of our system allowing both small and large museums to build a bespoke system suited to their needs [55]. Moreover, different multimodal interfaces have been developed for cultural heritage. The integration of these technologies provides a novel multimodal mixed reality interface that facilitates the implementation of more interesting digital heritage exhibitions. With such exhibitions, participants can switch dynamically between virtual web-based environments to indoor augmented reality environments as well as make use of various multimodal interaction techniques to better explore different applications such as virtual museums.

**Paper:** Sylaiou, S, Liarokapis, F., Kotsakis, K., Patias, P. Virtual museums, a survey and some issues for consideration, *Journal of Cultural Heritage*, Elsevier, 10(4): 520-528, 2009.

**Contribution (30%):** Collaboration on the collection of the material and write-up of the paper (full text on section 8.12).

**Paper:** Liarokapis, F. An Augmented Reality Interface for Visualizing and Interacting with Virtual Content, *Virtual Reality*, Springer, 11(1): 23-43, 2007.

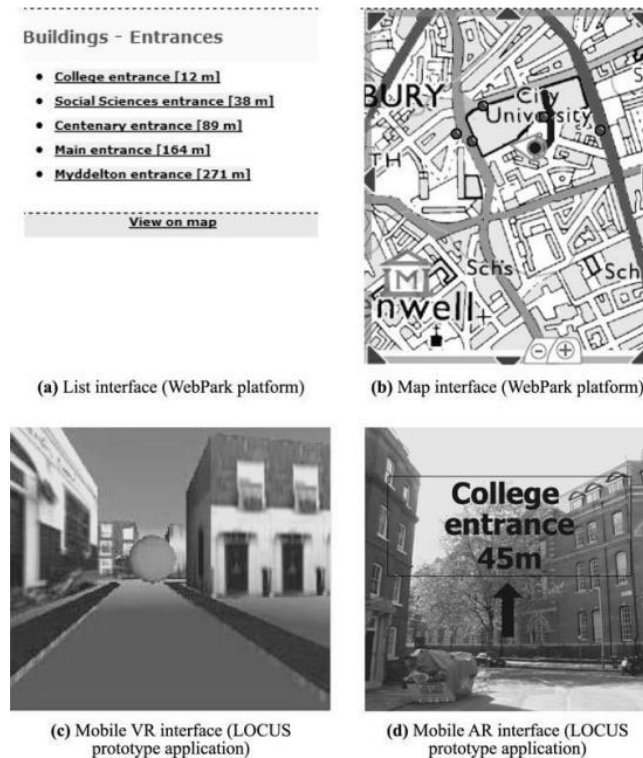
**Contribution (100%):** Design of the architecture and implementation of the AR interface. Write-up of the paper (full text on section 8.7).

**Paper:** White, M., Mourkoussis, N., Darcy, J., Petridis, P., Liarokapis, F., Lister, P.F., Walczak, K., Wojciechowski, R., Cellary, W., Chmielewski, J., Stawniak, M., Wiza, W., Patel, M., Stevenson, J., Manley, J., Giorgini, F., Sayd, P., Gaspard, F. ARCO-An Architecture for Digitization, Management and Presentation of Virtual Exhibitions, Proc. of the 22<sup>nd</sup> International Conference on Computer Graphics (CGI2004), IEEE Computer Society, Hersonissos, Crete, June 16-19, 622-625, 2004.

**Contribution (15%):** Collaboration on the design of the VR and AR architecture. Implementation of the most of the VR and AR interface. Write-up of parts of the paper (full text on section 8.5).

### 5.3 Urban Navigation

Up to now most attempts to develop pedestrian navigation tools for the urban environment have used GPS technologies to display position on two-dimensional digital maps (as in the classic 'satnav' systems on the market). Although GPS is the key technology for location-based services (LBS), it cannot currently meet all the requirements for navigation in urban environments. Specifically, GPS technologies suffer from multipath signal degradation and they cannot provide orientation information at low or zero speed, which is an essential component of navigation. It has also been demonstrated that maps are not always the most effective interfaces to pedestrian navigation applications on mobile devices. Orientation information is necessary to help the user self-localise in an unknown environment and can be provided by either sensors (i.e. accelerometers or digital compass) or through computer vision techniques.



**Figure 5-2 Interfaces for presenting information retrieved from a mobile information system**

The LOCUS project has developed alternative, mixed reality interfaces for existing mobile information system technology based upon the WebPark platform. The WebPark platform can assist users in formulating spatially referenced, mobile queries. The retrieved set of spatially referenced results can then be displayed using various alternative interfaces: a list, a map, VR or AR (Figure 5-2). An evaluation exercise was undertaken to assess appropriate levels of detail, realism and interaction for the mobile virtual reality interface. Virtual 3D scenes were found to have many advantages when compared to paper maps: the most positive feature was found to be the possibility to recognize the features in the surrounding environment, which provides a link between the real and virtual worlds. Overall, results showed that these technologies are helpful however, the most suitable interface is likely to vary according to the user and task in hand [56].

**Paper:** Liarokapis, F., Brujic-Okretic, V., Papakonstantinou, S. Exploring Urban Environments using Virtual and Augmented Reality, *Journal of Virtual Reality and Broadcasting*, GRAPP 2006 Special Issue, Digital Peer Publishing, 3(5): 1-13, 2006.

**Contribution (70%):** Collaboration on the design of the architecture. Implementation of the majority of the VR interface. Write-up of most of the paper (full text on section 8.6).



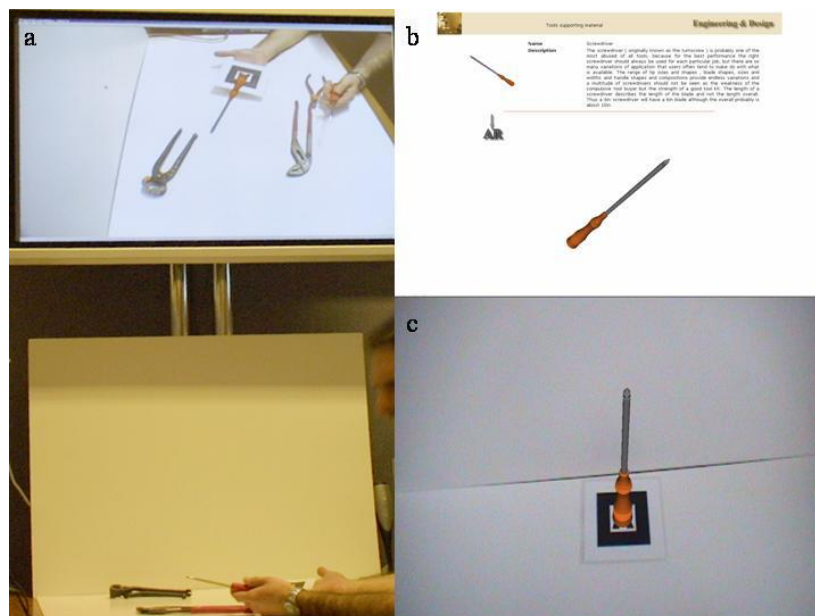
**Paper:** Mountain, D., Liarokapis, F. Mixed reality (MR) interfaces for mobile information systems, Aslib Proceedings, Special issue: UK library & information schools, Emerald Press, 59(4/5): 422-436, 2007.

**Contribution (50%):** Collaboration on the design of the architecture and implementation of the VR interface. Write-up of half of the paper (full text on section 8.8).

## 5.4 Higher Education

### 5.4.1 VR and AR in Education

Although current teaching methods work successfully, Universities are interested in introducing more productive methods for improving the learning experience and increasing the level of understanding of the students. The emergence of new technological innovations such as the Internet, multimedia, virtual and augmented reality technologies, was able to demonstrate the weaknesses of traditional teaching methods but also the potential of improving them. This work focuses on the use of high-level AR interfaces for the construction of collaborative educational applications that can be used in practice to enhance current teaching methods (Figure 5-3).



**Figure 5-3 Operation of the AR application (a) AR environment (b) Visualisation of educational content [59]**

A combination of multimedia information including spatial 3D models, images, textual information, video, animations and sound, can be superimposed in a student-friendly manner into the learning environment. In several case studies, different learning scenarios have been carefully designed based on HCI principles so that meaningful virtual information is presented in an interactive and compelling way. Collaboration between the participants is achieved through use of a tangible AR interface that uses marker cards as well as an immersive AR environment which is based on GUIs and sensors devices. The interactive AR interface has been piloted in the classroom of two UK universities in the Departments of Informatics and Information Science. Initial results indicated that students appreciated this type of tool for assisting the lecturer and improving the learning process [59].

**Paper:** Liarokapis, F., Mourkoussis, N., White, M., Darcy, J., Sifniotis, M., Petridis, P., Basu, A., Lister, P.F. Web3D and Augmented Reality to support Engineering Education, *World Transactions on Engineering and Technology Education*, UICEE, 3(1): 11-14, 2004.

**Contribution (80%):** Collaboration on the design of the architecture. Implementation of the most of the VR interface. Write-up of most of the paper (full text on section 8.4).

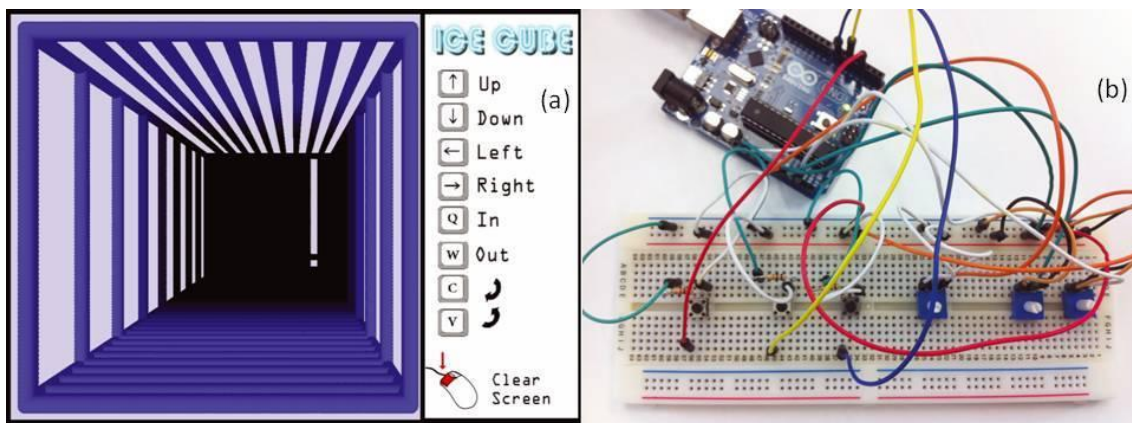
**Paper:** Liarokapis, F., Anderson, E. Using Augmented Reality as a Medium to Assist Teaching in Higher Education, *Proc. of the 31<sup>st</sup> Annual Conference of the European Association for Computer Graphics (Eurographics 2010)*, Education Program, Norrkoping, Sweden, 4-7 May, 9-16, 2010.

**Contribution (90%):** Implementation of the AR interface and collection of all the experimental data. Write-up of most of the paper (full text on section 8.13).

### 5.4.2 Activity-Led Introduction to First Year Creative Computing

One of the goals of higher education is to prepare students for life by enabling them to become independent learners. Independent learning does not come easy to students who have adapted to becoming passive participants in the learning process, where they are presented with all of the required learning material, a learning style that many of them acquired during their secondary education [60]. Activity Lead Learning (ALL) is focused on providing students with a specific problem, scenario, task or activity in order to motivate, engage and stimulate them for providing effective and efficient solutions. The range of activities and tasks has a wide range and according to the requirements, different activities have to be planned and disseminated. ALL is a student-centred approach that has its roots in problem-based learning (PBL) [61].

Misconceptions about the nature of the computing disciplines pose a serious problem to university faculties that offer computing degrees, as students enrolling on their programmes may come to realise that their expectations are not realistic. This frequently results in the students' early disengagement from the subject of their degrees, which in turn can lead to excessive 'wastage', i.e. reduced retention. This work, reports on our academic group's attempts within creative computing degrees at a UK university to counter these problems through the introduction of a six-week long project that newly enrolled students embark on at the very beginning of their studies (Figure 5-4).



**Figure 5-4 3D etch-a-sketch. (a) Student-based drawing application [60], (b) student group's hardware interface [60]**

This group project provided a breadth-first, activity-led introduction to their chosen academic discipline, aiming to increase student engagement while providing a stimulating learning experience with the overall goal to increase retention. The methods and results of two iterations of these projects in the 2009/2010 and 2010/2011 academic years were presented. Results indicate that the ALL approach worked well for these cohorts, with students expressing increased interest in their chosen discipline, in addition to noticeable improvements in retention following the first year of the students' studies [60].

**Paper:** Anderson, E.F., Peters, C., Halloran, J., Every, P., Shuttleworth, J., Liarokapis, F., Lane, R., Richards, M. In at the Deep End: An Activity-Led Introduction to First Year Creative Computing, Computer Graphics Forum, Wiley-Blackwell, 31(6): 1852-1866, September, 2012.

**Contribution (10%):** Collaboration on the teaching methods and write-up of the paper (full text on section 0).

## 5.5 Serious Games and Virtual Environments

### 5.5.1 Serious Games Technologies

The success of computer games, fuelled among factors such as the great realism that can be attained using modern consumer hardware, and the key techniques of games technology that have resulted from this, have given way to new types of games, including serious games, and related application areas, such as virtual worlds, mixed reality, augmented reality and virtual reality. All of these types of application utilise core games technologies (e.g. 3D environments) as well as novel techniques derived from computer graphics, human computer interaction, computer vision and artificial intelligence, such as crowd modelling. Together these technologies have given rise to new sets of research questions, often following technologically driven approaches to increasing levels of fidelity, usability and interactivity. The aim has been to use this state-of-the-art report to demonstrate the potential of serious games technology for cultural heritage, to outline key problems and to indicate areas of technology where solutions for remaining challenges may be found. However, the same technology can be easily applied to other application domains.

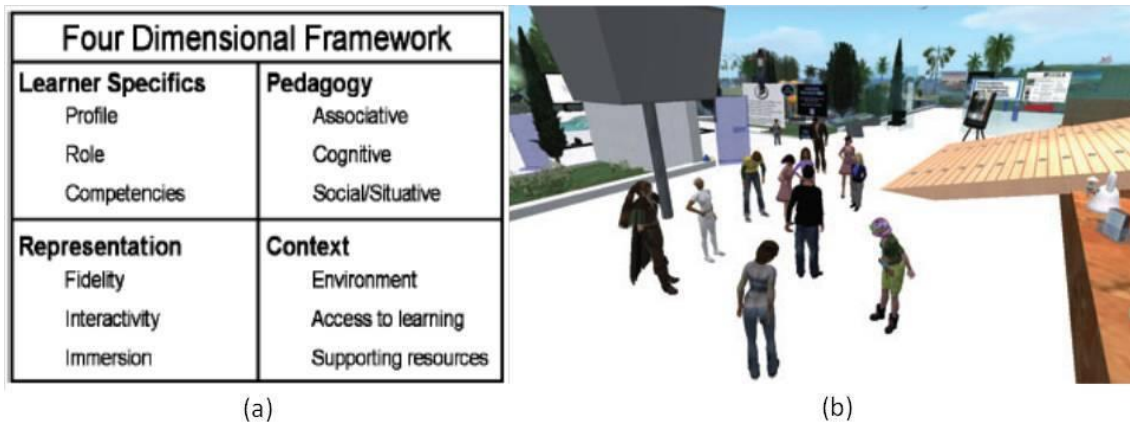
**Paper:** Anderson, E.F., McLoughlin, L., Liarokapis, F., Peters, C., Petridis, P., de Freitas, S. Developing serious games for cultural heritage: a state-of-the-art review, *Virtual Reality*, Springer, 14(4): 255-275, 2010.

**Contribution (20%):** Write-up of the serious games, virtual and augmented reality sections of the paper. Also co-written the introduction and conclusions (full text on section 8.15).

### 5.5.2 Learning as Immersive Experiences within Serious Games

Traditional approaches to learning have often focused upon knowledge transfer strategies that have centred on textually based engagements with learners, and dialogic methods of interaction with tutors. The use of virtual worlds, with text-based, voice-based and a feeling of ‘presence’ naturally is allowing for more complex social interactions and designed learning experiences and role plays, as well as encouraging learner empowerment through increased interactivity. To unpick these complex social interactions and more interactive designed experiences, this work considers the use of virtual worlds in relation to structured learning activities for college and lifelong learners [62]. This consideration necessarily has implications upon learning theories adopted and practices taken up, with real implications for tutors and learners alike. Alongside this is the notion of learning as an ongoing set of processes mediated via social interactions and

experiential learning circumstances within designed virtual and hybrid spaces. This implies the need for new methodologies for evaluating the efficacy, benefits and challenges of learning in these new ways.



**Figure 5-5 Learning as immersive experiences. (a) Four Dimensional Framework [63], (b) Meeting in-world in Second Life for virtual tour [62]**

Towards this aim, this work proposed an evaluation methodology for supporting the development of specified learning activities in virtual worlds, based upon inductive methods and augmented by the four-dimensional framework [63]. The approach was based upon an assumption that learning experiences need to be designed, used and tested in a multidimensional way due to the multimodal nature of the interface (Figure 5-5). The presented evaluation methodology may be used as a design tool for designing learning activities in-world as well as for evaluating the efficacy of experiences, due to its set of consistent criteria.

**Paper:** de Freitas, S., Rebolledo-Mendez, G., Liarokapis, F., Magoulas, G., Poulouvassilis, A. Learning as immersive experiences: Using the four-dimensional framework for designing and evaluating immersive learning experiences in a virtual world, *British Journal of Educational Technology*, Blackwell Publishing, 41(1): 69-85, 2010.

**Contribution (20%):** Collaboration on the design and evaluation of the serious game as well as the write-up of the paper (full text on section 8.16).

## Chapter 6

---

# Conclusions and Future Work

### **6.1 Conclusions**

In this habilitation thesis there was a presentation of several contributions to interactive virtual and augmented reality environments as well as various application domains. The thesis covered a number of different procedural generation techniques for generating content as well as human behaviour. Moreover, it provided contributions in virtual and augmented reality environments ranging from indoor to outdoor (mobile) solutions. It also covered a significant amount of contributions in the area of HCI, including more standard techniques using sensors to more advanced ones such as EEG methods. Finally, it showed how all the above mentioned methods can be applied in creating novel applications. It is worth mentioning that all these areas are fast evolving and the state-of-the-art research changes very fast.

### **6.2 Future Work**

In terms of future directions, it is realistic to expect contributions in all areas mentioned in this thesis. Firstly, to explore in more detail procedural approaches in different contexts. Secondly, to develop further the architecture of virtual and augmented reality allowing for more realistic computer graphics functionality. Thirdly, to improve the human computer interaction techniques by making use of multimodal approaches as well as more sensor devices. Finally, to apply the systems in different application domains such as medicine.

# Chapter 7

---

## References

- [1] Sutherland, I. The ultimate display, Proc. of the IFIP Congress, vol.2, 506-508, (1965).
- [2] Anderson, E.F., McLoughlin, L., Liarokapis, F., Peters, C., Petridis, P., de Freitas, S. Developing serious games for cultural heritage: a state-of-the-art review, *Virtual Reality*, Springer, 14(4): 255-275, 2010.
- [3] Pausch, R., Crea, T., Conway, M. A literature survey for virtual environments: military flight simulator visual systems and simulator sickness, *Presence: Teleoperators and Virtual Environments*, MIT Press, 1(3): 344-363, 1992.
- [4] Schuemie, M.J., Straaten, P.V.D., et al., Research on Presence in Virtual Reality: A Survey, *CyberPsychology & Behavior*, 4(2): 183-201, 2001.
- [5] Zhao, Q.P. A survey on virtual reality, *Science in China Series F: Information Sciences*, Springer, 52(3): 348-400, 2009.
- [6] LaViola, J.J. A discussion of cybersickness in virtual environments, *ACM SIGCHI Bulletin*, ACM Press, 32(1): 47-56, 2000.
- [7] Feiner, S.K. Augmented Reality: A New Way of Seeing. *Scientific American*, 286, 4, April 24, 48–55, (2002).
- [8] Liarokapis, F., *Augmented Reality Interfaces - Architectures for Visualising and Interacting with Virtual Information*, Sussex theses S 5931, Department of Informatics, School of Science and Technology, University of Sussex, Falmer, UK, 2005.
- [9] Azuma, R. A Survey of Augmented Reality, *Teleoperators and Virtual Environments*, 6(4): 355-385, 1997.
- [10] Mahoney, D. Better Than Real, *Computer Graphics World*, February 1999, 32-40, 1999.

- [11] Klinker, G., Ahlers, et al. Confluence of Computer Vision and Interactive Graphics for Augmented Reality, *PRESENCE: Teleoperations and Virtual Environments*, Special Issue on Augmented Reality, 6(4): 433-451, August 1997.
- [12] Azuma, R., Bailiot, Y., et al. Recent Advances in Augmented Reality, *Computers Graphics and Applications*, IEEE Computer Society, November/December, 21(6): 34-47, 2001.
- [13] Van Krevelen, D.W.F., Poelman, R. A survey of augmented reality technologies, applications and limitations, *International Journal of Virtual Reality* 9(2): 1-20, 2009.
- [14] Smelik, R.M., De Kraker, K.J., Tutenel, T., Bidarra, R., Groenewegen, S.A. A survey of procedural methods for terrain modelling, *Proc. of the CASA Workshop on 3D Advanced Media In Gaming And Simulation (3AMIGAS)*, 25-34, 2009.
- [15] Kelly, G., McCabe, H. A survey of procedural techniques for city generation, *ITB Journal*, 14, 87-130, 2006.
- [16] Smelik, R. M., Tutenel, T., Bidarra, R., Benes, B. A survey on procedural modelling for virtual worlds, *Computer Graphics Forum*, 33(6): 31-50, 2014.
- [17] Parish, Y.I.H. Muller, P. Procedural Modeling of Cities. *Proc. of the 28<sup>th</sup> Annual Conference on Computer Graphics and Interactive Techniques (SIGGRAPH '01)*, ACM Press, 301-308, 2001.
- [18] Greuter, S., Parker, J., Stewart, N., Leach, G., Real-time Procedural Generation of 'Pseudo Infinite' Cities, *Proc. of the 1<sup>st</sup> International Conference on Computer Graphics and Interactive Techniques in Australasia and South East Asia (Graphite)*, 87-95, 2003.
- [19] Wonka, P., Wimmer, M., Sillion, F. & Ribarsky, W. Instant Architecture, *ACM Transactions on Graphics*, 22(3): 669-677, 2003.
- [20] Muller, P., Wonka P., et al., Procedural Modeling of Buildings. *ACM Transactions on Graphics*, 25(3): 614-623, 2006.
- [21] Azahar, M.A.B.M., Sunar, M.S., Daman, D., Bade, A. Survey on Real-Time Crowds Simulation, *Technologies for E-Learning and Digital Entertainment, Lecture Notes in Computer Science*, Springer, Volume 5093, 573-580, 2008.
- [22] Zhou, S., Chen, D., et al. Crowd modeling and simulation technologies, *ACM Transactions on Modeling and Computer Simulation (TOMACS)*, ACM Press, 20(4), Article 20, 2010.
- [23] Luo, L., Zhou, S., Cai, W., Low, M., Lees, M. Toward a Generic Framework for Modeling Human Behaviors in Crowd Simulation, *Proc. of the IEEE/WIC/ACM Int'l Joint*



- Conference on Web Intelligence and Intelligent Agent Technology - Volume 02 (WI-IAT '09), Vol. 2. IEEE Computer Society, Washington, DC, USA, 275-278, 2009.
- [24] Cui, X., Shi, H. A\*-based Pathfinding in Modern Computer Games, *International Journal of Computer Science and Network Security*, 11(1): 125-130, 2011.
- [25] Reynolds C. Steering behaviours for autonomous characters, Proc. of game developers conference, Miller Freeman Game Group, San Francisco, California, 763-782, 1999.
- [26] Ondej, J., Pettre, J., Olivier, A.-H., Donikian, S. A Synthetic-Vision-Based Steering Approach for Crowd Simulation, *ACM Transactions on Graphics*, 29(4): 123, 2010.
- [27] Helbing, D., Molnar, P. Social force model for pedestrian dynamics, *Phys. Rev E, American Physical Society*, 51(5): 4282-4286, 1995.
- [28] Ennis, C., Peters, C., O'Sullivan, C. Perceptual effects of scene context and viewpoint for virtual pedestrian crowds, *ACM Transactions on Applied Perception (TAP)*, 8(2), Article 10, 2011.
- [29] O'Sullivan, C., Ennis, C. Metropolis: multisensory simulation of a populated city, *International Conference on Games and Virtual Worlds for Serious Applications (VS-Games)*, IEEE Computer Society, Athens, Greece, 1-7, 2011.
- [30] Rego, P., Moreira, P.M., Reis, L.P. Serious games for rehabilitation: A survey and a classification towards a taxonomy, Proc. of the 5<sup>th</sup> Iberian Conference on Information Systems and Technologies (CISTI), IEEE Computer Society, 1-6, 2010.
- [31] Paivio, A. *Mental representations: A dual coding approach*, Oxford University Press, New York, 1990.
- [32] Baddeley, A.D. The episodic buffer: a new component of working memory?, *Trends in Cognitive Science*, 4(11): 417-423, 2000.
- [33] Zyda, M. From visual simulation to virtual reality to games, *IEEE Computer*, 38(9): 25-32, 2005.
- [34] de Freitas, S., Oliver, M. How can exploratory learning with games and simulations within the curriculum be most effectively evaluated?, *Computers and Education*, Elsevier, 46(3): 249-264, 2006.
- [35] Dix, A., Finlay, J., Abowd, G., Beale, R. *Human-computer interaction*, 3<sup>rd</sup> edition, Prentice Hall, 2003.

- [36] Wright, P.C., Fields, R.E., Harrison, M.D. Analyzing Human-Computer Interaction as Distributed Cognition: The Resources Model, *Human-Computer Interaction*, Taylor and Francis, 15(1): 1-41, 2000.
- [37] Kjeldskov, J., Graham, C. A review of mobile HCI research methods. In *Human-computer interaction with mobile devices and services*, Springer Berlin Heidelberg, 317-335, 2003.
- [38] Rekimoto, J., Nagao, K. The World through the Computer: Computer Augmented Interaction with Real World Environments, *Proc. of UIST '95*, (ed B.A. Myers), ACM Press, Pennsylvania, 29-36, 1995.
- [39] Oviatt, S. Ten myths of multimodal interaction, *Communications of the ACM*, ACM Press, 42(11): 74-81, 1999.
- [40] Noghani, J., Liarokapis, F., Anderson, E.F. Randomly Generated 3D Environments for Serious Games, *Proc. of the 2nd IEEE International Conference in Games and Virtual Worlds for Serious Applications*, IEEE Computer Society, Braga, Portugal, 25-26 March, 3-10, 2010.
- [41] Noghani, J., Anderson, E., Liarokapis, F. Towards a Vitruvian Shape Grammar for Procedurally Generating Classical Roman Architecture, *Proc. of the 13th International Symposium on Virtual Reality, Archaeology and Cultural Heritage VAST 2012*, Short and Project Papers, Eurographics, Brighton, UK, 19-21 November, 41-44, 2012.
- [42] Barton, I.M. Palaces. In *Roman Domestic Buildings*, University of Exeter Press, 91-120, 1996.
- [43] Jacobson, J., Holden, L. The virtual egyptian temple, *ED-MEDIA: Proceedings of the World Conference on Educational Media, Hypermedia & Telecommunications 2005*.
- [44] O'Connor, S., Liarokapis, F., Peters, C. An Initial Study to Assess the Perceived Realism of Agent Crowd Behaviour in a Virtual City, *Proc. of the 5th International Conference on Games and Virtual Worlds for Serious Applications (VS-Games 2013)*, IEEE Computer Society, Bournemouth, UK, 11-13 September, 85-92, 2013.
- [45] Liarokapis, F. An exploration from virtual to augmented reality gaming, *Simulation and Gaming, Symposium: Virtual Reality Simulation*, SAGE Publications, December, 37(4): 507-533, 2006.
- [46] Vourvopoulos, A., Liarokapis, F. Evaluation of commercial brain-computer interfaces in real and virtual world environment: A pilot study, *Computers and Electrical Engineering*, Elsevier, 40(2): 714-729, 2014.

- [47] Liarokapis, F., Debattista, K., Vourvopoulos, A., Ene, A., Petridis, P. Comparing interaction techniques for serious games through brain-computer interfaces: A user perception evaluation study, *Entertainment Computing*, Elsevier, 5(4): 391-399, 2014.
- [48] Liarokapis, F., Macan, L., Malone, G., Rebolledo-Mendez, G., de Freitas, S. Multimodal Augmented Reality Tangible Gaming, *Journal of Visual Computer*, Springer, 25(12): 1109-1120, 2009.
- [49] Liarokapis, F., Mourkoussis, N., White, M., Darcy, J., Sifniotis, M., Petridis, P., Basu, A., Lister, P.F. Web3D and Augmented Reality to support Engineering Education, *World Transactions on Engineering and Technology Education*, UICEE, 3(1): 11-14, 2004.
- [50] Liarokapis, F., Brujic-Okretic, V., Papakonstantinou, S. Exploring Urban Environments using Virtual and Augmented Reality, *Journal of Virtual Reality and Broadcasting*, GRAPP 2006 Special Issue, Digital Peer Publishing, 3(5): 1-13, 2006.
- [51] Feiner S, MacIntyre B, et al. Windows on the World: 2D Windows for 3D Augmented Reality, *Proc. of the ACM Symposium on User Interface Software and Technology*, Atlanta, November 3-5, ACM Press, 145-155, 1993.
- [52] Haller M, Hartmann W, et al. Combining ARToolKit with Scene Graph Libraries, *Proc. of The 1<sup>st</sup> IEEE International Augmented Reality Toolkit Workshop*, Darmstadt, Germany, 29 September, 2002.
- [53] MacIntyre B., Gandy M., Dow S., Bolter J.D. DART: a toolkit for rapid design exploration of augmented reality experiences, *ACM Transactions on Graphics (TOG)*, 24(3): 932, 2005.
- [54] Liarokapis, F. An Augmented Reality Interface for Visualizing and Interacting with Virtual Content, *Virtual Reality*, Springer, 11(1): 23-43, 2007.
- [55] White, M., Mourkoussis, N., Darcy, J., Petridis, P., Liarokapis, F., Lister, P.F., Walczak, K., Wojciechowski, R., Cellary, W., Chmielewski, J., Stawniak, M., Wiza, W., Patel, M., Stevenson, J., Manley, J., Giorgini, F., Sayd, P., Gaspard, F. ARCO-An Architecture for Digitization, Management and Presentation of Virtual Exhibitions, *Proc. of the 22<sup>nd</sup> International Conference on Computer Graphics (CGI'2004)*, IEEE Computer Society, Hersonissos, Crete, June 16-19, 622-625, 2004.
- [56] Mountain, D., Liarokapis, F. Mixed reality (MR) interfaces for mobile information systems, *Aslib Proceedings*, Special issue: UK library & information schools, Emerald Press, 59(4/5): 422-436, 2007.

- [57] Goldsmith, D., Liarokapis, F., Malone, G., Kemp, J. Augmented Reality Environmental Monitoring Using Wireless Sensor Networks, Proc. of the 12<sup>th</sup> International Conference on Information Visualisation (IV08), IEEE Computer Society, 8-11 July, 539-544, 2008.
- [58] Sylaiou, S, Liarokapis, F., Kotsakis, K., Patias, P. Virtual museums, a survey and some issues for consideration, Journal of Cultural Heritage, Elsevier, 10(4): 520-528, 2009.
- [59] Liarokapis, F., Anderson, E. Using Augmented Reality as a Medium to Assist Teaching in Higher Education, Proc. of the 31<sup>st</sup> Annual Conference of the European Association for Computer Graphics (Eurographics 2010), Education Program, Norrkoping, Sweden, 4-7 May, 9-16, 2010.
- [60] Anderson, E.F., Peters, C., Halloran, J., Every, P., Shuttleworth, J., Liarokapis, F., Lane, R., Richards, M. In at the Deep End: An Activity-Led Introduction to First Year Creative Computing, Computer Graphics Forum, Wiley-Blackwell, 31(6): 1852-1866, September, 2012.
- [61] Savin-Baden, M., Major, C. Foundations of Problem Based Learning, Open University Press, Buckingham, UK, 2004.
- [62] de Freitas, S., Rebolledo-Mendez, G., Liarokapis, F., Magoulas, G., Poulouvassilis, A. Learning as immersive experiences: Using the four-dimensional framework for designing and evaluating immersive learning experiences in a virtual world, British Journal of Educational Technology, Blackwell Publishing, 41(1): 69-85, 2010.
- [63] de Freitas, S. Serious virtual worlds: a scoping study. Bristol: Joint Information Systems Committee, Report, 3<sup>rd</sup> November 2008. (Available at: <http://www.jisc.ac.uk/publications/publications/seriousvirtualworldsreport.aspx>, Accessed at: January 2015).

## Chapter 8

---

# Appendix – Paper Reprints

In the following sections, copies of the papers used for this habilitation thesis are provided. The selected conference papers concern topics which have not yet been published in journal papers, but this will happen in the future, since the work is on-going.

## **8.1 Paper #1**

Noghani, J., Liarokapis, F., Anderson, E.F. Randomly Generated 3D Environments for Serious Games, Proc. of the 2<sup>nd</sup> IEEE International Conference in Games and Virtual Worlds for Serious Applications, IEEE Computer Society, Braga, Portugal, 25-26 March, 3-10, 2010.

Contribution (40%): Design of the architecture, implementation of smoothing techniques and advice on evaluation. Write-up of most of the paper.

## Randomly Generated 3D Environments for Serious Games

Jeremy Noghani  
Interactive Worlds Applied  
Research Group  
Coventry University  
Coventry, UK  
noghanij@coventry.ac.uk

Fotis Liarokapis  
Interactive Worlds Applied  
Research Group  
Coventry University  
Coventry, UK  
F.Liarokapis@coventry.ac.uk

Eike Falk Anderson  
Interactive Worlds Applied  
Research Group  
Coventry University  
Coventry, UK  
Eike.Anderson@coventry.ac.uk

**Abstract**— This paper describes a variety of methods that can be used to create realistic, random 3D environments for serious games requiring real-time performance. These include the generation of terrain, vegetation and building structures. An interactive flight simulator has been created as proof of concept. An initial evaluation with two small samples of users (remote and hallway) revealed some usability issues but also showed that overall the flight simulator is enjoyable and appears realistic and believable.

**Keywords** – serious games; 3D terrain modeling; computer graphics; flight simulator.

### I. INTRODUCTION

The creation of realistic virtual environments is an important issue in the computer animation, computer games, digital film effects and simulation industries. In recent years, the computer and video games industry has overtaken both the film and music industries as the top revenue producers, and the cost for developing a commercial game now usually requires investments of several million dollars, involving large teams of developers that can number in the hundreds of workers, many of whom are artists and designers providing content for the decoration of rich virtual game worlds. While many games companies have the necessary budget to develop these expensive modern computer games that employ state of the art computer graphics, not all game developers have the same resources. Serious games refer to computer games that are not limited to the aim of providing just entertainment but which can be used for other purposes, such as education or training in a number of application domains. There are several game engines and online virtual environments that have been used to design and implement these games for non-leisure purposes [1]. The development of serious games using the same approach as used for entertainment games is not possible because their budget is usually limited to a few thousand dollars. The literature states that when games and simulations technologies are applied to non-entertainment domains, serious gaming applications can be created [2]. When classifying a game, the definition of the term ‘game’ does not necessarily require formalised criteria for success such as praising winners, totalling points or reaching certain areas in a level [3]. ‘Gaming is by no means a replacement for existing model and simulation building processes and practices but it has

tangible advantages that ultimately could result in wider, more flexible, and more versatile products” [4].

To overcome these problems, a variety of methods for automatically creating detailed but also randomised environments have been developed. The use of these procedural methods [5] saves time and reduces the budget for creating effective serious games. However, if a user wishes to interact with the environment in a meaningful way, such as in a flight simulator that has an expansive world and implements collision detection, then numerous problems arise that are often not dealt with during the creation stage.

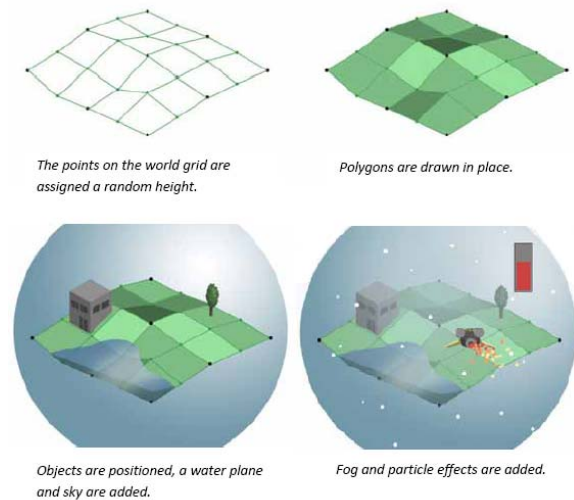


Figure 1 Randomly generated environment for an entertainment flight simulator

This paper explains some of the problems that can arise from this situation and describes a variety of methods that can be used to overcome them. These methods have been applied to a basic flight simulator (see Figure 1), so that the results could be observed and evaluated. Initial results with 2 types of user groups (remote and hallway) revealed some usability issues but also illustrated that overall the flight simulator is enjoyable, fun and looks realistic.

The rest of the paper is structured as follows. Section II provides past methods used in terrain generation. Section III presents how our flight simulator serious game was created to allow for navigation and interaction with the terrain whereas section IV describes

techniques used for creating infinite terrains. Section V provides an overview of procedural techniques for adding vegetation and buildings into randomised environments. Section VII presents a flight simulator as a case study and section VIII illustrates initial evaluation results. Finally section IX presents conclusions and future work.

## II. TERRAIN CREATION METHODS

The majority of the traditional methods used to create partially randomised terrain involve the use of fractals, such as fault formation [6] and noise algorithms [7]. Fractals are objects or shapes which, when split into smaller parts, result in shapes that are similar to the original shape as a whole [8] (self-similarity). Their use is advantageous from a computer graphics point of view, due to their ability to define complex geometry from a small set of instructions, and due to their ability to define shapes that are often difficult to define with simple Euclidian geometry.

Random one-dimensional midpoint displacement is a simple algorithm that can be used to create fractals that appear similar to the two-dimensional silhouette of mountain ranges. It is implemented by finding the midpoint of a single line, and displacing its height by a random offset value. This process is then repeated at the midpoints between these newly defined points with a reduced random number range. This algorithm is usually implemented recursively to allow the silhouette to be made as detailed as the user requires [9].

When the random midpoint displacement is applied to the centre of a terrain grid square only, this can be defined in terms of the displacements of the centre points of the square's sides. A more efficient way is to derive the same result by adding the four corners of the square and dividing them by four and adding the random value to the result.

The diamond-square algorithm can be considered an effective way of applying this one-dimensional method to a second dimension, creating three-dimensional terrain if the resulting lattice is used as a virtual heightmap [10]. The recursive algorithm works by refining a square area, whose four corner points' height values may be initialised randomly, and then calculating its centre point by calculating a mean of the corner points, to which a random value is added. Midpoints of the edges between the corners are then calculated in a similar manner and the original shape is then subdivided by generating new edges between the newly generated points, forming new squares for further subdivision, as well as diamond shapes within the squares. Using a smaller random value tends to result in the creation of smoother terrain, whereas larger offsets result in more jagged edges. The use of hexagonal and triangular shapes instead of a square grid has been proposed to reduce the problems of 'creasing' in the terrain [11].

There has been some work on modelling terrain based on realistic physical constraints. Kelley et al. [12] produced a system in which water drainage is simulated

to shape and constrain the landscape, in a similar manner to the way in which water erosion affects real terrain. Musgrave et al. [13] managed to achieve realistic results through a different method that took hydraulic and thermal erosion into account when creating a fractal terrain. Attempts to create more geographically accurate models have led to increased realism in some aspects, but have also increased the complexity of the design and rendering [14].

An alternative method to create randomised terrain is the use of Lindenmayer systems. L-systems were originally created to study organic growth, such as is the case with plants, but they can easily be adapted to cover other self-similar structures, such as mountainous terrain [15]. The distinctive feature of L-systems is the use of rules that rewrite strings, which can be called recursively to make a hierarchy of strings. When displayed visually, these may produce results similar to those of a mountain range silhouette, for example.

## III. SIMULATOR CREATION

For the purpose of this paper, a small, simple flight simulator was created to allow for navigation and interaction with the terrain. In the design, usability took precedence over realism, as a result of which the controls were deliberately kept simple; the mouse is used to alter the yaw and pitch of the aircraft, and two keys are used for acceleration and deceleration. Additionally, to allow for close examination of the terrain, the in-program physics were kept liberal; i.e. the plane was allowed to come to a complete stop in mid-air without gravity taking effect.

For the terrain itself, heightmaps that were generated using the diamond-square algorithm were chosen to provide surface detail. This method was chosen primarily due to the algorithm's simplicity and adaptability, meaning that the system itself could easily be altered to accommodate a more complex algorithm or to accept alterations to factors such as surface roughness without the need to be rewritten from scratch. Additionally, by choosing a recursive algorithm, the level of detail could be adjusted as necessary, which proved to be an advantage when dealing with different methods that required different levels of processing power.



Figure 2 Pyramid created by a random height-displacement of the centre point of the square base

Figure 2 illustrates how a vertical deformation of the centre of the square base connected to the neighbouring points can produce a pyramid. Instead of using this linear deformation of the neighbouring points, a two-dimensional Lorentz distribution [16] for the



height array was assumed. By adjusting the width of the Lorentzian shape, one could obtain a means for controlling the smoothness of the terrain. Another distribution that could be used is the Gaussian shape. After some trials, it was found that this bell-shaped distribution that creates a smooth terrain was the following:

$$height = y = (random\ number) \times \frac{D^2 / 8}{(x - x_0)^2 + (z - z_0)^2 + D^2 / 8}$$

where  $(x_0, z_0)$  is the position of the peak and  $D^2$  is the length of the square which is related with the width of the Lorentz distribution. The value of the width directly affects the smoothness of the terrain. By decreasing the width of the bell-shaped distribution, the terrain becomes steeper.

Water was added to the landscape in the form of a single translucent plane placed at an appropriate height (see Figure 1). Small buildings and trees were also added as decorations for the terrain, using pre-fabricated models. Their placement on the landscape was decided randomly, although rules were implemented to prevent their creation on top of mountain peaks, below the virtual world's water level, or on steep slopes.

#### IV. OUTER BOUNDARIES

In any simulator that requires travelling for a long time in one direction (flight simulators being the most notable example), the user may find that they see or pass an "outer boundary"; the terrain is only created within certain dimensions, so reaching an area where nothing is rendered can be a possibility, depending on the actual implementation. Three potential solutions were devised and implemented.

The first was to create a terrain of such large extent that the user would never reach the outer boundary (see Figure 3). The success of this method depended upon the scale of the landscape relative to the user's movement speed; if a user moved over the length of single terrain grid squares per second, then they would be less likely to reach a boundary than a user who moved at 5 grid squares per second, assuming that other factors remained equal. The implication of this is that landscapes must be scaled to be as large as possible to minimise the chance of the user discovering a boundary. Upon attempting this method, another problem arose in the form of the terrain looking bland and flat, due to the spread-out nature of the polygons. The solution to this was to increase the number of subdivisions, thus increasing the level of detail. However, it ought to be noted that if the designer were to keep increasing the scale and the level of detail of the terrain at the same time, then eventually the computer would reach the limitations of its hardware. For this reason, this method can be considered appropriate for small demonstration purposes or applications where the user moves slowly relative to the virtual landscape, but inappropriate for a full flight

simulator where a large detailed environment is desirable.

The second attempted method was to "loop" the old landscape when the user reaches a boundary. In order for the landscape to loop seamlessly, it was vital to ensure that the edges have equal height values; on an A1 to H9 grid, for example, B1 must equal B9, A5 must equal G5, and the corner values (A1, A9, G1, and H9) must equal each other. The landscape can then be kept in the computer memory as a single "tile", which can be duplicated to a new spot when needed. Tiles of terrain that are a particular distance from the player's avatar, i.e. the aircraft, can be deleted or moved to a more appropriate position, to keep memory usage to a minimum.

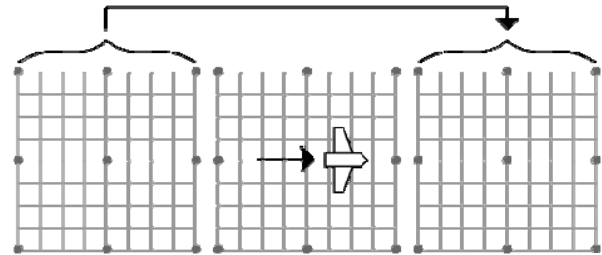


Figure 3 Upon reaching the right side of the landscape, a tile of terrain is moved in front of the player to provide the illusion of endless terrain

This solution could cause the problem of the user noticing the repetitive nature of the terrain (especially if the terrain contained a notable feature, such as a peak), but the severity of this issue would depend on several factors. For example, if the user intended on using the same terrain for a long period of time, then he or she would be more likely to notice the copied terrain tiles than a user who intended on playing for a short period of time. Additionally, if the terrain is seemingly large (i.e. if it took 60 seconds to travel across a single tile, for instance), then the repeating nature of the terrain tiles would be less noticeable than if the terrain were small (i.e. if it took 20 seconds to travel across a tile). The implications of this are that looping the landscape with the same terrain tile would work in a number of simulation scenarios, but it is difficult to assess whether this could be successfully applied to a particular flying simulator without some form of user testing. A final note on this method is that by making the four corner values identical, the deviation between the highest and lowest points of the landscape may be reduced. Alternate methods of increasing the stochasticity (such as using more random points) may be considered to avoid this problem.

The final method was to automatically generate a new, unique tile of terrain when the user reaches a boundary (see Figure 4). To ensure that the tiles matched seamlessly, one edge of the terrain would have its heights copied to the matching edge of the new tile. The rest of the heights can then be calculated via random values and midpoint displacement using the diamond-

square algorithm, in the same manner as was used to create the first tile.

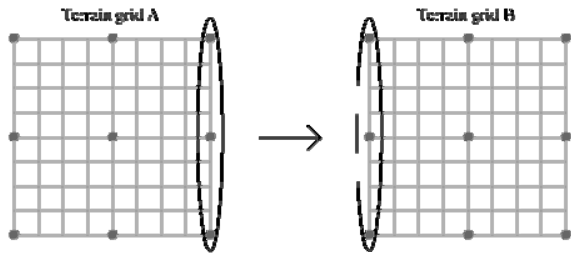


Figure 4 Upon reaching the right side of terrain grid A, the values of the far-right points are copied to the far left points of terrain grid B. The other points of terrain grid B are then calculated via randomisation and mid-point displacement, as done for grid A

The advantage of this method is that the terrain is genuinely infinite; from a user's perspective, the land would continue in all directions with no repetitions. However, there is the problem of memory usage. If a user were to continue travelling in a straight line, increasingly more tiles would have to be generated and stored in memory (even if they were not rendered), which could cause a memory overflow. The solutions to this are to either store previously visited tiles in a separate cache file, or to simply delete previously visited areas that are far away from the avatar. The former solution has the problem of requiring an efficient caching system (i.e. one that allows for fast writing and reading of large sets of coordinates), and the latter suffers from not allowing the user to backtrack to a previously-visited area that is a certain distance away.

## V. AUTOMATIC DECORATION OF VIRTUAL ENVIRONMENTS

Empty, featureless spaces resulting from terrain generation alone are insufficient for the creation of convincing virtual environments. To overcome this, the environment needs to be decorated with suitable vegetation and artificial structures, including buildings as well as settlements.

### A. Procedural Generation and Placement of Vegetation

There are different methods for the procedural creation of vegetation, many of which are based on fractal or simpler rule-based techniques. One of the latter methods has been used for on-the-fly generation of forests for real-time virtual environments [17], using a skeletal topology for procedurally generated and animated trees, which has also been combined effectively with on-the-fly generated grass to create a rich natural scenery [18]. A much more powerful, rule-based approach applying component-based modelling is the one by Lintermann and Deussen [19], which provides a more intuitive way for controlling plant modelling than the well-known L-Systems [15]. The decision of what type of plant needs to be placed into the virtual world

usually depends on a number of factors, including the elevation and slope of the terrain, as well as topographic features that dictate the probability of a specific plant's occurrence [20]. Once a position in the generated terrain has been decided, the generation of plant models can be followed with the placement of the vegetation.

For the proof of concept application, a simplified method was implemented that made use of randomised positioning of vegetation models, rather than random vegetation itself. A series of low-polygon trees and bushes were created and exported as 3D models, which were then loaded at the start of the program. Once the terrain had been created, the vegetation was randomly assigned to various places on the terrain grid. However, if a particular part of the terrain was too high, low, submerged in water, or on a heavily inclined slope, then that area was rejected and ignored, the purpose being to prevent vegetation appearing in unrealistic locations. To reduce the problem of repetition, the vegetation was also rotated and scaled by a random value. The result was surprisingly effective; the plants appeared to have stochastic properties despite being pre-defined models, and it was only when the density of the vegetation was increased to the level of woodland when the repetition became noticeable.

### B. Procedural Generation of Buildings

Most real-world environments include some sort of artificial structures. In a rural setting these might be scattered houses that make up only a fraction of the decorations of the terrain, with the majority of decorations being plants, whereas in urban settings this would be reversed with buildings providing the majority of virtual world decorations. If the level of detail required for buildings is relatively low, as would be the case in a flight simulator that depicts the virtual world from a high altitude, then simple geometric bodies can create adequate results if combined with suitable texture maps that hide the lack of actual detail in the geometry. The use of 'split grammars' [21] and 'shape grammars' [22] for describing architectural features allow the use of much more complex shapes and building structures, which can be intricately detailed [23].

At the greatest level of detail, even building interiors can be generated [24]. The placement of these artificial structures in the virtual world can reach great levels of complexity if the buildings form part of an urban environment [25]. These more complex settlements are created in a series of steps [26]: (a) *first a suitable road network is generated, effectively providing street maps that partition the terrain and to constrain the placement of buildings* (b) *this is then used to direct the division of the terrain into lots which may be partitioned further to generate building footprints, and* (c) *which are then used as input for the generation of the buildings themselves.*

Due to time restrictions, attempts at implementing a more complex urban generation system had to be simplified. The method of distributing buildings was therefore nearly identical to the method of distributing

plants, with a few distinct changes. Firstly, the building models were adjusted to have ‘foundations’, or basement levels. The purpose of this was to prevent the underside of the model showing, should the building be positioned on a slope. Secondly, rather than being rotated and scaled, which would be inappropriate for the majority of buildings, structures were assigned random ‘height’ and ‘extention’ values that copied parts of the building model above or to the side of the original, the purpose being to reduce repetition and to reduce the isolated feel that can be associated with solitary buildings.

The results were acceptable, and would be especially fitting if applied to a small settlement of village size, but the environment as a whole lacked the structure and density associated with urban areas. The solution to this problem would be to redesign the placement system, and possibly the terrain generation system, from scratch, whilst taking into account the architectural shape grammars and road networks used in previous city generation applications.

## VI. COLLISION DETECTION

Accurate detection of when an object hits the terrain surface is highly desirable in many applications, especially flight simulators. However, calculating precise polygon overlap between an aircraft and a landscape would be too computationally expensive, especially given the recursive and detailed nature of fractal terrain, which could result in thousands of checks per frame. Alternative methods were therefore implemented in an attempt to find a method that was both fast and accurate.

Traditionally, the most common method of calculating collisions is through the use of bounding volumes, such as spheres or orientated cuboids, which can be positioned in place of a complex model and then be checked for overlap. They can also be used hierarchally (i.e. checking a large bounding volume, followed by more precise checks), to give results that are both memory efficient and precise [27]. In this paper, a single axis-aligned bounding box was used to cover the aircraft. The trees, vegetation and buildings were covered by single oriented bounding boxes. Additionally, a single plane check was used to check whether the aircraft had hit the water surface. More complex and accurate methods, such as a series of bounding boxes, would be more appropriate for a final game or simulation, but simplicity was adhered to for the sake of shortening the debugging process and for achieving consistent results when testing the speed and accuracy of the collisions.

For the terrain in this instance, bounding spheres were quickly deemed as inappropriate, as they would fit awkwardly with the relatively flat polygons unless used at a high level of detail. Axis-aligned bounding boxes would be faster than spheres to calculate, due to the simpler calculations needed for every frame [27], and they could potentially be more accurate on less mountainous terrain due to the nearly-aligned nature of the landscape. The bounding boxes were applied by making use of the terrain data arrays; for every polygon

point, a box was applied that would match the point’s height, and have a polygon’s length and width. An advantage of this method was that, as with the fractal terrain itself, the checks could be called recursively to achieve a higher level of detail (collision accuracy), at the cost of efficiency.

For example, a bounding box could be applied every two polygons across, or bounding boxes of twice the width and length could be applied every four polygons across, resulting in significantly fewer calculations. During testing, it was found that, relative to the complexity of the fractal terrain, only a small number of bounding boxes were needed for the terrain collisions to be perceived as accurate, so program efficiency was not an issue. More bounding boxes were needed to maintain accuracy if the terrain was made to be mountainous. However, a scenario where more bounding boxes were needed than were possible with the terrain data array was not implemented. One observed point was that mountain peaks seemed to suffer occasionally from inaccurate collisions, probably due to the fact that this was the area with the most ‘space’ contained within the box. There did not seem to be a simple fix to this problem, as manually adjusting the bounding boxes used for mountain peaks was impractical and delivered mixed results. Nonetheless, axis-aligned bounding boxes were considered successful for this simulator due to their speed and overall accuracy.

Before oriented bounding boxes can be discussed, it ought to be noted that since we are working in three dimensions, there are two sets of rotation that can be implemented. The first would align the boxes according to the way the landscape is facing (the yaw); this would appear to be a rotated square, if viewed from above. The second would align the boxes according to the slope of a polygon or terrain face (the pitch and roll). Performing a check on whether an axis-aligned box (the aircraft) collides with a rotated square (part of the terrain after being rotated once) would require a simple series of checks for each of the oriented box’s points; the collision detection is still being performed in two dimensions. However, after applying the second rotation, checks must be performed between two bounding boxes aligned on separate axes, and consequently the number of calculations rises steeply. Considering the number of bounding boxes on the terrain, it was predicted that this could potentially become an issue.

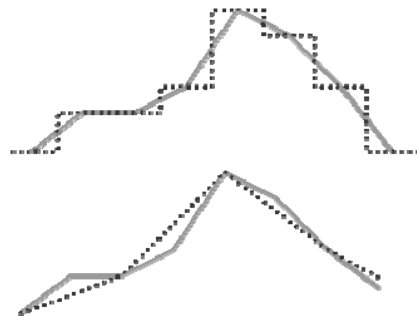


Figure 5 Comparison of axis-aligned (top) and oriented bounding box method (bottom). Recursive calls lead to a more accurate terrain match.

Upon testing, it was found that performing only a single rotation on the bounding boxes gave very similar results to the axis aligned boxes, both in terms of efficiency and accuracy. This method carried the same problems and advantages of the axis aligned method. Upon rotating the boxes a second time, the accuracy improved somewhat and the problems associated with the mountain peak inaccuracies disappeared. However, the program was notably less efficient; attempting to apply the fully oriented bounding boxes to every terrain polygon slowed the program down to an unusable level (although whether such a level of accuracy is actually required for a flight simulator is questionable). Deciding whether it would be more accurate and efficient to use a small number of fully oriented bounding boxes or a large number of axis-aligned boxes for fractal terrain is a matter that requires further research and testing.

The program could be further streamlined through the addition of Bounding Volume Hierarchies (BVHs). By splitting up the terrain area into large bounding volumes that in turn contain consecutively smaller bounding volumes, the number of collision checks carried out per frame could be substantially reduced. Since this particular program uses a square grid for the terrain, it would be logical for the checks to take the form of testing which side the aircraft falls on on an imaginary plane placed in the middle of the terrain grid. This is repeated to further divide the grid into quarters and eighths within the section in which the craft is located. Precise collision checks can then be carried out in the appropriate section. In this case, the efficiency of using BVHs depends upon the complexity of the landscape; a more complex landscape would benefit more from having BVHs implemented for collision detection, as a large number of calculations would be removed per frame, whereas if the landscape were only a few polygons in size, implementing BVHs would have little effect. If numerous aircraft are included, or buildings and objects overlap BVH boundaries, then the efficiency of BVHs becomes more difficult to calculate, and their use ought to be carefully considered.

## VII. FLIGHT SIMULATOR

The simple flight simulator game created as a case study for the random world generation allows navigation and interaction with the randomised 3D environment. Thematically the game is set to take place on alien worlds. This could potentially offer a wider range of potential scenarios (i.e. the online virtual world Second Life is based on this philosophy) and thus offer a higher level of entertainment and enjoyability compared to ‘Earth’ based scenarios. While of course this type of representation does not automatically lead to exploratory, challenging and problem-based learning experiences, the opportunities for players to “define their learning experiences or pathways, using the virtual mediations within virtual worlds, has the potential to invert the more hierarchical relationships associated with traditional

learning, thereby leading to more learner-led approaches based upon activities for example” [28].

Players can navigate intuitively inside the the alien worlds using mouse and keyboard inputs. As mentioned before, the controls were deliberately kept simple; the mouse is used to steer the aircraft, and two keys, ‘W’ and ‘S’, are used for acceleration and deceleration respectively. Players can also fire a weapon by pointing and clicking with the mouse. An overview of the flight simulator in operation is shown in Figure 6.



Figure 6 Alien world flight simulator in operation

A digital compass and a speedometer are implemented as overlays in the interface. The digital compass (see top left hand side of Figure 6) allows players to navigate inside the virtual environment using directional information. The speedometer representation (see top right hand side of Figure 6) was kept as simple as possible in order to leave maximum screen space available for the game. Additionally, a widget menu was implemented that allows players to change specific components of the 3D environment by right-clicking the mouse. The environmental components that can be modified include: *terrain re-generation*, *weather alteration* (i.e. *rain, fog, sunny, etc.*), and *colouring of the grass, water and sky*. An overview of the widget menu is shown in Figure 7.



Figure 7 Flight simulator widget menu

Options to change the music track or mute it entirely were also added to the widget menu. These controls were introduced to the user in the form of a simple menu

screen that appeared at the start of the game. The collision detection algorithm was based on axis-aligned bounding boxes for the terrain and the building structures (see section VI). In addition, based on the techniques described in section IV, the landscape creates the illusion that is infinite. An example screenshot of explosions generated when the aircraft collides with the ground is shown in Figure 8.

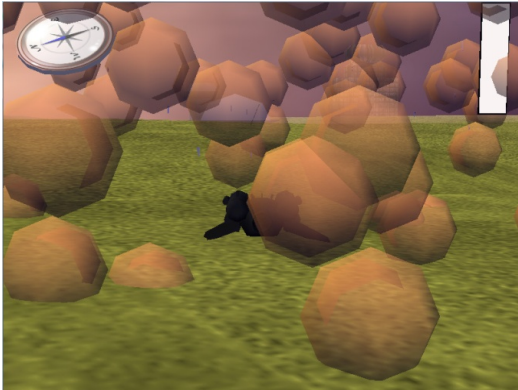


Figure 8 Collision between the aircraft and the terrain

To simulate the effect of collisions, explosions were incorporated into the game based on particle systems. Each particle effect (snow, rain, engine fumes and explosions) has a velocity value, rotation value, or transparency applied to it and for each collision, each particle is changed appropriately (such as a slight decrease in y position, for rain), and if it reaches a certain condition (such as rain falling too low), it is set to a new start position.

## VIII. INITIAL EVALUATION

To acquire feedback on the finished core of the application, a self-contained executable file was supplied to two sets of users: a small Internet general discussion forum (remote usability testing), and a group of Coventry University students (hallway usability testing). The intention of these tests was primarily to gather information on the playability and enjoyability of the game, but also to discover potential technical problems. All of the end-users had some experience with games, and the vast majority described themselves as ‘gamers’. A few of those involved also had experience with games programming, or had some knowledge of the architecture behind creating a game. For both sets of users, the aim of the flight simulator project was presented and it was explained that the players should not expect a complete game, but rather a prototype.

### A. Remote Testing

For the Internet forum users, the following set of questions was provided: (a) *do the collisions seem accurate?* (b) *would you be interested in playing a fuller version of this game?*, (c) *what would you like to see added?* (e.g. *larger variety of landscapes, different*

*controls*) and (d) *how large and varied were the environments?* – the final question testing, whether the attempts at making the different environments varied was a success. A qualitative analysis was done with five users. Feedback was received some in direct reply to the questions, and some raising additional issues. Recorded feedback was very encouraging and all users agreed that the methods used were very useful for the creation of serious games applications, although important issues were pointed out.

The answers to the second and third questions were positive and similar. Several users commented that they would play such a game, on the condition that further additions were made to the gameplay. “It needs more to do but the engine is cool”, one user noted. In regards to the final question, reactions were mixed. One user commented that they “enjoyed exploring the worlds”, and mentioned how the use of colour made a lot of difference, but another noted that “the buildings look samey. They need more variation”. Additional comments were also made. One player complimented the “atmosphere” of the game. However, the collision detection was criticised by another player. Specifically, he stated “I sometimes crash when I drive too close to mountains. [The collision detection] is fine for the water and flat land though”. The other users claimed that the collision detection was acceptable.

### B. Hallway Testing

Four students from the Faculty of Engineering and Computing, Coventry University were asked to partake in the second test group. Instead of asking the university students questions, they were asked to talk through what they were doing and how they felt as they played the game. Two students had some issues with the controls. Specifically, they found the delay between hitting a key and the aircraft movement difficult to get to grips with. It is worth-mentioning that after playing for long enough, the players adjusted to the issue. The object ‘popping’ due to terrain decorations being added to the scene was criticised by two players; one commented that “it’s nice that I can keep going forever, but it’s annoying that I can’t see the horizon properly”. Similar to the Internet test group, two players complimented the ‘feel’ of the virtual worlds. One admired the water and sky effects, and the other spent a fair amount of time recreating different landscapes. Despite making use of the repeated tile method for this test, none of the users were aware of the repetition, which meant that this method is successful and efficient for terrains that are explored by the user for less than five minutes. Further testing would be needed for the effectiveness of this method over longer periods of time, however.

## IX. CONCLUSIONS AND FUTURE WORK

This paper discussed a number of methods that can be used to create realistic but also randomised 3D environments for serious games. These methods referred

to the automated generation of fractal terrains, vegetation and building structures. To prove the feasibility of the techniques, an interactive flight simulator has been implemented and evaluated. Initial results with two different types of user groups (remote and hallway) showed that overall the flight simulator is enjoyable, looks realistic for a gaming scenario and thus also has the potential to be used for the development of serious games.

In the future, a classification regarding buildings and vegetation will be developed allowing for automatic random generation of larger urban environments. To improve the cognitive perception of the players, additional urban geometry will be generated automatically in the game such as: streets, pavements, signs etc., similar to the framework proposed by Smelik et al. [28]. Finally, scenarios will be developed and more evaluation studies will be performed with more users.

#### ACKNOWLEDGEMENTS

The authors would like to thank the Interactive Worlds Applied Research Group (iWARG) members for their support and inspiration. A video that illustrates the application in action can be found at: <http://www.youtube.com/watch?v=6G1NSALgSEY>

#### REFERENCES

- [1] Anderson, E.F., McLoughlin, L., Liarokapis, F., Peters, C., Petridis, P., de Freitas, S. Serious Games in Cultural Heritage, Proc. of the 10<sup>th</sup> Int'l Symposium on Virtual Reality, Archaeology and Cultural Heritage, VAST-STAR, Short and Project Proceedings, Eurographics, Malta, 22-25 September, 29-48, (2009).
- [2] Zyda, M. From visual simulation to virtual reality to games. *IEEE Computer* 38(9): 25-32, (2005).
- [3] Krause, D. Serious Games – The State of the Game, The relationship between virtual worlds and Web 3D, White Paper, Pixelpark Agentur, (2008).
- [4] Sawyer, B. Serious Games: Improving Public Policy through Game-based Learning and Simulation, Foresight and Governance Project, Available at: [<http://www.seriousgames.org/images/seriousarticle.pdf>], Accessed at: 26/10/2009.
- [5] Smelik, R.M., de Kraker, K.J., Groenewegen, S.A., Tutenel, T., Bidarra, R. A Survey of Procedural Methods for Terrain Modelling, Proc. of the CASA Workshop on 3D Advanced Media In Gaming And Simulation (3AMIGAS), (2009).
- [6] Shankel, J. Fractal Terrain generation – Fault Formation, Game Programming Gems, Charles River Media, 499-502, (2000).
- [7] Perlin, K. An Image Synthesizer. Proc. of ACM SIGGRAPH '85, 287-296, (1985).
- [8] Mandelbrot, B. The Fractal Geometry of Nature, W H Freeman, New York, (1983).
- [9] Fournier, A., Fussel, D., Carpenter, L. Computer Rendering of Stochastic Models. *Communications of the ACM*, 25(6): 371-384, (1982).
- [10] Miller, G.S.P. The definition and rendering of terrain maps, Proc. of ACM SIGGRAPH '86, ACM Press, 39-48, (1986).
- [11] Peitgen, H. Saupe, D. The Science of Fractal Images, Springer-Verlag, (1998).
- [12] Kelley, A.D., Malin, M.C., Nielson, G.M. Terrain simulation using a model of stream erosion. Proc. of ACM SIGGRAPH '88, ACM Press, 263-268, (1988).
- [13] Musgrave, F.K., Kolb, C.E., Mace, R.S. The synthesis and rendering of eroded fractal terrains. *Computer Graphics* 23(3): 41-50, (1989).
- [14] Belhadj, F. Terrain Modeling: A Constrained Fractal Model. Proc. of the 5<sup>th</sup> Int'l Conference on Computer Graphics, Virtual Reality, Visualisation and Interaction in Africa, Grahamstown, South Africa, 197-204, (2007).
- [15] Prusinkiewicz, P., Lindenmayer, A. The Algorithmic Beauty of Plants, Springer-Verlag New York, Inc, (1990).
- [16] Hecht, E. Optics, 2<sup>nd</sup> Edition, Addison Wesley, 603, (1987).
- [17] Di Giacomo, T., Capo, S. and Faure, F. An interactive forest. Proc. of the 2001 Eurographics Workshop on Computer Animation and Simulation, 65-74, (2001).
- [18] Guerraz, S., Perbet, F., Raulo, D., Faure, F., and Cani, M-P. A Procedural Approach to Animate Interactive Natural Sceneries. Proc. of the 16<sup>th</sup> Int'l Conference on Computer Animation and Social Agents (CASA 2003), 73-78, (2003).
- [19] Lintermann, B. and Deussen, O. Interactive Modeling of Plants. *IEEE Computer Graphics and Applications* 19 (1): 56-65, (1999).
- [20] Wells, W.D. Generating Enhanced Natural Environments and Terrain for Interactive Combat Simulations. Doctoral Dissertation, Naval Postgraduate School, Monterey (CA), (2005).
- [21] Wonka, P., Wimmer, M., Sillion, F., Ribarsky, W. Instant Architecture. *ACM Transactions on Graphics* 22(3): 669-677, (2003).
- [22] Müller, P. Wonka, P., Haegler, S., Ulmer, A. Van Gool, L. Procedural Modeling of Buildings. Proc. of ACM SIGGRAPH 2006, ACM Press, 614-623, (2006).
- [23] Havemann, S. Generative Mesh Modelling. Doctoral Dissertation, Technische Universität Braunschweig, (2005).
- [24] Hahn, E., Bose, P., Whitehead, A. Persistent Realtime Building Interior Generation. Proc. of Sandbox Symposium 2006, 179-186, (2006).
- [25] Greuter, S., Parker, J., Stewart, N., Leach, G. Real-time Procedural Generation of 'Pseudo Infinite' Cities. Proc. of GRAPHITE 2003, ACM SIGGRAPH, 87-94, (2003).
- [26] Parish, Y.I.H., Müller, P. Procedural Modeling of Cities. Proc. of ACM SIGGRAPH 2001, ACM Press, 301-308, (2001).
- [27] Gottschalk, S., Lin, M.C., Manocha, D. OBBTree: A Hierarchical Structure for Rapid Interference Detection, Proc. of SIGGRAPH '96, ACM Press, 171-180, (1996).
- [28] De Freitas, S. Serious Virtual Worlds - A scoping study, JISC, (2008), Available at: [<http://www.jisc.ac.uk/media/documents/publications/seriousvirtualworldsv1.pdf>], Accessed at: 26/10/2009.
- [29] Smelik, R.M., Tutenel, T., de Kraker, K.J., Bidarra, R. A procedural Terrain Modelling Framework, Poster Proc. of the Eurographics Symposium on Virtual Environments EGVE08, 39-42, (2008).

## **8.2 Paper #2**

Noghani, J., Anderson, E., Liarokapis, F. Towards a Vitruvian Shape Grammar for Procedurally Generating Classical Roman Architecture, Proc. of the 13<sup>th</sup> International Symposium on Virtual Reality, Archaeology and Cultural Heritage VAST 2012, Short and Project Papers, Eurographics, Brighton, UK, 19-21 November, 41-44, 2012.

Contribution (30%): Design of the architecture and advice on the implementation. Collaboration on writing of the paper.

# Towards a Vitruvian Shape Grammar for Procedurally Generating Classical Roman Architecture

Jeremy Noghani<sup>1</sup>, Eike Falk Anderson<sup>2</sup>, Fotis Liarokapis<sup>1</sup>

<sup>1</sup>Interactive Worlds Applied Research Group (iWARG), Coventry University, United Kingdom

<sup>2</sup>The National Centre for Computer Animation (NCCA), Bournemouth University, United Kingdom

## Abstract

*The ability to automatically generate urban virtual environments using procedural methods is important for anyone who needs to create plausible virtual representations of human settlements, and these techniques are used in simulation and reconstruction of archaeological sites as well as in education and the entertainment industry. In this paper we present an attempt at providing a shape grammar based on the writings of the ancient Roman architect Vitruvius, encoding rules for procedurally defining the make-up of Roman settlements. Our initial results allow the procedural generation of classical Roman Temples, which include many of the architectural elements found in Roman civic buildings.*

## 1. Introduction

Virtual recreations of cultural heritage sites are increasingly common, being found not just in archaeological reconstruction [Baw10], but also in educational serious games [AML\*10], as well as entertainment products. Due to the high costs that are associated with the manual creation of content for rich virtual worlds, automation of the content generation using procedural modelling techniques [NLA10] are becoming increasingly attractive. This is especially true in the case of cultural heritage visualization, where budgets are usually far smaller than those encountered in the entertainment industry. Procedural modelling enables the fast creation of large environments, including both urban and rural environments, based on a set of predefined rules, and over the past few years, virtual heritage has started utilizing these techniques in order to create heritage content in a fast but also accurate manner [HMVG09]. The reconstruction of ancient buildings, which often only exist as ruins or foundations, often relies on interpretation, making it impossible to define an accurate set of procedural generation rules. However, sometimes there are additional sources, such as architectural texts [HMVG09, DG11], which can be used to design procedural rules that can provide a ‘best guess’ attempt at reconstruction of lost buildings.

With his ten books on architecture [VPBC], the Roman ar-



**Figure 1:** Roman Settlement generated by our initial prototype.

chitect Vitruvius (Marcus Vitruvius Pollio) provides us with a “textual coding of the practice of building” [Pat97] from which one can extrapolate the rules for the construction of Roman settlements, which in turn can be used for the procedural generation of 3D Roman settlements using modern CGA (Computer Generated Architecture) techniques. These rules can be used to create plausible visual representations of typical Roman settlements following classical Roman architecture, as built in the late republican or early imperial era (Figure 1).

In this paper we present our attempt at developing a shape



grammar [WWSR03] inspired by ‘CGA Shape’ [MWH\*06] for describing Roman settlements derived from the writings of Vitruvius, initially with a focus on the description of classical Roman temples (meaning the main building of a religious site, excluding its courtyard). The construction of Roman temples included a large number of common elements found in Roman architecture, e.g. palaces shared many of these and often also incorporated temples themselves [Bar96]. Structures generated from these Vitruvian rules can provide an exemplar for archetypal Roman architecture in a similar manner as the “Virtual Egyptian Temple” by Jacobson and Holden [JH05] depicts architecture in ancient Egypt.

## 2. Ancient Rome in Virtual Heritage

In recent years, there have been a number of virtual heritage projects, depicting Roman urban environments (Figure 2). In many cases these projects use hand-crafted building models that are placed into the urban environment [MFPC07, WJD10], while other projects have employed range imaging techniques to capture virtual representations from models [GFD\*05] or real archaeological sites.

There are a few examples of the use of procedurally generated Roman architecture such as Müller et al. [MVUVG05], and some projects employed some procedural rules based on the writings of Vitruvius. De Paoli and Bogdan [DPB99] for instance translated the Vitruvian description of Roman theatres into a modelling language, and similarly, Vitruvian rules regarding proportions of temples were used for procedurally generating buildings about which little was known from the archaeological record in the “Rome Reborn 2.0” project [DFM\*10].

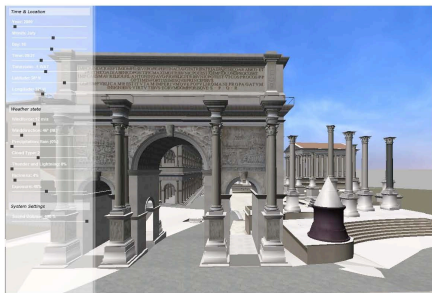


Figure 2: ‘Roma Nova’ serious game.

For any virtual heritage project depicting Roman urban environments, a method of generating the virtual world – including the archaeological sites and artefacts that will be depicted in it – must be chosen. Depending on the extent of the virtual environment, manual creation of sites and artefacts may be too costly or time consuming, which is where procedural content generation can provide solutions. The architectural rules and guidelines provided by Vitruvius, translated

into rules for procedural generation, can therefore be used to efficiently generate Roman buildings and settlements. For the procedural generation of architecture and urban environments, the use of shape grammars has proven very successful [MWH\*06, LWW08], and the availability of a shape grammar incorporating all of the writings of Vitruvius would be highly desirable.

### 2.1. Identifying Architectural Rules for Roman Temples according to Vitruvius

In his descriptions of Roman architecture, Vitruvius explains the makeup of buildings both in terms of proportions, as well as actual measurements, the latter usually expressed in terms of the pes, i.e. the Roman foot, which was a unit of length of approximately 29.6cm [DJ80, Ada94]. Two of Vitruvius’ books are devoted to the construction of temples (De Architectura III and IV), with rules that can be used to procedurally generate temples (Figure 3). There are several examples of classical ‘Vitruvian’ temples that can be found in the archaeological record, such as the Maison Carrée in Nîmes [Jon00], which is still standing, or the harbour temple of the Roman city Colonia Ulpia Traiana (Xanten) [HRG09].



Figure 3: Low-resolution Vitruvian Temple generated by our system prototype.

The dominant feature of a temple is its main building, which in most Roman settlements would be built within courtyard and enclosure wall. While the overall makeup of this usually followed the same pattern – the ‘cella’ (the temple building’s enclosed room), fronted or surrounded by a portico and raised on top of a podium – there is considerable architectural variation possible in Roman temple construction. Temple buildings can be built on a podium with steps only at the front (usually facing eastwards) or with steps on all four sides, with the number of steps in both cases being an odd number. The temples proportions would be such that the length of the main temple buildings would be twice the width of the temple with the length of the temple’s cella being 25 percent larger than the overall width of the temple.

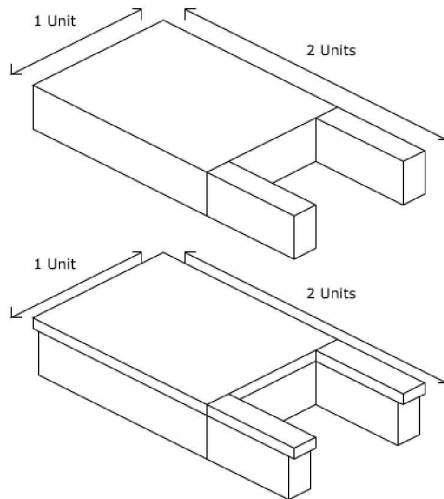


Figure 4: Proportions of Vitruvian Temple Pedestals.

### 2.1.1. Temple Base

The temple base consists of a simple podium that acts as a platform for its cella and portico, and 2 additional podium structures that act as banisters to the stairs. These structures can be seen geometrically as a set of 3 cuboid shapes, with 3 thinner, slightly wider cuboids resting on top. As per Vitruvius' rules, the length of the whole base is equal to twice its width (Figure 4).

### 2.1.2. Cella

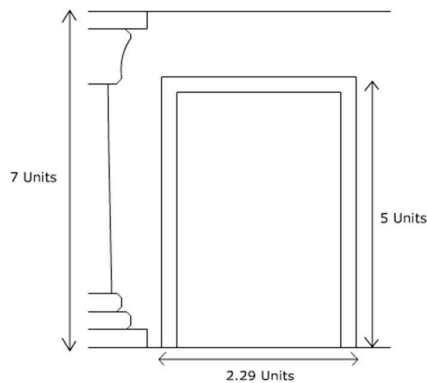


Figure 5: Proportions of the Cella front in Vitruvian Temples.

The temple's cella can be generated from a solid cuboid shape that rests between the temple base and its roof. It typically has a single doorway that faces towards the temple steps. The layout and shape of the doorway varies between temples, but Vitruvius specifies that if we assume that the

height of the cella is 3.5 units, the height of the doorway ought to be 2.5 of these units. Additionally, if the doorway height is divided into 12 equal parts, its width ought to be 5.5 of parts. Therefore, we can state that the ratio of the door width to door height to cella height is approximately 2.29:5:7 (Figure 5).

### 2.1.3. Temple Steps

Vitruvius notes that all temple steps would have to be consistent in size, with the step height being between  $5/6$  and  $3/4$  of a foot and the depth of steps being between 1.5 and 2 feet. These measurements indicate that the ratio of the step's depth to height is equal to approximately 2:1, which we used as the ratio for the generation of the model steps. During procedural generation, the stairway can be represented as a repeated set of 2-polygon steps (Figure 6). In addition to this, an altar is frequently positioned at the base of the stairs or else in front of them. The size, shape, and detail of the altar varies between temples, but the overall structure can be represented as a small cuboid upon a larger one.

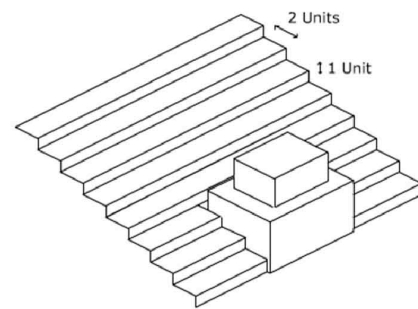


Figure 6: Proportions of Vitruvian Temple Steps.

### 2.1.4. Columns

In his third and fourth book, Vitruvius discusses at length the different styles of columns and how they are used in different temples. In general, the columns of a temple's portico have precisely calculated proportions; the diameter at the base is equal to  $1/7$  of the column height, and the diameter at the top of the column is equal to  $3/4$  of the diameter at the bottom. Additionally, the column base height is equal to half the diameter at the bottom of the column, and the whole column height should equal  $1/3$  of the temple width. The specific nature of these rules makes them simple to implement into a procedurally generated environment. However, the complex nature of the intricate details especially of Corinthian columns requires a much more detailed shape grammar than implemented in our initial prototype.

### 2.1.5. The Siting of Roman Temples

Vitruvius provides instructions on siting and orienting temples within a settlement. He “importantly defined the structural separation of the buildings in the city by indicating where each element falls” [Lei08], i.e. a differentiation of districts for domestic or public use. This separation was ideally governed by two major roadways of the city, intersecting at ninety degrees, with the ‘forum’ adjacent to this intersection (De Architectura I, 7). In Roman cities the forum played a central role in the location of communal facilities, with public buildings, including civic, religious and recreational buildings located in close proximity to the forum (De Architectura I, 7; De Architectura V), or along the major axes of the town. Vitruvius states that the main sacred site of a city with temples for the highest deities and the protectors of the settlement should be at its highest point of elevation (if there is a hill), easily reachable from the forum or the main roadways, and in most settlements, the location of this temple district would be adjacent to the forum. Vitruvius makes suggestions as to where the temples dedicated to different deities should be located. The temple orientation should be such that unless it is constricted by surrounding buildings, its front should face eastwards. If the makeup of the site makes this impossible then the doorway of the cella should face passing roads (the entrance being perpendicular to the road) or if possible allow as much of the extent of the city walls as possible to be seen from the cella entrance.

### 3. Conclusion

In this paper we have presented some of the rules of classical Roman architecture described in the writings of Vitruvius on the construction of temples, which provide guidelines for a shape grammar to allow the procedural generation of Roman buildings (Figure 3). This forms the core of a much larger grammar encoding rules for the procedural generation of Roman settlements (Figure 1), which we hope to refine in the future to allow the generation of much more intricate architectural features.

### References

- [Ada94] ADAM J.-P. F.: *Roman Building: Materials and Techniques*. Indiana University Press, 1994. 2
- [AML\*10] ANDERSON E. F., McLOUGHLIN L., LIAROKAPIS F., PETERS C., PETRIDIS P., DE FREITAS S.: Developing serious games for cultural heritage: A state-of-the-art review. *Virtual Reality* 14, 4 (2010), 255–275. 1
- [Bar96] BARTON I. M.: Palaces. In *Roman Domestic Buildings*. University of Exeter Press, 1996, pp. 91–120. 2
- [Baw10] BAWAYA M.: Virtual archaeologists recreate parts of ancient worlds. *Science* 327 (2010), 140–141. 1
- [DFM\*10] DYLLA K., FRISCHER B., MUELLER P., ULMER A., HAEGLER S.: Rome reborn 2.0: A case study of virtual city reconstruction using procedural modeling techniques. In *CAA 2009. Making History Interactive. 37th Proceedings of the CAA Conference* (2010), pp. 62–66. 2
- [DG11] DAS V. M., GARG Y. K.: Digital reconstruction of pavilions described in an ancient indian architectural treatise. *J. Comput. Cult. Herit.* 4, 1 (2011), 1:1–1:16. 1
- [DJ80] DUNCAN-JONES R. P.: Length-units in roman town planning: The pes monetalis and the pes drusianus. *Britannia* 11 (1980), 127–133. 2
- [DPB99] DE PAOLI G., BOGDAN M.: The front of the stage of vitruvius’ roman theatre – a new approach of computer aided design that transforms geometric operators to semantic operators. In *Proceedings of the Eighth International Conference on Computer Aided Architectural Design Futures* (1999), pp. 321–333. 2
- [GFD\*05] GUIDI G., FRISCHER B., DE SIMONE M., CIOCI A., SPINETTI A., CAROSSO L., MICOLI L. L., RUSSO M., GRASSO T.: Virtualizing ancient rome: 3d acquisition and modeling of a large plaster-of-paris model of imperial rome. In *Video-metrics VIII* (2005), vol. 5665 of *SPIE*, pp. 119–133. 2
- [HMGV09] HAEGLER S., MÜLLER P., VAN GOOL L.: Procedural modeling for digital cultural heritage. *J. Image Video Process.* 2009 (2009). 1
- [HRG09] HEIMBERG U., RIECHE A., GROTE U.: *Colonia Ulpia Traiana – Die römische Stadt – Planung, Architektur, Ausgrabung*. No. 18 in *Führer und Schriften des LVR-Archäologischen Parks Xanten*. Rheinland Kultur GmbH, 2009, pp. 71–73. 2
- [JH05] JACOBSON J., HOLDEN L.: The virtual egyptian temple. In *ED-MEDIA: Proceedings of the World Conference on Educational Media, Hypermedia & Telecommunications* (2005). 2
- [Jon00] JONES M. W.: *Principles of Roman Architecture*. Yale University Press, 2000, pp. 66–69. 2
- [Lei08] LEIGHT M.: Investigating urban centers in the ancient roman empire: City design based on the writings of vitruvius. Undergraduate thesis, West Virginia University, 2008. 4
- [LWW08] LIPP M., WONKA P., WIMMER M.: Interactive visual editing of grammars for procedural architecture. *ACM Trans. Graph.* 27, 3 (2008), 102:1–102:10. 2
- [MFPC07] MAGNENAT-THALMANN N., FONI A. E., PAPA-GIANNAKIS G., CADI-YAZLI N.: Real time animation and illumination in ancient roman sites. *The International Journal of Virtual Reality* 6, 1 (2007), 11–24. 2
- [MVUVG05] MÜLLER P., VEREENOGHE T., ULMER A., VAN GOOL L.: Automatic reconstruction of roman housing architecture. In *Proceedings of the International Workshop on Recording, Modeling and Visualization of Cultural Heritage* (2005), pp. 287–297. 2
- [MWH\*06] MÜLLER P., WONKA P., HAEGLER S., ULMER A., VAN GOOL L.: Procedural modeling of buildings. *ACM Trans. Graph.* 25, 3 (2006), 614–623. 2
- [NLA10] NOGHANI J., LIAROKAPIS F., ANDERSON E. F.: Randomly generated 3d environments for serious games. In *VS-GAMES 2010: Proceedings of the 2nd International Conference on Games and Virtual Worlds for Serious Applications* (2010), IEEE Computer Society, pp. 3–10. 1
- [Pat97] PATTERSON R.: What vitruvius said. *The Journal of Architecture* 2, 4 (1997), 355–373. 1
- [VPBC] VITRUVIUS POLLIO M.: De architectura, 27BC. 1
- [WJD10] WEIS A., JACOBSON J., DARNELL M.: The virtual theater district of pompeii. In *Computer Applications in Archaeology* (2010). 2
- [WWSR03] WONKA P., WIMMER M., SILLION F., RIBARSKY W.: Instant architecture. *ACM Trans. Graph.* 22, 3 (2003), 669–677. 2

### **8.3 Paper #3**

O'Connor, S., Liarokapis, F., Peters, C. An Initial Study to Assess the Perceived Realism of Agent Crowd Behaviour in a Virtual City, Proc. of the 5<sup>th</sup> International Conference on Games and Virtual Worlds for Serious Applications (VS-Games 2013), IEEE Computer Society, Bournemouth, UK, 11-13 September, 85-92, 2013.

Contribution (30%): Collaboration on the design of the architecture and advice on the experimental part. Collaboration on writing of the paper.

# An Initial Study to Assess the Perceived Realism of Agent Crowd Behaviour in a Virtual City

Stuart O'Connor  
Interactive Worlds Applied  
Research Group (iWARG)  
Coventry University  
Coventry, CV1 5FB, UK  
[oono13@uni.coventry.ac.uk](mailto:oono13@uni.coventry.ac.uk)

Fotis Liarokapis  
Interactive Worlds Applied  
Research Group (iWARG)  
Coventry University  
Coventry, CV1 5FB, UK  
[F.Liarokapis@coventry.ac.uk](mailto:F.Liarokapis@coventry.ac.uk)

Christopher Peters  
HPCViz, CSC  
KTH Royal Institute of  
Technology  
Stockholm, Sweden  
[chpeters@kth.se](mailto:chpeters@kth.se)

**Abstract**— This paper examines the development of a crowd simulation in a virtual city, and a perceptual experiment to identify features of behaviour which can be linked to perceived realism. This research is expected to feedback into the development processes of simulating inhabited locations, by identifying the key features which need to be implemented to achieve more perceptually realistic crowd behaviour. The perceptual experimentation methodologies presented can be adapted and potentially utilised to test other types of crowd simulation, for application within computer games or more specific simulations such as for urban planning or health and safety purposes.

*Keywords* – crowd simulation, perceptual studies, artificial intelligence, agent behaviour, virtual environments.

## I. INTRODUCTION

Simulating vast crowds of agents within a virtual environment is a challenging endeavour from a technical perspective [1]; however it becomes even more difficult when the subjective nature of viewer perception is also taken into account. Agent behaviour is the product of artificial intelligence systems working in tandem; nevertheless the sophistication of these systems is not a guarantee of achieving believable behaviour [2]. Within a medium such as computer games that require viewer immersion [3], the perceived realism of agent behaviour is a crucial factor for consideration. The specific features of implemented behaviours may have a great impact towards creating a believable scene.

Crowd simulation is the process of populating a virtual scene with a large number of intelligent agents that display behaviour in a manner not dissimilar from a real person within the same context [4]. Realism in the context of crowd simulations and computer games has been given multiple definitions over time, making it a difficult factor to define for measurement. There are two presented definitions [5], one considering plausibility in terms of the graphically quality and the other considering plausibility in terms of the similarities to reality. Recent research considers these types of realism within predefined virtual environments, as well as the perceptual effects [6]. Perceived realism is one definition for realism within the context of simulations and computer games. It is the plausibility of an aspect or a

feature when perceived by a human viewer, and can have a varying level of intensity depending on whether it is perceptually realistic or not.

As it is a general definition it can be applied to different features such as 3D models or lighting, but for the purposes of this research it is applied to aspects of agent crowd behaviour. Since this type of realism is perceptually based, it can be measured by applying psychophysical testing methodologies. In this paper, an overview of the types of perceptual experiments that can be utilised to assess the perceived realism of agent crowd behaviour within a virtual simulation are presented, in addition to preliminary results. The core research challenge is how to assess the perceived realism of agent crowd behaviour within a virtual urban environment.

To carry out the psychophysical experimentation a platform was developed in the form of the urban crowd simulation utilising the C++ programming language and the OpenGL graphics library, both of which are commonly used for developing computer games. For this simulation, behaviour features are added consistently through the methodology of analysis, synthesis and perception [7], allowing for the definition of parameter spaces and customisability within the platform specifically for perceptual experimentation.

As the perceived realism of agent crowd behaviour is evaluated through the features that shape behaviour traits, for example velocity type, behavioural annotation and so on, graphical complexity is not essential to the core of the research. It is highly important that a crowd simulation is perceived to be realistic by human viewers or else plausibility will be lost, which is especially true for computer games that require a level of immersion [8]. This research investigates the perceived realism of agent behaviour within an urban environment through perceptual experimentation, to identify features of behaviour that are the most effective for ensuring the perceptual plausibility of the virtual scene.

The paper is structured as follows. Section II presents related research on crowd simulation and computer games. The methodology is detailed in Section III and Section IV describes the implementation of the urban crowd stimulation. The perceptual experiments and the psychophysical methods, along with preliminary results are outlined in Section V. Finally, Section VI provides conclusions and the direction of future research.

## II. BACKGROUND

Crowds can be simulated for numerous purposes like health and safety, where the simulated agents are utilised to test for current or possible dangers [9]. The identification of these dangers is used to improve the current system or inform the development processes if the simulation is pre-emptive. There are professionally developed simulations for crowd management, evacuation procedures etc, and research [10] have been conducted into evaluating the movement paths and density of virtual crowds for urban planning purposes. These types of simulations require virtual realism whereby the simulation must be as close to reality as possible [11] or it causes inaccuracies that can lead to significant issues.

This is one of the many definitions of realism in the context of simulations and computer games. In others, such as [12], aesthetic realism is set apart from realism of representation. While this virtual realism is important for a serious simulation to achieve its purpose [13], other mediums (i.e. entertainment) require a different type of realism that is not entirely dependent on mimetic representation [14].

Computer games in particular require a sense of immersion [3] within the game world to be successful, as in [15]. It is also conveyed that for immersion, total photo and audio realism is not required for a sense of the world to be real and complete. This is where the idea of perceived realism is relevant, as it acts as a gauge for the perceptual plausibility of features within the simulation or computer game. Game designer Chris Crawford wrote that “games represent a subset of reality” [16] which can be considered true in terms of the subjective nature of perceived realism. This can help to ensure that the virtual scene is perceived to be real, potentially aiding immersion and flow [17].

Crowds have been a reoccurring theme in computer games over the past decade [18] and continue to be so in the present and foreseeable future. Computer games are at the forefront of consumer media, with game launches often breaking sales records. The game Call of Duty: Modern Warfare 3 [19] for example broke records when it sold over 6.5 million copies in the first twenty four hours of its release in 2011. Considering the platform, games offer intuitive yet technically advanced visual and interactive simulation. So it is not surprising that games have incorporated crowd based systems within their game play for innovation [20].

In fact it is often the case that technologies from the games industry are utilised for research purposes and vice versa. As is the case with agent-based crowd simulation in airports using game technology [21], research aimed at simulating the navigational traits of agents within a special public location while maintaining an interactive frame-rate. Simulating crowds in games is not specific to a single genre either, with the action adventure title Assassin's Creed 3 [8], the Stealth shooter game Hitman: Absolution [22], the city builder title Tropico 4 [23] and the open world shooter set within a city Grand Theft Auto IV [24], all having some crowd

elements in game play. In Hitman crowds are utilised for blending into the game world to assassinate targets or hide from pursuers (Figure 1).



Figure 1 Crowds in the China Town level of Hitman: Absolution [22]

In Tropico, the crowds of agents react to the changes made within the environment by the player or from other events. In Grand Theft Auto, crowds are a living part of the city, simulated to act as normal pedestrians. The latest Assassin's Creed title takes crowd simulation in games forward by utilising the perceived realism of the in-game crowds as a core game play mechanic for its online multiplayer. In a multiplayer match there are only several differing character models for the many agents that populate crowds within the map and the players must try to hunt each other, but the catch is that they each look like one of the different character models (Figure 2).



Figure 2 A multiplayer match showing crowds and assassinations in Assassin's Creed 3 [8]

This means that often a player is identified as they do not react in a manner that is perceptually realistic to other players, where as the agents typically do. As is described in research towards understanding realism in computer games through phenomenology [25] for the game world to be perceived as real it must react in a realistic way, an especially relevant statement when considering crowd behaviour. Assassin's Creeds game play takes this sentiment to its core showing the significance of perceived realism and crowd simulation within computer games.

In the art of computer game design [16] it is highlighted that the nature of human fantasy can turn an objectively unreal situation into a subjectively real situation, which in essence may indicate that the

perceived realism of a virtual scene is a highly important aspect that is set apart from the virtual realism. Furthermore, within research into immersion and presence in computer games [3], it is noted that areas for further research include the links between immersion and perception showing that investigating the perceived realism of agent crowd behaviour within a virtual city is a viable line of enquiry.

### III. METHODOLOGY

To assess the perceived realism of agent crowd behaviour, a general three stage methodology was employed. This allowed for the development of the urban crowd simulation as an interactive process, meaning features could be added over time and perceptually tested. This forms a cycle allowing for a corpus of data to be collected, while at the same time adding more sophistication to the simulation. There are three distinct aspects to this methodology including: *analysis*, *synthesis* and *perception* [7], as illustrated below:

- **Analysis:** Identify a feature and inform algorithm construction, by analysing real-world and similar instances of crowd behaviour.
- **Synthesis:** Synthesise a new simulation with further refinement and the behaviour impacting feature that was identified in analysis.
- **Perception:** Conduct the psychophysical experiment for gauging the perceived realism values of the added feature.

As the methodology was employed the most obvious features were identified as part of analysis. The first core feature distinguished in analysis was the varying velocity of agents, due to the fact that when one looks towards reality it is easily identified that pedestrians move at different rates. This may seem like a simplistic choice but within each feature there is the depth of parameter space for customisability. In this instance, what is the maximum velocity? What is the minimum velocity? Is there a specific velocity range that is most effective? Should the distribution of velocities be closer to the maximum or minimum? This type of methodology therefore allows each feature to be added and psychophysically tested individually, enabling perceptual study before other features are added to the system.

After the feature is identified in analysis, it is then implemented into the urban crowd simulation with the specific parameter spaces and customisability required. The urban crowd simulation is discussed in detail in Section IV. In the case of varying velocity parameter spaces the minimum and maximum velocity was added, along with a value to control the distribution of velocities within the agent populous. The output stimuli produced in this stage are video clips of the same virtual scene with the different configurations required for the experimentation methods. These configurations are generally the different parameter value set-ups for the newly added feature but can include different features as required.

Once the output from the synthesis stage is acquired the primary psychophysical experiment can be conducted as part of the perception stage. The aim of the experiment is to acquire the highest level of perceived realism for the newly implemented feature so that it can be linked to a specific configuration and set of values. This will allow for the setup to be replicated, helping to ensure perceptual plausibility can be obtained in other instances.

There are two other experiments that have different aims. The first, aims to identify the optimum number of features required in a simulation before it becomes overall perceptually plausible in terms of crowd behaviour. The other test will allow for the features to be ranked with regards to their overall effectiveness at implementing a sense of perceptual realism for crowd behaviour within a virtual scene. These two tests are not primary and require multiple features to be implemented to be successful. As such, they are not conducted at each iteration of the methodology but only after several new features have been added.

An online survey platform has been developed for the purposes of conducting the perceived realism experiments on large numbers of participants. It is currently in the prototype stage but has been utilised for a pilot study into the varying velocity feature, which is covered in Section V. These experiments are aimed at gauging the perceived realism values for features within the simulation to allow a set of guidelines to be shaped with the corpus of data that will allow the developers of other simulations or computer games to implement high quality of perceived realism in terms of the agent crowd behaviour.

### IV. URBAN CROWD SIMULATION

The main purpose for developing the urban crowd simulation was to create a platform with alterable parameters capable of customising agent behaviour for the purposes of perceptual evaluation. Since the primary aim is to probe human perception, the standard modelling and behaviour approaches had to be altered to accommodate the fact that more stimuli were needed than just the configuration that appeared most realistic to the developer. The real-time simulation of crowds has been conducted using a variety of approaches. The most common methods involve employing a series of models and algorithms working in tandem to animate each agent.

These include decision making [26], pathfinding navigation [27], local steering mechanics [28] and an agent perception system [29]. Social forces models [30] can also be utilised to enhance crowd believability under certain situations. The urban crowd simulation developed as part of this research implements a range of these techniques for simulating agents. While the general methodology adds behaviour orientated features to the system, a base containing the virtual urban environment and the core AI elements was still required. Development of the urban crowd simulation is conducted using the C++ programming language and the OpenGL graphics library. The three core components in the urban crowd

simulation are discussed in the following subsections.

#### A. Procedural City Generation

The first major aspect considered for the urban crowd simulation was the virtual urban environment. A procedural approach was used to generate the virtual city, allowing for the possibility of multiple layouts and setups. Procedural generation means that the city is automatically created based on a series of predefined rules for the general *shape*, *structure* and *layout*. It is possible to generate many different layouts, due to the parameterised nature of the approach. At the same time it is also possible to record a specific virtual environment if needed for experimentation purposes. The main benefit of procedurally generating the virtual city is that it allows for substantial complex geometry in terms of the buildings and roads, without the need to manually place them.

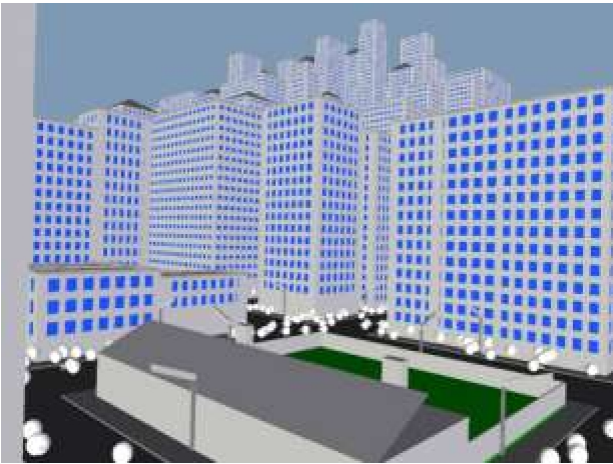


Figure 3 A procedurally generated city in the urban crowd simulation

The geometry for the architecture within the virtual city is rendered using OpenGL. For the urban crowd simulation, a procedural city modelling open source toolkit [31] written in C++ was utilised. The geometry for the architecture and layout within the virtual city is defined using the toolkit.

The layout of city is produced by land generation rules, which produce templates from subdividing quads and triangles. These templates are populated with the urban architecture models from the geometric generation rules in order to create the virtual city (see Figure 3). Other features initialised as part of the procedural generation routine include materials, light sources and camera controls. The generated virtual city is very large at around 100km<sup>2</sup> and includes three zone types: *commercial*, *residential* and *industrial*. Given the approach above for generating the urban environment, agents are introduced to populate it.

#### B. Core AI Components

There are four core AI components implemented in the urban crowd simulation: *decision making*, *pathfinding*, *steering* and *perception*. These support the real-time simulation of crowds of agents. The core AI

components are separate from the behaviour-oriented features, as they only allow for the most basic elements of operation. Using the core systems an agent can perceive, think and act [32], to select destinations and navigate through the virtual environment. Each agent is updated on a frame-by-frame basis and is modelled as an individual entity, with its own data structure containing key variables.

The decision making system is a highly important aspect in any artificial intelligence system, as it allows for the selection of a specific behaviour or action from a range of possible behaviours or actions. The decision making system discerns which of these is the most appropriate to choose at that given moment. Finite-state machines were implemented as the decision making mechanism for agents in the simulation. Currently agents follow the main paths within the environment. However the approach is extensible, so more states can be added to accommodate new behaviours.

The main purpose of pathfinding is to plan a path for an agent from its current location to the next selected location, as resolved by its decision making system. A\* pathfinding was implemented to achieve this [27]. Connected nodes are defined for the major paths within the virtual environment. The A\* algorithm calculates the path in the form of a list of nodes from the agents starting location to its destination location according to a number of heuristics.

A perception system allows an agent to sense its local environment. The sophistication of the perception system is highly dependent upon the AI systems and the features available to utilise the local data. A simple yet common approach is to designate a radius around the agent to act as its locally accessible neighbourhood. This is the type of system implemented for the urban crowd simulation.

Steering allows an agent to navigate in a local and reactive manner to dynamic obstacles. There are multiple types of steering mechanics that are suited to different purposes [28]. Here, crowd path-following behaviour has been implemented. This includes path-following and separation mechanics so that agents follow the short paths between nodes as part of the calculated A\* main path. They also, have a degree of separation to prevent them clustering and forming large masses.

#### C. Quantitative Evaluation

Given these four core AI components and the virtual environment, it is possible to simulate crowds of agents in real-time within an urban context (see Figure 4). Some behaviour-orientated features were specified as a focus of interest for the analysis. These features require parameter space and customisability for perceptual evaluation. In this work, agent velocity is the main feature, which has parameter space ranging from minimum to maximum velocities, as well as velocity distribution.

A mechanism for behavioural annotation has been implemented in preparation for the identification of future features placed in the environment, such as pedestrian crossings, roads and stationary positions.



These features will be embedded within cells in the virtual world in order to alter the behaviour of nearby agents or influence them when detected by their perception systems. For example, in the context of the current focus on agent velocities, the detection of a road could cause the agent to increase velocity in order to cross it faster. It is hoped that when these annotations are implemented they will allow a clearer contextual relationship between the agents and their local environment that can be studied as part of the perceived realism experiments.

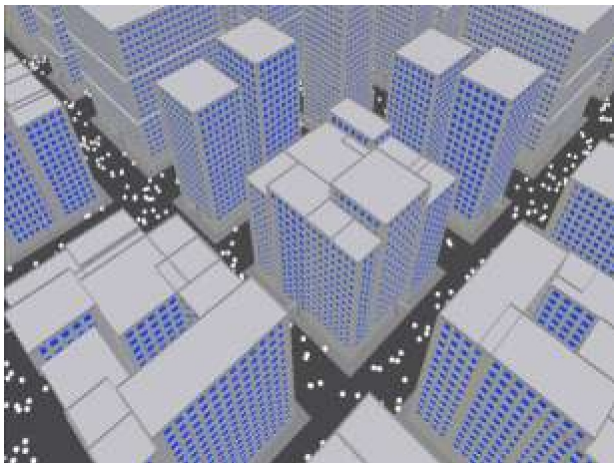


Figure 4 The urban crowd simulation displaying crowds of agents

While the next specific behaviour orientated feature to be implemented will be identified in the analysis step, there are several prominent feature considerations that will be considered for incorporation in future models. These could include physiological aspects such as age, height and weight, and psychological aspects such as emotions, internal states and predispositions. These features will be visually represented. Other features will almost certainly touch upon context related considerations and environmental awareness, adding further levels to the decision making processes in order to give agents refined objectives. These inclusions will add more depth and sophistication to agent behaviour but as a gradual process that can be examined, as outlined in the methodology.

## V. PERCEIVED REALISM EXPERIMENTS

At the core of the experimental methodology are a number of perceptual experiments that permit the exploration of perceived realism based on the parameter spaces for the implemented crowd simulation features. As the main purpose is to evaluate the plausibility of crowd behaviour, the aim is to identify thresholds for parameters and features that can produce credible virtual scenes. A corpus of data is generated in this stage which is analysed in order to rank the behaviour orientated features on their effectiveness, discern the optimum number of features for perceptual plausibility in terms of the crowds and discover the most effective configuration values for the parameters and customisability of the features. The perceptual experiments utilise

psychophysical methods for acquiring this threshold data.

“The art of psychophysics is to formulate a question that is precise and simple enough to obtain a convincing answer” [33], such that it is possible to study the perceptual effects of particular physical dimensions. To this end it is possible to investigate the limits of visual perception by parametrically varying the stimuli within a virtual scene, with the aim of measuring the thresholds and levels of realistic plausibility. The link between the level of stimuli and the subjective nature of the human response is known as the psychometric function.

### A. Psychophysical Methods

There are various psychophysical experimental methodologies ranging from the three classical methods of limits, constant stimuli and adjustment, to adaptive methods which include staircase and magnitude estimation. Here three main psychophysical methods are utilised: the constant stimuli procedure, the adjustment procedure and the staircase procedure.

The *constant stimuli procedure* is a classical psychophysical methodology where participants are asked to perceptually evaluate a stream of different levels of stimulus that are randomly shown rather than being presented in a given order such as ascending or descending. The *adjustment procedure* consists of participants taking control of the level of the stimulus in order to identify the detectable threshold. Finally, the *staircase procedure* is an adaptive type of psychophysical methodology in which the stimulus is constantly adapted to the individual participant. It involves starting at a high level of stimulus and then reducing the stimulus until the participant can notice the change, at which point there is a reversal of the staircase and the stimulus is increased until the participant notices again.

Research comparing the constant stimuli classical method to the staircase up-down adaptive method [34] has found that while both have the same accuracy of results, the staircase procedure has some advantages by automatically setting the dynamic range for the psychometric function. This is the reason why the staircase procedure is the primary experimental methodology employed in this work.

### B. Experiments

Three experiments are presented as part of this research. The first and primary experiment utilises the staircase procedure in order to establish the perceptual thresholds and perceived realism levels of a feature within the simulation. The second experiment utilises the adjustment procedure to rank the features based on their perceptual effectiveness. The third experiment utilises the constant stimuli procedure to determine the threshold and most effective number of features required for perceptual plausibility.

In these experiments the video clips obtained in the synthesis stage of the general methodology are utilised as stimuli and presented in the manner dictated by the respective psychophysical method. An online survey

platform has been developed (see Figure 5) in order for a large number of participants to perceptually evaluate the footage. The platform automates the display of video clips to the participant and provides a slider at the bottom of the page in order to collect ratings of realism. Data collected from the participants includes: initials, date of birth, gender and primary language. The number of video clips and the order in which they are shown varies according to the specific feature and type of experiment.

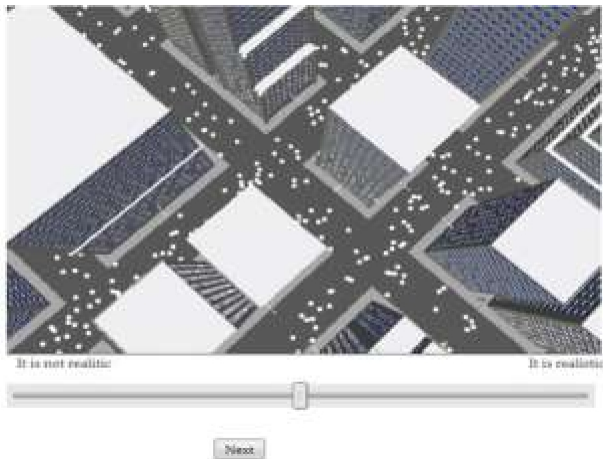


Figure 5 The survey platform prototype showing a video clip

Data is collected from the experiments in the form of a *perceived realism value* between 0 and 1 for each configuration shown, where a value of 1 maps to completely realistic and 0 maps to completely unrealistic. The manner in which the perceptual thresholds and perceptual plausibility levels are calculated varies between experiments.

Generally each configuration shown is on a scale from low to high stimuli. Here, when the perceived realism value is above 0.5 we consider the stimuli to be perceptually realistic. The general approach for calculating the thresholds is to consider the lowest and highest stimuli on the underlying scale that are perceptually plausible. From these thresholds the mean stimuli can also be calculated, which should be close to the optimum configuration. This is not necessarily always the case however. The optimum perceptually plausible level is identified as the highest average positive perceived realism response from the participants. This is the average perceived realism value closest to 1, which can then be linked to a specific configuration. Depending on the experiment type this configuration can be parameter space values, a specific feature or even a number of features. By completing these experiments, we endeavour to identify thresholds for features in addition to optimal configurations in order to create perceptually plausible crowd behaviour.

In the experiments the results will be treated fairly as perception is being tested. In terms of the bias, it may become apparent that some groups perceive the stimuli in different ways to other groups. This data will be noted and will become an important consideration in the final analysis. If this causes a skewer in the results then modifiers can be utilised for the groups that are not

represented equally, in order to combat the bias and ensure the optimum configuration is accurate as an average for all groups. A small number of outliers in later results with large pools of participants is to be expected and will not be considered an anomalous condition, unless presented in a relevant density. Results from the experiments will be statistically analysed using the general linear model, with X representing the configurations starting with low intensity stimuli and with Y representing the perceived realism value responses from participants.

### C. Perceived Realism Pilot Study

The staircase based primary experiment described above was conducted in pilot study form. Three participants were shown a series of video clips with different velocity configurations, using a prototype of the online survey platform. While within the crowd simulation system agent velocity is represented with a directional vector as well as a speed component, in the case of the results the agent velocity is essentially just the speed component as the direction would have no purpose for representation. The speed component is measured as a decimal of a single unit which is one. The speed is determined on a per second basis. This single unit can be altered depending on specific requirements, however within the urban crowd simulation a unit is the equivalent of 2.5 meters.

While the small test pool means that the results presented here are preliminary, the purpose of the study was to test the experimental approach and its viability for collecting larger, statistically valid samples. The underlying scale of stimuli for the velocity feature consisted of different configurations. The low end of the scale consisted of configurations with a small velocity range and distribution toward the minimum velocity. The high end of the scale consisted of configurations with a large velocity range and a distribution towards the maximum velocity. The order in which configurations were presented was adapted according to the psychophysical method but the general approach was to start at a high level and reduce it until the participant rated the behaviour as being unrealistic. At this point the stimuli would be increased until the participant found it to be realistic again. This was repeated several times to identify thresholds and perceived realism values.

As the velocity feature had two distinct factors, velocity range based on minimum and maximum velocities and velocity distribution, different passes were required in order to ensure both were evaluated properly. Firstly, the range was evaluated starting at a high stimuli with a large velocity range and being reduced to low stimuli of small velocity range and so on. Secondly, the velocity distribution was evaluated again starting at a high stimuli with distribution towards maximum velocity and then reduced and reversed again. The aggregate results from this experiment are as follows:

- The normalised velocity range was 0.3, based on an average perceived realism value of 0.82.
- Normalised velocity range thresholds were 0.2

and 0.5.

- The normalised velocity distribution was 0.5, based on an average perceived realism value of 0.85.
- Normalised velocity distribution thresholds were 0.3 and 0.7.

Even though these results cannot be considered accurate due to the small number of participants and therefore do not allow conclusions related to the agent velocity feature, the pilot study was appropriate for testing the viability of the experimental method. Our intention is to conduct a larger scale study to obtain statistically relevant data which can then be used to provide more useful information about the potential role of the features in viewer perception.

## VI. CONCLUSIONS AND FUTURE WORK

Details have been presented about ongoing research towards assessing the perceived realism of crowd behaviour in a virtual city. The methodology consists of the analysis of real-world and related instances of crowd behaviour, synthesis to replicate the crowd behaviour through features and output experimental stimuli, and perceptual experimentation to investigate participants subjective views of the realism of agent crowd behaviour within an urban context. It is based on the insight that there are multiple types of realism each with their own merits and that the realism of simulations and computer games should not be judged only by aesthetic means, but also by more in-depth methods that consider the content of the virtual world. Crowds not only add a sense of life and realism to virtual environments but can be used as tools for serious and entertainment purposes. It's therefore important that they can be properly evaluated to highlight specific features that make them so effective with respect to viewers.

The research is envisaged to feedback into improving the development processes of simulations and applications implementing virtual crowds, especially those within urban environments. It has been shown that computer games in particular are developing at an accelerated rate and it is hoped they could benefit from the outcomes of this research. In the broader sense, these may include identifying human mental models of crowd behaviour through perceptual experimentation, thus adding new perspectives to enable AI systems to better predict or understand real behaviour. Guidelines will be constructed through data from the perceptual experiments, which will be useful for identifying key features of crowd behaviour for implementation. The simulation thus far is a prototype in order to support the development of the perceptual experiments and allow refinements to take place to both the simulation and experimental approaches.

Currently the urban crowd simulation is in the process of enhancement to improve the structural and visual quality of stimuli. In particular, one possibility is to investigate a procedural annotation mechanism, for automatically generating information for supporting artificial behaviours within the environment when the

city is initialised. This would allow new behavioural annotation types to be added with relative ease and would enable the use of multiple city layouts since manual annotation efforts would no longer be required. Other planned improvements include more complex pedestrian models and new behavioural features such as collision avoidance and a social forces model. These additions will need to be carefully managed however, as experimentally they can influence participant expectation in relation to crowd behaviour plausibility. In relation to this aspect, the use of simple stimuli thus far in the experimentation is important and could be applied across other simulation and application contexts.

In the future, the experiment will be evaluated by a larger sample of participants. This will be achieved by launching the survey platform online. In addition, guidelines for implementing perceptual plausibility in terms of agent crowd behaviour will be investigated.

## ACKNOWLEDGEMENTS

The authors would like to thank the Interactive Worlds Applied Research Group (iWARG) at Coventry University, Faculty of Engineering and Computing, UK as well as Dr. Etienne Roesch for their support and inspiration. A video that illustrates the system in operation can be found at: <http://www.youtube.com/watch?v=BqnmMksd30I&feature=youtu.be>

## REFERENCES

- [1] Leggett, R. "Real Time Crowd Simulation: A Review", Retrieved, March 20, (2004).
- [2] McDonnell, R., Newell, F., O' Sullivan, C. "Smooth movers: Perceptually guided human motion simulation", Proc. of the 2007 ACM SIGGRAPH/ Eurographics symposium on computer animation, 259-269, (2007).
- [3] Jennett, C, I. Cox, A, L. Cairns, P. "Being In The Game". Proc. of the Philosophy of Computer Games, 210-227. (2008).
- [4] Peters, C., Ennis, C., McDonnell, R., O'Sullivan, C. "Crowds in context: Evaluating the perceptual plausibility of pedestrian orientations", Proc. of Eurographics Short Papers, Springer-Verlag, Berlin, 227-230, (2008).
- [5] Barlett, C. Rodeheffer, C. "Effects of Realism on Extended Violent and Nonviolent Video Game Play on Aggressive Thoughts, Feelings and Physiological Arousal", *Aggressive Behaviour*, 35, 213-224, (2009).
- [6] Peters, C. Ennis, C. "Modelling groups of plausible virtual pedestrians", *IEEE Computer Graphics and Applications*, Special Issue on Virtual Populace, IEEE Computer Society, 29(4): 54-63, (2009).
- [7] Ennis, C., Peters, C., O' Sullivan, C. "Perceptual effects of scene context and viewpoint for virtual pedestrian crowds", *ACM Transactions Applied Perception*, 8(2), Article Number 10, (2011).
- [8] Assassin's Creed 3, Ubisoft Montreal, (Available at: <http://assassinscreed.ubi.com/ac3/en-gb/index.aspx>), Accessed at: 28/03/2013.
- [9] Almeida, J.E., Rossetti, R., Coelho, A.L. "Crowd Simulation Modeling Applied to Emergency and Evacuation Simulations using Multi-Agent Systems",

- Proc. of 6<sup>th</sup> Doctoral Symposium on Informatics Engineering, Porto, Portugal, 93-104, (2011).
- [10] Aschwanden, G., Halatsch, J., Schmitt, G. "Crowd simulation for urban planning", eCAADe 2008, Antwerp, (2008).
- [11] Banerjee, B., Kraemer, L. "Evaluation and comparison of multi-agent based crowd simulation systems". Agents for games and simulations II, Frank Dignum (eds), Springer-Verlag, Berlin, Heidelberg, 53-66, (2011).
- [12] Galloway, A.R. "Social Realism in Gaming", The International Journal of Computer Game Research, Volume 4, Issue 1. (2004).
- [13] Chalmers, A. "Level of Realism for Serious Games", Proc. of the IEEE Int'l Conference on Games and Virtual Worlds for Serious Applications (VS-Games 09), IEEE Computer Society, 225-232, (2009).
- [14] Sommerseth, H. "Gamic Realism: Player, Perception and Action in Video Game Play", Situated Play, DiGRA Conference, (2007).
- [15] McMahan, A. "Immersion, Engagement and Presence", The Video Game Theory Reader, Mark J.P. Wolf and Bernard Perron, (eds), New York, NY: Routledge, 77-78. (2003).
- [16] Crawford, C. "The Art of Computer Game Design", McGraw-Hill/Glencoe, (1982).
- [17] Cowley, B., Charles, D., Black, M., Hickey, R. "Towards an Understanding of Flow in Video Games", Computers in Entertainment, ACM Press, 6(2), Article 20, (2008).
- [18] Ennis, C., McDonnell, R., O' Sullivan, C. "Seeing is believing: Body motion dominates in multi-sensory conversations", ACM Trans Graph 29(4):19, (2010).
- [19] Infinity Ward, Call of Duty: Modern Warfare 3, Activision, (Available at: <http://www.callofduty.com/mw3>), Accessed at: 28/03/2013.
- [20] Bernard, S., Therien, J., Malone, C., Beeson, S., Gubman, A., Pardo, R. "Taming the Mob: Creating believable crowds in Assassin's Creed", Game Developers Conference, San Francisco, CA. Feb 18-22. (2008).
- [21] Szymanczyk, O., Dickinson, P., Duckett, T. "Towards Agent-based Crowd Simulation in Airports using Games Technology", Agent and MultiAgent Systems Technologies and Applications, Volume: 6682, 524-533. (2011).
- [22] Hitman: Absolution, Square Enix, IO Interactive, (Avaliable at: <http://hitman.com/>), Accessed at: 28/03/2013.
- [23] Tropico 4, Kalypso Media, Haemimont Games, (Available at: <http://www.tropico3.com/en/T4/en/index.php>), Accessed at: 28/03/2013.
- [24] Grand Theft Auto IV, Rockstar Games, Rockstar North, (Available at: <http://www.rockstargames.com/IV/>), Accessed at: 28/03/2013.
- [25] Low, G. "Understanding Realism in Computer Games through Phenomenology", Stanford University, California, (2001).
- [26] Luo, L., Zhou, S., Cai, W., Low, M., Lees, M. "Toward a Generic Framework for Modeling Human Behaviors in Crowd Simulation", Proc. of the IEEE/WIC/ACM Inter'l Joint Conference on Web Intelligence and Intelligent Agent Technology - Volume 02 (WI-IAT '09), Vol. 2. IEEE Computer Society, Washington, DC, USA, 275-278. (2009).
- [27] Cui, X., Shi, H. "A\*-based Pathfinding in Modern Computer Games", International Journal of Computer Science and Network Security, 11(1): 125-130, (2011).
- [28] Reynolds C. "Steering behaviours for autonomous characters", Proc. of game developers conference, Miller Freeman Game Group, San Francisco, California, 763-782, (1999).
- [29] Ondej, J., Pettre, J., Olivier, A.-H., Donikian, S. "A Synthetic-Vision-Based Steering Approach for Crowd Simulation", ACM Transactions on Graphics, 29(4): 123. (2010).
- [30] Helbing, D. and Molnar, P. "Social force model for pedestrian dynamics", Phys. Rev E, American Physical Society, 51(5): 4282-4286, (1995).
- [31] Lance, F., Matheossian, D., Poli, A. City Procedural Modelling Open Source Tool-kit, GitHub, (Available at: <https://github.com/Akado/City-procedural-modeling>), Accessed at: 01/03/2013
- [32] Anderson, E. "Playing smart – artificial intelligence in computer games", Proc. of zfxCON03 Conference on Game Development, (2003).
- [33] Ehrenstein, W.H., Ehrenstein, A. "Psychophysical Methods", Modern techniques in neuroscience research, 1211-1241, (1999).
- [34] Dai, H. "On measuring psychometric functions: A comparison of the constant-stimulus and adaptive up-down methods", Journal of the Acoustical Society of America, 98(6): 3135-3139, (1995).

#### **8.4 Paper #4**

Liarokapis, F., Mourkoussis, N., White, M., Darcy, J., Sifniotis, M., Petridis, P., Basu, A., Lister, P.F. Web3D and Augmented Reality to support Engineering Education, World Transactions on Engineering and Technology Education, UICEE, 3(1): 11-14, 2004.

Contribution (80%): Collaboration on the design of the architecture. Implementation of the most of the VR interface. Write-up of most of the paper.

## Web3D and augmented reality to support engineering education

**Fotis Liarokapis, Nikolaos Mourkoussis, Martin White, Joe Darcy, Maria Sifniotis, Panos Petridis,  
Anirban Basu & Paul F. Lister**

University of Sussex  
Falmer, England, United Kingdom

**ABSTRACT:** In the article, the authors present an educational application that allows users to interact with 3D Web content (Web3D) using virtual and augmented reality (AR). This enables an exploration of the potential benefits of Web3D and AR technologies in engineering education and learning. A lecturer's traditional delivery can be enriched by viewing multimedia content locally or over the Internet, as well as in a tabletop AR environment. The implemented framework is composed in an XML data repository, an XML-based communications server, and an XML-based client visualisation application. In this article, the authors illustrate the architecture by configuring it to deliver multimedia content related to the teaching of mechanical engineering. Four mechanical engineering themes (machines, vehicles, platonic solids and tools) are illustrated here to demonstrate the use of the system to support learning through Web3D.

### INTRODUCTION

Traditional methods of educating students have well-proven advantages, but some deficiencies have also been detected. A typical problem has been how to engage students with appropriate information and communication technologies (ICT) during the learning process. In order to implement innovative interactive communication and learning paradigms with students, teachers should make innovative use of new ICT [1]. Although multimedia material is provided in a number of formats, including textual, images, video animations and aural, educational systems are not designed according to current teaching and learning requirements. That requirement is to efficiently integrate these formats in well-proven means, eg through the Web. The system described here does this by introducing Web3D, virtual and augmented reality (AR) in the same Web-based learning support application.

Research into educational systems associated with the use of Web3D technologies is very limited. Web3D has the potential for a number of different applications ranging from 2D to 3D visualisation [2]. One of the most appropriate means of presenting 2D information is through the WWW Consortium [3]. On the other hand, a promising and effective way of 3D visualisation is AR, which combines computer-generated information with the real world, and it can be used successfully to provide assistance to the user necessary to carry out difficult procedures or understand complex problems [4].

An overview of existing AR systems in education and learning has been presented elsewhere [5]. A more recent educational application is an experimental system that demonstrates how to aid teaching undergraduate geography students using AR technologies [6]. An educational approach for collaborative teaching targeted at teachers and trainees that makes use of AR and the Internet has been illustrated by Wichert [7].

An educational system is presented here for improving the understanding of the students through the use of Web3D and AR presentation scenarios. An engineering and design application has been experimentally designed to support the teaching of mechanical engineering concepts such as machines, vehicles, platonic solids and tools. It should be noted that more emphasis has been given to the visualisation of 3D objects because 3D immediately enhances the process of learning. For example, a teacher can explain what a camshaft is using diagrams, pictures and text, etc. However, it still may be difficult for a student to understand what a camshaft does. In the current system's Web3D pictures, text and 3D model (which can be animated) are visualised so that the student can manipulate and interact with the camshaft, and also see other related components such as the tappets, follower, etc, arranged as they might be with an engine.

In this article, the authors present four example themes to support the teaching of engineering design. These four themes may represent different courses or different teaching sessions as part of the same course. The remainder of this article describes the requirements for augmented learning, provides a brief discussion of the presented system's architecture, and illustrates how the system might be used to support teaching processes using Web3D and AR technologies. Finally, conclusions are made and future work suggested.

### THE REQUIREMENTS OF AUGMENTED LEARNING

The requirements for virtual learning environments have been already well defined [8]. However, in AR learning environments, they have not been systematically studied. In general, any educational application requires technological, pedagogical and psychological aspects to be carefully investigated before their implementation [9]. Especially when introducing new technologies, such as Web3D and AR into the

education process, many aspects need to be considered. The authors have classified some of the most important issues that are involved in AR learning scenarios.

To begin with, the educational system must be simple and robust and provide users with clear and concise information. This will increase the level of students' understanding and their skills. Moreover, the system must provide easy and efficient interaction between the lecturer, students and the teaching material. Apart from these issues, the digitisation of the teaching material must be carried out carefully so that all of the information is accurately and clearly presented to users. This digitisation or *content preparation* is usually an offline process and consists of many different operations, depending on the target application.

The authors believe that a combination of Web3D and AR technologies can help students explore the multidimensional augmentation of teaching materials in various levels of detail. Students can navigate through the augmented information and, therefore, concentrate and study in detail any part of the teaching material in different presentation formats, thus enhancing understanding. With Web3D environments traditional teaching materials may be augmented by high quality images, 3D models, single- or multi-part models, as well as textual metadata information. An image could be a complex diagram, a picture or even a *QuickTime* movie. The 3D model allows the student to understand aspects of the teaching material that is not evident in the pictures, because they are hidden. Finally, metadata can provide descriptive information about the teaching material that cannot be provided by the picture and the 3D model.

## SYSTEM ARCHITECTURE

The system presented here can be used to create and deliver multimedia teaching material using Web3D and AR technologies. The authors have already demonstrated this in other application domains, such as virtual museum exhibitions [10]. The architecture of this system is based on an improvement of the researchers' previously defined three-tier architecture [11]. The architecture, as shown in Figure 1, consists of content production, a server and visualisation clients.

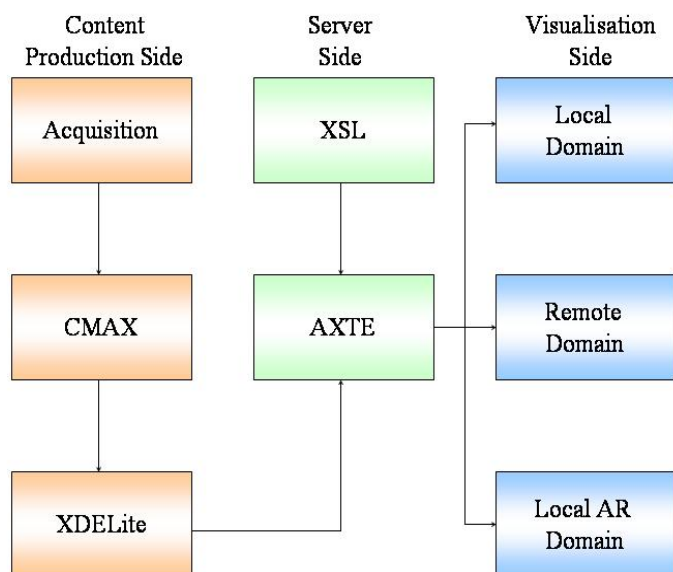


Figure 1: The three-tier architecture.

The first tier is the content production side, which consists of the content acquisition process – content can consist of 3D models, static images, textual information, animations and sounds – and a content management application for XML (CMAX) that gathers content from the file system and packages this content into an XML repository called XDELite. In the example illustrated in this article, most of the 3D models utilised were downloaded from the Internet [12]. This is quite important, because teachers should make best use of freely available content because generating 3D content can be expensive and time consuming.

The server side tier is based on XML and Java-Servlet technologies. The Apache Tomcat server was used and was configured with a Java Servlet, named the *ARCOLite* XML Transformation Engine (AXTE) [10]. The purpose of this server is to respond to user requests for data, stored in the XDELite repository, and dynamically deliver this content to the visualisation tier. XSL stylesheets are then utilised to render the content to the visualisation clients.

The client visualisation tier consists of three different visualisation domains, namely: the *local*, the *remote* and the *AR* domains. The local domain is used for delivering supporting teaching material over a Local Area Network (LAN), while the *remote domain* may be used to deliver the same presentations over the Internet, both utilising standard Web browsers.

The *AR domain* allows the presentation of the same content in a tabletop AR environment [10]. The authors have developed an application called *ARIFLite* that consists of a standard Web browser and an AR interface integrated inside a user friendly visualisation client built from Microsoft Foundation Class (MFC) libraries. The software architecture of *ARIFLite* is implemented in C++ using an Object-Oriented (OO) style. *ARIFLite* uses technologies, such as *ARToolKit's* tracking and vision libraries [13] and computer graphics algorithms based in the OpenGL API [14]. The only restriction of the AR system is that the marker cards and the camera are always in line of sight of the camera.

## USER OPERATION

The user, eg a student, accesses this system simply by typing a URL into a Web browser that addresses the index page of the presentation or launches the presentation from a desktop icon. In this case, the student will be accessing a Web3D presentation with 3D, but no AR view (see Figure 2), which illustrates the Web browser embedded in *ARIFLite*. This is the mode of operation for the Internet.

For local Web and AR use, eg in a university laboratory environment or a seminar room, the student would launch *ARIFLite* from an icon on the PC desktop. By using *ARIFLite*, the student can browse multimedia content as usual, but also extend the 3D models into the AR view. Switching to AR view causes the Web browser to be replaced with a video window in which the 3D model appears. The user can then interact with the 3D model and can compare it to real objects in a natural way, as illustrated in Figure 5.

## WEB3D PRESENTATION

Demonstration in seminars and lecture rooms is one of the most effective means of transferring knowledge to groups of

people. One of the capabilities of the presented system is to increase the level of understanding of students through interactive Web3D and AR presentation scenarios. The lecturer can control the sequence of the demonstration using the visualisation client [10]. One can imagine a group of students and the lecturer gathered around a table on which there is a computer and large screen display. The virtual demonstration starts by launching a Web browser (ie *Internet Explorer*) or *ARIFLite*. Figure 2 actually illustrates the Web browser embedded in *ARIFLite*.

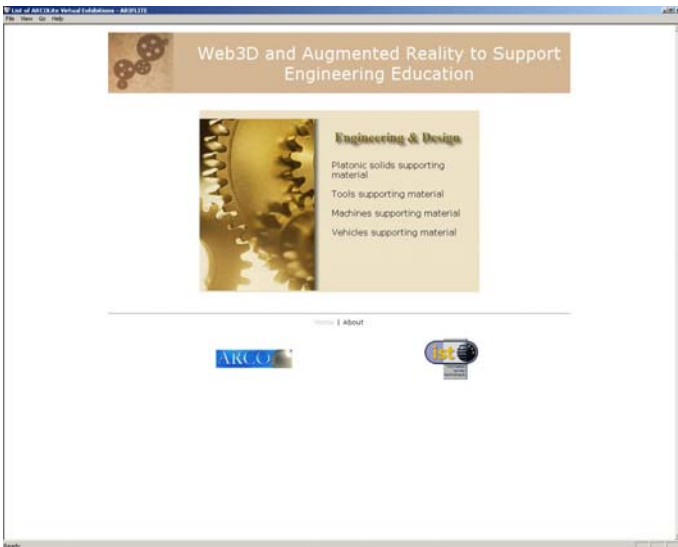


Figure 2: Web browser embedded in *ARIFLite* showing the presentation's homepage.

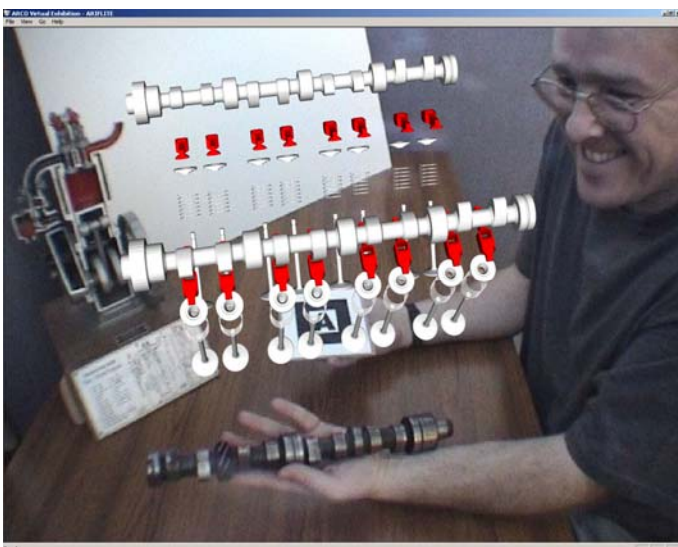


Figure 5: AR visualisation of a piston.

On the homepage, the user has the option to choose between four different supporting material themes, namely: platonic solids, tools, machines and vehicles. Each module contains a list of thumbnails representing links to relevant sub-categories, as shown in Figure 3.

Next, the user can access more specific information about any of the existing sub-categories. For example, in Figure 4, the user has clicked on the camshaft, which accesses a new Web page showing a thumbnail (that could access a larger picture or a *QuickTime* movie), description of the camshaft and an interactive 3D model displayed in an embedded VRML

browser. At this stage, the lecturer can describe the underlying theory of a camshaft while interacting with the 3D model, eg rotating, translating or scaling the model.

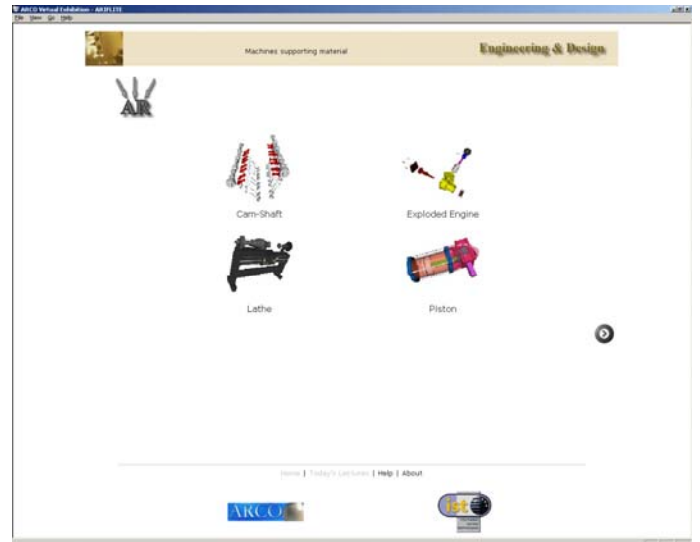


Figure3: Selection of machines.



Figure 4: Web3D visualisation.

Augmenting a Web-based presentation with 3D information (as shown in Figure 4) can enhance student understanding and allow the lecturer to present material in a more efficient manner.

## AUGMENTED REALITY PRESENTATION

By using *ARIFLite*, the authors could now extend the Web3D presentation into a tabletop AR environment. AR can be extremely effective in providing information to a user dealing with multiple tasks at the same time [15]. With *ARIFLite*, users can easily perceive visual information in a new and exciting way. In order to increase the level of understanding of the teaching material, 3D information is presented on the tabletop in conjunction with real objects. Figure 5 shows an AR view of a user examining a virtual 3D model of camshaft arrangement in conjunction with a set of real engine components.

Similarly with the system demonstrated by Kato et al, users can physically manipulate the marker cards in the environment by



just picking the markers and moving them into the real world [16]. In this way, students are able to visualise how a camshaft is arranged in relation to other engine components and examine the real components at the same time. Users can interact with the 3D model using standard I/O devices, such as the keyboard and the mouse.

In order to manipulate better the 3D model, haptic interfaces, such as 3D mouse (ie SpaceMouse XT Plus), are integrated within the system. The SpaceMouse provides an 11-button menu and a puck allowing six degrees of freedom, which gives a more efficient interface than the keyboard [17]. The user can zoom, pan and rotate virtual information as naturally as if they were objects in the real world.

## CONCLUSION AND FUTURE WORK

In this article, a simple and powerful system for supporting learning based on Web3D and AR technologies is presented. Students can explore a 3D visualisation of the teaching material, thus enabling them to understand more effectively through interactivity with multimedia content. It is believed that the presented experimental scenarios can provide a rewarding learning experience that would be otherwise difficult to obtain.

In the future, the authors plan to create more educational templates and add further multimedia content for the XML repository so as to apply the system in practice. In order to optimise the system's rendering capabilities, greater realism will be added into the augmented environment using augmented shadows. Finally, more work needs to be conducted in improving human-computer interactions by adding haptic interfaces so that the system will have a more collaborative flavour.

## ACKNOWLEDGEMENTS

This research was partially funded by the EU IST Framework V programme, Key Action III-Multimedia Content and Tools, Augmented Representation of Cultural Objects (ARCO) project IST-2000-28366.

## REFERENCES

1. Barajas, M. and Owen, M., Implementing virtual learning environments: looking for holistic approach. *Educational Technology and Society*, 3, 3 (2000).
2. Web3d Consortium (2004), <http://www.web3d.org>

3. World Wide Web Consortium (2004), <http://www.w3c.org>
4. Schwald, B. and Laval, B., An augmented reality system for training and assistance to maintenance in the industrial context. *J. of WSCG*, Science Press, 1 (2003).
5. Liarokapis, F., Petridis, P., Lister, P.F. and White, M., Multimedia Augmented Reality Interface for E-learning (MARIE). *World Trans. on Engng. and Technology Educ.*, 1, 2, 173-176 (2002).
6. Shelton, B.E. and Hedley, N.R., Using augmented reality for teaching earth-sun relationships to undergraduate geography students. *Proc. 1<sup>st</sup> IEEE Inter. Augmented Reality Toolkit Workshop*, Darmstadt, Germany (2002).
7. Wichert, R., A mobile augmented reality environment for collaborative learning and teaching. *Proc. World Conf. on E-Learning in Corporate, Government, Healthcare and Higher Education (E-Learn)*, Montreal, Canada (2000).
8. Dillenbourg, P., Virtual learning environments. *Proc. EUN Conf. 2000*, Workshop on virtual environments (2000).
9. Kaufmann, H., Collaborative augmented reality in education. *Proc. Imagina 2003 Conf. (Imagina03)*, Monaco (2003).
10. White, M., Liarokapis, F., Mourkoussis, N., Basu, A., Darcy, J., Petridis, P. and Lister, P.F., A lightweight XML driven architecture for the presentation of virtual cultural exhibitions (ARCOLite). *Proc. IADIS Inter. Conf. on Applied Computing 2004*, Lisbon, Portugal, 205-212 (2004).
11. Liarokapis, F., Mourkoussis, N., Petridis, P., Rumsey, S., Lister, P.F. and Whiet, M., An interactive augmented reality system for engineering education. *Proc. 3<sup>rd</sup> Global Congress on Engng. Educ.*, Glasgow, Scotland, UK, 334-337 (2002).
12. VRML Models (2004), <http://www.ocnus.com/models/>
13. Kato, H., Billinghamurst M. and Poupyrev, I., ARToolkit User Manual, Version 2.33, Human Interface Lab. Seattle: University of Washington (2000).
14. Woo, M., Neider, J. and Davis, T., OpenGL Programming Guide: *The Official Guide to Learning OpenGL, Version 1.2*, Addison Wesley, September, (1999).
15. Kalawsky, R.S. and Hill, K., *Experimental Research into Human Cognitive Processing in an Augmented Reality Environment for Embedded Training Systems*. London: Springer-Verlag (2000).
16. Kato, H., Billinghamurst, M., Poupyrev, I., Imamoto, K. and Tachibana, K., Virtual object manipulation on a table-top AR environment. *Proc. Inter. Symp. on Augmented Reality 2000*, Munich, Germany, 111-119 (2000).
17. SpaceMouse XT Plus (2004), <http://www.logicad3d.com/press/archive/2000/20001002.html>

## **8.5 Paper #5**

White, M., Mourkoussis, N., Darcy, J., Petridis, P., Liarokapis, F., Lister, P.F., Walczak, K., Wojciechowski, R., Cellary, W., Chmielewski, J., Stawniak, M., Wiza, W., Patel, M., Stevenson, J., Manley, J., Giorgini, F., Sayd, P., Gaspard, F. ARCO-An Architecture for Digitization, Management and Presentation of Virtual Exhibitions, Proc. of the 22<sup>nd</sup> International Conference on Computer Graphics (CGI'2004), IEEE Computer Society, Hersonissos, Crete, June 16-19, 622-625, 2004.

Contribution (15%): Collaboration on the design of the VR and AR architecture. Implementation of the most of the VR and AR interface. Write-up of parts of the paper.

# ARCO—An Architecture for Digitization, Management and Presentation of Virtual Exhibitions

Martin White<sup>1</sup>, Nicolaos Mourkoussis<sup>1</sup>, Joe Darcy<sup>1</sup>, Panos Petridis<sup>1</sup>, Fotis Liarokapis<sup>1</sup>,  
Paul Lister<sup>1</sup>, Krzysztof Walczak<sup>2</sup>, Rafał Wojciechowski<sup>2</sup>, Wojciech Cellary<sup>2</sup>,  
Jacek Chmielewski<sup>2</sup>, Mirosław Stawniak<sup>2</sup>, Wojciech Wiza<sup>2</sup>, Manjula Patel<sup>3</sup>, James Stevenson<sup>4</sup>,  
John Manley<sup>5</sup>, Fabrizio Giorgini<sup>6</sup>, Patrick Sayd<sup>7</sup>, Francois Gaspard<sup>7</sup>

*University of Sussex<sup>1</sup>, Poznan University of Economics<sup>2</sup>, UKOLN –University of Bath<sup>3</sup>, Victoria and Albert Museum<sup>4</sup>, Sussex Archaeological Society<sup>5</sup>, Giunti Publishing Group<sup>6</sup>, Commissariat a l'Energie Atomique<sup>7</sup>*

*{M.White, N.Mourkoussis, J.Darcy, P.Petridis, F.Liarokapis, P.F.Lister}@sussex.ac.uk  
{walczak, wojciechowski, cellary, jchmiel, stawniak, wiza}@kti.ae.poznan.pl  
M.Patel@ukoln.ac.uk, j.stevenson@vam.ac.uk, ceo@sussexpast.co.uk,  
f.giorgini@giuntilabs.com, {SAYD, francois.gaspard}@ortolan.cea.fr*

## Abstract

*A complete tool chain starting with stereo photogrammetry based digitization of artefacts, their refinement, collection and management with other multimedia data, and visualization using virtual and augmented reality is presented. Our system provides a one-stop-solution for museums to create, manage and present both content and context for virtual exhibitions. Interoperability and standards are also key features of our system allowing both small and large museums to build a bespoke system suited to their needs.*

## 1. Introduction

The concept of using virtual exhibitions in museums has been around for many years. Museums are keen on presenting their collections in a more appealing and exciting manner using the Internet to attract visitors both virtually and into the physical museum site. Recent surveys show that about 35% of museums have already started developments with some form of 3D presentation of objects [4].

Requirements related to the development of augmented reality (AR) applications in the Cultural Heritage field have been well documented [3]. Many museum applications based on VRML have been developed for the web [1][5][6]. An example of an interactive virtual exhibition is the Meta-Museum visualized guide system based on AR [2]. Another

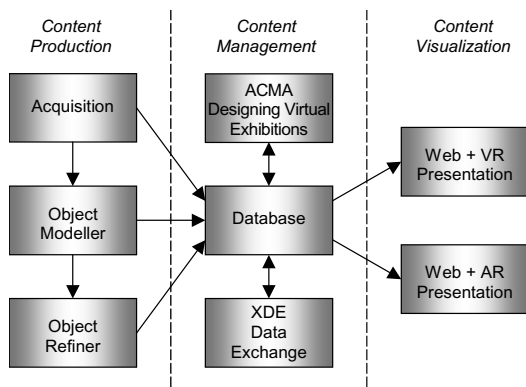
simple museum AR system is the automated tour guide, which superimposes audio on the world based on the location of the user [7].

The European Union has also funded many research projects in the field of cultural heritage and archaeology. For example, the SHAPE project [8] applies AR to the field of archaeology to educate visitors about the artefacts and their history. The 3DMURALE project [9] is developing and using 3D multimedia tools to record, reconstruct, encode and visualize archaeological ruins in VR. In the Ename 974 project [10] visitors can enter a specially designed on-site kiosk where real-time video images and architectural reconstructions are superimposed, and visitors can control the video camera and display images using a touch screen. The ARCHEOGUIDE project [11] provides an interactive AR guide for the visualization of outdoor archaeological sites. Similar to ARCHEOGUIDE is the LIFEPLUS project which additionally encompasses real-time 3D simulations of ancient fauna and flora [12].

The main advantage of the ARCO system over the projects described above are that ARCO offers a complete museum focused solution that can be configured for museum needs—we can build bespoke museum systems from interoperable ARCO components. But more importantly, ARCO offers methods for digitization, management and presentation of heritage artefacts in virtual exhibitions based on well understood metaphors that are also interactive and appealing [13].

## 2. ARCO System Overview

The ARCO functionality mentioned above defines the specification of the system architecture, illustrated in Figure 1. For the content production process ARCO provides two tools for 3D modelling of museum artefacts: the Object Modeller (OM) and the Model Refiner (MR). The OM tool is a 3D stereo photogrammetry system based on the principles of Image-based Modelling. The MR tool is a 3D reconstruction refinement tool based on the *3ds max* framework that complements the OM tool. Note that content production also includes acquiring other multimedia data such as images, movies, etc. for input to the content management process.

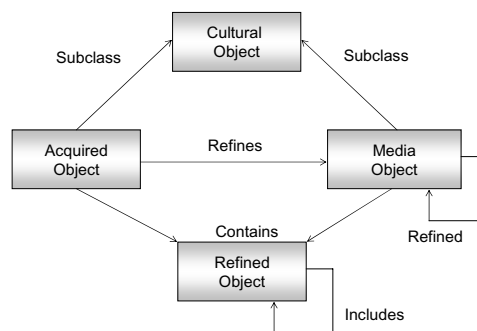


**Figure 1: ARCO System Architecture**

For the content management process ARCO provides a multimedia database management system based on Oracle9i and the ARCO content Management Application (ACMA). The database is the central component of the ARCO system in that it stores, manages and organises virtual artefacts into collections for display in virtual exhibitions.

The final part of the ARCO architecture is the content visualization process. The visualization of the virtual museum artefacts is performed by VR and AR browser. These browsers combine Web-based form of presentation with either VR or AR virtual exhibitions.

The end user is able to browse content stored in the database either remotely through the web, in a museum kiosk, or to interact with the virtual objects in an AR table-top environment using either a simple monitor display or HMD. The ARCO system is based on the data model [14] illustrated in Figure 2.



**Figure 2: ARCO Data Model**

A key element of the ARCO system is the specification of an appropriate metadata element set that underpins both the heritage and technical aspects of ARCO. We need both to describe museum artefacts and the technical processes that transform the artefacts from the physical to the virtual. Accordingly, we have designed a metadata element set called AMS [14], called the ARCO Metadata Schema.

## 3. Virtual Museum Exhibitions

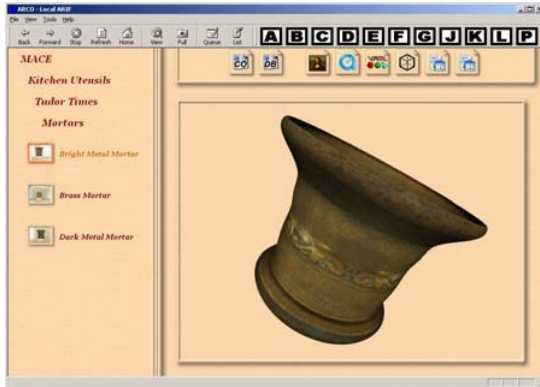
Virtual museum artefacts are displayed as virtual exhibitions through three presentation domains: *WEB\_LOCAL* for use on local web-based displays inside museums, *WEB\_REMOTE* for use on the Internet, and *WEB\_AR* for use in AR presentations.

The ARCO system provides two main kinds of user interfaces for browsing cultural heritage exhibitions: *Web-based interfaces* and *Augmented Reality interfaces*, see sections 3.1 and 3.2

### 3.1. Virtual Reality Exhibitions

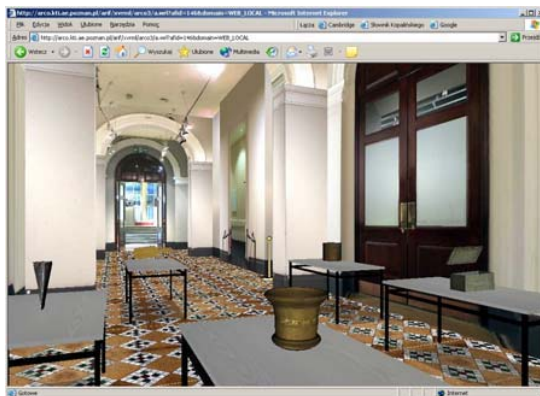
In the Web-based interface a user can browse information presented in a form of 3D VRML virtual galleries or 2D Web pages with embedded multimedia objects. The Web-based interface requires a standard Web browser such as Internet Explorer with a VRML plug-in. This kind of user interface can be used both on local displays inside a museum (*WEB\_LOCAL*) and remotely on the Internet (*WEB\_REMOTE*).

An example visualization of virtual exhibitions in a Web browser is presented in Figure 3. This visualization consists of Web pages with embedded 3D VRML models and other multimedia objects and can be used remotely over the Internet. Users can browse the hierarchy of virtual exhibitions and virtual museum artefacts by clicking on appropriate icons at the top of the page.



**Figure 3: Web-based visualization**

Virtual exhibitions can also be visualized in the Web browser in a form of 3D galleries, see Figure 4. In this visualization, users can browse objects simply by walking along the 3D room, which is a reconstruction of a real gallery—an exhibition corridor in the Victoria and Albert Museum in London.



**Figure 4: Example 3D virtual exhibitions**

### 3.2. Augmented Reality Exhibitions

To enable visualization of selected objects in an AR environment an AR application has been developed. The AR application is used instead of a typical Web browser used in the Web-based interfaces. The AR application integrates two components: a Web browser and an AR browser. For the AR visualization a camera and a set of physical markers placed in a real environment is used. Video captured by the camera is passed on to the AR browser that overlays virtual representations of virtual museum artefacts using the markers for object positioning [15].

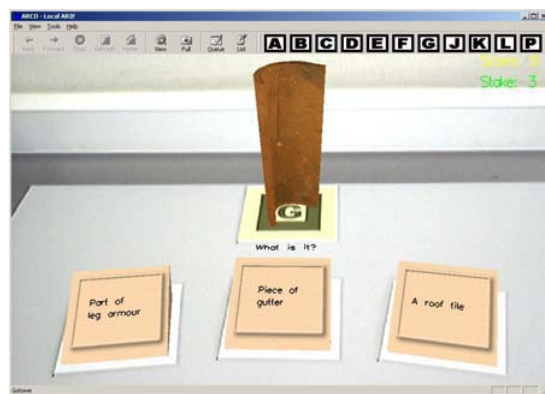
Users can interact with the displayed objects using both the markers and standard input devices, such as

the SpaceMouse®. In the first method, a user can manipulate a marker in front of a camera as it is presented in Figure 5 and look at overlaid objects from different angles and distances. This is a natural and intuitive method of interaction with virtual objects.



**Figure 5: Real scene augmented with superimposed virtual models**

The content and layout of the visualized scenes are determined by visualization templates that define which components of a virtual museum artefact are composed into one VRML scene. One of the important goals of the ARCO system is presenting virtual museum artefacts in an attractive manner that would make people, especially children, more interested in cultural heritage. ARCO enables museum curators to build interactive learning scenarios, where visitors can gain information not only by browsing it, but also by answering series of questions presented in the form of a quiz. As an example, we have implemented an interactive AR quiz based on Fishbourne Roman Palace [16], illustrated Figure 6.



**Figure 6: Example quiz scene**

In this quiz we use one of the markers to display the virtual museum artefact and a question, and three more markers to display potential answers. The user then chooses an answer, and depending on whether the answer is correct or not, an appropriate response in the AR scene appears, see Figure 7.



**Figure 7: Wrong and correct answers**

#### 4. Conclusions

The ARCO system provides a complete solution for digitization, management and presentation of virtual museum exhibitions. We have addressed digital acquisition, storage, management and visualization in interactive VR and AR interfaces by adopting a component based approach. Furthermore, mixing and matching of individual components is supported through the use of XML for interoperability purposes. A system such as ARCO has the potential to revolutionise the use of computer-based systems in museums in the future, so that they are no longer regarded as mere tools for cataloguing purposes, but rather as ways of engaging and enhancing the experience of their users.

#### 5. Acknowledgments

This research was funded by the EU IST Framework V programme, Key Action III-Multimedia Content and Tools, Augmented Representation of Cultural Objects (ARCO) project IST-2000-28336.

#### 6. References

[1] Martin White, Krzysztof Walczak, Nicholas Mourkoussis, 'ARCO—Augmented Representation of Cultural Objects' *Advanced Imaging*, Len Yencharis (ed.), June 2003, Vol 18, No. 6 pp14-15, 46.

[2] Mase, K., Kadobayashi, R., et al., *Meta-Museum: A Supportive Augmented-Reality Environment for Knowledge Sharing*, ATR Workshop on Social Agents: Humans and Machines, Kyoto, Japan, April 21-22, 1997.

[3] Brogni A., Avizzano C.A., Evangelista C., Bergamasco M., *Technological Approach for Cultural Heritage: Augmented Reality*, Proc. of 8th International

Workshop on Robot and Human Interaction (RO-MAN '99).

[4] ORION, Object Rich Information Network [<http://www.orion-net.org/index.asp>], [accessed 11/12/2003]

[5] Sinclair, P and Martinez, K., *Adaptive Hypermedia in Augmented Reality*, Proceedings of Hypertext, 2001.

[6] Gatermann, H., *From VRML to Augmented Reality Via Panorama-Integration and EAI-Java*, in *Constructing the Digital Space*, Proc. of SiGraDi'2000, 254-256, September 2000.

[7] Bederson, B.B., *Audio Augmented Reality: A Prototype Automated Tour Guide*, In the ACM Human Computer in Computing Systems conference (CHI'95), pp 210-211

[8] Hall, T., Ciolfi, L., et al. *The Visitor as Virtual Archaeologist: Using Mixed Reality Technology to Enhance Education and Social Interaction in the Museum*, Proc. of Virtual Reality, Archaeology, and Cultural Heritage (VAST 2001), (Ed Spencer, S.) New York, ACM SIGGRAPH, pp 91-96, Glyfada, Nr Athens, November 2001.

[9] Cosmas, J., et al. *3D MURALE: a multimedia system for archaeology*, In Proceedings of the 2001 conference on Virtual Reality, archaeology, and cultural heritage, pp 297-306 (2001).

[10] Pletinckx, D., Callebaut, D., Killebrew, A., Silberman, N., *Virtual-Reality Heritage Presentation at Ename*, IEEE Multimedia April-June 2000 (Vol. 7, No. 2), pp. 45-48.

[11] Gleue, T., Dähne, P., *Design and Implementation of a Mobile Device for Outdoor Augmented Reality in the ARCHEOGUIDE Project*, Virtual Reality, Archaeology, and Cultural Heritage International Symposium (VAST01), Glyfada, Nr Athens, Greece, 28-30 November 2001.

[12] LIFEPLUS, [<http://www.miralab.unige.ch/subpages/lifeplus/HTML/home.htm>], [accessed 11/12/2003]

[13] ARCO Consortium, 'Augmented Representation of Cultural Objects', [<http://www.arco-web.org>], [accessed 11/12/2003]

[14] Nicholas Mourkoussis, Martin White, Manjula Patel, Jeczek Chmielewski and Krzysztof Walczak, 'AMS-Metadata for Cultural Exhibitions using Virtual Reality', *DC-2003 Proc. of the International DCMI Metadata Conference and Workshop*, September 29-Oct 2, 2003, Seattle, Washington, USA, ISBN 0-9745303-0-1.

[15] Kato, H., Billingham, M., Poupayev, I., Imamoto, K. and Tachibana, K., *Virtual object manipulation on a table-top AR environment*, *Proceedings of International Symposium on Augmented Reality (ISAR'00)*, 111-119.

[16] Fishbourne Roman Palace, [<http://www.sussexpast.co.uk/fishbo/fishbo.htm>], [accessed 11/12/2003]

## **8.6 Paper #6**

Liarokapis, F., Brujic-Okretic, V., Papakonstantinou, S. Exploring Urban Environments using Virtual and Augmented Reality, Journal of Virtual Reality and Broadcasting, GRAPP 2006 Special Issue, Digital Peer Publishing, 3(5): 1-13, 2006.

Contribution (70%): Collaboration on the design of the architecture. Implementation of the majority of the VR interface. Write-up of most of the paper.

# Exploring Urban Environments Using Virtual and Augmented Reality

Fotis Liarokapis, Vesna Brujic-Okretic, Stelios Papakonstantinou

City University

giCentre, Department of Information Science, School of Informatics

London EC1V 0HB

email: fotisl, vesna, stelios@soi.city.ac.uk

www: [www.soi.city.ac.uk/organisation/is/research/giCentre/](http://www.soi.city.ac.uk/organisation/is/research/giCentre/)

## Abstract

In this paper, we propose the use of specific system architecture, based on mobile device, for navigation in urban environments. The aim of this work is to assess how virtual and augmented reality interface paradigms can provide enhanced location based services using real-time techniques in the context of these two different technologies. The virtual reality interface is based on faithful graphical representation of the localities of interest, coupled with sensory information on the location and orientation of the user, while the augmented reality interface uses computer vision techniques to capture patterns from the real environment and overlay additional way-finding information, aligned with real imagery, in real-time. The knowledge obtained from the evaluation of the virtual reality navigational experience has been used to inform the design of the augmented reality interface. Initial results of the user testing of the experimental augmented reality system for navigation are presented.

**Keywords:** Mobile Interfaces, Augmented and Virtual Environments, Virtual Tours, Humancomputer interaction.

## 1 Introduction

Navigating in urban environments is one of the most compelling challenges of wearable and ubiquitous computing. The term navigation which can be defined as the process of moving in an environment can be extended to include the process of wayfinding [DS93]. Wayfinding refers to the process of determining one or more routes (also known as paths). Mobile computing has brought the infrastructure for providing navigational and wayfinding assistance to users, anywhere and anytime. Moreover, recent advances in positioning technologies - as well as virtual reality (VR), augmented reality (AR) and user interfaces (UIs) - pose new challenges to researchers to create effective wearable navigation environments. Although a number of prototypes have been developed in the past few years there is no system that can provide a robust solution for unprepared urban navigation. There has been significant research in position and orientation navigation in urban environments. Experimental systems that have been designed range from simple location-based services to more complicated VR and AR interfaces.

An account of the user's cognitive environment is required to ensure that representations are not just delivered on technical but also usability criteria. A key concept for all mobile applications based upon location is the 'cognitive map' of the environment held in mental image form by the user. Studies have shown that cognitive maps have asymmetries (distances between points are different in different directions), that they are resolution-dependent (the greater the den-

Digital Peer Publishing Licence
Any party may pass on this Work by electronic means and make it available for download under the terms and conditions of the current version of the Digital Peer Publishing Licence (DPPL). The text of the licence may be accessed and retrieved via Internet at <a href="http://www.dipp.nrw.de/">http://www.dipp.nrw.de/</a> .

*First presented at the International Conference on Computer Graphics Theory and Applications (GRAPP) 2006, extended and revised for JVRB*



sity of information the greater the distance between two points) and that they are alignment-dependent (distances are influenced by geographical orientation) [Tve81]. Thus, calibration of application space concepts against the cognitive frame(s) of reference is vital to usability. Reference frames can be divided into the egocentric (from the perspective of the perceiver) and the allocentric (from the perspective of some external framework) [Kla98].

End-users can have multiple egocentric and allocentric frames of reference and can transform between them without information loss [MA01]. Scale by contrast is a framing control that selects and makes salient entities and relationships at a level of information content that the perceiver can cognitively manipulate. Whereas an observer establishes a 'viewing scale' dynamically, digital geographic representations must be drawn from a set of preconceived map scales. Inevitably, the cognitive fit with the current activity may not always be acceptable [Rap00].

Alongside the user's cognitive abilities, understanding the spatio-temporal knowledge users have is vital for developing applications. This knowledge may be acquired through landmark recognition, path integration or scene recall, but will generally progress from declarative (landmark lists), to procedural (rules to integrate landmarks) to configurational knowledge (landmarks and their inter-relations) [SW75]. There are quite significant differences between these modes of knowledge, requiring distinct approaches to application support on a mobile device. Hence, research has been carried out on landmark saliency [MD01] and on the process of self-localisation [Sho01] in the context of navigation applications.

This work demonstrates that the cognitive value of landmarks is in preparation for the unfamiliar and that self-localisation proceeds by the establishment of rotations and translations of body coordinates with landmarks. Research has also been carried out on spatial language for direction-giving, showing, for example, those paths prepositions such as along and past is distance-dependent [KBZ<sup>+</sup>01]. These findings suggest that mobile applications need to help users add to their knowledge and use it in real navigation activities. Holl et al, [HLSM03] illustrate the achievability of this aim by demonstrating that users who pre-trained for a new routing task in a VR environment made fewer errors than those who did not. This finding encourages us to develop navigational wayfinding and commentary support on mobile devices accessible to the cus-

tomers.

The objectives of this research include a number of urban navigation issues ranging from mobile VR to mobile AR. The rest of the paper is structured as follows. In section 2, we present background work while in section 4 we describe the architecture of our mobile solution and explain briefly the major components. Sections 5 and 6 present the most significant design issues faced when building the VR interface, together with the evaluation of some initial results. In section 8, we present the initial results of the development towards a mobile AR interface that can be used as a tool to provide location and orientation-based services to the user. Finally, we conclude and present our future plans.

## 2 Background Work

There are a few location-based systems that have proposed how to navigate through urban environments. Campus Aware [BGKF02] demonstrated a location-sensitive college campus tour guide, which allows users to annotate physical spaces with text notes. However, user-studies showed that navigation was not well supported. The ActiveCampus project [GSB<sup>+</sup>04] tests whether wearable technology can be used to enhance the classroom and campus experience for a college student. The project also illustrates ActiveCampus Explorer, which provides location aware applications that could be used for navigation. The latest application is EZ NaviWalk, a pedestrian navigation service launched in Japan in October 2003 by KDDI [oTI04] but in terms of visualisation it offers only the 'standard' 2D map.

From the other hand, many VR prototypes have been designed for navigation and exploration purposes. A good overview of the potential and challenges for geographic visualisation has been previously provided [MEH99]. One example is LAMP3D - a system for the location-aware presentation of VRML content on mobile devices, applied in tourist mobile guides [BC05]. Although the system provides tourists with a 3D visualization of the environment they are exploring, synchronized with the physical world through the use of GPS data, there is no orientation information available. Darken and Sibert [DS96] examined whether real world wayfinding and environmental design principles can be effective in designing large virtual environments that support skilled wayfinding behaviour.

Another example is the mobile multi-modal interaction platform [WSK03] which supports both indoor and outdoor pedestrian navigation by combining 3D graphics with synthesised speed generation. Indoor tracking is achieved through infra-red beacon communication while outdoor via GPS. However, the system does not use georeferenced or accurate virtual representations of the real environment, neither report on any evaluation studies. For the route guidance applications, 3D City models have been demonstrated as useful for mobile navigation [KK02], but studies pointed out the need for detailed modelling of the environment and additional route information. To enhance the visualisation, to aid navigation, a combination of 3D scene representation and a digital map were previously used in a single interface [RV01], [LGS03].

In terms of AR navigation, a few experimental systems have been reported on, until present. One of the first wearable navigation systems is MARS (Mobile Augmented Reality Systems) [FMHW97], which aimed at exploring the synergy of two promising fields of user interface research: AR and mobile computing. Thomas et al, [TD<sup>+</sup>98] proposed the use of a wearable AR system with a GPS and a digital compass as a new way of navigating into the environment. Moreover, the ANTS project [RCD04] proposes an AR technological infrastructure that can be used to explore physical and natural structures, mainly for environmental management purposes. Finally, Reitmayr, et al., [RS04] demonstrated the use of mobile AR for collaborative navigation and browsing tasks in an urban environment.

Although the experimental systems listed above focus on some of the issues involved in navigation, they cannot deliver a functional system capable of combining all accessible interfaces, consumer devices and web metaphors. The motivation for the research reported on in this paper is to address those issues, namely an integration of a variety of hardware and software components to provide effective and flexible navigational and wayfinding tool for urban environments. In addition, we compare potential solutions for detecting the user location and orientation in order to provide appropriate urban navigation applications and services.

To realise this we have designed a mobile platform based on both VR and AR interfaces. To understand in depth all the issues that relate to location and orientation-based services, first a VR interface was designed and tested on a personal digital assistant (PDA)

as a navigation tool. Then, we have incorporated the user feedback into the design of an experimental AR interface. Both prototypes require the precise calculation of the user position and orientation, for the registration purpose. The VR interface is coupled with the GPS and digital compass output to correlate the model with the location and orientation of the user, while the AR interface is only dependent on detecting features belonging to the environment.

### 3 Urban Modelling



Figure 1: Accurate modelling of urban environment (a) high resolution aerial image (b) 3D building extruding

The objectives of this research include issues, such as modelling the urban environment and using visualisation concepts and techniques on a mobile device to help navigation. Currently, the scene surrounding the user is modelled in 3D, and the output is used as a base for both VR and AR navigation scenarios. A partner on the project, GeoInformation Group (GIG), Cambridge, provided a unique and comprehensive data

set, containing the building height/type and footprint data, for the entire City of London. We are using 3D modelling techniques, ranging from manual to semi-automated methods, to create virtual representation of the users immediate environment. The first step of the process involves the extrusion of a geo-referenced 3D mesh using aerial photographs as well as building footprints and heights (Figure 1).

The data set is enhanced by texture information, obtained from the manually captured photographs of the building sides, using a standard, higher resolution digital camera. The steps in the semiautomated technique for preparing and texturing the 3D meshes include: *detaching the objects in the scene; un-flipping the mesh normals; unifying the mesh normals; collapsing mesh faces into polygons and texturing the faces*. An example screenshot of the textured model is shown in Figure 3. All 3D content is held in the GIG City heights database for the test sites in London. The geo-referenced models acquire both the orientation information and the location through a client API on the mobile device, and the application is currently fully functional on a local device. In the final version, the models will be sent to the server in the packet-based message transmitted over the used network. The server will build and render the scene graph associated with the location selected and return it to the client for portrayal.

## 4 Mobile Platform and Functionality

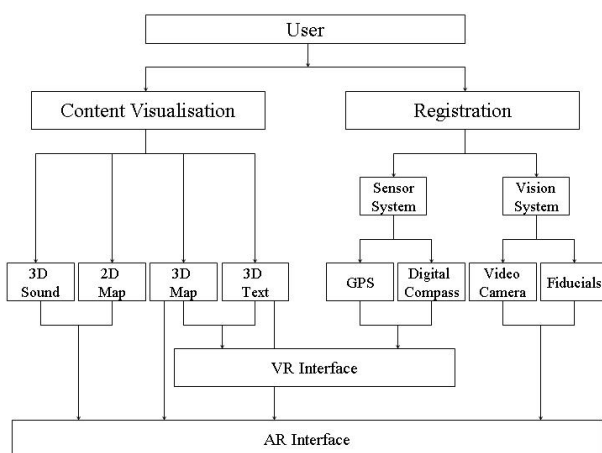


Figure 2: Architecture of our mobile interfaces

Based on these geo-referenced models as building blocks, a generic mobile platform architecture has been designed and implemented for urban navigation

and wayfinding applications and services (Figure 2).

### 4.1 System Configuration

Figure 2 illustrates the system architecture aimed at optimising navigation by using intelligent data retrieval inside an urban area and providing types of digital appropriately visualised information, suitable to be offered as a core of an enhanced location based service. The hardware configuration consists of two distinct sub-systems: i) the remote server equipment and ii) the client device (e.g. a PDA) enhanced with a selection of sensors and peripherals to facilitate the information acquisition, in real time. Both sides feed into the interface on a mobile device, in the form adequate for the chosen mode of operation.

### 4.2 System Functionality

Software applications are custom made and include the information retrieval application, clientserver communication software and a cluster of applications on the client side, which process sensory information, in real-time, and ensure seamless integration of the outputs into a unique interface. The calibration and registration algorithms are at the core of the client side applications ensuring all information is geo-referenced and aligned with the real scene. Registration, in this context, is achieved using two different methods: i) a sensor based solution, taking and processing the readings off the sensors directly, and ii) the image analysis techniques coupled with the information on user's location and orientation obtained from the sensors. The sensor system delivers position and orientation data, in real-time, while a vision system is used to identify fiducial points in the scene. All this information is used as input to the VR and AR interfaces. The VR interface uses GPS and digital compass information for locating and orientating the user.

### 4.3 Interface modalities

Information visualisation techniques used vary according to the nature of the digital content, and/or the navigational task in hand, throughout the navigation. In terms of the content to be visualised, the VR interface can present only 3D maps and textual information. On the other hand, the AR interface uses the calculated user's position and orientation coordinates from the image analysis to superimpose 2D and 3D

maps as well as text and auditory information on the 'spatially aware' framework.

#### 4.4 Notes on Hardware Components

Initially, the mobile software prototype was tested on a portable hardware prototype consisting of a standard laptop computer (equipped with 2.0 GHz M-processor, 1GB RAM and a GeForce FXGo5200 graphics card), a Honeywell HMR 3300 digital compass, a Holux GPS component and a Logitech web-camera (with 1.3 mega-pixel resolution). Then, the prototype system has been ported to a mobile platform based on a Personal Digital Assistant (PDA) and is currently being tested with users.

#### 4.5 Software infrastructure

In terms of the software infrastructure used in this project, both interfaces are implemented based on Microsoft Visual C++ and Microsoft Foundation Classes (MFC). The graphics libraries used are based on OpenGL, Direct3D and VRML. Video operations are supported by the DirectX SDK (DirectShow libraries).

### 5 Virtual Reality Navigation

Navigation within our virtual environment (the spatial 3D map) can take place in two modes: *automatic* and *manual*. In the automatic mode, GPS automatically feeds and updates the spatial 3D map with respect to the users position in the real space. This mode is designed for intuitive navigation. In the manual mode, the control is fully with the user, and it was designed to provide alternative ways of navigating into areas where we cannot obtain a GPS signal. Users might also want to stop and observe parts of the environment in which case control is left in their hands.

During navigation, there are minor modifications obtained continuously from the GPS to improve the accuracy, which results in minor adjustments in the camera position information. This creates a feeling of instability in user, which can be avoided by simply restricting minor positional adjustments. The immersion provided by GPS navigation is considered as pseudo-egocentric because fundamentally the camera is positioned at a height which does not represent a realistic scenario. If, however, the user switches to manual navigation, any perspective can be obtained, which is very helpful for decision-making purposes.

While in a manual mode, any model can be explored and analysed, therefore additional enhancements of the graphical representation are of vital importance.

One of the problems that quickly surfaced during the system evaluation is the viewing angle during navigation which can make it difficult to position the user. This can make it difficult to understand at which point the user is positioned. After informal observation of users during the development process, an altitude of fifty meters over the surface was finally adopted as adequate. In this way, the user can visualise a broader area plus the tops of the buildings, and acquire richer knowledge about their location, in the VR environment. The height information is hard-coded when the navigation is in the automatic mode because user testing (section 7) showed that it can be extremely useful in cases where a user tries to navigate between tall buildings, having low visibility.



Figure 3: FOV differences (a) low angle (b) high angle

Figure 3, illustrates to what extent the FOV is influenced by that angle and how much more information can be included from the same field-of-view, if the angle is favourable. In both Figure 3 (a) and Figure 3 (b), the camera is placed at exactly the same position

and orientation in the horizontal plane, with the only difference in the pitch angle. In Figure 3 (a), the pitch angle is very low and in the Figure 3 (b) it is set to maximum (90°). This feature was considered important to implement after initial testing. The obvious advantage is that, once in a position, no additional rotations are required from the user to understand the exact position of the camera. Taking into consideration the fact that the normal human viewing angle is about 60° and the application supports angles in the range from 0° to 90°, wide angles (including more objects of the landscape) can be interactively obtained. This can be extremely useful in cases where a user tries to navigate between tall buildings, having low visibility.

We are currently implementing two different technologies for presenting 3D maps on PDA interfaces, involving VRML and Managed Direct3D Mobile (MD3DM). The first solution operates as a stand-alone mobile application and uses VRML technology combined with GPS for determining the position and a digital compass for calculating orientation.

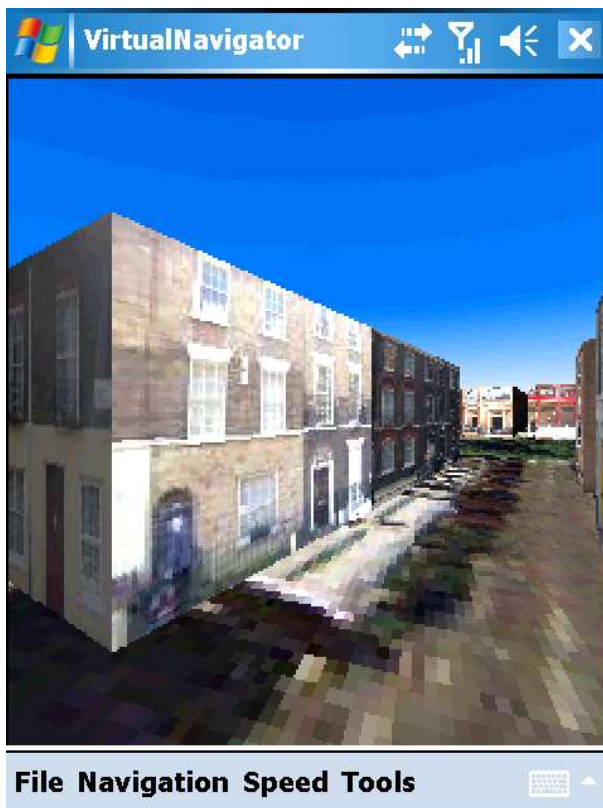


Figure 4: VR navigation in City University's campus

Figure 4 illustrates how the PDA-based navigation inside a virtual environment can be performed. Specifically, stylus interactions can be used to navigate in-

side a realistic virtual representation of City University's campus. Alternatively, menu interactions can be used as another medium for performing navigation and wayfinding tasks. In terms of performance, the frame-rate per second (FPS) achieved varies depending on the device capabilities. For example, using an HTC Universal device the efficiency ranges between 3 to 5 FPS while in a Dell Axim X51v PDA (with a dedicated 16 MB graphics accelerator) the efficiency ranges between 12 to 15 FPS.

The second interface is based on MD3DM that operates as a separate mode, with the aim of handling the output from the GPS/compass automatically providing sufficient functionality to generate mobile VR applications. Compared to the VRML interface, the major advantage of MD3DM is that it takes full advantage of graphics hardware support and enables the development of highperformance three-dimensional rendering [LRBO06]. On the other hand, the major disadvantage of MD3DM is that the Application Programming Interface (API) is low level and thus a lot of functionality which is standard in VRML has to be re-implemented.

## 6 Preliminary Evaluation

The aims of the evaluation of the VR prototype included assessment of the user experience with particular focus on interaction via movement, identification of specific usability issues with this type of interaction, and to stimulate suggestions regarding future directions for research and development. A 'thinking aloud' evaluation strategy was employed [DFAB04]; this form of observation involves participants talking through the actions they are performing, and what they believe to be happening, whilst interacting with the system. This qualitative form of evaluation is highly appropriate for small numbers of participants testing prototype software: Dix et al, [DFAB04] suggested that the majority of usability problems can be discovered from testing in this way. In addition, Tory and Möller [TM05] argued that formal laboratory user studies can effectively evaluate visualisation when a small sample of expert users is used.

The method used for the evaluation of our VR prototype was based on the Black Box technique which offers the advantage that it does not require the user to hold any low-level information about the design and implementation of the system. The user testing took place at City University campus which includes building structures similar to the surrounding area with

eight subjects in total (testing each one individually). All subjects had a technical background and some were familiar with PDAs. Their age varied between 25 and 55. For each test, each subject followed a pre-determined path represented by a highlighted line. Before the start of the walk, the GPS receiver was turned on and flow of data was guaranteed between it and the 'Registration' entity of the system. The navigational attributes that were qualitatively measured include the: *user perspective, movement with device and decision points.*

### 6.1 User Perspective

The main point of investigation was to test whether the user can understand where they are located in the VR scene, in correspondence to the real world position. An examination of the initial orientation and level of immersion was also evaluated after minimum interaction with the application and understanding of the available options. The information that was obtained by the users was concerning mainly four topics including: *level-of-detail (LOD), user-perspective, orientation and field-of-view (FOV).*

Most of the participants agreed that the LOD is not sufficiently high for a prototype navigational application. Some concluded that texture based models would be a lot more appropriate but others expressed the opinion that more abstract, succinct annotations would help, at a different level (i.e. A to Z abstract representations). Both groups of answers can fit in the same context, if all interactions could be visualised from more than one perspective. A suggested improvement was to add *geo-bookmarks* (also known as hotspots) that would embed information about the nature of the structures or even the real world functionality.

As far as the 'user-perspective' attribute is concerned, each user expressed a different optimal solution. Some concluded that more than one perspective is required to fully comprehend their position and orientation. Both perspectives, the egocentric and the allocentric, are useful during navigation for different reasons [LGM<sup>+</sup>05] and under different circumstances. During the initial registration, it would be more appropriate to view the model from an allocentric point of view (which would cover a larger area) and by minimising the LOD just to include annotations over buildings and roads. This proved easier to increase the level of immersion with the system but not being directly exposed to particular information such as

the structure of the buildings. In contrast, an egocentric perspective is considered productive only when the user was in constant movement. When in movement, the VR interface retrieves many updates and the number of decision points is increased. Further studies should be made on how the system would assist an everyday user, but a variation on the user perspective is considered useful in most cases.

The orientation mechanism provided by the VR application consists of two parts. The first maintains the user's previous orientation whilst the second restores the camera to the predefined orientation (which is parallel to the ground). Some users noted that when angle direction points towards the ground gives better appreciation of the virtual navigation. Another subject that the users agree in is the occurrence of fast updates. This can make it difficult to navigate, because the user needs to align the camera on three axes and not two. Based on our experiments we noticed that the used orientation mechanisms are inadequate for navigational purpose and it is imperative that the scene should be aligned in the same direction as the device in the real world.

Furthermore, all participants appreciated the user-maintained FOV. They agreed that it should be wide enough to include as much information, on the screen, as possible. They added that in the primary viewing angle, there should be included recognisable landmarks that would aid the user comprehend the initial positioning. One mentioned that the orientation should stay constant between consecutive decision points, and hence should not be gesturebased. Most users agreed that the functionality of the VR interface provides a wide enough viewing angle able to recognise some of the surroundings even when positioned between groups of buildings with low detail level.

### 6.2 Movement with the Device

The purpose of this stage was to explore how respondents interpreted their interaction with the device, whilst moving. The main characteristics include the *large number of updates* as well as the *change of direction* followed by the user. The elements, which are going to be discussed, are mainly considered with the issues of making the navigation easier, the use of the most appropriate perspective, and the accuracy of the underlying system as well as the performance issues that drive the application. One important issue is to consider the inheritance of a specific perspective for

use throughout the navigation process. Some participants mentioned the lack of accurate direction way-points that would assist route tracking. A potential solution is to consider the adoption of a user-focused FOV during navigation using a simple line on the surface of the model. However, this was considered partially inadequate because the user expects more guidance when reaching a decision point. Some participants suggested to use arrows on top of the route line which would be either visible for the whole duration of the movement or when a decision point was reached.

In addition, it was positively suggested that the route line should be more distinct, minimising the probability of missing it while moving. Some expressed the opinion that the addition of recognisable landmarks would provide a clearer cognitive link between the VR environment and the real world scene. However, the outcomes of this method are useful only for registering the users in the scene and not for navigation purposes. A couple of participants included in their answers that the performance of the system was very satisfactory. This is an important factor to consider, because in the change of the camera position occurs when new data is being retrieved from the external sensor. The characterisation, of the position transition, as smooth reflects that the main objective of any actor is to obtain new information about his position, at the time it is available. The latency that the system supports is equal to the latency the H/W receiver obtains meaning that the performance of the application is solely dependent on the quality of operating hardware. The adaptation to a mobile operating system (i.e. Windows Mobile 5.0) would significantly increase the latency of the system, since devices are not powerful enough to handle heavy operations.

Moreover, opinions, about the accuracy of the system, differ. One of respondents was convinced that the accuracy, provided by the GPS receiver, was inside the acceptable boundaries, which reflected the GPS specifications supporting that the level of accuracy between urban canyons was reflecting the correspondence to reality, in a good manner. A second test subject revealed that the occlusion problem was in effect due to GPS inaccuracy reasons underlining that when the GPS position was not accurate enough, the possibility to miss the route line or any developed direction system increased. Both opinions are equally respected and highlighted the need for additional feedback.

### 6.3 Decision Points

The last stage is concerned with the decision points and the ability of the user to continue the interaction with the system when it reaches them. A brief analysis of the users' answers is provided to identify ways forward with the design, but full analysis will be published in a separate publication. As described previously, the user has the feeling of full freedom to move at any direction, without being restricted by any visualisation limitations of the computergenerated environment. Nonetheless, participants may feel overwhelmed by the numerous options they may have available and be confused about what action to take next. We take into consideration that large proportion of users is not sufficiently experienced in 3D navigational systems and the appropriate time is given to them to familiarise with the system.

Preliminary feedback suggests that some users would prefer the application to be capable of manipulating the users perspective automatically, when a decision point (or, an area close to it) is reached. This should help absorb more information about the current position as well as supporting the future decision making process. Another interesting point relates to the provision of choice to the user in the future to accommodate sudden, external factors that may allow them to detour from a default path. Partially, some of these requirements would be met if the user could manually add geo-bookmarks in the VR environment representing points in space with supplementary personal context. The detailed analysis of the responses will be taken into account in further developments of the system, which is underway.

## 7 Augmented Reality Navigation

The AR interface is the alternative way of navigating in the urban environment using mobile systems. Unlike the VR interface, which uses the hardware sensor solution (a GPS component and a digital compass), the AR interface uses a webcam (or video camera) and computer vision techniques to calculate position and orientation. Based on the findings of the previous section and a previously developed prototypes [Lia05], [LGM<sup>+</sup>05], a high-level AR interface has been designed for outdoor use. The major difference with other existing AR interfaces, such as the ones described in [FMHW97], [TD<sup>+</sup>98], [RS04] and [RCD04], is that our approach allows for the combi-

nation of four different types of navigational information: 3D maps, 2D maps, text and spatial sound. In addition, two different modes of registration have been designed and experimented upon, based upon *fiducial points* and *feature* recognition. The purpose for the exercise was to understand some of the issues involved in two of the key aspects of urban navigation: *wayfinding* and *commentary*.

In the fiducial points recognition mode, the outdoor environment needs to be populated with fiducials prior to the navigational experience. Fiducials are placed in points-of-interest of the environment, such as corners of the buildings, ends of streets etc, and play a significant role in the decision making process. In our current implementation we have adopted ARToolKit's template matching algorithm [KB99] for detecting marker cards and we try to extend it for natural feature detection. Features that we currently detect can come in different shapes, such as square, rectangular, parallelogram, trapezium and rhomb [Lia05] similar to shapes that exist in the environment. In addition, it is not convenient, sometimes it is even impossible, to populate large urban areas with fiducials. Therefore, we have experimentally used road signs as fiducials to compute the users pose [LRBO06]. Road signs are most of the time printed in black colour on a white background. Also, they are usually placed at the decision points, such as beginning and ending of streets, corners and junctions. As a result, if a highresolution camera is used to capture the object, it is relatively easy to detect the road signs, as illustrated in Figure 5.

One of the known limitations of this technique is that, sometimes, road signs are not in a good condition, which makes it more difficult to recognise a pattern. Also, the size of the road signs is usually fixed (depending on the urban area) limiting severely the number of operations that can be done on it. An example screenshot of how road signs can be used in practice as fiducial points (instead of using pre-determined markers) during urban navigation is illustrated in 5.

Alternatively, in the feature recognition, the user is 'searching' to detect natural features of the real environment to serve as 'fiducial points and pointsof-interest, respectively. Distinctive natural features like door entrances, windows etc, have been experimentally tested to see whether they can be used as 'natural markers'. Figure 6 shows the display presented to a user navigating in City University's campus, to acquire location and orientation information using 'natural markers'.



Figure 5: Pattern recognition of road signs: (a) original image; (b) detected image

As soon as the user turns the camera (on a mobile device) towards these predefined natural markers, audio-visual information (3D arrows, textual and/or auditory information) can be superimposed on the real-scene imagery (Figure 7), thus satisfying some of the requirements identified in section 6.1. Userstudies for tour guide systems showed that visual information could sometimes distract the user [BGKF02], while audio information could be used to decrease the distraction [WAHS01]. With this in mind, we have introduced a spatially referenced sound into the interface, to be used simultaneously with the visual information. In our preliminary test case scenario, a pre-recorded sound file is assigned to the corresponding fiducial point, for each pointof-interest. As the user approaches a fiducial point, commentary information can be spatially identified; the closer the user to the object the louder the volume of the commentary audio information. Depending on the end-user's preferences,



or needs, the system allows for a different type of digital information to be selected and superimposed. For example, for visually impaired users audio information may be preferred to use over visual, or a combination of the two may be found optimal [Lia05]. A coarse comparison between the use of fiducial points and the feature recognition mode is shown in Table 1. Further testing is underway and the detailed analysis will be published in a separate publication.

Recognition Mode	Range	Error	Robustness
Fiducial	0.5~ 2m	Low	High
Feature	2~ 10m	High	Low

Table 1: Fiducial vs feature recognition mode

In the feature recognition mode, the advantage is that the range within which it may operate is much greater because it does not require preparation of the environment. Thus, it can be applied when wayfinding is the focus of the navigation. However, the natural feature tracking algorithm, which is used in this scenario, does require improved accuracy of the position and orientation information, which is currently limited. In contrast, the fiducial points recognition mode offers the advantage of a very low error during the tracking process (i.e. detecting fiducial points). However, the limited space of operation due to the need to populate the area with tags, makes it more appropriate for confined areas and commentary navigation modes. The research suggests, however, that the combination of fiducial and feature recognition modes allows the user to pursue both wayfinding and commentary based navigation into urban environments within a single application.



Figure 6: Road sign pedestrian navigation

## 8 Discussion

After completing the development of a portable prototype application (based on a laptop computer based) specific requirements to enhance the user interface and interaction mechanisms on a mobile device (PDA) were identified. Through this research, it was found obligatory to retrieve and visualise spatio-temporal content from a remote server in order to support real-time operation and meet the information needs of a user. This was accomplished by transmitting geographic coordinates (i.e. GPS input) to the server-side and automatically retrieving geo-referenced information in the form of VRML 3D maps. The 3D content was designed to cover an area encompassing the current position of the user and the position of one or more actors/points-of-interest in their proximity. The quality and accuracy of these models are proved good while the techniques used are customdeveloped and based on a semi-automated routine, developed in a specialised software development environment.

## 9 Conclusions

This paper addresses how virtual and augmented reality interface paradigms can provide enhanced location based services for urban navigation and wayfinding. The VR interface operates on a PDA and presents a realistic and geo-referenced graphical representation of the localities of interest, coupled with sensory information on the location and orientation of the user. The knowledge obtained from the evaluation of the VR navigational experience has been used to inform

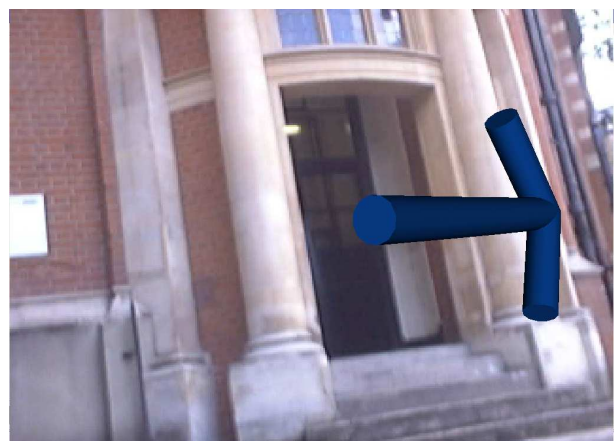


Figure 7: Detecting door entrances

the design of the AR interface which operates on a portable computer and overlays additional way-finding information onto the captured patterns from the real environment.

Both systems calculate the user's position and orientation, but using a different methodology. The VR interface relies on a combination of GPS and digital compass data whereas the AR interface is only dependent on detecting features of the immediate environment. In terms of information visualisation, the VR interface can only present 3D maps and textual information while the AR interface can, in addition, handle other relative geographical information, such as digitised maps and spatial auditory information.

Work on both modes and interfaces is in progress and we also consider a hybrid approach, which aims to find a balance between the use of hardware sensors (GPS and digital compass) and software techniques (computer vision) to achieve the best registration results. In parallel, we are designing a spatial database to store our geo-referenced urban data, which will feed the client-side interfaces as well as routing algorithms, which we are developing to provide more services to mobile users. The next step in the project is a thorough evaluation process, using both qualitative and quantitative methods. The results will be published in due course.

## 10 Acknowledgments

The work presented in this paper is conducted within the LOCUS project, funded by EPSRC, through the Location and Timing (KTN) Network. We would also like to thank our partner on the project, GeoInformation Group, Cambridge, for making the entire database of the City of London buildings available to the project. The invaluable input from David Mountain on resolving the sensor fusion issues and from Christos Gatzidis for generating components of the 3D content are greatly acknowledged.

## References

- [BC05] Stefano Burigat and Luca Chittaro, *Location-aware visualization of VRML models in GPS-based mobile guides*, Proceedings of the 10th International Conference on 3D Web Technology, ACM Press, 2005, ISBN 1-59593-012-4, pp. 57–64.
- [BGKF02] Jenna Burrell, Geri K. Guy, Kiyo Kubo, and Nick Farina, *Context-aware computing: test case*, Proceedings of UbiComp, Lecture Notes in Computer Science Vol. 2498, Springer, 2002, ISBN 3-540-44267-7, pp. 1–15.
- [DFAB04] Alain J. Dix, Janet E. Finlay, Gregory D. Abowd, and Russel Beale, *Human-Computer Interaction*, 3rd edition ed., Prentice Hall, Harlow, 2004, ISBN 0-13-046109-1.
- [DS93] Rudy P. Darken and John L. Sibert, *A toolset for navigation in virtual environments*, Proceedings of the 6th annual ACM symposium on User interface software and technology (New York, NY, USA), ACM Press, 1993, ISBN 0-89791-628-X, pp. 157–165.
- [DS96] Rudy P. Darken and John L. Sibert, *Navigating Large Virtual Spaces*, International Journal of Human-Computer Interaction **8** (1996), no. 1, 49–72, ISSN 1044-7318.
- [FMHW97] Steven Feiner, Blair MacIntyre, Tobias Höllerer, and Antony Webster, *A touring machine: Prototyping 3D mobile augmented reality systems for exploring the urban environment*, Proceedings of the 1st IEEE International Symposium on Wearable Computers, IEEE Computer Society, 1997, ISBN 0-8186-8192-6, pp. 74–81.
- [GSB<sup>+</sup>04] William G. Griswold, Patricia Shanahan, Steven W. Brown, Robert S. Boyer, Matt Ratta. R. Benjamin Shapiro, and Tan Minh Truong, *ActiveCampus: Experiments in Community-Oriented Ubiquitous Computing*, Computer **37** (2004), no. 10, 73–81, ISSN 0018-9162.
- [HLSM03] Doris Höll, Bernd Lepow, Robby Schöpfung, and Maximilian Mehdorn, *Spatial cognition III*, Lecture Notes in Computer Science Vol. 2685, ch. Is it possible to learn and transfer spatial information from virtual to real worlds?, pp. 143–156, Springer, Berlin, 2003, ISBN 3-540-40430-9.

- [KB99] Hirokazu Kato and Mark Billinghurst, *Marker Tracking and HMD Calibration for a Video-Based Augmented Reality Conferencing System*, Proceedings of the 2nd IEEE and ACM International Workshop on Augmented Reality, IEEE Computer Society, 1999, ISBN 0-7695-0359-4, pp. 85–94.
- [KBZ<sup>+</sup>01] Christian Kray, Jörg Baus, Hubert D. Zimmer, Harry R. Speiser, and Antonio Krüger, *Two path prepositions: along and past*, Proceedings of the International Conference on Spatial Information Theory: Foundations fo Geographic Information Science (London) (D. Montello, ed.), Lecture Notes in Computer Science Vol. 2205, Springer, 2001, ISBN 3-540-42613-2, pp. 263–277.
- [KK02] M. Kulju and E. Kaasinen, *Guidance Using a 3D City Model on a Mobile Device*, Workshop on Mobile Tourism Support Mobile HCI 2002 Symposium, Pisa, Italy, Sept. 17th, 2002.
- [Kla98] R. L. Klatzky, *Spatial cognition - An interdisciplinary approach to representation and processing of spatial knowledge*, ch. Allocentric and egocentric spatial representations: Definitions, distinctions, and interconnections, pp. 1–18, Springer, Berlin, 1998, ISBN 3-540-64603-5.
- [LGM<sup>+</sup>05] Fotis Liarokapis, Ian Greatbatch, David Mountain, Anil Gunesh, Vesna Brujic-Okretic, and Johnathan Raper, *Mobile Augmented Reality Techniques for Geo-Visualisation*, Proceedings of the 9th International Conference on Information Visualisation IV'05, IEEE Computer Society, 2005, ISBN 0-7695-2397-8, pp. 745–751.
- [LGS03] K. Laakso, O. Gjesdal, and J.R. Sulebak, *Tourist information and navigation support by using 3D maps displayed on mobile devices*, Workshop on Mobile Guides, Mobile HCI 2003 Symposium, Udine, Italy, 2003.
- [Lia05] Fotis Liarokapis, *Augmented Reality Interfaces Architectures for Visualising and Interacting with Virtual Information*, Ph.D. thesis, School of Science and Technology, University of Sussex, Department of Informatics, 2005, Sussex theses S 5931, ISBN/ISSN/CNM0426866US.
- [LRBO06] Fotis Liarokapis, Johnathan Raper, and Vesna Brujic-Okretic, *Navigating within the urban environment using Location and Orientation-based Services*, European Navigation Conference, 710 May, Manchester, UK, 2006.
- [MA01] Christy R. Miller and Gary L. Allen, *Spatial frames of reference used in identifying directions of movement: an unexpected turn*, Proceedings of the International Conference on Spatial Information Theory: Foundations of Geographic Information Science (London) (D. Montello, ed.), Lecture Notes in Computer Science Vol 2205, Springer, 2001, ISBN 3-540-42613-2, pp. 206–216.
- [MD01] Pierre-Emmanuel Michon and Michel Denis, *Proceedings of the international conference on spatial information theory: Foundations of geographic information science*, Lecture Notes in Computer Science Vol 2205, ch. When and why are visual landmarks used in giving directions, pp. 292–305, London, 2001, ISBN 3-540-42613-2.
- [MEH99] A.M. MacEachren, R. Edsall, and D. Hauq, *Virtual environments for Geographic Visualization: Potential and Challenges*, Proceedings of the ACM Workshop on New Paradigms in Information Visualization and Manipulation, ACM Press, 1999, ISBN 1-58113-254-9, pp. 35–40.
- [oTI04] DTI Department of Trade and Industry, *Location-based services: understanding the japanese experience, global watch mission report*, [http://www.oti.globalwatchonline.com/online\\_pdfs/36246MR.pdf](http://www.oti.globalwatchonline.com/online_pdfs/36246MR.pdf), 2004, visited: 10/02/2006.

- [Rap00] Johnathan F. Raper, *Multidimensional geographic information science*, Taylor and Francis, London, 2000, ISBN 0-7484-0506-2.
- [RCD04] T. Romão, N. Correia, and E. Dias, *ANTS-Augmented Environments*, Computers & Graphics **28** (2004), no. 5, 625–633, ISSN 0097-8493.
- [RS04] Gerhard Reitmayr and Dieter Schmalstieg, *Collaborative Augmented Reality for Outdoor Navigation and Information Browsing*, Proceedings of the Symposium of Location Based Services and TeleCartography, Vienna, Austria, January 2004, 2004, pp. 31–41.
- [RV01] I. Rakkolainen and T. Vainio, *A 3D City Info for Mobile Users*, Computer & Graphics **25** (2001), no. 4, 619–625, ISSN 0097-8493.
- [Sho01] M. Jeanne Sholl, *The role of self reference system in spatial navigation*, Proceedings of the International Conference on Spatial Information Theory: Foundations of Geographic Information Science (London) (D. Montello, ed.), Lecture Notes in Computer Science Vol. 2205, Springer, 2001, ISBN 3-540-42613-2, pp. 217–232.
- [SW75] A.W. Siegel and S.H. White, *The development of spatial representation of large scale environments*, Advances in child development and Behaviour **10** (1975), 9–55, ISSN 0065-2407.
- [TD<sup>+</sup>98] B.H. Thomas, , V. Demczuk, W. Piekarski, D. Hepworth, and B. Gunther, *A Wearable Computer System with Augmented Reality to Support Terrestrial Navigation*, Proceedings of the 2nd International Symposium on Wearable Computers, IEEE and ACM, 1998, ISBN 0-8186-9074-7, pp. 168–171.
- [TM05] M. Tory and T. Möller, *Evaluating Visualizations: Do Expert Reviews Work?*, Computer Graphics and Applications **25** (2005), no. 5, 8–11, ISSN 0272-1716.
- [Tve81] B. Tversky, *Distortions in memory for maps*, Cognitive Psychology **13** (1981), 407–433, ISBN 0010-0285.
- [WAHS01] Allison Woodruff, Paul M. Aoki, Amy Hurst, and Margaret H. Szymanski, *Electronic Guidebooks and Visitor Attention*, Proceedings of 6th International Cultural Heritage Informatics Meeting ICHIM'01, Milan, Italy, Sep. 2001, 2001, ISBN 1-885626-24-X, pp. 437–454.
- [WSK03] Rainer Wasinger, Christoph Stahl, and Antonia Krüger, *Mobile Multi-Modal Pedestrian Navigation*, Second International Workshop on Interactive Graphical Communication IGC 2003, London, 2003.

Citation
Fotis Liarokapis, Vesna Brujic-Okretic, Stelios Papakonstantinou, <i>Exploring Urban Environments Using Virtual and Augmented</i> , Journal of Virtual Reality and Broadcasting, 3(2006), no. 5, December 2006, urn:nbn:de:0009-6-7720, ISSN 1860-2037.

## **8.7 Paper #7**

Liarokapis, F. An Augmented Reality Interface for Visualizing and Interacting with Virtual Content, *Virtual Reality*, Springer, 11(1): 23-43, 2007.

Contribution (100%): Design of the architecture and implementation of the AR interface.  
Write-up of the paper.

# An augmented reality interface for visualizing and interacting with virtual content

Fotis Liarokapis

Received: 15 December 2004 / Accepted: 19 October 2006 / Published online: 9 November 2006  
© Springer-Verlag London Limited 2006

**Abstract** In this paper, a novel AR interface is proposed that provides generic solutions to the tasks involved in augmenting simultaneously different types of virtual information and processing of tracking data for natural interaction. Participants within the system can experience a real-time mixture of 3D objects, static video, images, textual information and 3D sound with the real environment. The user-friendly AR interface can achieve maximum interaction using simple but effective forms of collaboration based on the combinations of human–computer interaction techniques. To prove the feasibility of the interface, the use of indoor AR techniques are employed to construct innovative applications and demonstrate examples from heritage to learning systems. Finally, an initial evaluation of the AR interface including some initial results is presented.

**Keywords** Augmented reality · Human–computer interaction · Tangible interfaces · Virtual heritage · Learning systems

## 1 Introduction

Augmented reality (AR) is an increasingly important and promising area of mixed reality (MR) and user interface design. In technical terms, it is not a single

technology but a collection of different technologies that operate in conjunction, with the aim of enhancing the user's perception of the real world through virtual information (Azuma 1997). This sort of information is usually referred to as virtual, digital or synthetic information. The real world must be matched with the virtual in position and context in order to provide an understandable and meaningful view (Mahoney 1999). Participants can work individually or collectively, experiment with virtual information and interact with a mixed environment in a natural way (Klinker et al. 1997). In an ideal AR visualisation scenario, the virtual information must be mixed with the real world in real-time in such a way that the user can either understand or not, the difference (Vallino 1998). In case where virtual information looks alike the real environment, the AR visualisation is considered as the ultimate immersive system where participants cannot become more immersed in the real environment (RE).

The term AR usually refers to one of the following definitions (Milgram and Colquhoun 1999). A class of display systems that consist of a type of head mounted display (HMD) (Azuma 1997); those systems that utilize an equivalent of an HMD belong to the second class, encompassing both large screen and monitor-based displays (Milgram and Kishino 1994). A third classification refers to the cases that include any type of mixture of real and virtual environments. Overall, the majority of AR systems rely on electronic sensors or video input in order to gain knowledge of the environment (Haniff et al. 2000). All these variables make these systems more complex than systems that do not rely on sensors. Vision based systems on the other hand, often use markers as feature points so they can estimate the camera pose (position and orientation).

---

F. Liarokapis  
Department of Informatics, University of Sussex,  
Falmer, Brighton BN1 9QT, UK

F. Liarokapis (✉)  
Department of Information Science, City University,  
London EC1V 0HB, UK  
e-mail: fotisl@soi.city.ac.uk; f.Liarokapis@sussex.ac.uk

In the upcoming years, AR systems will be able to include a complete set of augmentation applied and exploiting all people's senses (Azuma et al. 2001). However, although there are many examples of AR systems where users can interact with and manipulate virtual content and even create virtual content within some AR environments, one of their major constraints is the lack of ability to allow participants control multiple forms of virtual information in a number of different ways. To a great extent, this deficiency derives mainly from the lack of robustness of currently existing AR interface systems. At this stage, this can be dealt by using a user-friendly interface to allow users position audio-visual information anywhere inside the physical world. Since the pose can be easily estimated through an existing vision based tracking system such as the well-known ARToolKit (Kato et al. 2000a, b), the focus of this research is to provide effective solutions for interactive indoor AR environments.

Vision-based AR interface environments highly depend on four key elements. The first two relate to marker implementation and calibration techniques. The latter are interrelated with the construction of software user interfaces that will allow the effective visualisation and manipulation of the virtual information. The integration of such interfaces into AR systems can reduce the complexity of the human-computer interaction using implicit contextual input information (Rekimoto and Nagao 1995). Human computer interaction techniques can offer greater autonomy when compared with traditional windows style interfaces. Although some work has been performed into, the integration of such interfaces into AR systems (Feiner et al. 1993; Haller et al. 2002; MacIntyre et al. 2005) the design and implementation of an effective AR system that can deliver realistically audio-visual information in a user-friendly manner is a difficult task and an area of continuous research. However, it is very difficult even for technologists to create AR experiences to eliminate these barriers (MacIntyre et al. 2005) that prevent users to create new AR applications. To address the above issues, a prototype AR interface for assisting users that have some virtual reality experience to create fast and effective AR applications is proposed. The main novel contributions of this paper include the following:

- Simultaneous and realistic 3D audio-visual augmentation in real-time performance;
- Implementation and combination of five different ways for interacting with the virtual content;
- Design and implementation of a high-level user centred interface that provides accurate and reliable control of the AR scene;

- Two innovative applications: one for cultural heritage and one for higher education and
- Initial evaluation regarding the overall effectiveness of the system;

In the remainder of this paper, we describe our system starting with Sect. 2 that gives a historical overview of the AR interfaces. In Sect. 3, the architecture of the prototype AR interface is presented in detail. Section 4, presents various calibration approaches followed to calibrate our camera sub-system accurately. Section 5 presents realistic augmentation techniques that can be applied in real-time performance. Section 6 proposes five different ways of interacting with the AR scene while in Sect. 7 two application scenarios are presented. In Sect. 8, the results from an initial evaluation are presented whereas Sect. 9 summarises the key findings and the current status of research and suggests future work.

## 2 Historical overview of AR interfaces

One of the earliest applications involved an experimental AR system that supports a full X11 server on a see-through HMD. The display overlays a selected portion of the X bitmap, on the user's view of the world, creating an X-based AR. Three different types of windows were developed: surround-fixed windows, display-fixed windows and world-fixed windows. The performance of the system was in the range of 6–20 frames-per-second (FPS). A fast display server was developed supporting multiple overlaid bitmaps having the ability to index into a display a selected portion of a larger bitmap (Feiner et al. 1993). EMMIE (Butz et al. 1999) is another experimental hybrid user interface designed for a collaborative augmented environment that combines various different technologies and techniques such as virtual components (i.e. 3D widgets) and physical objects (tracked displays, input devices). The objects in the system can be moved among various types of displays, ranging from see-through HMDs to additional 2D and 3D displays. These vary from palm-sized to wall-sized depending on the nature of the task.

The MagicBook (Billinghurst et al. 2001) and the Tiles system (Poupyrev et al. 2002) are two of the most well known AR interfaces based on the ARToolKit. The Tiles system proposes a way of creating an AR workspace blending together virtual and physical objects. The interface combines the advantages (power and flexibility) of computing environments with the comfort and awareness of the traditional workplace

(Poupyrev et al. 2002). On the other hand, the MagicBook uses a real book to transfer users from reality to virtuality. Virtual objects are superimposed on the pages of the book and users can interact with the augmented scene (Billinghurst et al. 2001). Another example of an AR tangible interface is a tabletop system designed for virtual interior design (Kato et al. 2000a). One or multiple users can interact with the augmented scene, which consists of virtual furniture and manipulates the virtual objects.

MARE (Grasset and Gascuel 2002) is a collaborative system that mixes together AR techniques with human-computer interaction techniques, in order to provide a combination of natural metaphors of communication (voice, gesture, expression) with virtual information (simulation, animation, persistent data). The architecture of the system is based on OpenGL Performer and XML configuration files and it can be easily adapted to many application domains. Another interesting workspace is a wearable AR generic platform that supports true stereoscopic 3D graphics (Reitmayr and Schmalstieg 2001). The system supports six degrees-of-freedom (DOF) manipulations of virtual objects in the near field using a pen and a pad interface. Slay et al. (2001) developed an AR system that extends interactions from a traditional desktop interaction paradigm to a tangible AR paradigm. A range of issues related to the rapid assembly and deployment of adaptive visualisation systems was investigated. Three different techniques, for the task of switching the attributes of the virtual information in AR views, were presented.

Furthermore, the AMIRE project (Haller et al. 2002) aims at developing fast rapid prototyping through vision-based AR for users without detailed knowledge of the underlying base technologies of computer graphics and AR skills. AMIRE uses a component-oriented technology consisting of a reusable GEM collection, a visual authoring tool and object tracking system based on the ARToolKit library. Another system that allows users to create AR experiences is the designer's augmented reality toolkit (DART) (MacIntyre et al. 2005). The system is based on the Macromedia Director multimedia-programming environment to allow a user to visually create complex AR applications as well as providing support for the trackers, sensors and camera.

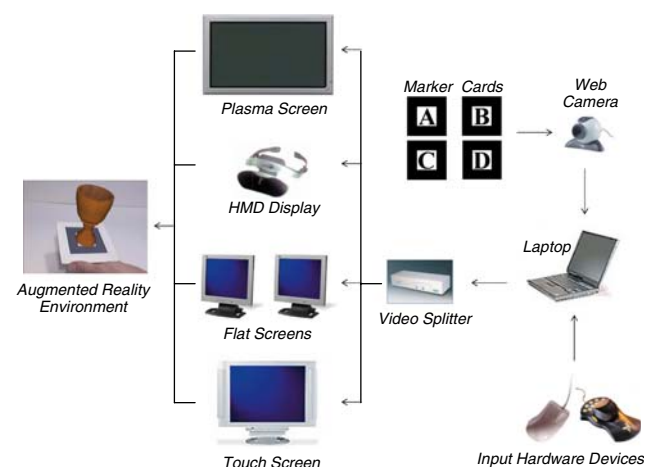
Although most of the above systems describe generic frameworks that allow for AR and/or MR applications, they have not focused on designing a high-level user-focused interface that can deliver audio-visual information. The DART system is the most similar to this approach but it is based on a commercial

multimedia package and thus it is addressed to designers and not general purpose developers. However, this sometimes limits the capabilities of the generated applications because they will be limited to the specific package (i.e. Director). On the contrary, this work is targeting developers who want to develop AR applications and use higher level tools than currently exist (i.e. ARToolKit).

### 3 Architecture of the system

The scope of the AR interface is to provide all the necessary tools for developers to generate user-specific AR applications (see Sect. 7). They will select which sort of functionality is useful, and either use it as it is or extend it to fit the needs of the application. Based on previous prototypes (Liarokapis et al. 2004a, b) a tangible AR interface focused on superimposing five different types of virtual information and allowing users to interact using a combination of five different interaction techniques was designed and implemented. The system allows for the natural arrangement of virtual information anywhere inside the interior of a building or any other type of indoor environment. A diagrammatic overview of the operation of the system is presented in Fig. 1.

In the simplest configuration, a laptop computer with a USB web-camera and a set of trained marker cards are employed. The most complex configuration performed for the purpose of this research included two cy-visor HMDs, four LCD monitors, an 18 in. iiyama touch screen and a 42 in. plasma screen (Sony PFM-42V1N). Depending on the capabilities of the



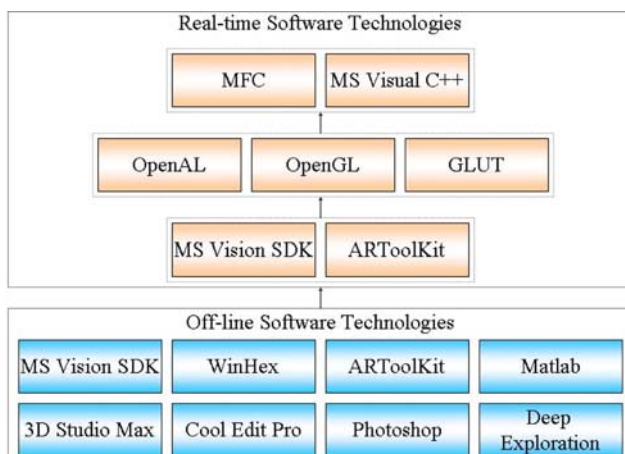
**Fig. 1** Overview of operation of the system



splitter different configurations can be supported depending on the level of immersion and collaboration required. For example, for some applications (i.e. museum environments) the plasma screen could provide an idealistic cognitive environment for collaborative while the touch screen could be preferred as an effective means for user-centred interaction. All displays have been used to present the capabilities of the system in various demonstrations and other dissemination events and the plasma screen found to be the most appealing one. To further increase the level of interaction, a 3D mouse is integrated into the system allowing users to manipulate the virtual information in a natural way in six DOF (see Sect. 6.5).

Audio–visual augmentation techniques have been also been implemented (see Sect. 5) in order to achieve a realistic visualisation such as, matching virtual lighting to real lighting, texture mapping techniques, shading and clipping. To further improve the quality of the visualisation, planar shadows and reflections are generated in real-time so that the user can get a more realistic perception of the augmented information in respect to the real world. It is worth-mentioning that the software and hardware infrastructure of the prototype AR interface developed in this research is based on off-the-self hardware components and low-priced software resources. The hierarchy of the software architecture is presented in Fig. 2.

The blue boxes represent the off-line tools used and which form the basis of the implementation. The technologies in the orange boxes show the software components implemented for the creation of the AR interface. A brief overview of how each technology was used is presented in the following sections.



**Fig. 2** Software technologies

### 3.1 Off-line technologies

The off-line software technologies include a number of commercial tools that must be used before the execution of the AR interface to prepare the content used in the augmentation (i.e. virtual information) as well as the AR environment. Specifically, the ARToolKit's tracking libraries were used for the calibration of the camera (see Sect. 4.2) as well as for the training of new markers designed for the needs of our research. Image processing (Adobe Photoshop) was appropriate for creating appropriate 2D images that were used as part of the visualisation process (see Sect. 5.2) and for generating textures for the 3D models.

To create professional-quality 3D models, 3ds max employed to digitise the models and export them into 3ds format. Next, deep exploration utilised to convert 3ds models into a number of formats including VRML and ASCII. CoolEdit Pro served as a useful off-line tool to record and processes all the necessary wave samples required for the augmentation. WinHex was helpful to analyse the robustness of the markers existing in the AR environment. Finally, the Calibration Toolbox for Matlab was used to improve the camera parameters calculated from ARToolKit (Sect. 4.2).

### 3.2 Real-time technologies

Real time software technologies consist of all the software libraries that have been integrated into a single application that comprise the AR interface. The Microsoft vision software development kit (SDK) was used as a basic platform to develop an interface between the video input (from and video or web cameras) and the rest of the AR application. Based on this, only ARToolKit's (Kato et al. 2000b) tracking library (AR32.lib) was integrated to calculate the camera pose in real-time. On top of the tracking library a high-level computer graphics rendering engine was implemented based on C++ that can perform mathematical operations between 3D vectors and matrices. Standard graphics functionalities like shading, lighting and colouring were based on the OpenGL API (Woo et al. 1999) while more advanced functions like shading and reflection were implemented in the rendering engine to provide a platform for the rapid development of AR applications (Sect. 7).

GLUT (OpenGL utility toolkit) (Angel 2003) was initially used to create a user-interface and to control the visualisation window of the AR interface. In addition, it was used for the textual augmentations (Sect. 5.4) because it provides sufficient support for

bitmap and stroke fonts. However, GLUT provides only a minimum set of functions for the user to control the visualisation and therefore a more advanced solution was implemented based on MFC (Microsoft foundation classes). The advantage of implementing a windows-based interface is that it allows users to familiarise quickly with the GUI (graphical user interface) as well as it provides menus and toolbars to implement any type of user interaction. Finally, OpenAL (open audio library) API was employed to generate audio in a simulated 3D space (Sect. 5.5) because it is similar to OpenGL coding style and it can be considered as an extension of it.

#### 4 Tracking

A key objective of this research was to provide a robust platform for developing innovative AR interfaces. However, to achieve the best tracking (with commercial web-cameras) accurate calculation of the camera parameters is required. As mentioned before, ARToolKit's tracking library was preferred because it seems to provide accurate results with regards to the estimation of the location of the object especially at small distances and in cases where the camera is not moving fast. However, the major flaw of this approach is that all fiducials must be visible continuously. Also in un-calibrated environments, with poor lighting condition, tracking might not work at all. In this section, the results obtained from measuring ARToolKit's error and the algorithms used for calibrating the camera (calculating the camera parameters) are briefly analysed.

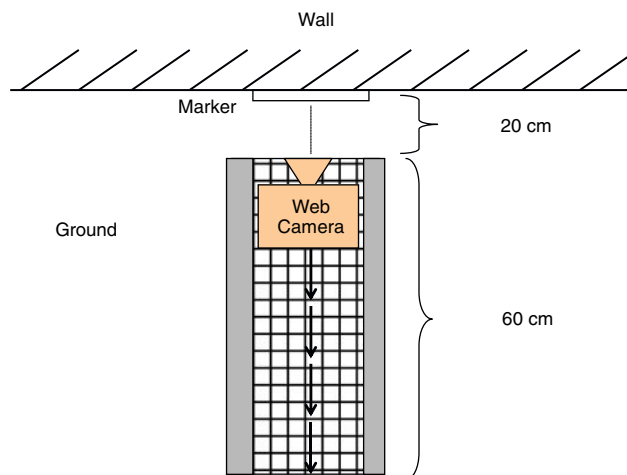
##### 4.1 Measuring ARToolKit's error

ARToolKit was originally designed for small applications working on a limited range of operation, usually around one meter. In these applications the distance between the marker and the user is often small so most of the errors occurred are not easily detectable. But in wide area applications, its positioning accuracy is not very robust. In distances between 1 and 2.5 m the error in the  $x$  and  $y$  values increases proportionally with the distance from the marker (Malbezin et al. 2002). Because this research is focused on indoor environments, it is very important to work accurately in small distances ranging between 1 m and reasonably well for up to 3 m. For this reason an experimental measurement of the accuracy of ARToolKit's tracking libraries was performed in the laboratory environment. The aim of the

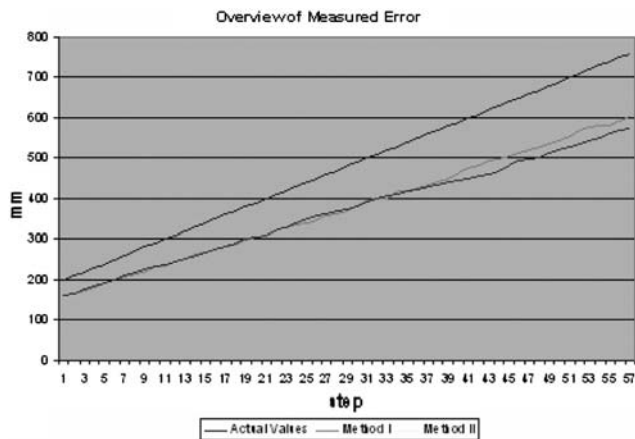
experiment was to evaluate the error in distances ranging between 20 and 80 cm under normal lighting conditions. The experimental apparatus of this procedure is illustrated in Fig. 3.

The optimal area, which contains the least error, is the one that is perpendicular to the marker card. To allow placing the camera on specific points with high precision a grid is positioned on the ground (Fig. 3). Besides, a rigid path was designed so that the camera cannot lose its direction while moving backwards. For each point on the grid, numerous measurements of the location of the web camera in a local coordinate system were taken. The camera is setup in the shortest operating distance (20 cm) and after completing measurements on its position it moves backwards on a step of 1 cm. When the camera moves 60 cm (60 different positions) the program exits. For each position 20 measurements were taken and they were averaged. Figure 4 illustrates the results of this experiment (purple line) showing that the error is proportional to the distance. In very small distances the error in the detection of the marker is small while in larger distances the error becomes considerably bigger. It increases proportionally to the angle of rotation when the camera does not change position, but it is rotated around the Y-axis.

To verify that the best location is on the perpendicular axis of the camera and the marker, another set of measurements were recorded with the camera facing the marker at variable angle (yaw) having the other two (pitch, roll) stable. In this case, the camera was setup again in the same plane (ground plane) but the



**Fig. 3** Experimental setup for the measurement of ARToolKit's error



**Fig. 4** Comparison of measured values

measurements were taken when the  $x$  and  $y$  values tended to zero values. It was measured that the angle in the initial position (20 cm from the wall, Fig. 3) is approximately  $12^\circ$  while for the final position the angle is approximately  $4^\circ$ . It is worth mentioning, that the second sets of measurements were not done automatically, so on each step the camera had to be manually adjusted to provide values as close as possible to values of  $x$  and  $y$  to zero. Figure 4 shows the results from the second experiment and illustrates that the measured error, when the camera is not pointing directly to the marker, is proportional to the distance. However, the difference is of minimal significance. This means that when the camera lies with a certain area and does not change its orientation the error is quite small. In contrast, if the camera changes direction the error increases considerably. Figure 4 illustrates differences in the errors produced from the experiments compared with the actual value (top dark line).

Figure 4 shows that the best results can be obtained when the camera is oriented to point at the centre of the marker. Even if the camera has a small offset, then the error increases linearly with the distance. Nevertheless, the tracking results are acceptable in the area of less than 1 m since the error is hardly noticed.

## 4.2 Camera calibration

This section describes the procedures used in order to calculate the intrinsic and extrinsic camera parameters. The purpose for this was to define an accurate camera model that can be effectively applied into indoor AR environments. Although there are a few camera calibration techniques available (Weng et al. 1992; Shi and Tomasi 1994) for calculating the intrinsic camera parameters, ARToolKit's calibration library (Kato et al. 2000b) was preferred since it works reasonable

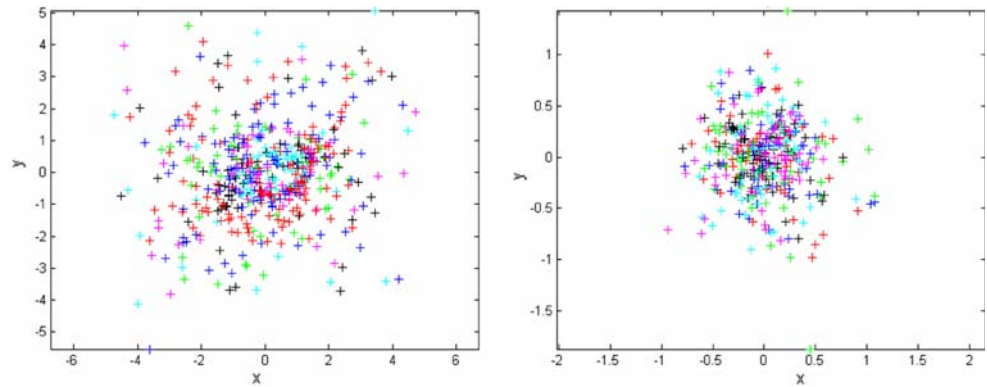
good in small distances and in cases where the camera is not moving fast. This method was originally applied to measure the camera model properties such as: the *center point* of the camera image; the *lens distortion*; and the camera's *focal length*. ARToolKit provides two software tools that can be used to calculate these camera properties, one to measure the lens distortion and the image center point, while the second to compute the focal length of the camera (Kato et al. 2000b). Based on this, the initial calibration was performed since it produces reasonable results for the calculation of the intrinsic camera parameters.

However, the greatest limitations of this vision solution include the tracking accuracy and the range of operation (Malbezin et al. 2002). To minimise some of the errors produced in the tracking of the markers, the extrinsic camera used had to be accurately estimated. The virtual objects will only appear when the tracking marks are in view. The size of the predefined patterns influences the effectiveness of the tracking algorithms. For instance, if the pattern is large then the pattern is detected further away. To calculate the extrinsic camera parameters the camera calibration toolbox (Camera Calibration Toolbox for Matlab 2003) was used which provides a user-friendly interface and it is very convenient when working with a large number of images. Another advantage over the previous method is that it provides very accurate results.

Before the camera calibration begins two steps need to be initially followed. In the first step the calibration rig must be generated while in the second all the calibration images must be collected. When done, the grid corners are easily extracted. The ToolKit offers an automatic mechanism for counting the number of squares in each grid and all calibration images used are searched and focal and distortion factors are automatically estimated. However, similarly to ARToolKit method, in most occasions the algorithm may not predict the right number of squares and thus provides a poor result. This can be clearer by observing the results of the calculation of the re-projection error. As it is clearly observed from Fig. 5a, the re-projection error is quite big compared to the scale. The reason behind this is because the extraction of the corners is not acceptable on some highly distorted images. However, the advantage of this technique is that it allows the user to improve the calibration. Specifically, the whole procedure can be repeated until the error is minimised up to a certain point.

After repeating the procedure for five times the error is reduced from a scale of five to a scale of one as illustrated in Fig. 5b. Furthermore, because ARToolKit accepts only binary data format for the calibration, a simple way to do this is to estimate the extrinsic

**Fig. 5** Calculation of camera error **a** re-projection error **b** minimisation of error



parameters and then save the computed parameters in the data structure replacing the old values. The old data structure that holds the calculated camera parameters (ARParam struct) is shown below:

```
typedef struct {
    int xsize, ysize;
    double mat [3][4];
    double dist_factor [4];
} ARParam;
```

In this structure the *xsize*, *ysize* and *dist\_factor* have been experimentally replaced with the new values calculated from the above. Specifically, the camera parameters including the focal length (*fc*), the principal point (*cc*), the skew (*s<sub>k</sub>*) and the distortion (*k<sub>c</sub>*) have been computed and based on these values the intrinsic matrix can be defined as shown in Eq. (1):

$$\begin{pmatrix} fc_0 & s_k fc_0 & cc_0 \\ 0 & fc_1 & cc_1 \\ 0 & 0 & 1 \end{pmatrix} \tag{1}$$

Since ARToolKit does not take into account the skew factor and makes use of the following matrix:

$$\begin{pmatrix} s_x f & 0 & x_0 \\ 0 & s_y f & y_0 \\ 0 & 0 & 1 \end{pmatrix} \tag{2}$$

To match the outputted camera matrix from Matlab and fit it into ARToolKit’s matrix, the following matrix can be derived:

$$\begin{pmatrix} fc_0 & 0 & cc_0 \\ 0 & fc_1 & cc_1 \\ 0 & 0 & 1 \end{pmatrix} \tag{3}$$

After testing the new camera model a small improvement was succeeded in the distortion in the magnitude of 3–4%. As a further improvement, it was

decided to add the skew parameters to the matrix, thus the skew parameter was used instead of zero in the matrix as shown below:

$$\begin{pmatrix} fc_0 & s_k fc_0 & cc_0 \\ s_k & fc_1 & cc_1 \\ s_k & s_k & 1 \end{pmatrix} \tag{4}$$

Although the last modification provided us with a more correct camera model with an estimated improvement of about 1%, the effectiveness of the tracking system was not significantly improved. This is due to the fact that the optics used in the camera system (web camera) is really poor compared to professional video cameras. Other environmental issues that influence tracking include lighting conditions and range of operation.

### 5 Audio–visual augmentations

Each type of virtual media information is designed for specific purposes and as a result produces different outcomes. For instance, textual explanation can be utilized much more effectively than auditory description when communicating verbal information. On the other hand, pictures work better than text, for recalling or explaining diagrammatically a procedure. To describe a sequence of events video seems to be one of the most efficient techniques. In this section, the methodology used for the simultaneously multimedia visualisation of virtual information into an AR indoor environment is presented.

#### 5.1 Object augmentation

An ideal AR system must be able to mix the virtual information with the real in a physical way. The participants should not realize the difference between the real and the augmented visualisation. The focus of this research is to present and implement methods of

realistically rendering 3D representations of real objects in an easy and interactive manner. The selection of the most appropriate 3D format is a crucial task in order to achieve a high level of realism in the system. In this research, both 3ds and VRML file formats have been used as shown in Table 1.

In any case, one of the first problems derived when displaying a 3D representation of a real model is the correct alignment on the required position. Virtual

must be calculated based only on the plane equation coefficients and the position of the light (Moller 1999). Say that  $L$  is the position of the point light source;  $P$  the position of a vertex of the AR object where the shadow is cast; and  $n$  the normal vector of the plane. The projection matrix of the shadow can be calculated by solving the system, which consists of the equation of the plane and a straight that passes from the plane point in the direction of the light source (see Eq. 5).

$$P_s = \begin{pmatrix} Lp \bullet Pc - Lp_0 \times Pc_0 & 0 - Lp_1 \times Pc_0 & 0 - Lp_2 \times Pc_0 & 0 - Lp_3 \times Pc_0 \\ 0 - Lp_0 \times Pc_1 & Lp \bullet Pc - Lp_1 \times Pc_1 & 0 - Lp_2 \times Pc_1 & 0 - Lp_3 \times Pc_1 \\ 0 - Lp_0 \times Pc_2 & 0 - Lp_1 \times Pc_2 & Lp \bullet Pc - Lp_2 \times Pc_2 & 0 - Lp_3 \times Pc_2 \\ 0 - Lp_0 \times Pc_3 & 0 - Lp_1 \times Pc_3 & 0 - Lp_2 \times Pc_3 & Lp \bullet Pc - Lp_3 \times Pc_3 \end{pmatrix} \quad (5)$$

objects may appear to float on the marker and the user will be easily confused. This usually occurs because the 3D model is not registered correctly into the scene. For example, when a 3D object is transformed into the real scene it may appear below the origin as illustrated in the left image of Fig. 6.

To correct the problem of misalignment, Fig. 6a, a sorting algorithm for registering 3D objects precisely onto the top of the markers was implemented. To achieve a correct registration the virtual information need to be first sorted and then initialised to exactly the same level, as the marker is located in the  $Z$ -axis. An efficient way to align objects is by using a two-stage process. In the first part, the vertices of the object are sorted by the  $Z$ -axis. Upon completion, the vertices are translated to the minimum value, which is the origin of the marker cards, resulting in a proper object registration.

Next, to improve the realism of the AR scene a fast algorithm for planar shadows and reflections was implemented. The location of the shadow can be calculated by projecting all the vertices of the AR object to the direction of the light source. To generate augmented shadows an algorithm that creates a  $4 \times 4$  projection matrix ( $P_s$ ) in homogeneous coordinates

where  $L_p \cdot P_c$  is the dot product of plane and light position. The projection matrix has a number of advantages compared with other methods (i.e. fake shadows) but the most important is that it works fast and it is generic so that it can generate hard shadows in real-time for any type of objects independently of their complexity (Liarokapis 2005). An example screenshot that illustrates planar shadows is shown in Fig. 7.

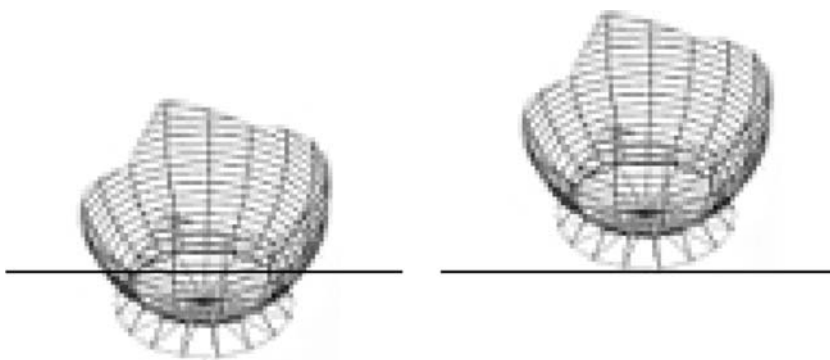
The main disadvantage of this algorithm is that it renders the virtual information twice for each frame: once for the virtual object and another one for its shadow. Another obvious flaw is that it can cast shadows only into planar surfaces but with some modifications, it can be extended to be applied to specific cases such as curved surfaces (Liarokapis 2005).

To realistically model reflections in AR environments, many issues must be taken into account. Although in reality the light is scattered uniformly in all directions depending on the material of the object in this work, the effect of mirror reflections has been implemented. An example screenshot of a virtual object casting a shadow and a reflection on a virtual plane is illustrated in Fig. 8.

**Table 1** Categorisation of 3D file formats

Advantages	Disadvantages
3ds	
Includes pre-vertex texture coordinates	3D can have 216 vertices maximum
Unknown parts can be skipped	Poor normal information
VRML	
Easy to read	Contains less information than 3ds
Standard for 3D internet presentations	Does not support advanced lighting and texturing
Contains animation and collision detection	

**Fig. 6** Object augmentation **a**  
misalignment of object **b**  
correct registration



**Fig. 7** Illustration of planar shadows



**Fig. 8** Planar shadows and reflections

Based on the OpenGL's stencil buffer a reflection of the object is performed onto a user-defined virtual ground. The stencil buffer is initially set to sixteen bits in the pixel format function. Then, the buffer is emptied and finally the stencil test is enabled.

## 5.2 Image augmentation

Images are widely used as a means to increase realism and in the past, they have been used with success for educating purposes. The augmentation of images is a highly cost effective means to present simple 2D information in the real world. The use of their operation may be performed in a number of different ways depending on the learning scenario applied. The digital image augmentation can be either static or dynamic. Dynamic image augmentation is widely used for achieving video augmentation (see Sect. 5.3). With static augmentation, a single image only is rendered into the scene. Based on the theoretical framework provided by Smith (1994), images used for AR environments have been categorised into description, symbolic, iconic and functional as shown in Table 2.

The algorithm used is simple but very efficient and can be applied into two types of image formats (BMP and TGA). First, it loads an image file and checks if it is a valid image format. In the next step, textures are generated using data from the image file. Following this, the texture is created and the parameters are set based on the OpenGL API. Finally, the texture is bound to the target texture, which is a quadrilateral.

## 5.3 Video augmentation

The mode of operation within the video AR system is to read an AVI file, decompose it into  $256 \times 256 \times 24$  bit images, mix it with the dynamic video (coming from the camera) and finally display it on the selected visualisation display (Liarokapis 2005). When

**Table 2** Categorisation of images augmentation

Purpose	Usage
Description	
Most popular format	Explain a 3D real object
Describe the real world	Textual information have a useful meaning
Image itself can tell a self-explanatory story	
Symbolic	
Identify a basic principle or symbol	Concerns images that represent various types of well known symbols
Allow both simple and complex symbolism	
Interpretation can change over time	
Iconic	
Identify a case of a multinational meaningful icon that is not related to a specific language	Image contains different types of iconic representations that can illustrate something useful
For example the “exit” or “danger” sign	
Functional	
A single operation can be expressed	Functional images act as virtual buttons and a specific operation is assigned on each
Multiple operations can be also supported	

the video file is loaded, the program automatically counts the number of frames so that its size is known. Then all frames are decomposed into 2D images and each image is applied to a square quad, exactly in the same way as textures are wrapped to objects. It is worth mentioning here that because each animation has a specific length (in seconds) and its own frame rate, the time required for each frame is calculated.

Moreover, the augmented video starts automatically when two things occur: a marker is detected and user has loaded a particular file from the filling system using the interface menu. When the animation is completed it repeats itself until the user decides to stop it (by pressing a keyboard key or using the interface menu). To increase the feasibility of the system, if the camera is not in line of view with the marker card, then the video augmentation will continue playing until the video sequence is finished. This was designed on purpose to prevent cases where the user changes position or orientation rapidly and thus loses the perceived visualisation. The augmented animation can be controlled in a number of different means including the cease of the animation, resize the animation window or even manipulate the augmented video animation into six DOF (see Sect. 6).

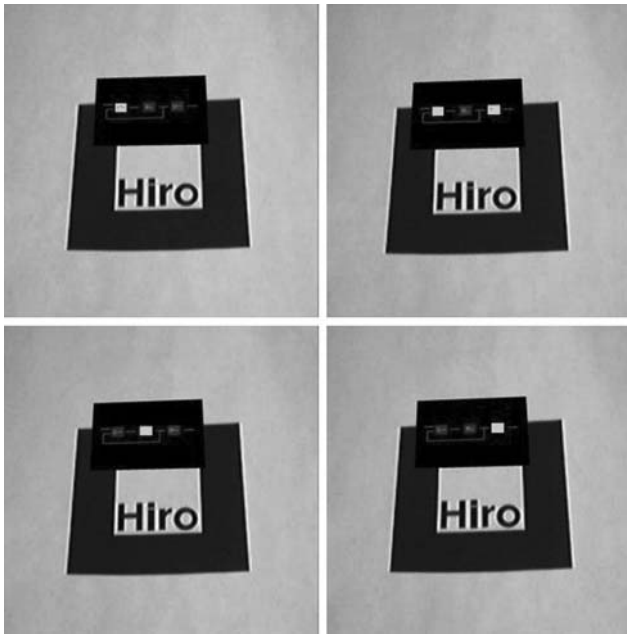
Figure 9 shows four frames of a video sequence superimposed into a marker card, describing a complex concept in electronics (i.e. Moore Diagram). In terms of efficiency, the video augmentation results range between 20 and 35 FPS depending on the resolution of the videos. However, the drawback of this method is that when the animation is augmented the overall performance of the system is significantly reduced. In particular, the performance was experimentally

measured to be reduced by approximately 20–50% the FPS of the system. For instance, if the performance of the system is in real-time (i.e. 25 FPS) then the AR video algorithm would drop the performance to 12–20 FPS. Another limitation of the proposed video augmentation is that it can currently decompose videos into only  $256 \times 256 \times 24$  bit images.

#### 5.4 Textual augmentation

Textual annotations are the simplest form of information that can be easily augmented in any type of AR environments. This can be either presented as a *label* or as a *description*. Label text has been used in the past (Klinker et al. 1997; Sinclair and Martinez 2001), to point out specific parts of a complex system using the minimum textual information. In this case, the most important aspect is to ensure that the augmented labels do not obscure each other and that the information is clearly presented to the user. Description text requires a much more demanding process because it needs to provide complete information about an object or about a virtual operation. The problems begin in cases where the magnitude of textual information needs to be augmented on a display is large.

In this research, label and descriptive textual information was performed by dynamically loading ASCII text files. Each file contained a very different level of information depending on the reasons for utilising it. For example, label text files were defined to specify the type of visualisation (i.e. image augmentation) or the name of an object. The main advantage of this method is that the textual information, which will be augmented on the real environment, is stored on a txt file. Text files are widely used and can be easily transferred



**Fig. 9** Video augmentation

over all types of networks. Users of the system can position textual augmentations anywhere in the real environment using standard transformations. In addition, they can change their appearance in terms of colour, size and font type (Bitmap, Times Roman and Helvetica).

### 5.5 Audio augmentation

In the real world, audio is a process that is heard spatially and thus it is a very important aspect for any simulation scenario. The most important issue when designing 3D sound is to “see” the sound source (Yewdall 1999). However, most AR applications have not incorporated 3D sound component even if it can contribute to the sense of immersivity. The augmented sound methodology followed in this work, has some similarities with the ASR approach (Dobler et al. 2002) in the way virtual sounds are augmented in the real environment. Unlike this experimental approach, which is based on a Creative EAX API, the implemented 3D audio system is based on OpenAL, which has many similarities with OpenGL and was originally designed for generating 3D sounds around a listener. The recording of speech sounds was done in mono format, using a standard microphone and the mono samples were converted into stereo format. Furthermore, each sound source in the system has been specified to have the following three properties: *position*, *orientation* and *velocity*. The spatial audio system can handle multiple sound sources and mix them together.

The user can move the sources in 3D space using the keyboard and menu interaction techniques illustrated in Sect. 6.

The spatial sound algorithm first initialises all the necessary OpenAL variables (position, orientation and velocity) and then loads them into the appropriate buffers (format, length and frequency) for further processing. Next, sound sources and buffers are initialised and the sources are assigned to the buffers. The pitch and gain are set to one and the sources are set into a continuous loop unless stopped by the user. Each time the camera detects the marker the transformation matrix is inverted to estimate the position of the camera. In the context of this research, the distance model experimentally applied used to simulate the distance followed the linear equation as illustrated below:

$$y = \alpha x + \beta \quad (6)$$

where  $\alpha$  represents distance between the camera and the marker and  $\beta$  the offset position of the marker card. Although this cannot accurately represent the distribution of sound in 3D space it provides very good results. To provide more freedom to the listener the values of the linear function may change depending on the requirements of the visualisation. If the sound source is positioned in the origin then the above equation may be re-written as shown below:

$$Listener = \frac{camera\_position}{distance\_factor} \quad (7)$$

where *camera\_position* refers to the inverse transformation of the camera and *distance\_factor* to a constant number. To achieve a realistic simulation of the sound different values have been tried to simulate the *distance\_factor*. However, this constant value may change off-line depending on the requirements of the visualisation. For example, some users may prefer to perceive the auditory information louder than others do. In addition, the system is capable of loading and mixing music sound files. This option can be extremely useful for simulating surround music audio. The sound files may be overlaid into the same marker or onto a different marker depending on the needs of the application.

## 6 Human–computer interactions

Human–computer interactions are one of the most important issues when designing a robust real-time system. They have to be performed in a natural way so that inexperienced participants familiarise quickly in the AR environment (Liarokapis et al. 2004a). The proposed

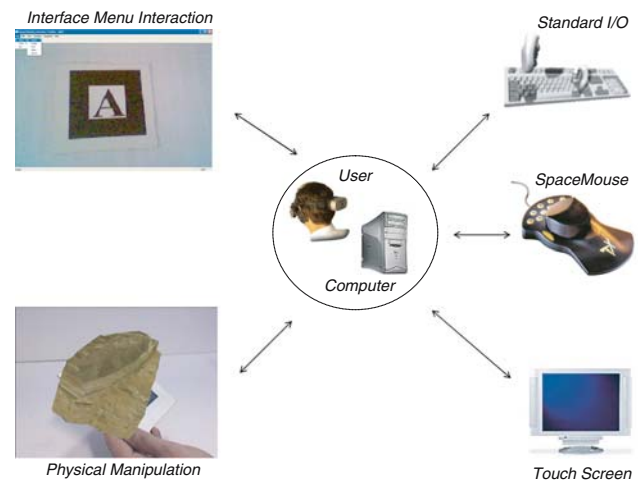


interface allows users tangible interaction with various types of multimedia information such as 3D models, images, textual information and 3D sound, using a number of interaction techniques. Interactions controlled by the user-computer can be distinguished into five different categories including physical manipulation, interface menu interaction, standard I/O, Touch Screen and SpaceMouse interaction as illustrated in Fig. 10.

Although, some types of interactions proposed in Fig. 10 are not novel (i.e. physical manipulation), the novelty comes in the way they are used by the participants. Participants can combine two or more types and experience a novel form of interaction with great flexibility. For example, the most significant combination of human-computer interactions is the use of intuitive methods like the physical manipulation with sophisticated devices such as the SpaceMouse. Users can hold in one hand a marker card with a virtual object superimposed and on the other hand use the SpaceMouse to perform graphics operations like virtual lighting. In the following sections, all the types of interactions are explained in detail.

### 6.1 Standard interactions

The first method is addressed to users with some computer experience and is based on standard



**Fig. 10** Interactions within the system

the requirements of this research because it provides only the minimum functionality to rotate an object around  $X$ ,  $Y$  or  $Z$ -axis. However, in a tabletop AR environment this is constraining the user when rotating the virtual information as well as it restricts the use of simultaneous interactions. To tackle this problem a generic rotational matrix that takes as input three angles and rotates the object around an arbitrary axis is specified in Eq. 8:

$$\begin{bmatrix} \cos \psi & -\cos \sin \psi & \sin \psi \\ \sin \varphi \sin \psi + \cos \varphi \sin \psi & -\sin \varphi \sin \psi + \cos \varphi \cos \psi & -\sin \varphi \cos \psi \\ -\cos \varphi \sin \psi + \sin \varphi \sin \psi & \cos \varphi \sin \psi + \sin \varphi \cos \psi & \cos \varphi \cos \psi \end{bmatrix} \quad (8)$$

interaction input devices like the keyboard and the mouse. For example, by pressing buttons (hot keys) the visual parameters of the virtual objects can be changed faster instead of using the menu dialogues. Some of the most characteristic are described in (Liarokapis et al. 2004a, b; Liarokapis 2005) and include the change of lighting conditions (ambient, diffuse, specular and shininess); the texturing information (standard and environmental); the switch from solid mode to wireframe mode; and others (see Sect. 6.2). Moreover, the keyboard is also employed for changing the position (translation), orientation (rotation) and scaling of the virtual information in six DOF. Initially, the above transformations were implemented based on the OpenGL functionality but soon it became obvious that OpenGL could not meet

Based on the above rotational matrix, it became possible for users to rotate virtual information around an arbitral-defined axis. The above matrix was also implemented to the standard mouse providing a quick way to perform intuitive rotations. Although, it provides the means to perform a rotation around all three axes simultaneously if one interaction device is used, problems occur when more than one device is used (i.e. keyboard and mouse). An alternative way of performing transformations is by using quaternions. To specify multiple rotations, many intermediate control points are required where a quaternion interpolation depends only on the relation between the initial and final rotations. The easiest way to prove the link between a rotation matrix and a quaternion is by linking them in three dimensions. Say that  $q = s + v \cdot I$  a unit

quaternion and defined  $Q$ , where  $v = (u_x, u_y, u_z)^T$ , it can be shown that there is a  $3 \times 3$  matrix that represents a rotation matrix of the form (Eq. 9):

$$\begin{aligned}
 &vv^T + (sI_{3 \times 3} + C_u)^2 \\
 &= \begin{pmatrix} s^2 + u_x^2 - u_y^2 - u_z^2 & 2(u_x u_y - s u_z) & 2(u_x u_z + s u_y) \\ 2(u_x u_y + s u_z) & s^2 + u_x^2 + u_y^2 - u_z^2 & 2(u_y u_z - s u_x) \\ 2(u_x u_z - s u_y) & 2(u_y u_z + s u_x) & s^2 - u_x^2 - u_y^2 + u_z^2 \end{pmatrix} \tag{9}
 \end{aligned}$$

To obtain a quaternion corresponding to a given rotation matrix we first define an arbitrary rotation matrix  $R$  and then the corresponding quaternion  $q = s + u_x i + u_y j + u_z k$  to the rotation matrix. Using the above equation it is easy to solve the equation and derive the values for  $u_x$ ,  $u_y$  and  $u_z$ , respectively. In OpenGL, rotations are specified as matrices since homogeneous matrices are the standard 3D representations. By combining the property of unit quaternion with the above rotation quaternion matrix we can deduce the following equation (Eq. 10):

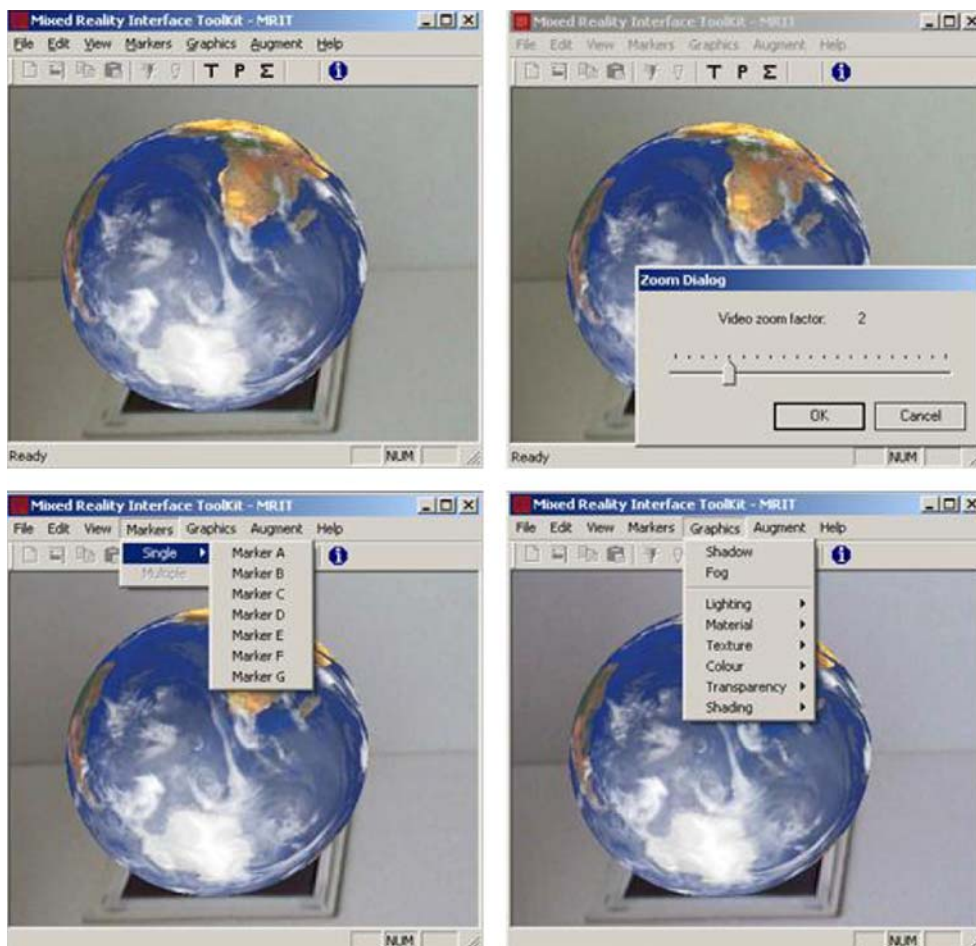
$$\begin{aligned}
 &vv^T + (sI_{3 \times 3} + C_u)^2 \\
 &= \begin{pmatrix} 1 - 2(u_y^2 - u_z^2) & 2(u_x u_y - s u_z) & 2(u_x u_z + s u_y) \\ 2(u_x u_y + s u_z) & 1 - 2(u_x^2 + u_z^2) & 2(u_y u_z - s u_x) \\ 2(u_x u_z - s u_y) & 2(u_y u_z + s u_x) & 1 - 2(u_x^2 - u_y^2) \end{pmatrix} \tag{10}
 \end{aligned}$$

Other functions that were integrated to the mouse include translations and scaling. On the other hand, using the mouse users can access the carefully designed GUI. This allows users to have full access to the superimposed virtual information. An example is presented in Fig. 11, where users can select the information that is going to be augmented on the real environment.

### 6.2 GUI Interactions

On the other hand, using the mouse or the Touch Screen users can access the functionality that has been carefully integrated into a novel GUI. The GUI consists of a menu, a toolbar, a status bar and a number of

Fig. 11 GUI functionality

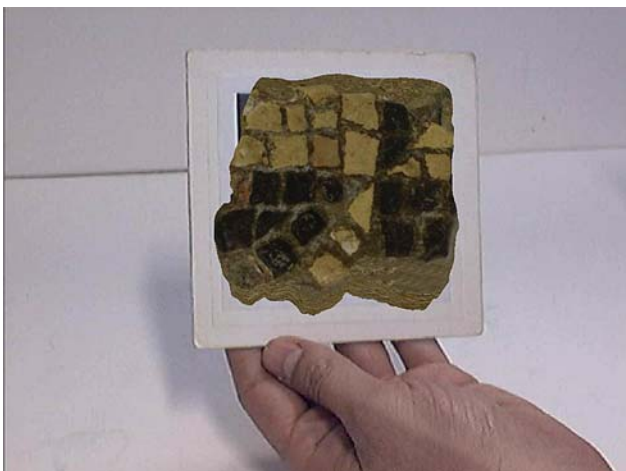


dialog boxes. This allows participants to have the same access to the augmented virtual information as if they were using standard interaction techniques. Four example screenshots that illustrate some of the functionalities of the GUI is presented in Fig. 11.

The greatest advantage of the proposed GUI is that it allows participants to perform complex operations very accurately. Specifically, sometimes it is of crucial importance to transform a virtual object in a specific location in the real environment. Using other methods it could take a great amount of time and effort (depending on the experience of the user) to achieve this and it will definitely not be very accurate. However, the GUI interaction techniques offer the solution to the problem using double point precision.

Next, the “Edit” category consists of three basic operations including video (start or stop), a zoom dialog box and a scale dialog box. The “View” category consists of two sets of operations. Firstly, a Toolbar and a Status bar, which is commonly found in windows based applications. It is worth-mentioning here that the GUI has been built on top of the windows API so that full compatibility with windows based operating systems is ensured. As far as this research is concerned this is the only true windows based interface that can superimpose five different types of multimedia content into the real environment.

The second set of operations consists of three functions called axis (to insert a Cartesian set of axis indicating the origin of the AR environment), debug (to threshold the live video sequence and thus check whether a marker is detectable) and clip (to clip the graphics geometry). The rest of the menu categories (graphics and augment) are used to control visualisation properties of the augmented information. Func-



**Fig. 12** Natural manipulation of virtual object

tions that have been implemented include shadows, fog, lighting, material, texturing, colouring, transparency and shading. Finally, the “help” category provides some information about the release version of the AR interface as well as the date and the author name.

### 6.3 Physical manipulation

Physical manipulations were specifically designed for users with no computer experience and refers to a physical manipulation of the marker cards (Kato et al. 2000a; Billinghurst et al. 2001). As illustrated in Fig. 12, users can manipulate freely the marker cards in six DOF to receive a different perception of the superimposed information.

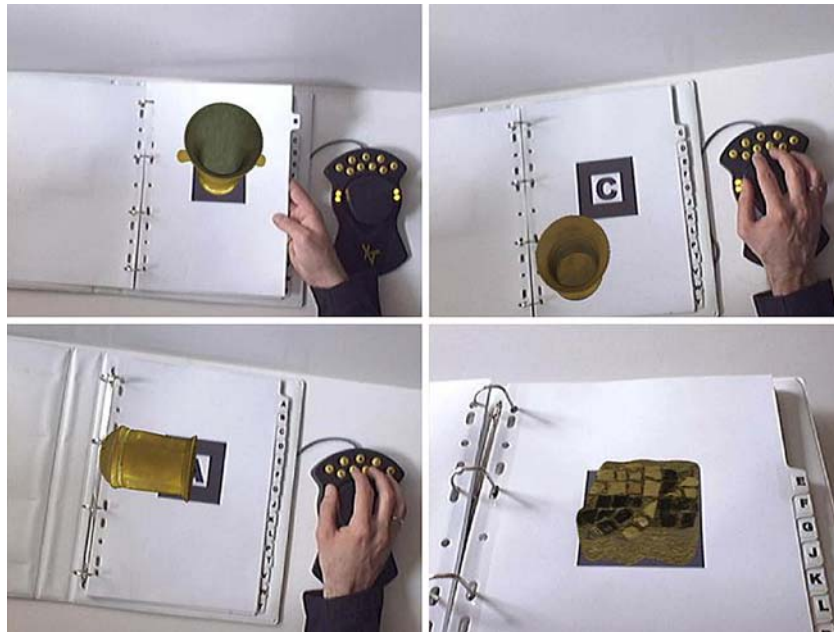
Another benefit of natural interactions is that they can be used with the other types of interactions described in this section. This allows producing unique combinations (see Fig. 14) that can provide solutions for specific AR applications that require a high-level of interaction. In addition, apart from using the marker cards for just superimposing virtual information, they have been used to perform some basic operations such as: assign an object into a marker, de-assign an object from a marker, scale, rotate and translate. The advantage of this method is that users can use only physical objects (marker cards) to visualise and interact with the virtual information. However, the disadvantage is that when multiple markers are used the overall efficiency of the system is reduced. Specifically, the template matching algorithm used operates very

```

Read Event from hardware device
if Translation button
    if Tx is positive then Translate forward in X axis
    if Tx is negative then Translate backward in X axis
    if Ty is positive then Translate forward in Y axis
    if Ty is negative then Translate backward in Y axis
    if Tz is positive then Translate forward in z axis
    if Tz is negative then Translate backward in Z axis
if Rotation button
    if Rx is positive then Pitch right in X axis
    if Rx is negative then Pitch left in X axis
    if Ry is positive then Yaw right in Y axis
    if Ry is negative then Yaw left in Y axis
    if Rz is positive then Roll right in Z axis
    if Rz is negative then Roll left in Z axis
if Scaling button
    if S is positive then Scale up by a predetermined factor
    if S is negative then Scale down by a predetermined factor
if Lighting button
    Ambient Lighting is turned on/off
if Ground button
    A virtual ground drawn to occlude the marker card is turned on/off
if Clipping button
    An invisible plane that clips the AR scene is turned on/off
if Reset button
    Initialise the transformations to default values
  
```

**Fig. 13** Pseudo code for SpaceMouse

**Fig. 14** SpaceMouse interactions



effectively in real-time performance with one marker but it starts to decrease drastically as more markers are added. The reason behind this is because for each marker the algorithm is aware; it creates four templates, one at each orientation. Each marker has to be compared to all known templates until the best match is detected. To calculate the number of comparisons performed by the algorithm the following equation is illustrated:

$$N_c = 4 \times N_m \times N_t \quad (11)$$

where  $N_c$  is the number of comparisons,  $N_m$  is the number of known markers and  $N_t$  corresponds to the number of known templates. If in the scene there are 10–20 markers and the application knows about 250 markers then the system performs around 10,000–20,000 comparisons. This makes any system to run much slower and makes the application operate in less than 25 FPS. Thus, to achieve a fast AR application in the final system it was preferred to use as few markers as possible having as limit ten markers.

#### 6.4 Touch screen interactions

An alternative way of interacting with the virtual information is to make use of interaction devices such as Touch Screens. This is ideal for some application scenarios where, the use of other interaction devices is not possible. For example, in museum environments, Touch Screens are the most appropriate means of interacting with the virtual exhibitions. Besides, although it was easy to integrate the Touch Screen to the

AR interface, many problems arose when users tried to interact with the GUI menu. The reason for this is because the menus in the GUI were too small and it was difficult for some users to select. To tackle the problem, large toolbar buttons and dialog boxes were designed and associated with appropriate functionality. The main advantage of using the Touch Screen is that it can serve both the visualisation and interaction all in one device. However, the major drawback is that the effectiveness of the interactions is dependent on the effectiveness of the GUI. If the GUI is not user-friendly, it will affect the usefulness of the Touch Screen interactions.

#### 6.5 SpaceMouse interactions

Finally, users can manipulate virtual information using sophisticated VR sensor devices such as SpaceMouse (Liarokapis et al. 2004b) and InertiaCube. SpaceMouse allows the programmer to assign functionality to provide a customised nine button-menu interface. This method has the advantage manipulating virtual information in six DOF in a natural way using only one hand. A combination of C++ functions, SpaceMouse commands and OpenGL allowed the integration of the 3D mouse into the system. Important functionalities that have been implemented and assigned to the menu buttons include either standard graphics transformations for easier manipulation, or more advanced graphics operations (Fig. 13).

In Fig. 13,  $S$  represents the scaling operations,  $T_x$ ,  $T_y$  and  $T_z$  represent the translations and  $R_x$ ,  $R_y$ , and  $R_z$  the rotations. To perform one of the above operations

the user has to press one of the buttons (the translation button for example) and then use the bar to translate the object in 3D space. Depending on which direction force is applied, the object will move respectively. Furthermore, the ambient lighting, the clipping of superimposed geometry through an infinite plane and the augmentation of a virtual plane can be switched on and off using the remaining SpaceMouse buttons. Four example screenshots of a user interacting with 3D information using the SpaceMouse is illustrated in Fig. 14.

It illustrates how a user can adapt the MagicBook approach (Billinghurst et al. 2001) in conjunction with the SpaceMouse to visualise and interact with the virtual artefacts. On the top left image, the user is only visualizing the virtual artefact (marker B) while on the top right image, the user translates the artefact using the SpaceMouse. On the bottom left image, the user interacts (rotates) with another artefact (marker A) and on the bottom right image the user visualises another artefact (belonging on marker E). The most important limitation of this tangible interface is the use of a single marker for tracking by the computer vision based tracking system.

## 7 Application scenarios

To test the functionality of the proposed AR interface system two application scenarios have been designed. The first section (see Sect. 7.1) presents an educational application used to support and simplify teaching and learning techniques currently applied in the higher education sector. The second section (see Sect. 7.2) illustrates a museum application with the aim of facilitating access to museums and other cultural heritage galleries. In the following sections, each application is briefly analysed and the most important findings of the research are presented.

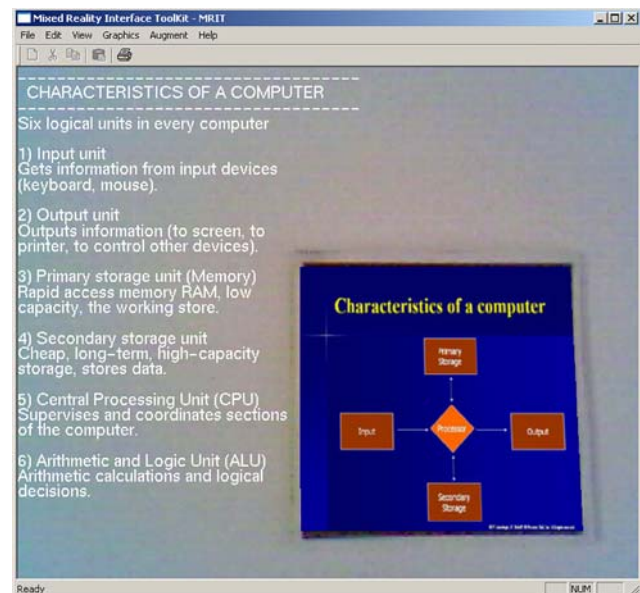
### 7.1 Educational application

Most educational AR applications operate in indoor environments (Begault 1994; Fuhrmann and Schmaltieg 1999) and the scenarios proposed in this section are focused on enhancing the teaching and learning process for higher education institutions like colleges and universities. With this purpose in mind, AR educational scenarios have been designed to assist teachers to transfer knowledge to the students in other ways than traditionally has been the case (Liarokapis 2005). The aim is to provide a rewarding learning experience that is otherwise difficult or impossible to obtain by

offering the ability to achieve better user interaction (with teaching material and complex tools) while the provision of an interactive augmented presentation provides students a high degree of flexibility and understanding of the teaching material. All scenarios are specifically engaged with the improvement of learning and teaching techniques in the fields of engineering and informatics at the University of Sussex.

Based on the functionality of the AR interface described in the above sections, a lecture was prepared introducing students on how computers work. This application has in practice some similarities with the experimental application proposed by Fernandes and Miranda (2003). However, the higher education application offers a very powerful user interface that allows audio–visual augmentation as well as simultaneous interactions. From a visualisation point of view, the system displays the data in a single window and the lecturer can describe basic IT principles with the use of AR technology in a number of different ways. In Fig. 15, a PowerPoint slide presentation that describes the characteristics of a computer system as well as relative textual information is augmented onto the appropriate marker card.

Learners can zoom into the diagram in two ways. Firstly, by using the predefined functionality (scale and translate) existing in the keyboard, the menu and the SpaceMouse interfaces. Alternatively, learners can either move the marker card intuitively closer to the camera and vice versa. In both ways, potential users can clearly observe and understand the theoretical



**Fig. 15** Teaching IT using AR

operation of a computer. The textual information describes the diagram in detail providing a more complete learning presentation simultaneously. Learners can now get simultaneously appropriate audio–visual information that helps them to acquire a deeper knowledge about the characteristics of a computer. In the same way, to increase the level of understanding of the teaching material presented to the students, 3D information can be presented to deepen the level of knowledge transfer. Along these lines, learners can have a more rounded idea of what are the main characteristics of a computer, what are the main parts and how they look like in reality.

The main advantage of the educational application over the traditional teaching methods is that learners can actually “see” and “listen” the virtual information superimposed in the real world (Liarokapis et al. 2002). Students can naturally manipulate the virtual information using standard or sophisticated VR devices and they can repeat a specific part of the augmentation as many times as they want. Another benefit of the system is that it does not require students to have any previous experience to operate it. Finally, even AR has been experimentally applied for teaching engineering and information technology (IT) courses it has been designed in such a way that it can be easily adapted and applied very easily to other educational courses.

## 7.2 Cultural heritage application

The concept of virtual exhibitions in museums has been around for many years and researchers have designed and developed several applications (Liarokapis et al. 2004a, b; Hall et al. 2001; Gatermann 2000). In addition, a number of museums hold innumerable archives or collections of artefacts, which they cannot exhibit in a low cost and efficient way. Museums simply do not have the space to exhibit all the artefacts in an educational and learning manner. Augmented Representations of Cultural Objects (ARCO) was an EU-funded research project (completed in September 2004) in order to analyse and provide innovative but simple to use technical solutions for virtual cultural object creation and visualisation. In short, ARCO provides museums with a set of tools that allow them to digitize, manage and present artefacts in virtual exhibitions. To evaluate the usability of the system properly ARCO collaborated with Victoria and Albert Museum and the Sussex Archaeological Society.

The work illustrated in the previous sections has been applied in ARCO to explore the potential of AR in a museum environment by mixing virtual information

in an environment comprised of real objects. The success of an AR exhibition is highly related to the level of realism achieved. In general, there are a few AR applications that do not require a high level of realism, but within the cultural heritage field realistic visualisation is an important issue (Liarokapis et al. 2004a). The scenarios illustrate how virtual museum visitors can visualise archaeological information such as virtual artefacts or even whole virtual museum galleries providing an educational narration for the preservation of cultural heritage. An example screenshot of four different virtual galleries from Victoria and Albert Museum are illustrated in Fig. 16.

In theory, this technique can be extended to as many markers as long as the camera can detect them within the field-of-view. The major drawback of this method is that the frame rate drops analogous to the number of markers used (as illustrated in Sect. 6.3) but the overall effectiveness in all galleries was between 25–30 FPS. Furthermore, the realism of the system highly depends on the 3D modelling procedure and for this reason the 3D models used in this scenario are the very high resolution models. The visualisation of an expedition to large groups of people can be considered as a collaborative activity. By looking at and interacting with the artefact visualisation visitors can communicate with each other by expressing their thoughts about any aspects that relate to the history of the artefact. This results in an exchange of opinions amongst the visitors in an implicit and explicit way. By zooming into the artefact more contained arguments can be made about the nature of the material used for its construction. On the other hand, in a perspective view more verbal communication is possible. By using the configuration setting of the collaboration in the AR interface, visitors can use HMDs and obtain a completely immersed view.

## 8 Preliminary evaluation

The knowledge gained from reviewing the literature and the experimental results, enabled an initial dissemination of the prototype AR interface. Even if this work is still on an experimental status, the results can be taken into consideration to improve the effectiveness of the presented system as well as to design future high-level AR interfaces. An expert-based evaluation approach was followed that argues that formal laboratory user studies can effectively evaluate visualisation when a small sample of expert users is used (Tory and Möller 2005). In terms of evaluating the system, some initial empirical research was conducted based on a two stage human-centred questionnaire. The first part

**Fig. 16** Virtual museum gallery visualisation



is generic but aims at evaluating the usability of the system in the learning process while the second part is more technical and refers to the effectiveness of the visualisation and interaction techniques of the interface. An educator would design the questionnaire in a different way taking primarily into consideration educational aspects whereas in this case, the purpose was to obtain a number of useful conclusions regarding the technicalities and practicalities of the system. This pilot study was disseminated to a five research staff from Sussex University that had experience in working with VR applications. Four were men and one was woman. Subjects were between the ages of 24 and 28 and the average time of the evaluation was 30 min.

### 8.1 General questions

The feedback received following the completion of the evaluation process varied but in general lines was encouraging. As far as the first part of the questionnaire is concerned, all the users thought that the system has the potential to be used as a learning tool in the future although it currently lacks from interoperability

**Table 3** General questions about the system

General questions	Mean (max = 5)	SD (yEr±)
Rate quality of performance?	4	0.7071
Rate the overall usage?	3.6	0.5477
Can aid the education process?	3.2	0.4472
Create new AR applications?	4.6	0.5477

issues. Specifically, they argued that the application scenarios were really interesting and exciting but for teaching purposes more comprehensive learning scenarios have to be implemented. The findings from this study are summarised in Table 3.

Results illustrate that 92% of the users believe that the system has the potential to be used as a basic platform to create AR scenarios and applications whereas 80% rate the quality of the system good. On the other hand, 72% of the users liked the overall usage of the system and 64% feel that educational applications could benefit from this technology. Moreover, two users mentioned that the system would be much more useful if a multimedia database system with a content management system is used to increase interoperability issues. Another one stated that a print function would help to capture and store into images the different views of the AR environment.

### 8.2 Technical questions

Regarding the second part, all users agreed that the system is very easy to use and that the visualisation process is more than satisfactory. Surprisingly, most of the users preferred the HMD-based visualisation versus the monitor-based visualisation (Fig. 17).

Figure 17 shows the user-response in comparing the monitor-based AR (mean = 6.8, SD = 2.77489, SE = 1.24097) versus video see through HMD-based AR (mean = 8.2, SD = 2.48998, SE = 1.11355). Similar studies have shown the exact opposite result but in this study all users were computer literature and all had

previously used VR prototypes that make use of HMDs. Moreover, many difficulties were observed when participants tried to move around with the camera mounted on the HMD because they could not keep it in line with the sight of view. Also because the resolution of the HMD is limited to  $800 \times 600$  and the quality of the overlaid graphics into the optics system is not very good, two participants felt nausea and motion sickness after a 10 min usage. However, even if these problems seem to restrict the use of HMDs, participants appreciated the level of immersion provided and thus preferred it. As far as the interaction techniques are concerned, the natural interaction techniques based on the marker cards were found to be very effective and intuitive to use compared to the other interaction techniques. Figure 18 illustrates a comparison based on the user-response between the most important interaction techniques implemented.

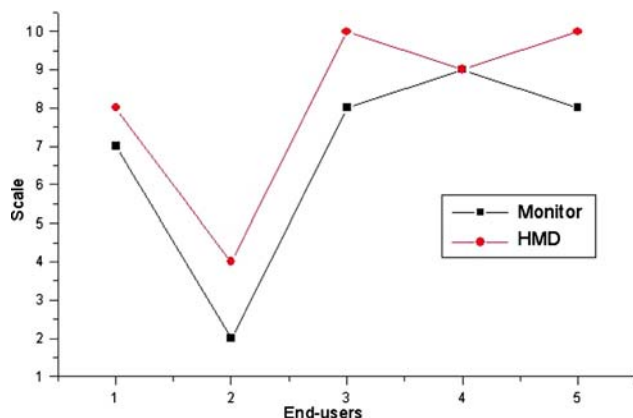
The I/O interaction techniques got the second highest score (mean = 6.2, SD = 2.16795, SE = 0.96954) since they are the standard way for interacting with computers and the end-users feel more familiar with. Surprisingly, the SpaceMouse interactions (mean = 5.4, SD = 2.50998, SE = 1.1225) received the most variable responses. Some participants argued that it is extremely useful to manipulate the virtual information using only one hand but others recorded that a lot of time is required to fully familiarise with the device and even then it is not as easy to use other means such as the I/O devices and the marker cards.

Moreover, the GUI interactions (mean = 4.4, SD = 2.19089, SE = 0.9798) got the worst score from all other types of interaction. One of the end-users argued that it is difficult to understand how to alter the orientation of the virtual objects since it was specified as yaw, pitch and roll. Other users stated that it takes the most time to perform a single rotation compared to

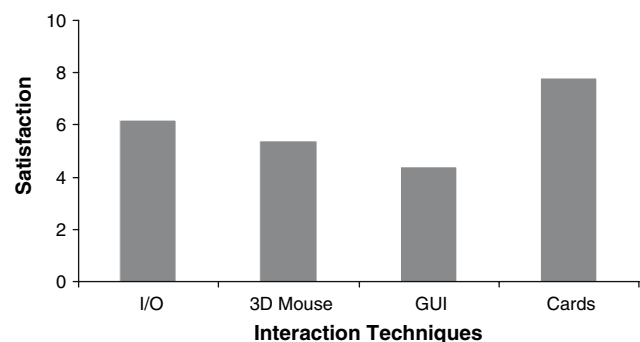
the rest of the methods. For example, the keyboard keys replicate the functionality and as soon as the user becomes familiar with the “shortcuts” it is much faster. On the contrary, the marker cards interaction (mean = 7.8, SD = 2.58844, SE = 1.15758) received the most positive feedback of all other types of interaction and although it was pretty much expected, an initial comparison between different techniques has been made. All participants agreed that it very easy and intuitive to manipulate the virtual information in 3D space using any type of physical interface but they also proposed to use in the future a physical interface that consists of a handle. Overall, the preliminary evaluation was a profitable experience to complete the first cycle of this research but more user-studies need to be performed in the future.

## 9 Conclusions and future work

In this paper the design and implementation of effective AR interfaces for indoor-environments was presented and analysed. The proposed framework can be used as a generic tool to create high-level AR applications. The final visualisation can be performed either on a variety of display technologies ranging from monitor based to video see-through display technologies. A series of visualisation and interaction techniques were investigated in order to create the illusion that the virtual information coexists with the real world. In addition, two innovative AR case studies have been implemented: one for higher education purposes (university environments) and the second for archaeological and cultural heritage purposes (museum environments). Finally, an initial evaluation was performed to obtain useful critique concerning the overall technicalities and practicalities of the system.



**Fig. 17** Monitor versus HMD user response



**Fig. 18** Interaction techniques user response



The main advantages of the AR architecture are the low cost and the multimedia augmentation in real-time. The structure of the architectures is based on the philosophy that the most appropriate tool/device must be used for the task ones seeking to achieve. This, however, does not imply that the best tool/device is the most expensive one. The two different experimental setups successfully tested for this research clearly demonstrate this. One cost effective setup has been constructed comprising of off-the-self hardware and a second one based on state of the art expensive hardware components (i.e. Spacemouse, Touch Screen).

Although the system is designed for indoor environments it can be easily extended to operate in outdoor environments. The current status of the research is focused in various mobile devices such as personal digital assistants (PDAs) and third-generation (3G) phones as well as positioning technologies (such as GPS). This will create a robust mobile AR environment that will be integrated with the rest of the interface framework to provide prototype applications for outdoor environments.

**Acknowledgments** Part of this research work was funded by the EU IST Framework V programme, Key Action III- Multimedia Content and Tools, Augmented Representation of Cultural Objects (ARCO) project IST-2000-28366.

## References

- Angel E (2003) Interactive computer graphics: a top-down approach using OpenGL, 3rd edn. Addison-Wesley, Reading, pp 17–18, 69, 107, 322–349, 472
- Azuma R (1997) A survey of augmented reality. *Teleoper Virtual Environ* 6(4):355–385
- Azuma R, Bailiot Y et al (2001) Recent advances in augmented reality. *IEEE Comput Graph* November/December 21(6):34–47
- Begault DR (1994) 3D Sound for virtual reality and multimedia, Academic, New York, 1, 17–18
- Billinghurst M, Kato H, Poupyrev I (2001) The magicbook: a traditional AR interface. *Comput Graph* 25:745–753
- Butz A, Höllner T et al (1999) Enveloping users and computers in a collaborative 3D augmented reality. In: Proceedings of the 2nd IEEE and ACM international workshop on augmented reality '99. San Francisco, October 20–21
- Camera calibration toolbox for Matlab, available at: [[http://www.vision.caltech.edu/bouguetj/calib\\_doc/](http://www.vision.caltech.edu/bouguetj/calib_doc/)], Accessed at 14/01/2003
- Dobler D, Haller M, Stampfl P (2002) ASR—augmented sound reality, ACM SIGGRAPH 2002 conference abstracts and applications, San Antonio, p 148
- Feiner S, MacIntyre B et al (1993) Windows on the world: 2D Windows for 3D augmented reality. In: Proceedings of the ACM symposium on user interface software and technology, Atlanta, November 3–5, Association for Computing Machinery, pp 145–155
- Fernandes B, Miranda JC (2003) Learning how computer works with augmented reality. In: Proceedings of the 2nd international conference on multimedia and information and communication technologies in education, Badajoz, December 3–6
- Fuhrmann A, Schmalstieg D (1999) Concept and implementation of a collaborative workspace for augmented reality, *GRAPHICS '99*, 18(3)
- Gatermann H (2000) From VRML to augmented reality via panorama-integration and EAI-Java, in constructing the digital space. In: Proceeding of the SiGraDi, September, 254–256
- Grasset R, Gascuel J-D (2002) MARE: multiuser augmented reality environment on table setup. ACM SIGGRAPH conference abstracts and applications
- Hall T, Ciolfi L et al (2001) The visitor as virtual archaeologist: using mixed reality technology to enhance education and social interaction in the museum. In: Spencer S (ed) Proceedings of the virtual reality, archaeology, and cultural heritage (VAST 2001), New York, ACM SIGGRAPH, Glyfada, Nr Athens, November, pp 91–96
- Haller M, Hartmann W et al (2002) Combining ARToolKit with scene graph libraries. In: Proceedings of the first IEEE international augmented reality toolkit workshop, Darmstadt, Germany, 29 September
- Haniff D, Baber C, Edmondson W (2000) Categorizing augmented reality systems. *J Three Dimens Images* 14(4):105–109
- Kato H, Billinghurst M, et al (2000a) Virtual object manipulation on a table-top AR environment. In: Proceedings of the international symposium on augmented reality 2000, Munich, 5–6 Oct, pp111–119
- Kato H, Billinghurst M, Poupyrev I (2000b) ARToolkit user manual, version 2.33, Human Interface Lab, University of Washington
- Klinker G, Ahlers KH et al (1997) Confluence of computer vision and interactive graphics for augmented reality, *PRESENCE: teleoperations and virtual environments*. special issue on augmented reality, August 6(4):433–451
- Liarokapis F, White M, Lister PF (2004a) Augmented reality interface toolkit. In: Proceedings of the international symposium on augmented and virtual reality, London, pp 761–767
- Liarokapis F, Sylaiou S, et al (2004b) An interactive visualisation interface for virtual museum. In: Proceedings of the 5th international symposium on virtual reality, Archaeology-Cultural Heritage, pp 47–56
- Liarokapis F (2005) Augmented reality interfaces—architectures for visualising and interacting with virtual information. PhD thesis. University of Sussex, Falmer
- Liarokapis, Petridis P, Lister PF, White M (2002) Multimedia augmented reality interface for E-learning (MARIE). *World TransEng Technol Educ* 1(2):173–176
- MacIntyre B, Gandy M, Dow S, Bolter JD (2005) DART: a toolkit for rapid design exploration of augmented reality experiences. *ACM Trans Graph (TOG)*, 24(3):932
- Mahoney D (1999b) Better than real, computer graphics world, pp 32–40
- Malbezin P, Piekarski W and Thomas B (2002) Measuring ARToolKit accuracy in long distance tracking experiments. In: Proceedings of the 1st international augmented reality toolkit workshop, Germany, Darmstadt, September 29
- Milgram P, Colquhoun H (1999) A Taxonomy of real and virtual world display integration, mixed reality merging real and virtual worlds. Ohta Y, Tamura H (eds) Ohmsha Ltd, Chapter 1, pp 5–30

- Milgram P, Kishino F (1994) A taxonomy of mixed reality visual displays, *IEICE Trans Inf Syst E77-D(12)*:1321–1329
- Moller T (1999) *Real-time rendering*. AK Peters Ltd, Natick, 23–38, 171
- Poupyrev I, Tan D et al (2002) Developing a generic augmented reality interface. *Computer* 35(3):44–50
- Reitmayr G, Schmalstieg D (2001) A wearable 3D augmented reality workspace. In: *Proceedings of the 5th international symposium on wearable computers*, October 8–9
- Rekimoto J, Nagao K (1995) The world through the computer: computer augmented interaction with real world environments. In: Myers BA (ed) *Proceedings of UIST '95*. ACM, Pennsylvania, pp 29–36
- Shi J, Tomasi C (1994) Good features to track, *IEEE conference on computer vision and pattern recognition*, Seattle, June, pp 593–600
- Sinclair P, Martinez K (2001) Adaptive hypermedia in augmented reality. In: *Proceedings of the third workshop on adaptive hypertext and hypermedia at the twelfth ACM conference on hypertext and hypermedia*, Denmark, August 2001, pp217–219
- Slay H, Phillips M et al (2001) *Interaction modes for augmented reality visualization*, Australian symposium on information visualization, Sydney, December
- Smith GC (1994) The art of interaction. In: MacDonald L, Vince J (eds) *Interacting with virtual environments*. Wiley, New York, pp 79–94
- Tory M, Möller T (2005) Evaluating visualizations: do expert reviews work? *IEEE Comput Graph Appl* 25(5):8–11
- Vallino J (1998) *Interactive augmented reality*. PhD thesis, Department of Computer Science, University of Rochester, pp 1–25
- Weng J, Cohen P, Herniou M (1992) Camera calibration with distortion models and accuracy evaluation, *IEEE transactions on pattern analysis and machine intelligence*, 14(10)
- Woo M, Neider J, Davis T (1999) *OpenGL programming guide: the official guide to learning OpenGL, Version 1.2*, Addison–Wesley, Reading
- Yewdall D (1999) *Practical art of motion picture sound*. Focal Press, Boston

## **8.8 Paper #8**

Mountain, D., Liarokapis, F. Mixed reality (MR) interfaces for mobile information systems, *Aslib Proceedings, Special issue: UK library & information schools*, Emerald Press, 59(4/5): 422-436, 2007.

Contribution (50%): Collaboration on the design of the architecture and implementation of the VR interface. Write-up of half of the paper.



# Mixed reality (MR) interfaces for mobile information systems

David Mountain

*giCentre, Department of Information Science, City University London,  
London, UK, and*

Fotis Liarokapis

*Coventry University, Coventry, UK and Department of Information Science,  
City University London, London, UK*

Received 15 December 2006  
Accepted 12 June 2007

## Abstract

**Purpose** – The motivation for this research is the emergence of mobile information systems where information is disseminated to mobile individuals via handheld devices. A key distinction between mobile and desktop computing is the significance of the relationship between the spatial location of an individual and the spatial location associated with information accessed by that individual. Given a set of spatially referenced documents retrieved from a mobile information system, this set can be presented using alternative interfaces of which two presently dominate: textual lists and graphical two-dimensional maps. The purpose of this paper is to explore how mixed reality interfaces can be used for the presentation of information on mobile devices.

**Design/methodology/approach** – A review of relevant literature is followed by a proposed classification of four alternative interfaces. Each interface is the result of a rapid prototyping approach to software development. Some brief evaluation is described, based upon thinking aloud and cognitive walk-through techniques with expert users.

**Findings** – The most suitable interface for mobile information systems is likely to be user- and task-dependent; however, mixed reality interfaces offer promise in allowing mobile users to make associations between spatially referenced information and the physical world.

**Research limitations/implications** – Evaluation of these interfaces is limited to a small number of expert evaluators, and does not include a full-scale evaluation with a large number of end users.

**Originality/value** – The application of mixed reality interfaces to the task of displaying spatially referenced information for mobile individuals.

**Keywords** Reality, Mobile communication systems, Information systems, Geography

**Paper type** Research paper

## 1. Introduction

Two of the most significant technological trends of the past 15 years have been the increased portability of computer hardware – such as laptop computers and personal digital assistants (PDAs) – and the increasing availability of wireless networks such as mobile telecommunications, and more recently wireless access points (Brimicombe and Li, 2006). The convergence of these technological drivers presents opportunities

The work presented in this paper is conducted within the LOCUS project, funded by EPSRC through the Pinpoint Faraday Partnership. The authors would also like to thank their partner on the project, GeoInformation Group, Cambridge, for contributing the building geometry and heights data used by the project. The authors are also grateful to Hulya Guzel for her assistance in the expert user evaluation.



---

within the emerging field of mobile computing. Increasingly there is ubiquitous access to information stored via a variety of media (for example, text, audio, image and video) via mobile devices with wireless network connections. Advances in software development tools for mobile devices have resulted in the implementation of user-friendly interfaces that aim to appeal to a wide audience of end users. A key challenge for researchers of mobile information systems is to decide the type of interface to adopt when presenting this information on mobile devices. Additionally, developers should assess whether the most suitable interface is dependent upon the audience, the task-in-hand and geographic context in which the mobile information system is likely to be used (Jiang and Yao, 2006).

The LOCUS project (Location Context tools for UMTS Services) being conducted within the Department of Information Science at City University is addressing some of the research challenges described above (LOCUS, 2007). The main aim of the project is to enhance the effectiveness of location-based services (LBS) in urban environments by investigating how mixed reality interfaces compare with the current map- and text-based approaches used by the majority of location-based services for the tasks of navigation and wayfinding (Mountain and Liarokapis, 2005). To satisfy this aim, LOCUS is tackling a number of issues including the three-dimensional representation of urban environments, the presentation of spatially referenced information – such as the information retrieved as the result of a user query, and navigational information to specific locations – and advanced visualisation and interaction techniques (Liarokapis *et al.*, 2006).

The LOCUS system is built on top of the WebPark mobile client-server architecture (WebPark, 2006) which provides the basic functionality associated with LBS including the retrieval of information based upon spatial and semantic criteria, and the presentation of this information as a list or on a map (see Figures 1a and b). In common with the majority of LBS, the basic architecture provides no mechanism for the display of information in a three-dimensional environment, such as a mixed reality interface.

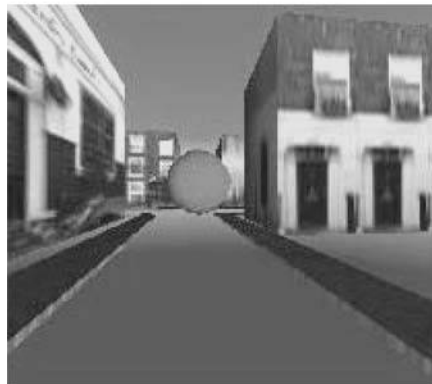
Mixed reality environments occupy a spectrum between entirely real environments at one extreme and entirely virtual environments on the other. This mixing of the real and the virtual domain offers great potential in terms of displaying information retrieved as a result of a location-based search, since this requires the presentation of digital information relative to your location in the physical world. This presentation may on the one hand be entirely synthetic, for example, placing virtual objects representing individual results within a virtual scene as a backdrop. Alternatively, an augmented reality interface can superimpose this information over the real world scene in the appropriate spatial location from the mobile user's perspective. Both interfaces can present the location of information within the scene as well as navigation tools that describe the routes to the spatial locations associated with retrieved information. The LOCUS project is extending the functionality of the WebPark architecture to allow the presentation of spatially referenced information via these mixed reality interfaces on mobile devices (see Figures 1c and d).

The rest of the paper is structured as follows. First, a review of relevant background literature in mobile computing and mixed reality is presented. Next, candidate interfaces for mobile information provision into mobile devices are suggested: these include the list, the map, and virtual and augmented reality interfaces. The paper closes with a discussion and conclusions.



(a) List interface (WebPark platform)

(b) Map interface (WebPark platform)



(c) Mobile VR interface (LOCUS prototype application)



(d) Mobile AR interface (LOCUS prototype application)

**Figure 1.**  
Interfaces for presenting information retrieved from a mobile information system

## 2. Background

### 2.1 Mobile computing

Just as the evolution of the internet has had a profound impact upon application development, forcing a change from a stand-alone desktop architecture to a more flexible, client-server architecture (Peng and Tsou, 2003), researchers in mobile computing are currently having a similar impact, forcing the development of web resources and applications that can be run on a wider range of devices than traditional desktop machines. According to Peng and Tsou (2003), mobile computing environments have three defining characteristics:

- (1) mobile clients that have limited processing and display capacity (e.g. PDAs and smart phones);

- (2) non-stationary users who may use their devices whilst on the move; and
- (3) wireless connections that are often more volatile, and have more constrained bandwidth, compared to the “fixed” internet.

These three characteristics suggest that mobile devices have both specific constraints and unique opportunities when compared to their desktop counterparts. First, screen real estate is limited; typically screens are small (usually less than 60 mm by 80 mm) with low resolution (typically 240 pixels width), and a relatively large proportion of this space may be taken up with *marginalia* such as scroll bars and menus; hence every pixel should be used wisely. Next, the outdoor environment is a more unpredictable and dynamic environment than the typically familiar indoor home and office environments in which desktop machines are used; hence, user attention is more likely to be distracted in the mobile context. Mobile computer usage tends to be characterised by multiple short sessions per day, as compared with desktop usage, which tends to be for relatively few, longer durations (Ostrem, 2002). Given these constraints, there is a clear need for information to be communicated concisely and effectively for mobile users.

Despite constraints, the mobile computing environment offers a unique opportunity for the presentation of information, in particular taking advantage of location sensors to organise information relative to the device user’s position, or their spatial behaviour (Mountain and MacFarlane, 2007).

While spatial proximity is perhaps the most intuitive and easily calculated measure of geographic relevance, it may not be the most appropriate in all situations and a variety of other measures of geographic relevance (Mountain and MacFarlane, 2007; Raper, 2007) have been suggested. Individuals may be more interested in the relative accessibility of results, which can be quantified by travel time and can take account for natural and manmade boundaries (Golledge and Stimson, 1997) or the transportation network, to discount results that are relatively inaccessible despite being physically close (Mountain, 2005). Geographic relevance can also be quantified as the results are most likely to be visited in the future (Brimicombe and Li, 2006), or those that are most visible from the current location (Kray and Kortuem, 2004). However geographic relevance is quantified, there are opportunities to use this property to retrieve documents from document collections. Given a set of spatially referenced results that are deemed to be geographically relevant according to some criterion, there are a variety of different approaches to presenting this information.

Various mobile information systems have been developed. Kirste (1995) developed one of the first experimental mobile information systems based on wireless data communication. A few years later, Afonso *et al.* (1998) presented an adaptable framework for mobile computing information dissemination systems called UbiData. This model adopts a “push” model where relevant information is sent to the user, without them making a specific request, based upon their location. There are now a host of commercial and prototype mobile information systems that can present information dependent upon an individual’s semantic and geographic criteria (Yell Group, 2006; WebPark, 2006), the majority of which present results either as a list or over a backdrop map.

## 2.2 Mixed reality

The mixed reality spectrum was proposed by Milgram and Kishino (1994), who depicted representations on a continuum with the tangible, physical (“real”) world at

one extreme and entirely synthetic virtual reality (VR) at the other. Two classes were identified between these extremes. Augmented reality (AR) refers to virtual information placed within the context of the real world scene, for example, virtual chess pieces on a real chessboard. The second case – augmented virtuality – refers to physical information being placed in a virtual scene, for example, real chess pieces on a virtual board. The resulting reality-virtuality continuum is shown in Figure 2.

The first VR system was introduced in the 1950s (Rheingold, 1991), and since then VR interfaces have taken two approaches:

- (1) immersive head-mounted displays (HMDs); and
- (2) through the window approaches.

HMDs are very effective at blocking the signals from the real world and replacing this natural sensory information with digital information. Navigation within the scene can be controlled by mounting orientation sensors on top of the HMD, a form of gesture computing whereby the user physically turning their head results in a rotation of the viewpoint in the virtual scene. The ergonomic limitations of HMDs proved unpopular with users and this immersive interface has failed to be taken up on a wide scale (Ghadirian and Bishop, 2002). In contrast to HMDs, the through the window (Bodum, 2005) – or monitor-based VR/AR – approach exploits monitors on desktop machines to visualise the virtual scene, a far less immersive approach since the user is not physically cut-off from the physical world around them. This simplistic form of visualisation has the advantage that it is cost-effective (Azuma, 1997). Interaction is usually realised via standard input/output (I/O) devices such as the mouse or the keyboard but also more sophisticated devices (such as spacemouse, inertia cube, etc.) may be employed (Liarokapis, 2005).

Both HMDs and through the window approaches of VR aim to replace the physical world with the virtual. The distinction of AR is that it aims to seamlessly combine real and virtual information (Tamura and Katayama, 1999) by superimposing digital information directly into a user’s sensory perception (Feiner, 2002) (see Figure 3). Whilst VR and AR can process and display similar information (for example three-dimensional buildings) the combination of the “real” and the “virtual” in the AR case is inherently more complex than the closed virtual worlds of VR systems.

This combination of real and virtual requires accurate tracking of the location of the user (in three spatial dimensions:  $x$ ,  $y$  and  $z$ ) and the orientation of their view (around three axes of orientation: yaw, pitch and roll), in order to be able to superimpose digital information at the appropriate location with respect to the real world scene, a procedure known as registration. In the past few years, research has achieved great advances in tracking, display and interaction technologies, which can improve the effectiveness of AR systems (Liarokapis, 2005). The required accuracy of the AR

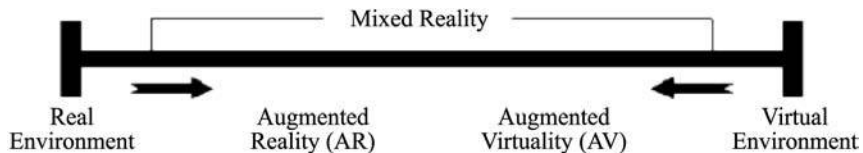
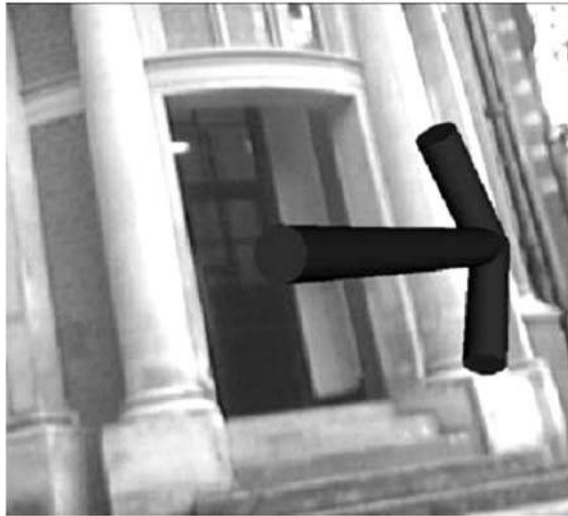


Figure 2.  
The reality-virtuality continuum

Source: After Milgram *et al.* (1994)





**Figure 3.**  
Augmented reality  
representation: a computer  
vision sensor recognises  
the doorway outline, and  
augments the video  
stream with virtual  
information (the direction  
arrow). Developed as part  
of the LOCUS project

---

tracking depends to a degree upon the scenario of use. In order to correctly superimpose an alternative building façade (for example, a historic or planned building façade) over an existing building, highly accurate tracking is required in terms of position and orientation, else the illusion will fail since the real and virtual façades will not align, or may drift apart as the user moves or turns their head (Hallaway *et al.*, 2004). However, if simply augmenting the real world scene with annotations in the forms of text or symbols, for example, an arrow indicating the direction to turn at an upcoming junction, this tracking may not be required to be so accurate.

The two most common tracking techniques used in AR applications include computer vision and external sensor systems. The visual approach uses fiducial reference points, where a specific number of locations act as links between the real and virtual scenes (Hallaway *et al.*, 2004). These locations are usually marked with distinctive high contrast markers to assist identification, but alternatively can be distinctive landmarks within the real-world scene. Computer vision algorithms first need to identify at least three reference points in real time from a video camera input, then calculate the distance and orientation of the camera with respect to those reference points. Tracking using a computer vision-based system therefore establishes a relative spatial relationship between a finite number of locations in the real-world scene and the observer, via a video camera carried or worn by that observer (Hallaway *et al.*, 2004), which can allow very accurate registration between the real and virtual scenes in a well-lit indoor environment. This computer vision approach nevertheless has significant constraints. First, the system must be trained to identify these fiducial reference points, and may further require the real-world scene to have markers placed within it. It requires both good lighting conditions (although infrared cameras can be also used for night vision) and significant computing resources to perform real-time tracking, and therefore has usually been conducted in an indoor, desktop environment (Liarokapis and Brujic-Okretic, 2006).

An alternative to the vision-based approach is to use external sensors to determine the position of the user and the orientation of their view. Positioning sensors such as

the Global Positioning System (GPS) can determine position in three dimensions and digital compasses, gyroscopes and accelerometers can be employed to determine the orientation of the user's view. These sensor-based approaches have the advantage that they are not constrained to specific locations, unlike computer vision algorithms, which must be trained to recognise specific reference points within a scene. Also, the user's location is known with respect to an external spatial referencing system, rather than establishing relative relationships between the user and specific reference points. A major disadvantage is the accuracy of the positioning systems, which can produce errors measured in tens of metres and can produce poor results when attempting to augment the real-world scene with virtual information. While advances in GPS systems such as differential GPS and real-time kinematic GPS can bring down the accuracy to one metre and a few centimetres respectively, GPS receivers still struggle to attain a positional fix where there is no clear view of the sky, for example in doors.

Digital compasses also have limitations; the main flaw is that they are prone to environmental factors such as magnetic fields.

Having identified a spatial relationship between the real-world scene and the user location, virtual information needs to somehow be superimposed upon the real world scene. Traditionally there have been two approaches to achieving this:

- (1) video see-through displays; and
- (2) optical see-through displays.

Video see-through displays are comprised of a graphics system, a video camera, a monitor and a video combiner (Azuma, 1997). They operate by combining a HMD with a video camera. The video camera records the real environment and then sends the recorded video to the graphics system for processing. There the outputted video and the generated graphics images, by the graphics system, are blended together. Finally, the user perceives the augmented view in the closed-view display system. Using the alternative approach, optical see-through displays are usually comprised of a graphics system, a monitor and an optical combiner (Azuma, 1997). They work by simply placing the optical combiners in front of the user's view. The main characteristic of the optical combiners is that they are partially transmissive and reflective. That is because the combiners operate like half-silvered mirrors, permitting only a portion of the light to penetrate. As a result, the intensity of the light that the user finally sees is reduced.

A novel approach to augmenting the real-world scene with virtual information, emerging from within the field of mobile computing, is to use the screen of a handheld device to act as a virtual window on the physical world. Knowing the position and orientation of the device, the information displayed on screen can respond to movements and gestures of a mobile individual, for example, presenting the name of a building as text on the screen when a user points their mobile device at it, or updating navigational instructions via symbols or text as a user traverses a route.

MARS is one of the first outdoor AR systems and a characteristic example of a wireless mobile interface system for indoor and outdoor applications. MARS was developed to aid navigation and to deliver location-based information to tourists in a city (Höllerer *et al.*, 1999). The user stands in an outdoor environment wearing a prototype system consisting of a computer, a GPS system, a see-through head-worn display and a stylus-operated computer. Interaction is via a stylus and display is via a tracked see-through head-worn display. MARS like most current mobile AR systems

has significant ergonomic restrictions which stretch the definitive of mobile and wearable computing beyond what is acceptable for most users (the system is driven by a computer contained in a backpack).

Tinmith-Hand AR/VR is a unified interface technology designed to support outdoor mobile AR applications and indoor VR applications (Piekarski and Thomas, 2002). This system employs various techniques including 3D interaction techniques, modelling techniques, tracked input gloves and a menu control system, in order to build VR/AR applications that can be applied to construct complex models of objects in both indoor and outdoor environments. A location-based application that was designed for a mobile AR system is ARLib: it aims to assist the user in typical tasks that are performed within a library environment (Umlauf *et al.*, 2002). The system follows a wide area tracking approach (Hedley *et al.*, 2002) based on fiducial-based registration. Many distinct markers are attached to bookshelves and walls so that the book's positions are superimposed on the shelves as the user navigates inside the library. To provide extra support to the user, a simple interface and a search engine are integrated to provide maximum usability and speed during book searches.

### 3. Interfaces for mobile information systems

There are many candidate interfaces for the presentation of the results of an information retrieval query on mobile devices (Mannings and Pearson, 2003; Schofield and Kubin, 2002; Mountain and Liarakapis, 2005). This section describes how the interfaces described previously can be applied to the task of presenting information retrieved as the result of a mobile query.

As described in the introduction, the LOCUS project has developed alternative, mixed reality interfaces for existing mobile information system technology based upon the WebPark platform. The WebPark platform can assist users in formulating spatially referenced, mobile queries. The retrieved set of spatially referenced results can then be displayed using various alternative interfaces: a list, a map, virtual reality or augmented reality. Each interface is described in more detail in the rest of this section.

#### 3.1 List interface

The most familiar interface for the presentation of the results of an information retrieval query is a list; this is the approach taken by the majority of internet search engines where the most relevant result is placed at the top of the list, with relevance decreasing further down the list (see Figure 1a). In the domain of location-aware computing, results that are deemed to be particularly geographically relevant (Mountain and MacFarlane, 2007; Raper, 2007) will be presented higher up the list (Google, 2006; WebPark, 2006). While familiar, this approach of simply ordering the results does not convey their location relative to your current position.

#### 3.2 Map interface

The current paradigm in the field of LBS is to present information relevant to an individual's query or task over a backdrop map (see Figure 1b). This information may include the individual's current position (and additionally some representation of the spatial accuracy), the locations of features of interest that were retrieved as the result of a user query (e.g. the results from a "find my nearest" search), or navigation information such as a route to be followed. This graphical approach has the advantage

of displaying the direction and distance of results relative to the user's location (a vector value), as opposed to just an ordering results based on distance. The viewpoint is generally allocentric (Klatzky, 1998), adopting a bird's eye view looking straight down on a flat, two-dimensional scene (see Figure 1c). The backdrop contextual map used is usually an abstract representation and may choose to display terrain, points or regions of interest, transportation links, or other information, alternatively a degree of realism can be included by using aerial photography (WebPark, 2006; Google, 2006).

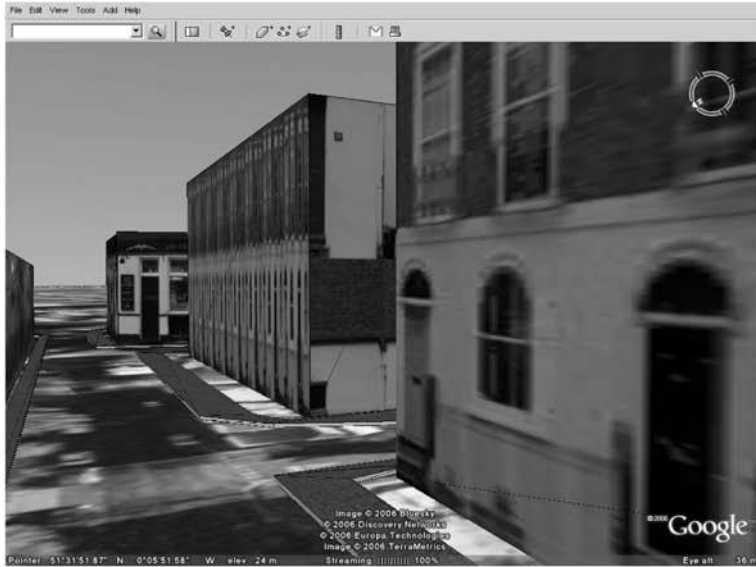
### 3.3 VR interface

An alternative to the allocentric viewpoint of a two-dimensional, abstract scene is to choose an egocentric viewpoint within a three-dimensional scene (see Figure 4). Such a perspective is familiar from VR discussed in section 2.2. While the concept of VR has existed for many decades, only during the past few years has it been used on handheld mobile devices. Traditionally, VR applications have been deployed on desktop devices and have attempted to create realistic looking models of environments to promote a feeling of immersion within a virtual scene. This has resulted in less opportunity for individuals to compare the virtual scene with its real-world counterpart. This separation of the real and the virtual is due in part to the static nature of desktop devices, and in addition that the appeal of many virtual scenes is that they allow the viewing of locations that cannot be visited easily, for example, virtual fly-throughs on other planets (NASA Jet Propulsion Laboratory, 2006) and imagined landscapes (Elf World, 2006).

In a location-aware, mobile computing context, the position of the user's viewpoint within a VR scene can be controlled from an external location sensor such as GPS, and the orientation of the viewpoint can be controlled by sensing the direction of movement (from GPS heading), or an orientation sensor to gauge the direction an individual is facing (e.g. a digital compass). The VR scenes themselves can adopt different levels of detail and realism (Bodum, 2005). A particular building may be represented with an exact three-dimensional geometric representation, and graphics added as textures to the façades of the building to create as true a representation as possible – known as a verisimilar representation (Bodum, 2005). Alternatively, the building may be modelled with a generalised approximation of the geometry within specific tolerances. For texturing the building facades, generic images may be applied that are typical of that class of building. The building block can be left untextured, but more abstract information conveying using shading, icons, symbols or text (Bodum, 2005). The level of detail and realism required by different users for different tasks is an open question currently under investigation (Liarokapis and Brujic-Okretic, 2006).

Traditionally, for VR applications deployed in a static, desktop context, there has been greater emphasis placed upon scenes looking realistic than ensuring that the content of these scenes is spatially referenced. However, in a mobile context, accurate spatial referencing of VR scenes is required when setting the viewpoint within that scene (using position and orientation sensors) to ensure that the viewpoint in the virtual scene is registered accurately with the user's location in the real world scene. Realism is still important since this can help the user make associations between objects in the virtual scene and those in the real world.

For the applications developed as part of the LOCUS project, within this VR backdrop, additional, non-realistic visual information can be included to augment the scene. Such information can include nodes representing documents retrieved from a



(a) Egocentric perspective



(b) An oblique, allocentric perspective

**Note:** Scene rendered in Google Earth

**Figure 4.**  
A virtual representation of  
a London neighbourhood

---

spatially referenced document collection (see Figure 1c), or navigational information and instructions (i.e. 3D textual directions). This approach has the advantage of promoting a feeling of immersion, and creating a stronger association between the physical world and relevant geo-referenced information, but is potentially less effective than a map in providing a quick synopsis of larger volumes of information relative to your location. There are opportunities to adopt multiple viewpoints within the VR scene that fall between the extremes of the allocentric-egocentric spectrum, for example an oblique perspective several metres higher than the user's viewpoint (see Figure 4).

### *3.4 AR interface*

A fourth approach to the display of information in mobile computing is to use the device to merge the real world scene with relevant, spatially referenced information by using an AR interface – the virtual window approach described in section 2.2. Just as for the mobile VR case described above, knowing the location and orientation of the device is an essential requirement for outdoor AR, in order to superimpose information in the correct location.

As described in the literature review, a GPS receiver and digital compass can provide sufficient accuracy for displaying points of interest in the approximate location relative to the user's position. At present, however, these sensor solutions lack the accuracy required for more advanced AR functionality, such as aligning an alternative façade on the front of a building in the real world scene. In the LOCUS system, the handheld mobile device presents text, symbols and annotations in response to the location and orientation of the device. There is no need for a HMD, since the screen on the device can be aligned with the real world scene. On the screen of the device, information can either be overlaid on imagery captured from the device's internal camera, or the screen can display just the virtual information with the user viewing the real world scene directly.

The information displayed is dependent upon the task in hand. When viewing a set of results, as the user pans the device around them, the name and distance of each result is displayed in turn as it coincides with the direction that the user is pointing the device, allowing the user to interrogate the real world scene by gesturing. By adopting an egocentric perspective to combine real and virtual information in this way, users of the system can base their decisions of which location to visit on more quantifiable criteria – such as the distance to a particular result, and the relevance on semantic criteria – but also the more subjective criteria that could never be quantified by an information system. For example, following a mobile search for places to eat conducted at a crossroads, by gesturing with a mobile device, users can see the distance and direction of candidate restaurants, and make an assessment based upon the ambience of the streets upon which different restaurants are located.

Having selected a particular result from the list of candidates, the AR interface can then provide navigational information, in the form of distance and direction annotations (see Figure 1d), to guide the user to the location associated with those results. Although most examples from location-based service suggest “where's my nearest” shop or service, there is no reason that this information could not be the location of breaking news stories from a news website, or spatially referenced HTML pages providing historical information associated with a particular era or event.

---

#### 4. Discussion

An evaluation exercise was undertaken to assess appropriate levels of detail, realism and interaction for the mobile virtual reality interface. Whilst there has been extensive evaluation of these requirements in a static desktop context (Dollner, 2005), relatively little attention has been paid to the specific needs of mobile users. In order to gauge these specific requirements, an expert evaluation was conducted. Two common evaluation techniques were applied:

- (1) think aloud; and
- (2) cognitive walkthrough (Dix *et al.*, 2004).

Think aloud is a form of observation that involves participants talking through the actions they are performing, and what they believe to be happening, whilst interacting with a system. The cognitive walkthrough technique was also used where a prototype of a mobile VR application and scenario of use were presented to expert users: evaluating in this way allows fast assessment of early mock-up, and hence can influence subsequent development and the suitability of the final application. Both forms of evaluation are appropriate for small numbers of participants testing prototype software and it has been suggested that the majority of usability problems can be discovered from testing in this way (Dix *et al.*, 2004).

The expert user testing took place at City University with a total of four users with varied backgrounds: one human-computer interaction expert, one information visualization expert, one information retrieval expert and one geographic information scientist. Each user spent approximately one hour performing four tasks. The aims of the evaluation of the VR prototype included assessment of the expert user experience with particular focus on:

- the degree of realism required in the scene;
- the required spatial accuracy and level of detail of the building outlines; and
- a comparison of 3D virtual scenes with 2D paper maps.

A virtual reality scene was created of the University campus and surrounding area, and viewpoints placed to describe trajectories of movement through the scene. The expert-evaluation process covered two tasks, including mobile search and navigation. The first scenario was in relation to searching for, then locating, specific features. For example, a user searching on a mobile system for entrances to the City University campus from a nearby station. The second scenario was in relation to navigation from one point to another, for example, from the station to the University. Starting and target locations were marked in the 3D maps, and sequences of viewpoints were presented, to mimic movement through the scene.

There was a great deal of variation in terms of the level of photorealism required in the scene, and whether buildings should have image textures placed over the building faces, or whether the building outlines would be sufficient alone. Opinions varied between evaluators and according to the task in hand. Plain, untextured buildings are hard to distinguish from each other and, in contrast, buildings with realistic textures were considered easy to recognise in a micro-scale navigation context (for example, trying to find the entrance to a particular building). However, many evaluators thought that much of this realism would not be required or visible on a small screen device when an overview

of the area was required, for example, when considering one's present location in relation to information retrieved from a mobile search. Expert users also suggested various departures from the realism traditionally aspired to within the field of virtual reality. These included transparency, to allow users to see through buildings as an aid to navigation, since this will allow the identification of the location of a concealed destination point. Other suggestions included labelling of objects in the scene (for example, building and street names). The inclusion of symbology in the scene to represent points, and routes to those points, was considered to be beneficial to the task of navigation.

In terms of the level of detail and spatial accuracy, some users thought that it was not important to have very detailed models of building geometry. Building outlines that are roughly the right size and shape are sufficient, especially when considering an overview of an area, as often required in the mobile search task. For micro-navigation, a higher degree of accuracy may be required.

Virtual 3D scenes were found to have many advantages when compared to paper maps: the most positive feature was found to be the possibility to recognize the features in the surrounding environment, which provides a link between the real and virtual worlds. This removes the need to map-read, which is required when attempting to link your position in the real world with a 2D map, hence the VR interface offers an effective way to gauge your initial position and orientation. A more intangible response was the majority of the users enjoyed interacting with the VR interface more than a 2D map. However, the 3D interface also has significant drawbacks. Some users said that they are so used to using 2D maps that they do not really need a 3D map for navigating, however they thought this attitude may change with the next generation. The size, resolution and contrast of the device screen were also highlighted as potential problems for the VR interface.

## 5. Conclusions

This paper has presented some insights on how mixed reality interfaces can be used in conjunction with mobile information systems to enhance the user experience. We have explored how the LOCUS project has extended LBS through different interfaces to aid the tasks of urban navigation and wayfinding. In particular, we have described how virtual and augmented reality interfaces can be used in place of text- and map-based interfaces, which can provide an egocentric perspective to location-based information which is lacking from map- and text-based representations.

Expert user evaluation has proven to be a useful technique to aid development, and suggests that the most suitable interface is likely to vary according to the user and task in hand. Continued research, development and evaluation is required to provide increasingly intuitive interfaces for location-based services that can allow users to make associations between spatially referenced information retrieved from mobile information systems, and their location in the physical world.

## References

- Afonso, A.P., Regateiro, F.S. and Silva, M.J. (1998), "UbiData: an adaptable framework for information dissemination to mobile users", *Object Oriented Technology, ECOOP'98 Workshop on Mobility and Replication, Brussels, July 20-24*, p. 1543.
- Azuma, R. (1997), "A survey of augmented reality", *Teleoperators and Virtual Environments*, Vol. 6 No. 4, pp. 355-85.



- Bodum, L. (2005), "Modelling virtual environments for geovisualization: a focus on representation", in Dykes, J.A., Kraak, M.J. and MacEachren, A.M. (Eds), *Exploring Geovisualization*, Elsevier, London, pp. 389-402.
- Brimicombe, A. and Li, Y. (2006), "Mobile space-time envelopes for location-based services", *Transactions in GIS*, Vol. 10 No. 1, pp. 5-23.
- Dix, A., Finlay, J.E., Abowd, G.D. and Beale, R. (2004), *Human-Computer Interaction*, Prentice-Hall, Harlow.
- Dollner, J. (2005), "Geovisualization and real-time computer graphics", in Dykes, J.A., Kraak, M.J. and MacEachren, A.M. (Eds), *Exploring Geovisualization*, Elsevier, London, pp. 325-44.
- Elf World (2006), "Elven Forest, 3D", available at: [www.allelves.ru/forest/](http://www.allelves.ru/forest/) (accessed 12 December 2006).
- Feiner, S.K. (2002), "Augmented reality: a new way of seeing", *Scientific American*, Vol. 4 No. 24, pp. 48-55.
- Ghadirian, P. and Bishop, I.D. (2002), "Composition of augmented reality and GIS to visualise environmental changes", *Proceedings of the Joint AURISA and Institution of Surveyors Conference, Adelaide, 25-30 November*.
- Golledge, R.G. and Stimson, R.J. (1997), *Spatial Behaviour: A Geographic Perspective*, The Guildford Press, New York, NY.
- Google (2006), Google Local, available at: <http://local.google.co.uk/> (accessed 10 December 2006).
- Hallaway, D., Hollerer, T. and Feiner, S. (2004), "Bridging the gaps: hybrid tracking for adaptive mobile augmented reality", *Applied Artificial Intelligence*, Vol. 18 No. 6, pp. 477-500.
- Hedley, N.R., Billingham, M., Postner, L., May, R. and Kato, H. (2002), "Explorations in the use of augmented reality for geographic visualization", *Presence*, Vol. 11 No. 2, pp. 119-33.
- Höllerer, T., Feiner, S.K., Terauchi, T., Rashid, G. and Hallaway, D. (1999), "Exploring MARS: developing indoor and outdoor user interfaces to a mobile augmented reality system", *Computers and Graphics*, Vol. 23 No. 6, pp. 779-85.
- Jiang, B. and Yao, X. (2006), "Location-based services and GIS in perspective", *Computers Environment and Urban Systems*, Vol. 30 No. 6, pp. 712-25.
- Kirste, T. (1995), "An infrastructure for mobile information systems based on a fragmented object model", *Distributed Systems Engineering Journal*, Vol. 2 No. 3, pp. 161-70.
- Klatzky, R.L. (1998), "Alloentric and egocentric spatial representations: definitions, distinctions, and interconnections", in Freksa, C., Habel, C. and Wender, K.F. (Eds), *Spatial Cognition – An Interdisciplinary Approach to Representation and Processing of Spatial Knowledge*, Springer, Berlin, pp. 1-18.
- Kray, C. and Kortuem, G. (2004), "Interactive positioning based on object visibility", in Brewster, S. and Dunlop, M. (Eds), *Mobile Human-Computer Interaction*, Springer, Berlin, pp. 276-87.
- Liarokapis, F. (2005), "Augmented reality interfaces – architectures for visualising and interacting with virtual information", DPhil thesis, Department of Informatics, School of Science and Technology, University of Sussex, Brighton.
- Liarokapis, F. and Brujic-Okretic, V. (2006), "Location-based mixed reality for mobile information services", *Advanced Imaging*, Vol. 21 No. 4, pp. 22-5.
- Liarokapis, F., Mountain, D., Papakonstantinou, S., Brujic-Okretic, V. and Raper, J. (2006), "Mixed reality for exploring urban environments", *Proceedings of the 1st International Conference on Computer Graphics Theory and Applications, Setúbal, 25-28 February*, pp. 208-15.
- LOCUS (2007), Homepage, available at: [www.locus.org.uk](http://www.locus.org.uk) (accessed 22 January 2007).

- Mannings, R. and Pearson, I. (2003), "Virtual air: a novel way to consider and exploit location-based services with augmented reality", *Journal of the Communications Network*, Vol. 2 No. 1, pp. 29-33.
- Milgram, P. and Kishino, F. (1994), "A taxonomy of mixed reality visual displays", *IEICE Transactions on Information Systems E Series D*, Vol. 77 No. 12, pp. 1321-9.
- Milgram, P., Takemura, H., Utsumi, A. and Kishino, F. (1994), "Augmented reality: a class of displays on the reality-virtuality continuum", *Telemanipulator and Telepresence Technologies*, Vol. 2351, pp. 282-92.
- Mountain, D.M. (2005), "Exploring mobile trajectories: an investigation of individual spatial behaviour and geographic filters for information retrieval", PhD thesis, Department of Information Science, City University London, London.
- Mountain, D.M. and Liarokapis, F. (2005), "Interacting with virtual reality scenes on mobile devices", *Human Computer Interaction with Mobile Devices and Services*, University of Salzburg, Salzburg, 19-22 September.
- Mountain, D.M. and MacFarlane, A. (2007), "Geographic information retrieval in a mobile environment: evaluating the needs of mobile individuals", *Journal of Information Science*, forthcoming.
- NASA Jet Propulsion Laboratory (2006), "Lander and Rover on Mars", available at: <http://mars.sgi.com/worlds/pathfinder/pathfinder.html> (accessed 12 December 2006).
- Ostrem, J. (2002), "Palm OS user interface guidelines", available at: [www.palmos.com/dev/support/docs/ui/UIGuide\\_Front.html](http://www.palmos.com/dev/support/docs/ui/UIGuide_Front.html) (accessed 10 April 2006).
- Peng, Z.-R. and Tsou, M.-H. (2003), *Internet GIS: Distributed Geographic Information Services for the Internet and Wireless Networks*, Wiley, New York, NY.
- Piekarski, W. and Thomas, B.H. (2002), *Unifying Augmented Reality and Virtual Reality User Interfaces*, University of South Australia, Adelaide.
- Raper, J.F. (2007), "Geographic relevance", *Journal of Documentation*, forthcoming.
- Rheingold, H.R. (1991), *Virtual Reality*, Summit Books, New York, NY.
- Schofield, E. and Kubin, G. (2002), "On interfaces for mobile information retrieval", *Lecture Notes in Computer Science*, Vol. 2411, pp. 383-7.
- Tamura, H. and Katayama, A. (1999), "Steps toward seamless mixed reality", in Ohta, Y. and Tamara, H. (Eds), *Mixed Reality Merging Real and Virtual Worlds*, Ohmsha, Tokyo, pp. 59-84.
- Umlauf, E., Piringer, H., Reitmayr, G. and Schmalstieg, D. (2002), "ARLib: the augmented library", *Proceedings of the First IEEE International Augmented Reality ToolKit Workshop, Darmstadt*.
- WebPark (2006), "Geographically relevant information for mobile users in protected area", available at: [www.webparkservices.info](http://www.webparkservices.info) (accessed 12 December 2006).
- Yell Group (2006), The UK's Local Search Engine, available at: [www.yell.com](http://www.yell.com) (accessed 12 December 2006).

**Corresponding author**

David Mountain can be contacted at: [dmm@soi.city.ac.uk](mailto:dmm@soi.city.ac.uk)

**This article has been cited by:**

1. David M. Mountain From Location-Based Services to Location-Based Learning: Challenges and Opportunities for Higher Education 327-343. [[CrossRef](#)]
2. Jonathan Raper. 2009. Geographic information science. *Annual Review of Information Science and Technology* 43:1, 1-117. [[CrossRef](#)]
3. Jonathan Raper, Georg Gartner, Hassan Karimi, Chris Rizos. 2007. Applications of location-based services: a selected review. *Journal of Location Based Services* 1:2, 89-111. [[CrossRef](#)]
4. Yu-Horng Chen, Yih-Shyuan Chen A Study of Mobile Guide Applications in Wayfinding Context 230-246. [[CrossRef](#)]

## **8.9 Paper #9**

Liarokapis, F., Macan, L., Malone, G., Rebolledo-Mendez, G., de Freitas, S. Multimodal Augmented Reality Tangible Gaming, *Journal of Visual Computer*, Springer, 25(12): 1109-1120, 2009.

Contribution (30%): Contribution on the design of the architecture. Implementation of parts of the AR interface. Write-up of most of the paper.

# Multimodal augmented reality tangible gaming

Fotis Liarokapis · Louis Macan · Gary Malone ·  
Genaro Rebolledo-Mendez · Sara de Freitas

Published online: 27 August 2009  
© Springer-Verlag 2009

**Abstract** This paper presents tangible augmented reality gaming environment that can be used to enhance entertainment using a multimodal tracking interface. Players can interact using different combinations between a pinch glove, a Wiimote, a six-degrees-of-freedom tracker, through tangible ways as well as through I/O controls. Two tabletop augmented reality games have been designed and implemented including a racing game and a pile game. The goal of the augmented reality racing game is to start the car and move around the track without colliding with either the wall or the objects that exist in the gaming arena. Initial evaluation results showed that multimodal-based interaction games can be beneficial in gaming. Based on these results, an augmented reality pile game was implemented with goal of completing a circuit of pipes (from a starting point to an end point on a grid). Initial evaluation showed that tangible interaction is preferred to keyboard interaction and that tangible games are much more enjoyable.

**Keywords** Serious games · Pervasive computing · Augmented reality · Multimodal interfaces

## 1 Introduction

Computerized games which have learning or training purposes demonstrate a popular trend in training due to the wide availability and ease of use of virtual worlds. The use of serious games in virtual worlds not only opens up the possibility of defining learning game-based scenarios but also of enabling collaborative or mediated learning activities that could lead to better learning [1]. An added benefit of using serious games in combination with virtual worlds is that learners engage with these in a multimodal fashion (i.e. using different senses) helping learners to fully immerse in a learning situation [2] which might lead to learning gains [3]. The multimodal nature of virtual worlds [4] and the facilities they offer to share resources, spaces and ideas greatly support the development and employment of serious games and virtual worlds for learning and training. The use of games as learning devices is not new. The popularity of video games among younger people led to the idea of using them with educational purposes [5]. As a result, there has been a tendency to develop more complex serious games which are informed by both pedagogical and game-like, fun elements. One common example of these combinations is the use of agents [6]: The idea behind agents is to provide pedagogical support [7] while providing motivating environments in the form of agents [8]. However, the use of agents is not the only motivating element in serious games as the use metaphors [9] and narratives [10] have been used to support learning and training in game-like scenarios.

Tangible games can sometimes have an educational aspect. The whole idea of playability in tangible games is the

---

F. Liarokapis (✉) · L. Macan · G. Malone  
Interactive Worlds Applied Research Group, Coventry University,  
Coventry, UK  
e-mail: [F.Liarokapis@coventry.ac.uk](mailto:F.Liarokapis@coventry.ac.uk)

L. Macan  
e-mail: [macanl@coventry.ac.uk](mailto:macanl@coventry.ac.uk)

G. Malone  
e-mail: [maloneg@coventry.ac.uk](mailto:maloneg@coventry.ac.uk)

G. Rebolledo-Mendez · S. de Freitas  
Serious Games Institute, Coventry University, Coventry, UK

G. Rebolledo-Mendez  
e-mail: [GRebolledo-Mendez@coventry.ac.uk](mailto:GRebolledo-Mendez@coventry.ac.uk)

S. de Freitas  
e-mail: [s.defreitas@coventry.ac.uk](mailto:s.defreitas@coventry.ac.uk)

player's interaction with the physical reality. In addition, the accessibility space that is the key to the oscillation between embedded and tangible information [11]. On the contrary, augmented reality (AR) has existed for quite a few years and numerous prototypes have been proposed mainly from universities and research institutes. AR refers to the seamless integration of virtual information with the real environment in real-time performance. AR interfaces have the potential of enhancing ubiquitous environments by allowing necessary information to be visualized in a number of different ways, depending on the user needs. However, only a few gaming applications combined them together to offer a very enjoyable and easy to use interface.

This paper presents tangible augmented reality gaming environment that can be used to enhance entertainment using a multimodal tracking interface. The main objective of the research is to design and implement generic tangible augmented reality interfaces that are user-friendly in terms of interaction and can be used by a wide range of players including the elderly or people with disabilities. Players can interact using different combinations between a pinch glove, a Wiimote, a six-degrees-of-freedom tracker, through tangible ways as well as through I/O controls. Two tabletop augmented reality games have been designed and implemented including a racing game and a pile game. The goal of the augmented reality racing game is to start the car and move around the track without colliding with either the wall or the objects that exist in the gaming arena. Initial evaluation results showed that multimodal-based interaction games can be beneficial in gaming. Based on these results, an augmented reality pile game was implemented with goal of completing a circuit of pipes (from a starting point to an end point on a grid). Initial evaluation showed that tangible interaction is preferred to keyboard interaction and that tangible games are much more enjoyable.

## 2 Background

In the past, a number of AR games have been designed in different areas including education, learning, enhanced entertainment and training [12]. A good survey of tracking sensors used in ubiquitous AR environments [13], as well as a taxonomy of mobile and ubiquitous applications, was previously documented [14]. This section presents an overview of the most characteristic applications and prototypes that integrate tracking sensors into AR tabletop and gaming environments. One of the earliest examples of educational AR was the MagicBook [15]. This is a real book which shows how AR can be used in schools for educational purposes and is an interesting method of teaching. MagicBook was also used as a template for a number of serious applications in numerous AR games.

One of the earliest pervasive AR prototypes is NaviCam [16], which has the ability to recognize the user's situation by detecting color-code IDs in real-world environments by displaying situation-sensitive information by superimposing messages on its video see-through screen. Another early important work refers to the Remembrance Agent [17], a text-based AR wearable system which allows users to explore over a long period of time augmented representations and provide better ways of managing such information. EMMIE [18] is a hybrid user interface to a collaborative augmented environment which combines a variety of different technologies and techniques, including virtual elements such as 3D widgets, and physical objects such as tracked displays and input devices. Users share a 3D virtual space and manipulate virtual objects that can be moved among displays (including across dimensionalities) through drag-and-drop. A more recent prototype is DWARF [19] which includes user interface concepts, such as multimedia, multimodal, wearable, ubiquitous, tangible, or augmented reality-based interfaces. DWARF covers different approaches that are all needed to support complex human-computer interactions. Higher level functionality can be achieved allowing users to manage any complex, interrelated processes, using a number of physical objects in their surroundings. The framework can be used for single-user as well as multi-user applications. In another prototype, the combination of AR and ubiquitous computing can lead to more complex requirements for geometric models that are appearing [20]. For such models a number of new requirements appear concerning cost, ease of reuse, inter-operability between providers of data, and finally use in the individual application.

In terms of enhanced entertainment, outdoor AR gaming plays a significant role. A characteristic example is the Human Pacman project [21] that is built upon position and perspective sensing via GPS, inertia sensors and tangible human-computer interfacing with the use of Bluetooth and capacitive sensors. The game brings the computer gaming experience to a new level of emotional and sensory gratification by embedding the natural physical world ubiquitously and seamlessly with a fantasy virtual playground. AR Tennis [22] is the first example of a face-to-face collaborative AR application developed for mobile phones. Two players sit across the table from each other, while computer vision techniques are used to track the phone position relative to the tracking markers. When the player points the phone camera at the markers, they see a virtual tennis court overlaid on live video of the real world.

Another interesting project is STARS [23] which focused on the nature of state representation in augmented game designs and developed several games based on these principles. Moreover, Mixed Fantasy [24] presents a MR experience that applies basic research to the media industries of entertainment, training and informal education. As far as

training is concerned, the US Army paid more than \$5 million to design an educational game based on the Xbox platform to train troops in urban combat [25]. Another example is the MR OUT project [26] which uses extreme and complex layered representation of combat reality, using all the simulation domains such as live, virtual, and constructive by applying advanced video see-through mixed reality (MR) technologies. MR OUT is installed at the US Army’s Research Development and Engineering Command and focuses on a layered representation of combat reality.

### 3 Architecture

The architecture of our system has been based on an earlier prototype [29] but it provides similarities with AR interfaces such as [18, 19, 27, 28]. In the current system, interaction is performed using a pinch glove, a six-DOF tracker and Wiimote (2 DOF). The processing unit can be wearable (or mobile) and thus the Sony VAIO UMPC was selected (1.3 GHz, two 1.3 mega-pixel cameras, VGA port, Wi-Fi, USB ports, Bluetooth and keyboard/mouse). The rest of the hardware devices used and integrated to the UMPC included a pinch glove (5DT Data glove), a Wii Remote, a 6-DOF tracker (Polhemus Patriot) and an HMD (eMagin Z800). An overview of the system is shown in Fig. 1.

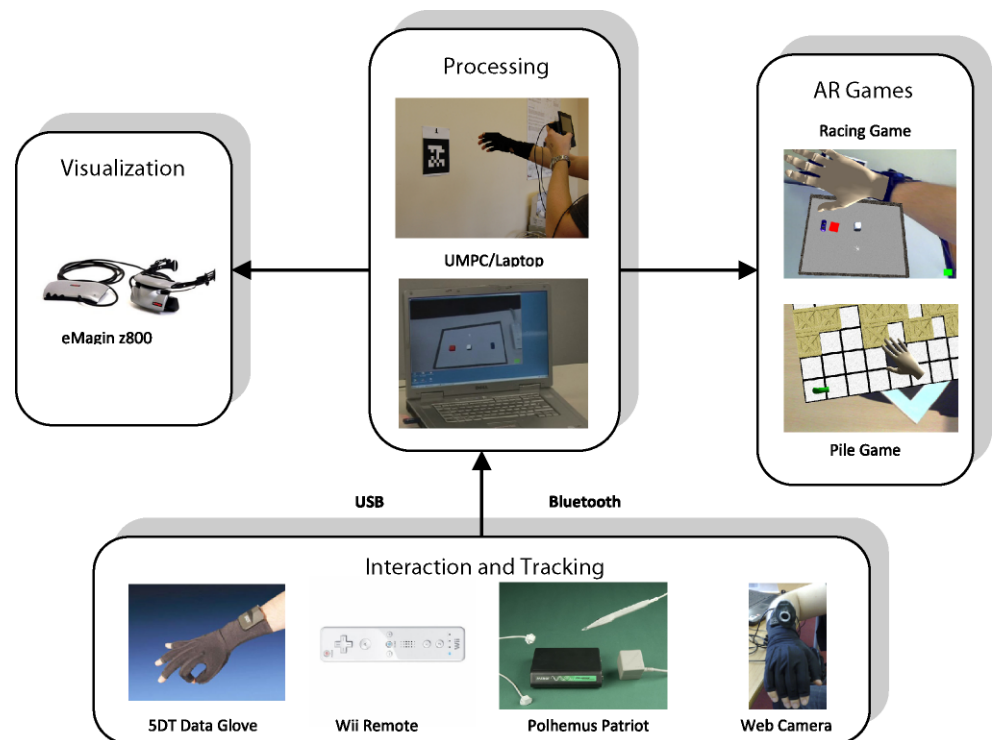
Visualization is enhanced through the use of a head-mounted display (HMD) which includes a three-DOF orientation tracker. Visual tracking is based on multiple markers which provide better robustness and range of operation,

based on ARToolKit [30] and ARTag [31] libraries. To retrieve multimodal tracking data in real time, a socket was created which was constantly waiting for input. The input comes in a structured form so data structures were set up in order to grab and store this information for the visualization. This socket server function was placed inside a thread of its own to stop it affecting the whole process while it waits for data to be retrieved. Once the data are received, attributes are assigned to a particular marker or to multiple markers. When that marker comes into sight of the camera the rendering part recognizes that the current marker has further data attached to it.

### 4 Tracking

An issue that arose early on with using the player’s hand to interact with virtual objects in an AR environment was occlusion. Once the hand moves to interact with the AR scene, it obscures a real-world visual marker. As a result, the on-screen objects disappear; the AR system relies on the markers to accurately register the virtual information on screen. This being the case, one of the objectives early on was to create a system that would use multiple markers to represent a single object or set of objects. This would mean that even if one or several markers were blocked by a user’s hand, for example, the objects would still be displayed based on where the visible markers were placed. The game uses a method of detecting which marker of the available markers

**Fig. 1** Multimodal architecture for tangible gaming



on a sheet of paper is currently the most visible, based on a confidence rating assigned to it by a class in ARToolKit. This marker becomes the origin from which all other objects are drawn on the game board. If this marker becomes obscured, the program automatically switches to the next highest confidence-rated marker. The confidence is based on a comparison between the marker pattern stored in memory and what is detected by the camera in the current frame.

The first problem encountered with a tangible interface incorporated into an augmented reality application was that of marker occlusion. In order for an object to be drawn on screen, the camera must have direct line of sight of a recognized marker. If that marker is even partially obscured, the program cannot recognize it as a square nor read the information on it, so the object will not be shown. This presented an important challenge for the project; if the game is to be controlled via a tangible interface, then a user must be able to physically interact with the graphics which will mean frequently putting their hand between the camera and the marker. Thus, a method of preventing marker occlusion that would be simple for a user to set up and would still allow 3D movement was developed. Moreover, inspired by the currently unavailable ARTag [31] a multiple marker system was implemented. Using several markers to represent the game playing area, one marker at any given time is selected as the basis to draw many objects. This marker is selected through the use of a confidence value as explained in [29].

## 5 Multimodal interaction

The main objective of this work was to allow for seamless interaction between the users and the superimposed environmental information. To achieve this, a number of custom interaction devices have been researched, such as the PS3 controller, 3D mouse, etc. However, since usability and mobility were crucial, only a few interaction devices were finally integrated to the final architecture. In particular, six different types of interaction were implemented including: hand position and orientation, pinch glove interaction, head orientation, Wii interaction and UMPC I/O manipulation. A brief overview of these techniques is presented below.

### 5.1 Polhemus

Once integrated into the architecture, there were several issues that prevented it from being as effective as previously hoped. The sensors often suffer from inversion problems, meaning the user's hand is displayed in the wrong position and disrupting the interface for the game. A fix employed was to place the base sensor above and to the left of the game board, a position where the hand would always be expected to appear on the positive side of each axis. By taking

the absolute value of each of the position vector's values, the inversion problem was resolved, though this does reduce the area in which movement is tracked. Moreover, Polhemus Patriot offers two ways in which the data can be captured: single mode and continuous mode. In the Pile game, the single mode of capture was used, as the program is not multi-threaded and would stop functioning when the Polhemus's continuous method of data capture was selected. In testing, however, the single mode method was very slow in reporting the data, causing massively detrimental effects to the frame rate of the game.

### 5.2 Hand tracking

Detecting the orientation of the user's hand plays a large part in this project. The intention is to move around environments with ease. A separate measurement is needed from the player's body position due to the hand being free to move in a different orientation to the player's body. The tracking data were obtained by attaching a small USB web camera into the pinch glove. Based on ARTag tracking libraries, hand's pose was combined with the data of the player's position and orientation in the environment and then used to compute where the hand is located in the real environment. Based on those readings, it is easy to define different functionalities that may be used for different configurations. As an exemplar, a 'firing' function was implemented based on localization of the hand (see the next section). Another function that was experimentally implemented is multiple camera viewports (one originating from the UMPC camera and another one from the mounted web camera) to provide a more immersive view to the user.

### 5.3 Head orientation

Head orientation was achieved through the capabilities of the HMD since it included a three-DOF orientation tracker. The advantage of using head orientation is that it can illuminate the use of computer vision methods for head tracking. However, when used with monitor-based AR, it can provide a distracting effect. Another problem that occurred after experimentation is that if it is used with conjunction with the rest of the sensors (Wiimote and pinch glove), it can confuse the user. For this reason the tracking capabilities of the HMD were not used in the application scenarios.

### 5.4 Wiimote interaction

It was decided to implement the Wii remote as a device to obtain positional data of the user's hand for an alternative to mouse controls. Implementation was based on an extensive library written to manage the actual communication with the Wiimote, called Wiimuse. This takes care of all of



the Bluetooth communication between the Wiimote and the computer. It also recognizes events and data received from the Wiimote accelerometers giving orientation information. When the Wiimote was implemented into the system, another thread was added to continuously retrieve data without affecting the rest of the application. When directional or action buttons are activated, different operations may be performed (i.e. start the game, help screen, etc.). The Wiimote was also found to be very useful since it is a very 'mobile' piece of equipment. It is battery-powered and can work for roughly 35 hours without needing a replacement and it emits data via Bluetooth. The only disadvantage of the Wiimote is that it provides 2-DOF tracking, so it is not a complete-orientation device. However, for a number of tabletop games (i.e. puzzles, racing, etc.) it is a very useful device since only the yaw rotation is useful.

### 5.5 Pinch glove

The pinch glove has internal sensors that give the system data on each finger's position. If a user has placed their index finger into a curled-up position, any event could be triggered. In some circumstances this is an ideal choice for user input; however, if the user has to hold any other piece of hardware then it would become difficult to make use of the glove's data because their hand position would be set by whatever is in their hand. The pinch glove allows up to 15 different combinations. However, only 5 flexures have been implemented at this stage corresponding to a finger (translate X-axis, translate Y-axis, translate Z-axis, rotate clockwise and scale). In terms of operation, the glove is initialized and then a thread is created for the constant monitoring of the glove; this thread is responsible for grabbing the data for each finger and assigning it to a variable which can be used throughout the application.

### 5.6 I/O interaction

The I/O interaction (mouse/keyboard) is adequate only when the HMD was not used. However, it allows users to perform more accurate manipulations of the superimposed information and thus it was explored only as backup option. On the other hand, the camera mounted on the rear could be used for marker detection and as the hand was holding the camera, the value returned from the ARToolKit would be a position and orientation of the hand.

## 6 Gaming techniques

In the following subsections, an overview of the main functionality of the generic AR multimodal system is presented.

### 6.1 Picking and firing

Once it has been established that a user is interacting with a particular object, the program checks the state of the sensors on the glove. If it is detected that the user is bending all five sensors over a certain threshold, then the object adopts the same position and orientation as the user's hand. This gives the impression that the object has been picked up and is now held by the user. If the sensors running along the glove's fingers are detected to straighten, then the item is dropped and falls to the plane representing the virtual ground. One method of interaction that was not fully integrated into the game, but the framework was created, is a way of firing from the virtual hand. Figure 2 provides an overview of how firing is performed.

By making a predetermined gesture, detected by the glove, a user is able to fire a virtual projectile into the scene. From the marker on the user's hand, we can apply a transformation from its orientation to that of the virtual world and thereby determine the direction a projectile would travel from the hand and whether it would intersect with other objects in the scene. Once integrated into the game, this would allow a user to destroy obstacles or non-player characters (NPCs).

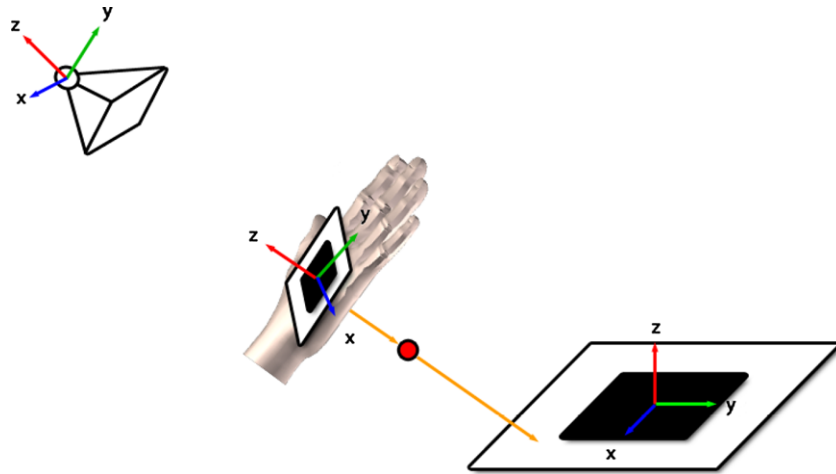
### 6.2 Collision detection and spatial sound

Using bounding boxes around each of the items in the scene and one to encompass the user's hand, intersection testing was used as a simple method of collision detection. Much as in any game, collision detection plays a vital role in gameplay; in the AR Racing game, for example, the car is not able to cross the boundaries of the game board and its progress is impeded by the other obstacles. Importantly for this particular application, the collision detection also enables the program to determine when the user's hand in the real world is intersecting with an object in the virtual scene.

To enhance the level of immersion of the application, features from the OpenAL and AL Utility Toolkit (ALUT) APIs were added. In our system, the camera is always defined as the listener, making the levels of all sounds in the augmented part of the environment relative to the camera's position and orientation. For the AR Racing game, examples of sound sources defined by the game include 'engine' and 'collision' sounds. The 'engine' sound is assigned in the virtual car whereas the 'collision' sound represents the noise created when the car collides with other virtual objects in the scene, such as the movable cubes. As the car is directed away from the center of the camera's view, the sound of its engine gradually reduces in volume and, depending on the direction of movement taken, the balance of stereo sound is altered accordingly.

Collisions between the car and virtual objects on-screen will trigger the playback of a ".wav" file. The volume of this

**Fig. 2** Augmented reality firing technique

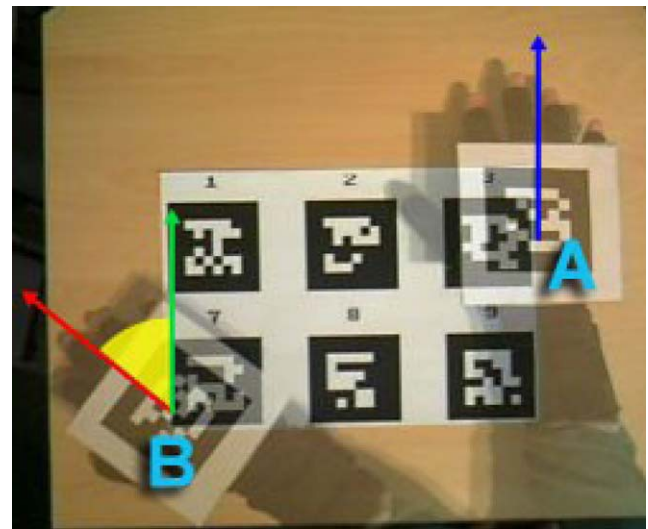


file in each audio channel will again be relative to the positions of both the camera and the collision. This functionality creates a base upon which a more complex system of sound could be developed. For example, the speed of the car affects the sound of the engine, by using different samples depending on the current velocity. Also, given the vehicle's velocity and the perceived material of another object in the scene, different sound files were tested representing varying levels of collision. A fast impact into a hard surface could sound completely different than a slight glance against a soft, malleable object.

### 6.3 Gestures

The main idea using gestures in AR gaming was to perform appropriate transformations (i.e. translations, rotations or scaling). Several possible solutions were considered. Firstly, using threshold values for rotation, wrist movements could be interpreted into larger rotations. By monitoring if rotations occur in a particular axis within a certain number of frames, it is inferred that the user wished to perform an operation (i.e. rotate a piece in that direction, Fig. 3). Depending on the AR gaming scenario, appropriate functionality is assigned. For instance, for the Pile game (see the next section) the pieces in the game can only be placed at right angles; it is reasonable to conclude that if a user wishes to rotate the pipe pieces in a certain direction, then they wish to do so in increments of 90 degrees. Similar gesture operation presents a more comfortable way of playing the game when compared to carrying out the full rotations each time.

There were, however, several issues that this functionality presented, which needed to be addressed. Firstly, the speed and extent of the rotations that would trigger the function were very difficult to define. Different users would move at different speeds and have variable ranges of motion. Setting the speed that triggered the rotation function too high would prevent certain users from accessing it, but too low would

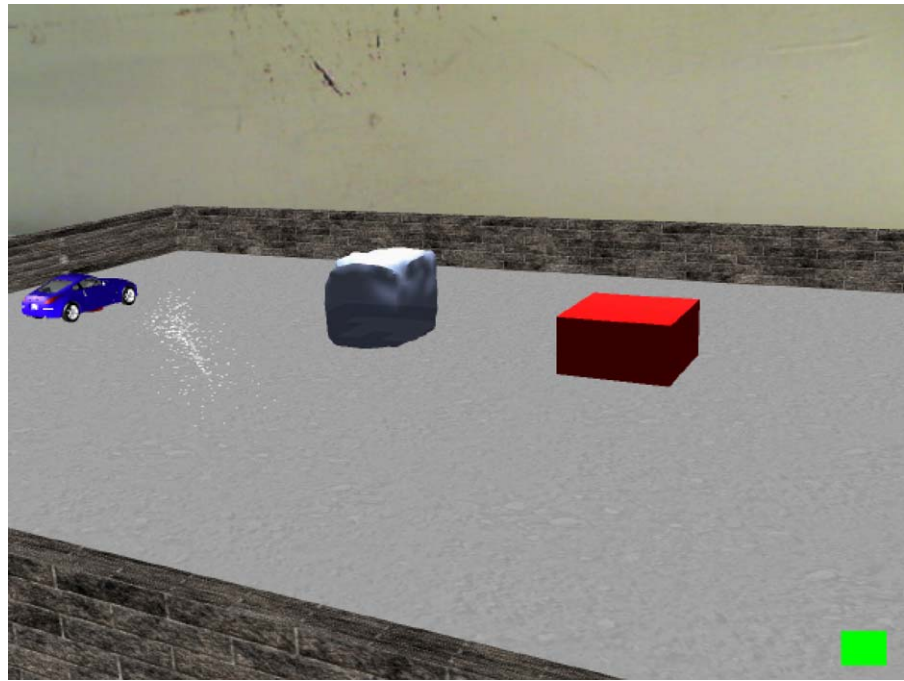


**Fig. 3** As the hand moves from position A to B, the unintentional anticlockwise rotation created is shown highlighted in yellow

cause it to be triggered when not required (i.e. by moving the hand to reach different parts of the board).

## 7 Tangible racing AR game

To illustrate the effectiveness of the multimodal interface, the interaction techniques presented above have been combined with a tabletop AR car gaming application. It was decided to use a simple gaming scenario and focus more on the reaction of the players during interaction. The goal of the game is to start the car and move around the track without colliding with either the wall or the objects that exist in the gaming arena. In addition, the objects can be rearranged in real time by picking and dropping them anywhere in the arena. A screenshot of the starting stage of the car game is shown in Fig. 4.

**Fig. 4** AR racing game

The main aim of the game is to move the car around the scene using the Wiimote without colliding with the other objects or the fountain. However, alternative interaction techniques may be used such as picking using the pinch glove. Players can interactively change the sound levels (of the car engine as well as the collisions), the speed of the simulation, and finally the color and the size of the car. In addition, they can interact with the whole gaming arena in a tangible manner by just physically manipulating the multi-markers. Interaction is performed in a far more instinctive and tangible way than is possible using a convention control system (for example, keyboard and mouse or a video game controller). The pinch glove was used to move objects in the scene by grabbing them as illustrated in Fig. 5.

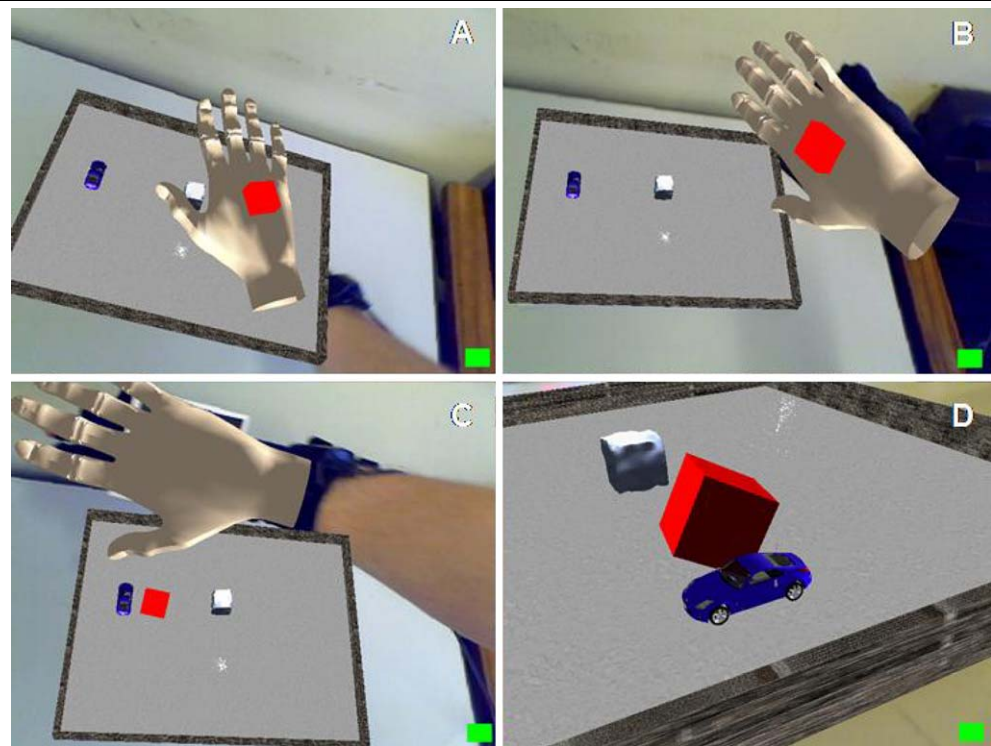
It is worth mentioning that the game can be played in a collaborative environment by eliminating the use of the HMD. One player can be in charge of the Wiimote interaction and another one for the pinch glove manipulation. Moreover, the game has only been qualitatively evaluated in two demonstrations at ‘Cogent Computing Applied Research Centre’ [32] and ‘Serious Games Institute’ (SGI) [33]. At Cogent, the basic functionality of the game was tested based on ‘think aloud’ evaluation technique [34]. Think aloud is a form of observation that involves participants talking through the actions they are performing, and what they believe to be happening, whilst interacting with a system. Overall, the feedback received was encouraging but certain aspects need to be improved in the future. The three tasks that were examined include: Wiimote interaction and pinch glove interaction. For the first task, a virtual sword was superimposed with a Wiimote placed next to it as shown

in Fig. 6(a). It must be mentioned that yaw detection cannot be detected as the motion sensor chip used in the Wiimote is sensitive to gravity, but rotation around the yaw axis is parallel to the earth.

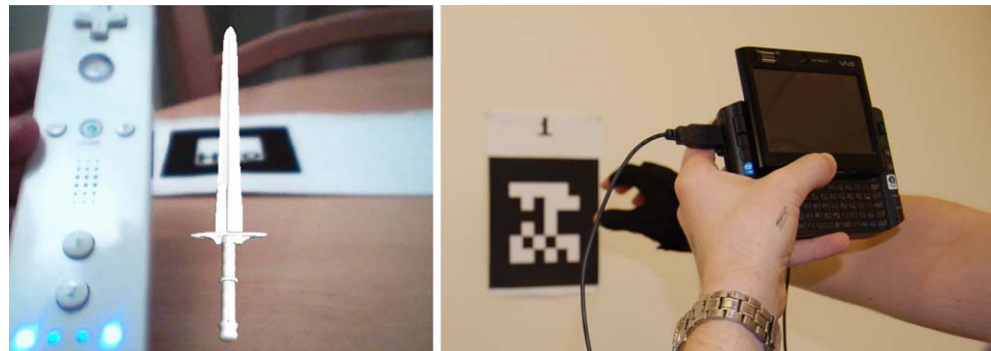
Users were asked to angle the sword upwards as if to point to an object in the sky. The feedback received was positive from four users. However, one user stated that although it is possible to detect the yaw rotation against one area, it is impossible to identify different IR sensors placed around the whole environment. As a result, interaction using the Wiimote would get confused as one IR sensor went out of site. For the second task, users were presented with the data glove placed on their right hand as illustrated in Fig. 6(a). Then they were asked to manipulate in 3D the virtual information, in this case the virtual sword used in the previous task. Virtual manipulation included scaling, rotations and translations using the fingers of the pinch glove. All users managed to interact with the pinch glove without any problems as soon as they were briefed with its operations. Four users agreed that it was very intuitive to perform the pre-programmed operations. Two users mentioned that they would like to have more combinations such as change or color and activate/deactivate the textual augmentations. One user stated that it can be tiring to control the UMPC with one hand only.

At the SGI the game was demonstrated in an internal event with 20 visitors. Initial feedback received stated that the game is very realistic in terms of interactions and enjoyable to play. Especially the idea of picking virtual objects and placing them in arbitrary positions was very enthusiastic. Most visitors felt that tangible games presented potential for

**Fig. 5** Pinch glove interaction scenarios: (a) the user is picking up a 3D object (in this case a 3D cube) that exists in the AR game; (b) shows how the user can manipulate the 3D object in three dimensions; (c) the user is dropping the 3D object in the gaming arena; (d) the object is placed in the gaming arena in a random position



**Fig. 6** (a) Wiimote interaction test; (b) data glove interaction test



the next generation of gaming. On the negative side, they preferred to experience a more complete gaming scenario including a score indicating successful achievements. In addition, some users requested more objects in the scene (i.e. obstacles), multi-player capabilities (i.e. more racing cars) and more tracks with different levels of difficulty.

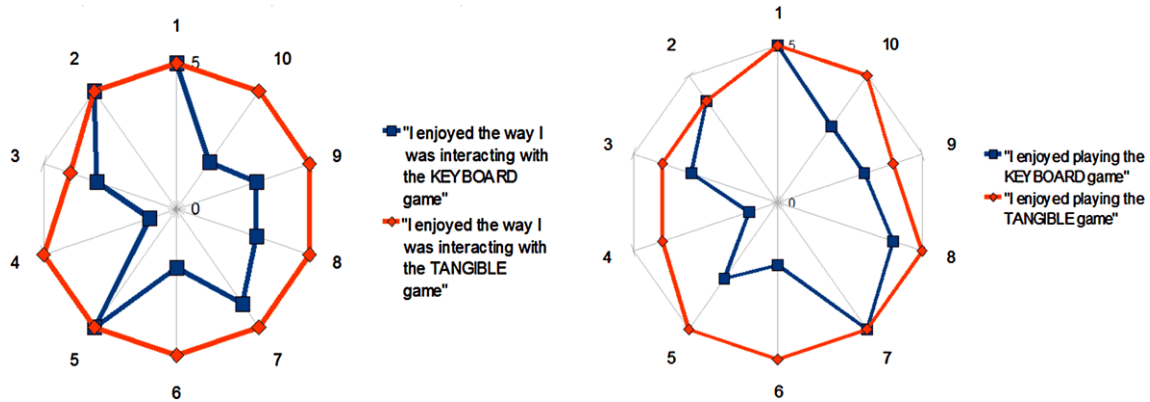
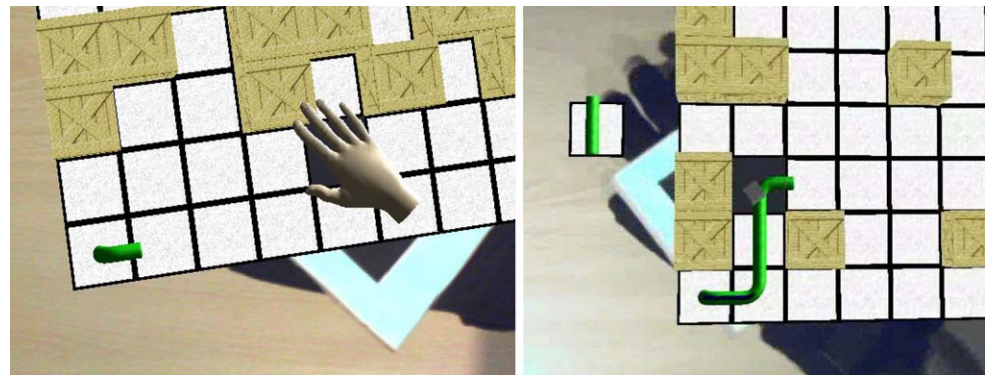
## 8 Pile AR game

The game created for this project is based on a simple game from 1989, 'Pipe Mania' [35] with sales of over 4 million units in the past twenty years. The goal of the game is to complete a circuit of pipes from a starting point to an end point on a grid. While a player is laying these pipes, there is a liquid that is gradually flowing through the pipework. If a player does not connect the pipes quickly enough, and the

liquid spills out, the game is lost. There were specific motivations behind using the template of a game already available. To begin with, it allowed for rapid development of the final product, which in itself was primarily a testing ground for the method of interaction that the project is proposing. In the version of 'Pipe Mania' created for this project, the interaction is entirely carried out through movements of a user's hand. They can reach over to pipe pieces, grip their hand into a fist to hold the piece, then move to a new position and open their hand to drop the piece. On the right-hand side of the board, there is a supply point for pipe pieces, which is automatically replenished when the player picks a piece from there.

Crates are positioned over some of the game board's squares, preventing pipes from being placed in some areas and blocking certain paths from the starting pipe to the end pipe (Fig. 7). If a user attempts to release the pipe piece on

**Fig. 7** AR pile game: (a) initial setup of the game; (b) pile game in process



**Fig. 8** Interaction test: (a) graph showing users' enjoyment of the tangible interaction method versus their enjoyment of the keyboard interaction; (b) graph showing the users' level of enjoyment of the keyboard and tangibly controlled versions of the game

squares blocked by crate or another pipe piece, they will find the pipe will remain in their hand until they move over an unoccupied square. The pipe pieces can be rotated by turning the hand in relation to the game board or visa versa. The game automatically corrects the rotation to increments of 90 degrees when the piece is placed on the board. Using this technique, a player can change any curved pipe piece to make any turn, and straight pieces can be made to run either left to right or top to bottom. The goal of the game remains the same: place the pieces to complete the pipe from the fixed start to the end before the pipe fills with water.

To obtain their opinions on particular aspects of the game's functionality, users were asked to rate their agreement with certain statements on a Likert scale. For the purposes of this project, the most important aspects of the game were related to its controls, so many of the questions related to this. The set of questions relating to both the keyboard and the tangibly controlled versions of the game were exactly the same, to attempt to find the different ways in which users perceived them. Specifically, the project aimed to discover which type of control the users found the most intuitive, the easiest to use and most enjoyable to interact with. The answers to the questionnaire, along with the observations from the tests and the notes from the post-test interview, form the basis for the conclusions drawn about the effectiveness of

the method of control, as well as the quality of the game developed. After playing the two versions of the pipe game, nine users were asked to indicate how strongly they agreed or disagreed with a number of statements, to gauge their enjoyment of the different types of game. The graphs in Fig. 8 show their responses to some of the questions.

By observing the players while playing the two versions of the game, there were several general points that were noted. Firstly, players with a background in gaming and particularly those with experience of PC games, were much faster to pick up the keyboard controls. People who had little experience of games or solely played console games, were slower to understand the controls. Several people in this category were observed to forget the controls and move the hand in ways they did not intend, slowing down the game. They were also noticeably frustrated at times, rotating the hand in the wrong direction, then back pressing several keys to find the correct movement through trial and error. Whilst playing the tangible version, all players were able to quickly understand the nature of the controls, even after little to no explanation from me. The games were generally completed more slowly, however, as the players became used to the interface and also explored the limits of the interaction. Several players had to be prompted to move the board to assist them with rotation, struggling to complete certain moves.

## 9 Conclusions and future work

Tangible AR gaming has the potential to change a number of applications that we perform in our day-to-day activities. This paper has presented a generic tangible augmented reality gaming environment that can be used to enhance entertainment using a multimodal tracking interface. The main objective of the research is to design and implement generic tangible interfaces that are user-friendly in terms of interaction and can be used by a wide range of players, including the elderly or people with disabilities. Players can interact using different combinations between a pinch glove, a Wiimote, a six-degrees-of-freedom tracker, through tangible ways as well as through I/O controls. Two tabletop augmented reality games have been designed and implemented, including a racing game and a pile game. The goal of the augmented reality racing game is to start the car and move around the track without colliding with either the wall or the objects that exist in the gaming arena. Initial evaluation results showed that multimodal-based interaction games can be beneficial in gaming. Based on these results, an augmented reality pile game was implemented with goal of completing a circuit of pipes (from a starting point to an end point on a grid). Initial evaluation showed that tangible interaction is preferred to keyboard interaction and that tangible games are much more enjoyable. From the research proposed many potential gaming applications could be produced such as strategy, puzzles and action games.

Future development will include more work on the graphical user interface to make it more user-friendly, and speech recognition is considered as an alternative option to enhance the usability of interactions. A potentially better solution that will be tested for the glove in the future on this system is to use a model for the hand that has separate sections for each finger, the position of which would be determined by the current readings on each of the finger sensors on the glove. This model could then use an alpha color value of 1.0, meaning that it is entirely transparent in the scene. As a result the model would become an alpha mask for the live video of the user's hand; the real-world hand would then appear to be above any virtual objects that the depth testing determined were further away from the camera. Finally, a formal evaluation with 30 users is currently under way and results will be used to refine the architecture.

**Acknowledgements** The authors would like to thank 'Interactive Worlds Applied Research Group (iWARG)' as well as 'Cogent Computing Applied Research Centre (Cogent)' for their support and inspiration. Videos of the AR racing game and pile game can be found at: [http://www.youtube.com/watch?v=k3r181\\_GW-o](http://www.youtube.com/watch?v=k3r181_GW-o) and <http://www.youtube.com/watch?v=0xPipinN4r8> respectively.

## References

1. Tudge, J.R.H.: Processes and consequences of peer collaboration: a Vygotskian analysis. *Child Dev.* **63**, 1364–1379 (1992)
2. Csikszentmihalyi, M.: *Flow: The Psychology of Optimal Experience*. Harper and Row, New York (1990)
3. Craig, S., Graesser, A., et al.: Affect and learning: an exploratory look into the role of affect in learning. *J. Educ. Media* **29**, 241–250 (2004)
4. de Freitas, S.: *Serious Virtual Worlds: A Scoping Study JISC*. Serious Games Institute Coventry University, London (2008)
5. Malone, T., Lepper, M.: Making learning fun. In: Snow, R., Farr, M. (eds.) *Aptitude, Learning and Instruction: Co-native and Affective Process Analyses*, pp. 223–253. Erlbaum, Lawrence (1987)
6. Lester, J.C., Towns, S.G., et al.: Deictic and emotive communication in animated pedagogical agents. In: Cassell, J., Prevost, S., Sullivant, J., Churchill, E. (eds.) *Embodied Conversational Agents*, pp. 123–154. MIT Press, Boston (2000)
7. Lester, J.C., Converse, S.A., et al.: Animated pedagogical agents and problem-solving effectiveness: a large-scale empirical evaluation. In: *Proc. of the 8th Int'l Conference on Artificial Intelligence in Education*, pp. 23–30. IOS Press, Kobe (1997)
8. Yoon, S.-Y., Blumberg, B.M., et al.: Motivation driven learning for interactive synthetic characters. In: *Fourth International Conference on Autonomous Agents*, Barcelona (2000)
9. Laurel, B.: *Interface agents: Metaphors with characters*. In: Bradshaw, J.M. (ed.) *Software Agents*, pp. 67–78. AAAI Press/MIT Press, London (1997)
10. Iuppa, N., Weltman, G., et al.: Bringing Hollywood storytelling techniques to branching storylines for training applications. In: *3rd Narrative and Interactive Learning Environments*. Edinburgh, Scotland, pp. 1–8 (2004)
11. Walther, B.K.: Reflections on the methodology of pervasive gaming. In: *Proc. of the 2005 ACM SIGCHI Int'l Conference on Advances in Computer Entertainment Technology*, pp. 176–179. ACM Press, Valencia (2005),
12. Oda, O., Lister, L., J., White, S., Feiner, S.: Developing an augmented reality racing game. In: *Proc. of the 2nd Int'l Conference on Intelligent Technologies for Interactive Entertainment*, January 8–10. Cancun, Mexico, Article No. 2 (2008)
13. Beigl, M., Krohn, A., Zimmer, T., Decker, C.: Typical sensors needed in ubiquitous and pervasive computing. In: *Proc. of the 1st Int'l Workshop on Networked Sensing Systems (INSS)*, Tokyo, Japan, June, pp. 153–158 (2004)
14. Dombroviak, K.M., Ramnath, R.: A taxonomy of mobile and pervasive applications. In: *Proc. of the 2007 ACM Symposium on Applied Computing*, pp. 1609–1615. ACM Press, Seoul (2007)
15. Billingham, M., Kato, H., Poupyrev, I.: The MagicBook: a transitional AR interface. *Comput. Graph.* **25**(5), 745–753 (2001)
16. Rekimoto, J., Nagao, K.: The world through the computer: computer augmented interaction with real-world environments. In: *Proc. of the 8th Annual Symposium on User Interface Software and Technology (UIST'95)*, Pittsburgh, Pennsylvania, USA, November, pp. 29–36. ACM Press, New York (1995)
17. Starner, T., Mann, S., et al.: Augmented reality through wearable computing. *Presence: Teleoper. Virtual Environ.* **6**(4), 386–398 (1997)
18. Butz, A., Höllerer, T., et al.: Enveloping users and computers in a collaborative 3D augmented reality. In: *Proc. of the 2nd IEEE and ACM International Workshop on Augmented Reality*, October, pp. 35–44. IEEE Computer Society, San Francisco (1999),
19. Sandor, C., Klinker, G.: A rapid prototyping software infrastructure for user interfaces in ubiquitous augmented reality. *Personal Ubiquitous Comput.* **9**(3):169–185 (2005)

20. Reitmayr, G., Schmalstieg, D.: Semantic world models for ubiquitous augmented reality. In: Proc. of Workshop Towards Semantic Virtual Environments' (SVE 2005), March (2005)
21. Cheok, A.D., Fong, S.W., et al.: Human pacman: a sensing-based mobile entertainment system with ubiquitous computing and tangible interaction. In: Proc. of the 2nd Workshop on Network and System Support for Games, pp. 106–117. ACM Press, California (2003)
22. Henrysson, A., Billinghurst, M., Ollila, M.: AR tennis. In: International Conference on Computer Graphics and Interactive Techniques Archive ACM SIGGRAPH 2006 Sketches, Article No. 13. ACM Press, New York (2006)
23. Magerkurth, C., Engelke, T., Memisoglu, M.: Augmenting the virtual domain with physical and social elements: towards a paradigm shift in computer entertainment technology. *Comput. Entertainment* 2(4), 12 (2004)
24. Stapleton, C.B., Hughes, C.E., Moshell, J.M.: MIXED FANTASY: exhibition of entertainment research for mixed reality. In: Proc. of the 2nd Int'l Symposium on Mixed and Augmented Reality, pp. 354–355. IEEE Computer Society, Tokyo (2003)
25. Korris, J.: Full spectrum warrior: How the institute for creative technologies built a cognitive training tool for the xbox. In: 24th Army Science Conference. Florida, Orlando, December (2004)
26. Hughes, C.E., et al.: Mixed reality in education, entertainment, and training. *IEEE Comput. Graph. Appl.*, November/December: 24–30 (2005)
27. Benford, S., Magerkurth, C., Ljungstrand, P.: Bridging the physical and digital in pervasive gaming. *Commun. ACM* 48(3), 54–57 (2005)
28. Magerkurth, C., Engelke, T., Grollman, D.: A component-based architecture for distributed, pervasive gaming applications. In: Proc. of the 2006 ACM SIGCHI International Conference on Advances in Computer Entertainment Technology. ACM Press, California (2006). Article No 15
29. Goldsmith, D., Liarokapis, F., et al.: Augmented reality environmental monitoring using wireless sensor networks. In: Proc. of the 12th Int'l Conference on Information Visualisation, pp. 539–544. IEEE Computer Society, Los Alamitos (2008)
30. ARToolKit. Available at: <http://www.hitl.washington.edu/artoolkit/>. Accessed at: 21/09/2008
31. ARTag. Available at: <http://www.artag.net/>. Accessed at: 21/09/2008
32. Cogent Computing Applied Research Centre. Available at: <http://www.coventry.ac.uk/researchnet/cogent>. Accessed at: 11/10/2008
33. Serious Games Institute. Available at: <http://www.seriousgamesinstitute.co.uk/>. Accessed at: 11/10/2008
34. Dix, A., Finlay, J., Abowd, G., Beale, R.: *Human-Computer Interaction*. Prentice Hall, Harlow (2004)
35. Empire Interactive Entertainment: Pipemania 2008 New \_2\_ With Consoles. Available at: [www.empireinteractive.co.uk/corporate/files/PDFs/Pipemania%202008%20New%20\\_2\\_%20With%20Consoles.pdf](http://www.empireinteractive.co.uk/corporate/files/PDFs/Pipemania%202008%20New%20_2_%20With%20Consoles.pdf). Accessed at: 19/05/2009



**Fotis Liarokapis** is the Director of Interactive Worlds Applied Research Group (iWARG) at the Faculty of Engineering and Computing, Coventry University and a Research Fellow at the Serious Games Institute (SGI), Coventry University. He is also a Visiting Lecturer at the Centre for VLSI and Computer Graphics, University of Sussex and a visiting research fellow at the giCentre, City University. His research interests include virtual and augmented reality, computer graphics, human-computer interaction and serious games. Furthermore, he is a member of IEEE, IET, ACM and BCS and he is on the editorial advisory board of The Open Virtual Reality Journal published by Bentham.



**Louis Macan** graduated this year with a first class honors B.Sc. degree in Creative Technologies from Coventry University. His previous research involving augmented reality has been published as part of VAST 2007 and IEEE VS-GAMES 2009, at the latter of which he also presented the work during the conference proceedings. Louis has recently worked as a consultant on a communications technology project in the Midlands and is preparing to start development on a serious game with a company in Milan. He will begin his Ph.D. with Coventry University towards the end of 2009.



**Gary Malone** obtained his Bachelor of Arts degree in Creative Computing at Coventry University in 2008. He is currently studying a Master of Science degree at The University of Newcastle-upon Tyne in Mobile and Pervasive Computing. His research interests include computer vision, 3D visualization, augmented reality, biometrics, serious games and virtual worlds.



**Genaro Rebolledo-Mendez** is a Senior Lecturer and Researcher at the Faculty of Informatics University of Veracruz, Mexico. Previously, he was a Senior Researcher at the Serious Games Institute, University of Coventry, UK. He has also been a Research Fellow at the London Knowledge Lab, University of London and the IDEAS Lab, Sussex University, UK. Genaro's interest is the design and evaluation of educational technology that adapts sensitively to affective and cognitive differences among students. To do so,

he studies how cognitive and affective differences impact students' behavior while interacting with educational technology and how, in turn, technology impacts students' learning and affect. To that end, he uses techniques from artificial intelligence, computer science, education and psychology.



**Sara de Freitas** is Professor of Virtual Environments and Director of Research at the Serious Games Institute (SGI)—an international hub of excellence in the area of games, virtual worlds and interactive digital media for serious purposes. Situated on the Technology Park at the University of Coventry, Sara leads an interdisciplinary and cross-university applied research group with expertise in AI and games, visualization, mixed reality, augmented reality and location-aware technologies. The Research Group

works closely with international industrial and academic research and development partners. Sara is a Visiting Fellow of the London Knowledge Lab, London, and a Fellow of the Royal Society of Arts.



### **8.10 Paper #10**

Goldsmith, D., Liarokapis, F., Malone, G., Kemp, J. Augmented Reality Environmental Monitoring Using Wireless Sensor Networks, Proc. of the 12<sup>th</sup> International Conference on Information Visualisation (IV08), IEEE Computer Society, 8-11 July, 539-544, 2008.

Contribution (30%): Collaboration on the design of the architecture. Advice on the implementation of the majority of the VR interface. Write-up of most of the paper.

# Augmented Reality Environmental Monitoring Using Wireless Sensor Networks

Daniel Goldsmith, Fotis Liarokapis, Garry Malone, John Kemp  
Cogent Computing Applied Research Centre, Coventry University  
Department of Computer Science, Coventry CV1 5FB  
{goldsmid, F.Liarokapis, maloneg, kempj}@coventry.ac.uk

## Abstract

*Environmental monitoring brings many challenges to wireless sensor networks: including the need to collect and process large volumes of data before presenting the information to the user in an easy to understand format. This paper presents SensAR, a prototype augmented reality interface specifically designed for monitoring environmental information. The input of our prototype is sound and temperature data which are located inside a networked environment. Participants can visualise 3D as well as textual representations of environmental information in real-time using a lightweight handheld computer.*

**Keywords---** Augmented Reality, Handheld Interfaces, Human-Computer Interaction, Wireless Sensor Networks.

## 1. Introduction

Augmented Reality (AR) is a subset of a Mixed Reality (MR) that allows for seamless integration of virtual and real information in real-time. Other important characteristics of AR include real-time and accurate representation in three-dimensions (3D) as well as being interactive. However, AR is not limited to vision but can be applied to all senses including touch, and hearing [1]. Although many applications of AR have emerged, they are usually concerned with tracking. This is achieved using computer vision techniques, sensor devices, or multimodal interactions to calculate position and orientation of a camera/user. However, AR has not been actively employed for the visualisation of environmental information originating from a Wireless Sensor Network (WSN).

Research within the WSN community has led to the development of new computing models ranging from distributed computing to large-scale pervasive computing environments [2]. This rapid evolution of pervasive computing technologies has allowed the development of novel interfaces which are capable of interacting with sensory information originating from the environment with little or no manual intervention. Although a number of technologies are able to perform natural interactions, pervasive AR is one of the strongest candidates.

WSN technology uses networks of sense enabled miniature computing devices to gather information about the world around them. Common applications include

environmental monitoring, military, health, home, and education [3], [4]. While the gathering of data within a sensor network is one challenge, another of equal importance is presenting the data in a useful way to the user. Using low cost, low power computing devices equipped. A sensor network is composed of a large number of sensor nodes, with wireless communication and sensing hardware. These are deployed within the area of interest to monitor and measure phenomenon and collaboratively processes the data before relaying information to a base station or sink node.

The constrained nature of the data gathering platform has lead to much of the active WSN research to be focused on network concerns such as data communication and energy efficiency. Recent initiatives such as Nokia's Sensor Planet [5] aim to incorporate sensor networks, mobile phones, and other devices into a large scale ad-hoc multipurpose sensor network [6] with sensor information available via a web based Application Programming Interface (API). The use of these commonly available and familiar devices is envisaged to allow WSN to become part of the pervasive computing mainstream, requiring new approaches to information visualisation to process the vast amount of information available.

This paper presents SensAR - an environmental monitoring prototype that uses WSN to gather temperature and audio data about the user's surroundings. SensAR displays the environmental information in an understandable format using a real-time handheld AR interface. Participants can visualise 3D as well as textual representations of the sound and temperature information in a tangible manner. To our knowledge, SensAR is the first to embody the idea of combining sound and temperature data in a handheld AR environment.

The remainder of this paper is organised as follows: Section 2 describes some of the most important related work. Section 3 gives an overview of our systems architecture including a brief description of its main components. Section 4 presents the operation of the handheld AR interface in an indoor networked environment. Section 5 illustrates how the environmental information coming from the sound and temperature sensors is visualised in an AR environment through 3D objects and textual annotations. Finally, section 6 concludes by presenting our plans for future work

## 2. Related Work

A number of WSN applications have been proposed in the past and some of the most characteristic systems are presented here. iPower [7] utilises a WSN to provide intelligent energy conservation for buildings. The system is composed of a sensor network that gathers data on light levels, temperature, and sound to activate appliances based on the likelihood of a room being occupied. If the system detects low temperature or high brightness in a room that is unlikely to be occupied, a signal can be sent to turn off the air-conditioning or reduce lighting levels. If the network receives a signal that the area is still occupied (for instance detection of a noise) the system returns the light and temperature levels to values suitable for comfortable use of the room. Aside from a system overview provided by the user interface iPower has no data visualization, it nevertheless presents a practical application of wireless sensor networks in environmental monitoring.

SpyGlass [8] is concerned with the provision of a visualization framework for WSNs. Data is gathered on a gateway node within the network then passed to the visualization application on a remote machine. The data is passed using the TCP/IP suite of protocols and therefore can be carried over many network types including Local Area Network (LAN), Wireless Local Area Network (WLAN), and General Packet Radio Service (GPRS). Network visualisation is provided by a Graphical User Interface (GUI) allowing an overall view of the network to be displayed. This visualization component is comprised of a relation layer to display relationships between nodes and a node layer to draw the nodes themselves.

The 'Plug' sensor network [9] is a ubiquitous networked sensing platform ideally suited to broad deployment in environments where people work and live. The backbone of the Plug sensor network is a set of 35 sensor, radio, and computation enabled power strips. A single Plug device fulfils all the functional requirements of a normal power strip and can be used without special training. Additionally, each Plug has a wide range of sensing modalities (e.g., sound, light, electrical voltage and current, vibration, motion, and temperature) for gathering data about how it is being used and its nearby environment.

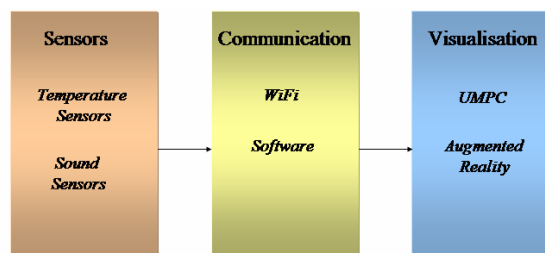
In terms of handheld AR sensing applications, most prototypes that exist focus on multimodal interactions using tracking sensors. An interesting approach to 3D multimodal interaction in immersive AR environment that accounts for the uncertain nature of the information sources was proposed by [10]. The multimodal system fuses symbolic and statistical information from a set of 3D gesture, spoken language, and referential agents. The referential agents employ visible or invisible volumes that can be attached to 3D trackers in the environment, and which use a time stamped history of the objects that intersect them to derive statistics for ranking potential referents.

Another approach proposed an architecture for handling events from different tracking systems and maintaining a consistent spatial model of people and objects [11]. The principal distinguishing feature is the automatic derivation of dataflow network of distributed sensors, dynamically and at run-time, based on requirements expressed by clients. This work also classifies sensor characteristics for AR and Ubicomp. Moreover, a grid of sensors was used to synthesize images in AR by interpolating the data and mapping them to colour values [12]. This application used an optically tracked mobile phone as a see-through handheld AR display allowing for interaction metaphors already familiar to most mobile phone users. The sensor network is interfaced by visualizing its data within its context, taking advantage of the spatial information.

Furthermore, techniques for creating indoor location based applications for mobile augmented reality systems using computer vision and sensors have been also well documented [13]. An indoor tracking system was proposed that covers a substantial part of a building. It is based on visual tracking of fiducial markers enhanced with an inertial sensor for fast rotational updates. To scale such a system to a whole building, a space partitioning scheme was introduced to reuse fiducial markers throughout the environment.

## 3. System Architecture

SensAR follows an experimental prototype recently presented [14]. However, there are many differences with the earlier prototype. Firstly, sound and temperature sensors are populated inside the environment (see Figure 2). Secondly, WiFi is used instead of Bluetooth providing a much faster method of communication, although Bluetooth can be enabled for connecting other hardware devices. Finally, the mobile client side provides enhanced visualisation options including textual and 3D information. SensAR uses a three-tier architecture consisting of a sensor layer, communication layer, and visualisation layer. A diagrammatic overview of the pipeline of our system is presented in Figure 1.



**Figure 1 Architecture of SensAR**

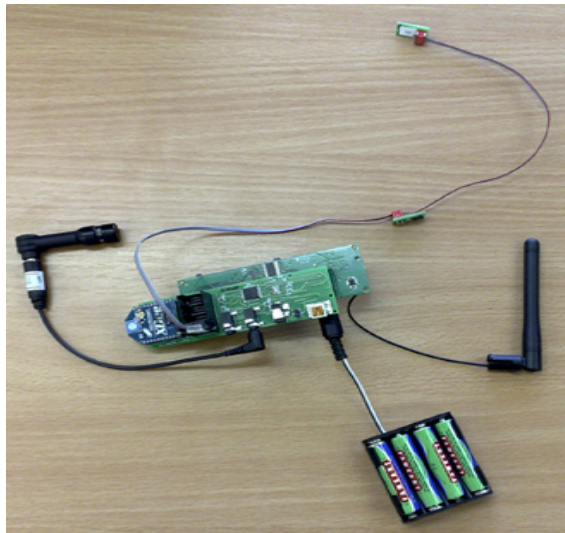
The sensor layer handles multimodal data from temperature and sound sensors, positioned at fixed locations within an indoor environment. These sensors are attached to a WSN node, which is capable of performing the initial processing before passing the data up the protocol stack. In the case of the WSN node, the

data is formatted ready for transmission by the communication layer. The data is transferred over a WiFi link via User Datagram Protocol (UDP) to a dedicated server on the visualisation machine. This link is bidirectional, and allows control packets to be sent between each device. The visualisation layer contains the handheld device running the AR software. The data received is represented using visual information such as 3D objects and textual information.

### 3.1 Hardware

There are a variety of available embedded platforms for sensing applications. Communication technologies such as Bluetooth, WiFi and ZigBee [15] allow for network collection and transfer of environmental data to wearable devices. The hardware choice decision for the network discussed here was based on the available platforms' sensing capability, ease of software development and size.

Gumstix Verdex XM4-bt boards were selected as the main processing platform. Although not as popular as Mica2 motes for wireless sensing applications, they are becoming more prevalent [16]. These devices offer more processing power and memory (in terms of both RAM and flash) than many similarly sized platforms. The particular model chosen includes an Intel XScale PXA270 400MHz processor, 16MB of flash memory, 64MB of RAM, a Bluetooth controller and antenna, 60-pin and 120-pin connectors for expansion boards, and a further 24-pin flex ribbon connector. There are no on-board sensors provided, though a variety of interface methods are available.



**Figure 2 Sound and Temperature Sensors**

Commercially available expansion boards for the Gumstix platform include communications options such as WiFi and Ethernet, along with additional storage provided by Compact Flash (CF) cards. An expansion board developed in house additionally provides an I<sup>2</sup>C

bus for the connection of sensors, along with a ZigBee compatible module. The sensors used for temperature sensing were the Analog Digital ADT75A chip [17], which performs sampling and conversion internally, providing the sensed temperature values via an I<sup>2</sup>C bus.

For visualisation, a VAIO UX Ultra Mobile PC (UMPC) was used, which is one of the smallest fully functioning PCs ever made. Comparable to PDAs in size, but with more powerful processing capabilities, it is able to run complex AR applications. VAIO UX includes an Intel® Core™ Solo Processor at 1.3MHz, wireless 802.11a/b/g, 32GB hard drive, 1GB SDRAM, 4.5" touch panel LCD, a Graphics Accelerator and 2 built-in digital cameras. This makes it a suitable device to handle our WSN configuration and display the visualisation with real-time performance

### 3.2 Software

At the heart of the sensing system is a collection of software libraries developed as part a software support system for WSN. The provision of a generic interface to common sensor network tasks allows the implementation details of complex tasks to be hidden, thereby offering the systems designer a cleaner workflow. Software abstractions of sensing and communication tasks have been created, allowing the user to plug functionality into the application.

A generic interface to the I<sup>2</sup>C bus has been implemented to allow access to data from the temperature modules. The API allows other I<sup>2</sup>C enabled devices such as digital compasses, pressure sensors, accelerometers, and light meters to be supported. Using an abstraction model for sensing interfaces, the process of gathering data is simplified, as similar function calls are used to retrieve information from different devices. This in turn allows a modular approach to application development.

The framework supports a range of communication protocols and interfaces, offering the choice of Bluetooth, WiFi and Ethernet based data transfer. Support is also provided for network protocols offered by each communications stack. As an example, WiFi offers connection orientated TCP and connectionless UDP allowing the user to balance the requirements of the application with the quality of service received. In keeping with the modular theme of the framework, the communication modules are interchangeable. This allows the user to swap between radio devices by simply changing the software module used. In the instance of WiFi and Ethernet this is a straight swap as the two communications mediums use the IP suite of protocols, and addressing schemes. However if the user wishes to switch to Bluetooth communication, the alternative hardware addressing would scheme would need used, all other communication calls are handled in the same way regardless of communication medium.

The sensing layer was developed using the above framework. Using the high-level Python programming language for development has allowed the algorithms for

the gathering of data to be prototyped with a development cycle much shorter than that associated with compiled languages such as C. Although Python offered ease of development, the framework has also been implemented in a collection of C libraries, allowing the final application to be transferred to this faster executing compiled language for deployment. Whilst Python and C have differences in syntax, the framework has been designed to take account of the similarities in functionality and programming methodology afforded by both languages. This allows the code developed to be transferred between each language making only small syntactical changes.

The visualisation layer used the OpenGL API for the rendering of the 3D environmental representations. The textual augmentations were implemented based on GLUT API which provides support for bitmap fonts. The six-degrees-of-freedom tracking of the user inside the environment was based on ARToolKit library [18] and the rest of the coding of the handheld interface was performed in C programming language. Finally, the 3D models used in the visualisation were designed using an open source modelling tool (Blender) and exported in VRML file format.

#### 4. Handheld AR Interface

A handheld AR interface has been implemented in order to allow a user to experience the environmental information gathered. Sensors collect sound and temperature level data at various points in space and relay this information to SensAR. A user interface is then used to seamlessly superimpose computer generated representations of sound and temperature based on the readings of these sensors. Figure 3 illustrates how a user operating the handheld interface would perceive a 3D representation of environmental information (in this case temperature and sound) in a mobile AR environment.

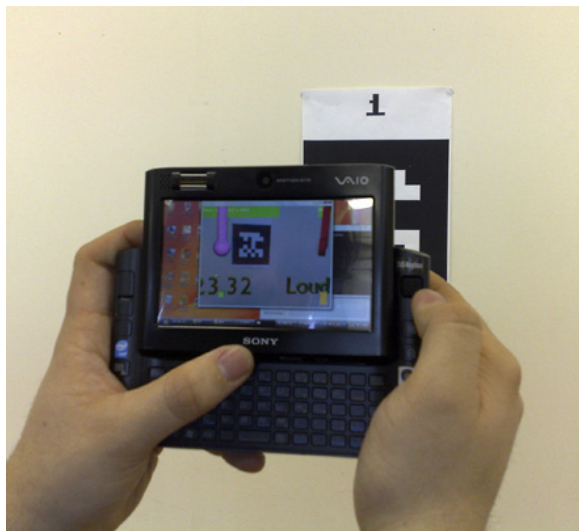


Figure 3 Handheld AR Visualisation

Users can navigate inside the room by moving the UMPC and detecting different markers. SensAR checks each video frame for predetermined patterns that are included in the environment. These are squares containing a unique black and white image that the program can be programmed to recognize. The markers used in this project have been specifically selected from the ARTag library [19] to be distinct from one another regardless of orientation or reflection. The current version of the system uses patterns numbered 1 to 12, taken from the ARTag implementation of the ARToolKit as shown in Figure 4.

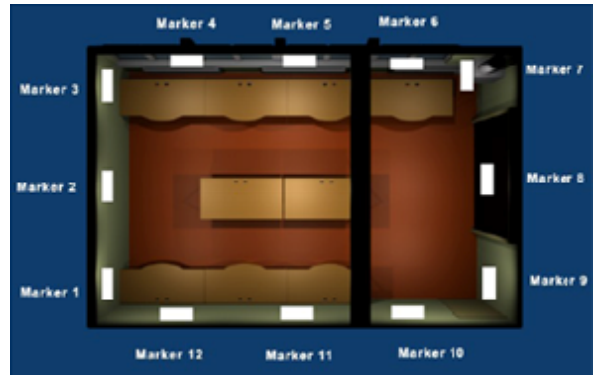


Figure 4 Marker Setup

The markers are placed so that the centre of the pattern is halfway up the height of the wall (142.5cm from the floor). For each marker different sound and temperature sensors are attached as close to the markers as possible to give accurate localisation readings. The markers are enlarged as much as possible whilst still fitting on a single sheet of A4 size sheet of paper. When the program detects a marker within a video frame, it overlays a 3D model of a thermometer and a music note onto the video image (Figure 6 and Figure 7). One of the versions of our program also includes a 3D representation of the entire room, which is projected over the real room in AR.

In order for this to line up with the real image sent from the camera, we have to attach the model to one of the 12 markers (Figure 4), much in the same way as the 3D virtual sensors as illustrated in Figure 5.



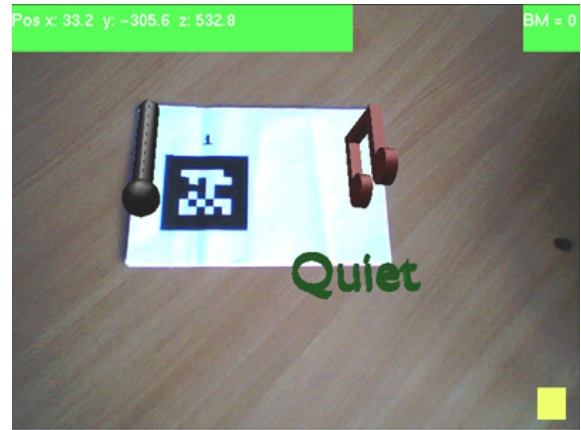
**Figure 5 Virtual Representation of Environment**

However, if there are several markers in view, we don't want the program to draw multiple versions of the virtual room. To prevent this, we exploit the confidence value that is used in marker detection. Each detected pattern is then checked for correlation with the markers detected by the program and a confidence value is generated to show the level of similarity. SensAR compares the confidence values of the patterns that have been established as being markers. The marker that has the greatest confidence value is used as the point from which to draw the virtual room. One advantage with using this system is that the room will automatically revert to the next best marker in sight should the most visible marker become obscured.

## 5. Environmental Data Visualisation

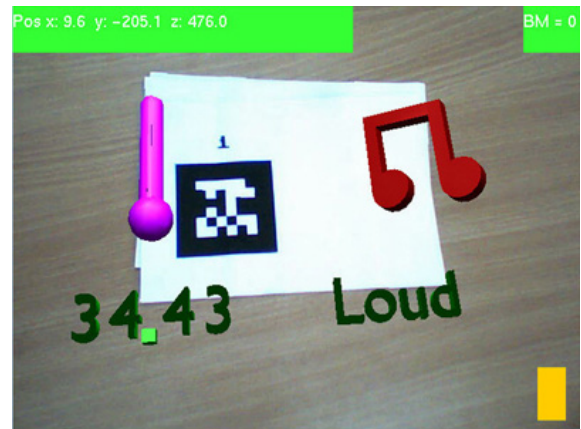
There is an open issue of how to visually represent environmental data coming from the WSNs. One of the aims of this work was to select an appropriate metaphor to assist users in rapid interpretation of the information. After some informal evaluation, it was decided to represent the environmental information through the use of a 3D thermometer and a 3D music note. In the previous prototype (which included only sound data) a 3D microphone was used.

In terms of operation, as soon as the temperature and sound sensors are ready to transmit data, visual representations including a 3D thermometer and a 3D music note as well as textual annotations are superimposed onto the appropriate marker. This is the neutral stage of SensAR where no sensor readings are actually inputted to the AR interface. An example screenshot of the neutral stage is illustrated in Figure 6.



**Figure 6 Low Levels of Sound and Temperature**

When environmental data is transferred to the AR interface, the color of the 3D thermometer and the 3D music note change according to the temperature level and sound volume accordingly. In addition, textual annotations indicate the sensor readings. For the temperature data, the readings from the sensors (which have an error of  $\pm 0.1$ ) are superimposed as text next to the 3D thermometer. For the sound data, a different measure was employed based on a scale 0 to 4, where 0 corresponds to 'quiet', 1 corresponds to 'low', 2 corresponds to 'medium', 3 corresponds to 'loud' and 4 corresponds to 'very loud'. This choice of banding has been based on user input, to provide a clearer representation of sound levels than a raw value could. Also note that the bottom right side displays the intensity of the sound level. A screenshot of the above configuration is shown in Figure 7.



**Figure 7 High Levels of Sound and Temperature**

It is worth-mentioning that the camera position is also displayed on the top left side of the interface. This feature is useful for calculating the position of the user in respect to the rest of the environment. Moreover, users can interact with the superimposed information using the keyboard or the mouse of the UMPC. In this way, it is possible to translate, rotate or scale the visual augmentations in real-time. In addition, it is possible to

hide the various elements of the interface such as the camera position, the textual annotations and the 3D objects.

## 6. Conclusions and Future Work

This paper describes SensAR, a prototype mobile AR system for visualising environmental information including temperature and sound data. Sound and temperature data are transmitted wirelessly to our client which is a handheld device. Environmental information is represented graphically as 3D objects and textual information in real-time based. Participants visualise and interact with the augmented environmental information using a small but powerful handheld computer. The main advantage of SensAR is the visual representation of wireless sensor data in a meaningful and tangible way. We believe that SensAR design principle is essential for the effective realisation of ubiquitous computing.

In the future we are planning to integrate more sensors to SensAR including light, pressure and humidity. On the visualization side, we are currently working with a head-mounted display that includes orientation tracking to provide a greater level of immersion to the users. In terms of interaction other forms of interaction will be added to the prototype such as a digital compass, a virtual reality glove and the Wii controller. Finally we plan to do user extensive studies to test the feasibility of SensAR application.

## Acknowledgements

The authors would like to thank Dr. Elena Gaura and the rest of the team in the Cogent Computing Applied Research Centre for their support and inspiration as well as Louis Macan, Sarah Mount and Prof. Robert Newman who worked so hard during the development of the first prototype.

## References

- [1] Azuma, R., Baillot, Y., et al. Recent Advances in Augmented Reality, *IEEE Computer Graphics and Applications*, **21**(6): 34-47, 2001.
- [2] Harihar, K., Kurkovsky, S. Using Jini to Enable Pervasive Computing Environments, In *Proc. of the 43<sup>rd</sup> Annual Southeast Regional Conference - Volume 1, Architecture and distributed systems*, ACM Press, Kennesaw, Georgia, 188-193, 2005.
- [3] Akyildiz, I.F., Su, W., Sankarasubramaniam, Y., Cayirci, E. Wireless sensor networks: a survey. *Computer Networks*, **38**(4): 393-422, 2002.
- [4] Culler, D., Estrin, D., Srivastava, M. Guest Editors' Introduction: Overview of Sensor Networks, *Computer* **37**(8): 41-49, 2004.
- [5] SensorPlanet, Available at: [http://www.sensorplanet.org/], Accessed at: 29/02/2008.
- [6] Tuulos, V.H., Scheible, J., Nyholm, H. Combining Web, Mobile Phones and Public Displays in Large-Scale: Manhattan Story Mashup, In *Proc. of the 5<sup>th</sup> Int'l Conference on Pervasive Computing*, Canada, 37-54, 2007.
- [7] Yeh, L.W. Wang, Y.C, Tseng, Y.C. iPower: An Energy Conservation System for Intelligent Buildings by Wireless Sensor Networks, To appear in *Int'l Journal of Sensor Networks*, **5**(2), 2009.
- [8] Buschmann, C., Pfisterer, D., et al. Spyglass: a wireless sensor network visualiser, *SIGBED Review*, **2**(1): 1-6, 2005.
- [9] Lifton, J., Feldmeier, M., et al. A platform for ubiquitous sensor deployment in occupational and domestic environments, In *Proc. of the 6<sup>th</sup> Int'l Conference on Information Processing in Sensor Networks*, ACM Press, Cambridge, Massachusetts, USA, 119-127, 2007.
- [10] Kaiser, E., Olwal, A., et al. Mutual Disambiguation of 3D Multimodal Interaction in Augmented and Virtual Reality, In *Proc. of the 5<sup>th</sup> Int'l Conference on Multimodal Interfaces*, ACM Press, November 5-7, Vancouver, British Columbia, Canada, 12-19, 2003.
- [11] Newman, J., Schall, G., Schmalstieg, D. Modelling and Handling Seams in Wide-Area Sensor Networks, In *Proc. of the 10<sup>th</sup> Int'l Symposium on Wearable Computers*, IEEE Computer Society, Montreux, Switzerland, 51-54, 2006.
- [12] Rauhala, M., Gunnarsson, A.S., Henrysson, A. A novel interface to sensor networks using handheld augmented reality, In *Proc. of the 8<sup>th</sup> Int'l Conference on Human-Computer Interaction with Mobile Devices and Services*, ACM Press, Helsinki, Finland, 145-148, 2006.
- [13] Reitmayr, G., Schmalstieg, D. Location based Applications for Mobile Augmented Reality, In *Proc. of the 4<sup>th</sup> Australasian User Interface Conference*, Adelaide, Australia, 65-73, 2003.
- [14] Liarokapis, F., Newman, R., et al. Sense-Enabled Mixed Reality Museum Exhibitions, In *Proc. of the 8<sup>th</sup> Int'l Symposium on Virtual Reality, Archaeology and Cultural Heritage*, Eurographics, Brighton, UK, 26-30 November, 31-38, 2007.
- [15] IEEE 802.15.4. IEEE Standard for Information technology Part 15.4: Specifications for Low-Rate Wireless Personal Area Networks (LR-WPANs), 2003.
- [16] Keoh, S.L., Dulay, N., et al. Self managed cell: A middleware for managing body sensor networks. In *Proc of the 4<sup>th</sup> Int'l Conference on Mobile and Ubiquitous Systems: Computing, Networking and Services (MobiQuitous)*, Philadelphia, USA, August, 1-5, 2007.
- [17] ADT75, Available at: [http://www.analog.com/en/prod/0%2C2877%2CADT75%2C00.html], Accessed at: 25/02/2008.
- [18] ARToolKit, Available at: [http://www.hitl.washington.edu/artoolkit/], Accessed at: 25/02/2008.
- [19] ARTAG, Available at: [http://www.artag.net/], Accessed at: 29/02/2008.

### **8.11 Paper #11**

Liarokapis, F., Debattista, K., Vourvopoulos, A., Ene, A., Petridis, P. Comparing interaction techniques for serious games through brain-computer interfaces: A user perception evaluation study, *Entertainment Computing*, Elsevier, 5(4): 391-399, 2014.

Contribution (40%): Collaboration on the design of the architecture. Advice on the implementation of the serious game as well as in the BCI interface. Write-up of most of the paper.





# Comparing interaction techniques for serious games through brain–computer interfaces: A user perception evaluation study <sup>☆</sup>



Fotis Liarokapis <sup>a,d,\*</sup>, Kurt Debattista <sup>b</sup>, Athanasios Vourvopoulos <sup>c</sup>, Panagiotis Petridis <sup>d</sup>, Alina Ene <sup>d</sup>

<sup>a</sup> Human Computer Interaction Laboratory, Masaryk University, Faculty of Informatics, Botanická 554/68a, Brno, Czech Republic

<sup>b</sup> WMG, University of Warwick, Gibbet Hill Road, Coventry CV32 5LD, UK

<sup>c</sup> Madeira Interactive Technologies Institute, University of Madeira, Funchal, Portugal

<sup>d</sup> Serious Games Institute, Coventry Innovation Village, Coventry University Technology Park, Cheetah Road, Coventry, West Midlands CV1 2TL, UK

## ARTICLE INFO

### Article history:

Received 11 April 2014

Revised 25 September 2014

Accepted 10 October 2014

Available online 22 October 2014

### Keywords:

Serious games

Brain–computer interfaces

Virtual environments

Human–machine interaction

## ABSTRACT

This paper examines the application of commercial and non-invasive electroencephalography (EEG)-based brain–computer (BCIs) interfaces with serious games. Two different EEG-based BCI devices were used to fully control the same serious game. The first device (NeuroSky MindSet) uses only a single dry electrode and requires no calibration. The second device (Emotiv EPOC) uses 14 wet sensors requiring additional training of a classifier. User testing was performed on both devices with sixty-two participants measuring the player experience as well as key aspects of serious games, primarily learnability, satisfaction, performance and effort. Recorded feedback indicates that the current state of BCIs can be used in the future as alternative game interfaces after familiarisation and in some cases calibration. Comparative analysis showed significant differences between the two devices. The first device provides more satisfaction to the players whereas the second device is more effective in terms of adaptation and interaction with the serious game.

© 2014 Elsevier B.V. All rights reserved.

## 1. Introduction

The past decade has seen a huge proliferation of commercial interaction devices for video games. Each of these new devices offers a diverse way of interacting with games and computer generated simulations. Typical technologies that these devices use include: optical, auditory, magnetic and inertia sensors. Some can operate as autonomous controllers while others work in hybrid mode (with standard I/O devices such as mouse and keyboard). However, only the hybrid approaches appear to be functional, the rest require a lot of physical effort. This restricts users' expressive capabilities as well as the information transferred from the user to the computer [1]. Nowadays, non-invasive brain–computer interfaces (BCIs) are getting a lot of attention as alternative human–computer interaction devices for games and virtual environments [2,3].

Non-invasive BCIs operate by recording the brain activity from the scalp with Electroencephalography (EEG) sensors attached to the head on an electrode cap or headset without being surgically implanted. However, they still have a number of problems and they cannot function as accurately as other natural user interfaces (NUIs) and traditional input devices such as the standard keyboard and mouse [4]. The Information Transfer Rate (ITR) of this kind of BCIs is still around 25 bits per minute [5], which makes them much slower compared to traditional input devices such as keyboard (which have typical speed of over 300 characters per minute, roughly 400 bits per minute) [6]. The main reasons behind this are due to bad classification, long training procedures, latency issues and cumbersome hardware [7]. Also, because of lack of training and accessibility in using BCI devices, some people find it difficult to use at all [8].

The majority of BCI studies are performed in laboratory environments under controlled conditions. However this is not always possible in real-life applications and makes current BCI technology not quite suitable for practical applications and widespread use [9]. Game designers and researchers must make sure that BCIs used for gaming environments should not become a barrier in terms of the interaction [10] but on the contrary a more effective medium. Although non-invasive BCI technologies seem to have the potential of providing an environment where “thoughts are not constrained

<sup>☆</sup> This paper has been recommended for acceptance by Christos Gatzidis, Ph.D.

\* Corresponding author at: Human Computer Interaction Laboratory, Masaryk University, Faculty of Informatics, Botanická 554/68a, Brno, Czech Republic. Tel.: +420 549493948.

E-mail addresses: [liarokap@fi.muni.cz](mailto:liarokap@fi.muni.cz) (F. Liarokapis), [K.Debattista@warwick.ac.uk](mailto:K.Debattista@warwick.ac.uk) (K. Debattista), [athanasios.vourvopoulos@m-iti.org](mailto:athanasios.vourvopoulos@m-iti.org) (A. Vourvopoulos), [PPetridis@cad.coventry.ac.uk](mailto:PPetridis@cad.coventry.ac.uk) (P. Petridis), [enea@uni.coventry.ac.uk](mailto:enea@uni.coventry.ac.uk) (A. Ene).

by what is physically possible" [7], they are still not ready for commercial use.

The main aim of the paper is to examine the effectiveness of two different BCI devices for fully controlling an avatar inside a serious game. The objectives of the paper are threefold. Firstly, to enable a user to fully control an avatar in real-time using only EEG data. Secondly, to qualitatively examine the behaviour and different reactions of the users while playing the game and, thirdly, to test each device in terms of: learnability of the interface using the game, satisfaction of the player, performance of the interfaces and effort expended by the player. Two different EEG-based BCI devices were used; one which requires no calibration (NeuroSky MindSet) and another one that requires the training of the classifiers (Emotiv EPOC). The user is visually stimulated by fully controlling an avatar in the serious game (see Section 3). Two different types of EEG-based BCIs were used: the NeuroSky MindSet and the Emotiv EPOC. All tests ( $N = 62$ ) were conducted using the same serious game, which was integrated with the devices; participants were divided equally across the devices.

The rest of the paper is structured as follows. Section 2 provides background information for serious games and BCIs. Section 3, presents the serious game that was used as a case study called Roma Nova. Section 4 demonstrates how the two different BCI devices were used for controlling the same serious game. Finally, Section 5 presents the evaluation results and Section 6 the conclusions and future work.

## 2. Background

Early non-invasive BCI research methods for serious games interaction were usually oriented towards the medical domain rather than entertainment. This kind of research was targeting locked-in patients where haptic and linguistic interfaces fail. The first BCI game was created in 1977. In this game, the user could move in four directions in a maze by fixating on one of four diamond-shaped points periodically flashed. The methodology used was far ahead of its time using online artefact rejection and adaptive classification. The information transfer rate (ITR) was remarkable even for today's standards being above 170 bits/min [11].

Two survey papers regarding BCI systems have been recently published [12,13]. In another recent paper the opportunities and challenges posed by neuroscientific methods when capturing user feedback and using the data to create greater user adaptivity in game are explored [35]. Berta et al. [36] provides an electroencephalogram and physiological signal analysis for assessing flow in games. The paper defines flow in games as a measure of keeping the player fully immersed and engaged in the process of activity within the game. The evaluation of flow involves a 4 electrode EEG, using the low beta frequency bands for discriminating among gaming conditions. Using simple signals from the peripheral nervous system three levels of user states were branded using a Support Vector Machine classifier. The user states were identified using 3 levels of a simple plane battle game identifying states of boredom, flow and anxiety. The paper argues that a personalised system could be implemented in a consumer context allowing for more flowing gameplay in consumer games.

Moreover, there are a number of experimental applications of using BCIs with computer games. Some of these prototypes are very simplistic and just allow the users to select 3D objects in games based on their attention levels [9]. Nowadays, many different techniques are currently used in BCI systems for user interaction and control. Steady State Visual Evoked Potentials (SSVEP) using flashing lights for visual stimulation, the P300 BCI is measuring the brain evoked response after stimulus onset with a positive curve on the EEG after 300 ms and the ERS/ERD

which stands for event related synchronisation/desynchronisation through the imaginary limp movement.

An example of BCI-based input devices using motor-control is the mu ( $\mu$ ) rhythm based first person shooter game proposed by Pineda et al. [14] which uses information from the motor cortices of the brain to steer left/right, while forward/backwards movement is controlled by physical buttons. Another similar BCI system by Krepki [15] uses motor-control based on lateralized readiness potential (LRP) – a slow negative shift in the EEG over the activated motor cortex – for controlling the Pacman game. An example of P300 based games are Bayliss virtual driving task and a virtual apartment [16], [17] with highlighted red objects evoking a P300 when the user wants to make a selection.

Another recent P300 game is Finkes MindGame [18] where the P300 events are translated into movements of a character on a three-dimensional game board. SSVEP based games have been also designed based on subjects' attention to a visual stimulus. In the Mind Balance game [19], a SSVEP is evoked by two different checkerboards with the participant's attention focused in one of the two checkerboards to balance an avatar on a piece of string. An advantage of SSVEP over induced BCI paradigms is the multiple option selection by focusing attention on a stimulus. An example is the 2D racing game using SSVEP to control four different directional movements [20] in a similar way to how the FPS game was controlled in SSVEP BCI [21].

There are a lot of prototypes that use a multimodal approach by combining BCIs with other gaming controllers (i.e. keyboard, mouse, Wii controller, etc) [7]. A typical example of multimodal BCI system is the Bacteria Hunt game where the goal is of 'eating' as many bacteria as possible [22]. The user avatar is controlled using the keyboard whereas the amoeba is modulated by the user's alpha activity (higher alpha results in more control). In the game 'FindingStar', users control the entities of the game using emotional signals coming from the BCI and use the mouse and the keyboard to defeat monsters and solve puzzles [23]. The 'NeuroWander' game uses the emotional and attentional states of the users to perform various quests and the navigation is performed using mouse and keyboard [24].

'Affective Pacman' was developed to investigate the frustration of users while playing a BCI game [25]. The game is controlled with two buttons that rotate Pacman. Frustration is caused by malfunctioning controls of the game. In another study, a steady-state visual evoked potential (SSVEP) based BCI, was used to control an avatar in the game 'World of Warcraft' [26]. To control the avatar, users had to control four icons. Three were used to have the avatar turn left, turn right and move forward and an-other one to perform certain actions such as grasping objects and attacking other objects.

In another study, researchers used BCI technology to interact with mobile games [10]. A maze game with three different levels utilising meditative and attentive states was tested with 22 participants. Results indicated that the use of BCI technology with mobile games has the potential to offer new and exciting ways of game interaction. The MindFlex game is a commercial EEG-based BCI game played as a social activity at home [27]. The aim of the game is to interact with a floating ball around an obstacle course assembled on the game board. While these kinds of BCI techniques for controlling games are quite interesting, most of the games so far are proofs of concept. The interaction with these games is really slow, often with decreased game-play speed to allow for BCI control. This has an obvious impact on the average gamer, resulting in potential frustration and loss of engagement and interest.

## 3. The Roma Nova serious game

The aim of the Rome Reborn project was to create highly realistic 3D representations illustrating the urban development of

ancient Rome from the first settlement in the late Bronze Age (ca. 1000 B.C.) to the depopulation of the city in the early Middle Ages (ca. A.D. 550) [28]. The project includes hundreds of buildings, thirty-two of which are highly detailed monuments reconstructed on the basis of reliable archaeological evidence. The remainder of the 25–30 km<sup>2</sup> model are filled with procedurally-generated buildings based on accurate historical knowledge. Fig. 1 (left image) illustrates the northern side of the Roman Forum with the Curia Iulia in the foreground, next to which is the Basilica Aemilia. In the distance is seen the Temple of the Divine Julius Caesar on the east side of the Forum plaza. The Curia was the principal place where the Senate met.

The interactive game (Fig. 1, right image) is built upon Rome Reborn and it is termed the Roma Nova project. It builds on previous work at Coventry University [29,30] and it is a serious game that aims at teaching history to children (11–14 years old). The game allows for exploratory learning by immersing the learner/player inside a virtual heritage environment where they learn different aspects of history through their interactions with a crowd of virtual Roman avatars. The game was built on the ‘Unity 3D’ game engine using parts of the realistic reconstruction of ancient Rome. For the purpose of this work, both navigation and interaction is performed using brain-wave technology from the two headsets.

Players can interact with the serious game in three different ways according to the Levels of Interaction (LoI) framework [29,30]. LoI aims in simplifying the interactions between the players and automated ‘non-player’ characters. The first level of interaction offers a living background enhancing the player’s experience and immersion. At the second level of interaction characters are automatically provided with a more realistic graphic representation and more complex behaviours. Finally, at the last level of interaction, a conversational agent at the dialogue level appears as a more traditional way to teach the learner [30].

The implementation of the Roma Nova game that was integrated and tested with the two BCI systems includes: (a) a crowd of Roman characters in the Forum and (b) a highly detailed set of buildings that belong to the Rome Reborn model. Intelligent agents were wandering around in the gaming environment between predefined points of interest, whereas the player is able to move freely and is controlled via the BCI devices. To interact with the intelligent agents, the BCI-controlled player needs to approach them. Then, the intelligent agents change their current target to the position of the player, and hence start walking towards the player. When the intelligent agent is close enough to enter the dialogue level, a series of actions are triggered by the engine [30]: (a) an animation is triggered to change the camera from a wide angle to a close-up perspective; (b) the smoothed highly detailed version of the Roman character mesh is loaded to replace the low polygon version, along with the corresponding animations and (c) the steering controller attributed to every background character is dropped and replaced by a straightforward engine developed to play the scenario.

## 4. BCI interactions

This section presents the two different BCI devices that were used (NeuroSky MindSet and the Emotiv), the hardware configuration as well as the methods used for interacting with the serious game.

### 4.1. Hardware configuration

In terms of the hardware configuration, both prototypes used off-the-shelf hardware components. For the first prototype a laptop with a 64 bit Intel(R) Core(TM) 2 Duo processor T6600 at 2.2 GHz and 4 GB of memory was used in conjunction with the NeuroSky MindSet neuro-headset (i.e. one electrode at the FP1 position). The laptop was equipped with an NVIDIA GeForce GT 240 M graphics card. Standard laptop display technology, a 16’ inch wide LCD has been used to display the serious game. The NeuroSky MindSet device is a complete headset with speakers and microphone transmitting data on Bluetooth. It allows extracting the ‘Attention’ and ‘Meditation’ levels of the user called “eSense Meters” [31] through the different EEG frequency bands by a black-boxed NeuroSky algorithm into an integer range of 0–100. The Attention is extracted through the modulation of the frequency band that is triggered by the intensity of a user’s level of mental “focus” when user focuses on a single thought or an external object and decreases when distracted. On the other hand, Meditation levels increase when the user relaxes his/her mind and decreases when he/she is uneasy or stressed. The NeuroSky headset uses a single dry sensor attached to the forehead outside the cerebral cortex in the frontal lobe of the brain being responsible for the attention level and short-term memory tasks [32]. As soon as a connection between the Roma Nova game and the headset is established, it initiates a connection with the avatar and the simulation to send movement instructions.

Similar hardware components were used for the second prototype – with the main component being the Emotiv EPOC neuro-headset. A laptop with a 64-bit Intel(R) Core(TM) 2 Duo processor T7700 at 2.4 GHz and 4 GB of memory was used for the evaluation study. The laptop is equipped with an NVIDIA GeForce 8700 M GT graphics card. The Emotiv headset is a neuro-signal acquisition and processing wireless neuro-headset with 14 wet sensors (and 2 reference sensors) which is capable of detecting brain signals as well user’s facial expressions [33]. This requires a unique user profile to be created and train a classifier for the movement of the four directions based on user’s brain-patterns. The prototype system uses a combination of Cognitive and Facial/Muscular functions. The Emotiv Development Kit was used connecting the Emotiv EPOC headset to the Emotiv control panel to create and train a new user profile. The user is visually stimulated by controlling an avatar in the serious game (Roma Nova). The raw data is calculated on the headset’s chip and sent to the dedicated computer in a 128 Hz frequency. Afterwards the software generates the keystroke events moving the player based on the users’ brain-activity and facial expressions.

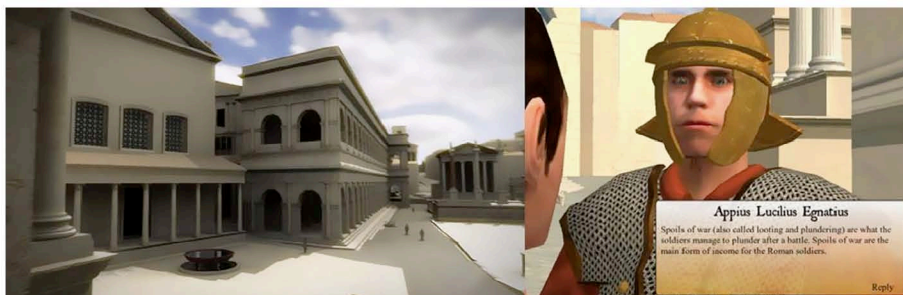


Fig. 1. Left image: Rome Reborn reconstruction [28], Right image: Roma Nova game in action.

#### 4.2. Avatar interaction

Both systems are capable of fully controlling the avatar even if they have a completely different configuration and numbers of sensors. As mentioned before, the NeuroSky MindSet uses only one sensor so that main challenge was to ‘map’ the output of the device to the standard movements of an avatar within a computer game. Moreover, even if no prior training is required to capture data from the device, new users need to familiarise themselves with the device, to control their attention and meditation. As soon as the Roma Nova serious game establishes connection with the headset it initiates the simulation. Both ends of the connection can ‘open’ input and output streams and read/write data. The headset sends two integer values (in the range 0–100) to the game with a frequency of 128 Hz depending on the attention and meditation levels of the user. The meditation component indicates the level of a user’s mental ‘calmness’ or ‘relaxation’. Table 1 illustrates how these values were mapped to represent direction and speed of the avatar.

Players control the avatar by changing cognitive states such as meditation and attention. For example, to go right, they will try to concentrate as hard as possible, while to take a left, users have to defocus their attention. Going straight ahead is possible only by maintaining a balance between the two states. High levels of meditation will prompt the avatar to run, medium levels will cause it to walk, low levels will make it go backwards and extremely low levels of meditation will cause it to stand still Table 1 illustrates the values that were assigned for controlling the direction and the speed of the players.

On the other hand, the Emotiv prototype (Fig. 2) requires prior training to operate the device. Training an effective new user profile takes approximately 30–60 min depending on the adaptability of the user. However, training the profile is not an easy task and requires practise and familiarisation, especially when the user needs to train more than two actions as it is easy to get distracted from outside stimuli and ‘confuse’ the training process of the user’s real ‘intentions’. In a training session of no more than 1 h, the

classifier can be trained adequately for the forward & backward movements by using the Emotiv control panel [34].

New players can gain control over a single action quite quickly. Learning to control multiple actions typically requires practise and becomes progressively harder as additional actions are added. As players learn to train reproducible mental states for each action, the detection becomes increasingly precise. Most players typically achieve their best results after training each action several times. Overtraining can sometimes produce a decrease in accuracy – although this may also indicate a lack of consistency and mental fatigue. Practise and experience will help determine the ideal amount of training required for each individual user to successfully interact with the serious game [34]. As soon as the profile is created, then the combination of Cognitive and Facial/Muscular functions were used to control the avatar in the Roma Nova game. In particular to go forwards or backwards cognitive functions were used. To turn left or right the Facial/Muscular functions were used such as eye blinking.

#### 5. User study

This section presents the procedure followed for the evaluation together with the qualitative and quantitative results of both studies as well as a comparative study.

##### 5.1. Participants and procedure

The testing was performed in open-space laboratory conditions and conference demonstration areas. Sixty-two participants took part in the user study. It was ensured that each participant was comfortable and at ease prior to the start of the experiment. Even if pulse rates or stress levels were not evaluated, each participant was asked to sit-down and a briefing was provided allowing them to be as calm and relaxed as possible.

For the NeuroSky prototype participants were allowed five minutes to accommodate with controlling the avatar and around three minutes to complete the task of arriving at a particular destination. The prototype was used by 31 users, across the span of two months (see Fig. 2). In particular, it was tested at Coventry University computer games laboratory, at the third Phoenix Partner Annual conference (which took place at Coventry University) as well as at the Archeovirtual 2012 International conference (which took place at Paestum, Italy, 15–18 November 2012). The participants were selected via random sampling to generate a non-biased set of results from a section that is more representative of the population. The dominant age group of the sample was 18–25 (74.19% of the total sample), consisting of 60.65% male participants and 39.35% females.

For the Emotiv prototype, 31 users were asked to provide comments on a questionnaire anonymously after playing the serious

**Table 1**  
Controlling the direction and speed of the player (avatar).

Attention (0–100)	Direction
0–40	Turn left
40–75	Go straight ahead
75–100	Turn right
Meditation (0–100)	Speed
0–30	Stand still
30–50	Go backwards
50–70	Walk
70–100	Run



**Fig. 2.** Ability to control events using the Emotiv device.

game. The testing was performed in an open-space environment at two locations at Coventry University: the Serious Games Institute (SGI) and the Department of Computing. Summarising the data samples from the users, descriptive statistics have been generated to quantify the experience. The sample consisted of 67.7% male and 32.3% females with a dominant age group of 18–25. The questionnaire used a Likert scale ranging from 1 to 5.

All users were given a participant information leaflet to read beforehand and a consent form to fill in and sign. The user survey was designed to gather both quantitative and qualitative data. Quantitative data has been collected on questions that require the participants to answer using a Likert scale, whereas qualitative data was gathered through the means of open-ended questions. After going through the gaming scenario, each user was asked to rate their experience taking into consideration learnability of the interface using the game, satisfaction of the player, performance of the interfaces and effort put by the player. These questions were closely followed by open-ended questions that aim to collect data on the overall experience and potentially suggest improvements.

## 5.2. Qualitative results

All participants were asked to provide detailed comments on the questionnaire anonymously. Also an unstructured short interview took place in a light mood after the trial. These comments/suggestions may contribute towards future improvements of BCI.

With respect to the first prototype, participants were drawn by the ease of use of the BCI device for controlling the virtual character (avatar) via EEG technology. They viewed the concept as an interesting approach for future gaming scenarios, categorising the whole experience as challenging, enjoyable and engaging. In particular, most participants were immediately engaged with the game and they wanted not only to perform the task (move from one position to another), but to explore the whole virtual city of the Roma Nova game and interact with the artificial intelligent agents (which were controlled by the computer) to become more knowledgeable about the history of Rome.

Even though the game-play was generally seen as ‘fun and fascinating’, some were dissatisfied with the avatar movement accuracy, and this affected learnability. It is worth-mentioning that some participants were able to recognise some movement types better than others. Turning left and right (switching from a strong state of concentration to a more relaxed state) was reported as the most difficult part to manage. Some reported a lag of around 2–3 s between their intentions and the actual output. While some participants had some difficulty with estimating the degree to which they reached their goal, others were running the test later during the day compared to people playing the game in the morning hours which may have affected their ability to concentrate fully.

Some participants mentioned the need for an initial training, user-profiling period and in-game guide. They found it easy to

use the device but much harder to adapt to it and felt it was hard to ‘train’ their brains to concentrate and meditate. Going backwards was not seen as a popular movement type amongst players and its removal was advocated. Seeing relevant feedback on the screen, apart from the actual character movement, was suggested as being highly beneficial to the experience. Also, they wanted the interface to show the actual measured data, to get real-time feedback and know what to do to attempt self-regulating the attention/meditation levels. A couple proposed the introduction of hotkeys that should facilitate what attention can measure and last but not least, the implementation of the concept using a different game was advised.

For the second prototype a very useful suggestion was made about the Graphical User Interface (GUI) of the game. Participants found it easier to focus on the GUI components instead of the virtual character to perform the required action. This might be a result of the training trial, in which the users had to push/pull a virtual cube and when entering the virtual environment they had to re-adapt to the new elements. This can be confirmed by reports that it was easier to move the avatar by concentrating on the cube from the training trial, than actually visualising the character movement through space.

In contradiction to NeuroSky, navigating into the game was much easier. It is a clear indication that it would be better for the training trial to include the components from the game for the user to get familiarised. Alternatively, assistive GUI components might be a useful addition. Overall the experience was reported as quite engaging and interesting regardless of certain issues of response time and accuracy that other Natural User Interfaces (NUI's) might have. As soon as the participants had an ‘effective’ trained profile, then the level interaction was very satisfactory thus improving the learning process in respect to the serious game. Furthermore, it was reported that people with more experience in computer games will have an easier time learning to use the interface due to the simulation and interaction required for a computer game. That experience makes it much easier to learn how to operate the interface.

In terms of possible improvements participants did not feel very confident with the initial training of the device to create a personalised user-profile. Some found it hard to train, especially when using the cognitive functions. This was expected since the Emotiv device requires training but it varies on the effectiveness of each person. Another issue that was reported had to do with the contact of the sensors in relation to the user's heads. Moreover, some participants reported that after a few minutes of interaction they were getting tired. The tiredness increased after more than 30 min of usage. Some participants proposed the introduction of hotkeys that should facilitate what attention and meditation can measure and last but not least, the implementation of the concept using a different game was advised. Additional recommendations comprise of the incorporation of more sensors and maybe eye tracking technology in order to enhance movement accuracy (see Fig. 3).

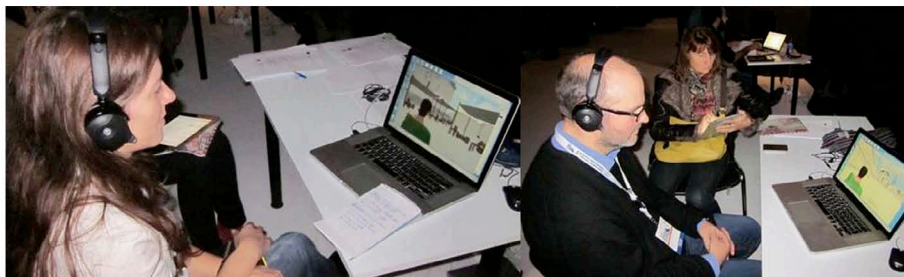


Fig. 3. User testing at archeovirtual, Paestum, Italy 2012.

5.3. Quantitative results

In this section the results of testing each device in terms of: learnability of the interface using the game (Learnability), satisfaction of the player (Satisfaction), performance of the interfaces (Performance) and effort required by the player (Effort) are presented. Table 2 presents an overview of all the relevant statistics which are presented in further detail below.

**Table 2**  
Relevant statistics for the NeuroSky and Emotiv based on questionnaire response.

	Learnability	Satisfaction	Performance	Effort
Mann–Whitney <i>U</i>	185.0	207.5	211.5	259.5
<i>z</i>	−4.376	−4.046	−3.957	−3.271
<i>P</i>	<0.001	<0.001	<0.001	<0.01
Mean NeuroSky	2.5161	4.4516	2.4516	3.8065
Mean Emotiv	3.6774	3.4516	3.5806	3.1290

Fig. 4 shows distribution of responses for Learnability. For Learnability, Kolmogorov–Smirnov was significant for both groups  $D(31) = 0.25, p < 0.001$ , for NeuroSky and  $D(31) = 0.21, p < 0.001$  for Emotiv indicating a non-normal distribution. Due to the lack of normal distribution and the relatively low number of samples ( $N = 62$ ) the non-parametric Mann–Whitney test was used to test for significance in this case and those below (Kolmogorov–Smirnov is reported for all cases). For Learnability the NeuroSky data (mean = 2.561) was considered significantly different from Emotiv data (mean = 3.667)  $U = 185, z = -4.376, p < 0.001$ . Emotiv is considered to be performing better in this case by a significant amount.

Fig. 5 shows distribution of responses for Satisfaction. For Satisfaction, Kolmogorov–Smirnov was again significant for both groups  $D(31) = 0.25, p < 0.001$  for NeuroSky and  $D(31) = 0.25, p < 0.001$  for Emotiv indicating the distribution of this variable was not normal. For Satisfaction the NeuroSky data (mean = 4.451) was considered significantly different from Emotiv data (mean =

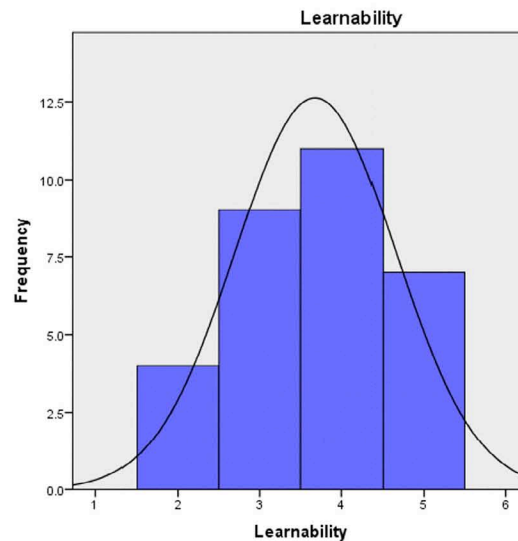
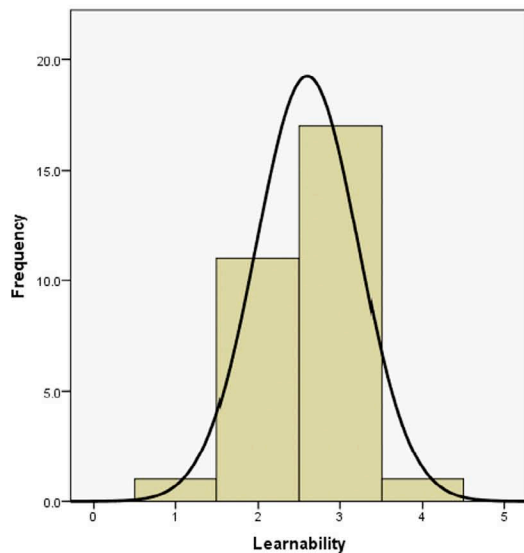


Fig. 4. Learnability of the interface using the game (left NeuroSky, right Emotiv).

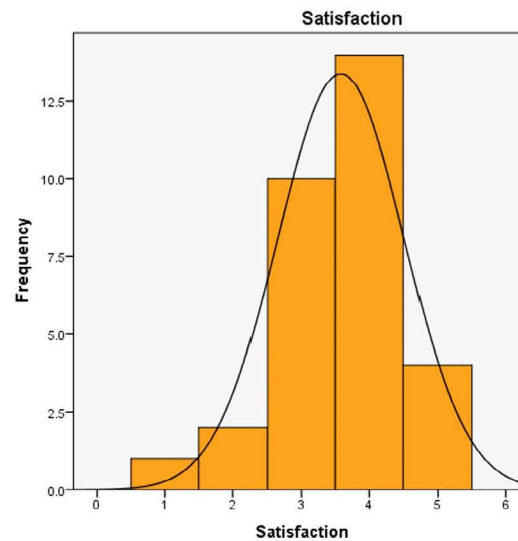
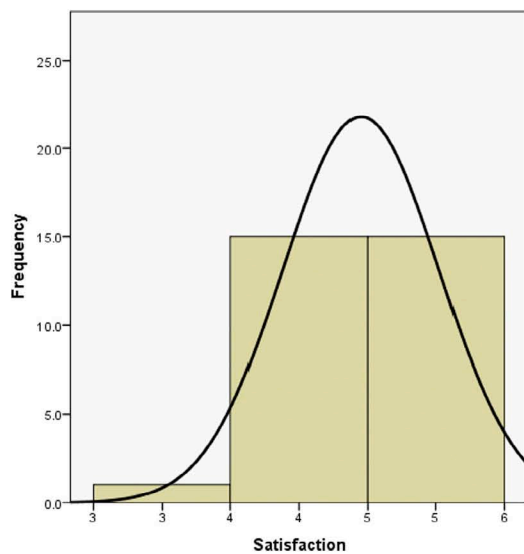


Fig. 5. Player satisfaction (left NeuroSky, right Emotiv).

3.451)  $U = 207.5$ ,  $z = -4.046$ ,  $p < 0.001$ . For Satisfaction the NeuroSky performs better.

Fig. 6 demonstrates the distribution of responses for Performance. For Performance, Kolmogorov–Smirnov was significant  $D(31) = 0.23$ ,  $p < 0.05$  for NeuroSky and  $D(31) = 0.26$ ,  $p < 0.001$  for Emotiv, indicating the data did not belong to a normal distribution. For Performance the NeuroSky data (mean = 2.451) was considered significantly different from Emotiv data (mean = 3.580)  $U = 211.5$ ,  $z = -3.957$ ,  $p < 0.001$ . For Performance, these results indicate that the Emotiv performs significantly better than the NeuroSky.

Fig. 7 illustrates the distribution of responses for Effort. For Effort, Kolmogorov–Smirnov was significant  $D(31) = 0.30$ ,  $p < 0.001$  for NeuroSky and  $D(31) = 0.27$ ,  $p < 0.001$  for Emotiv, indicating the data was not normally distributed. For Effort the NeuroSky data (mean = 3.807) was considered significantly different from Emotiv data (mean = 3.129)  $U = 259.5$ ,  $z = -3.271$ ,  $p < 0.01$ . For Effort, the participants considered the NeuroSky to require significantly more effort in the overall.

## 6. Conclusions and future work

This paper presented two different ways of fully interacting with the same serious game using non-invasive BCIs. Two different EEG-based BCI devices were used, one single dry electrode which requires no calibration and a second that needs some calibration classifier training in order to create a user profile using also 14 wet electrodes. Overall the results indicate that both BCI technologies offer the potential of being used as alternative game interfaces prior to some familiarisation with the device and in some cases some sort of calibration. It was clear that the Emotiv device was better for gaining knowledge about the History of Rome (which is the goal of the serious game) and thus preferred for serious games.

Specifically, as far as the qualitative feedback is concerned, both categories of participants enjoyed the interaction experience. They were all in favour of using EEG technology for controlling and interacting with games even if they stated that the technology is

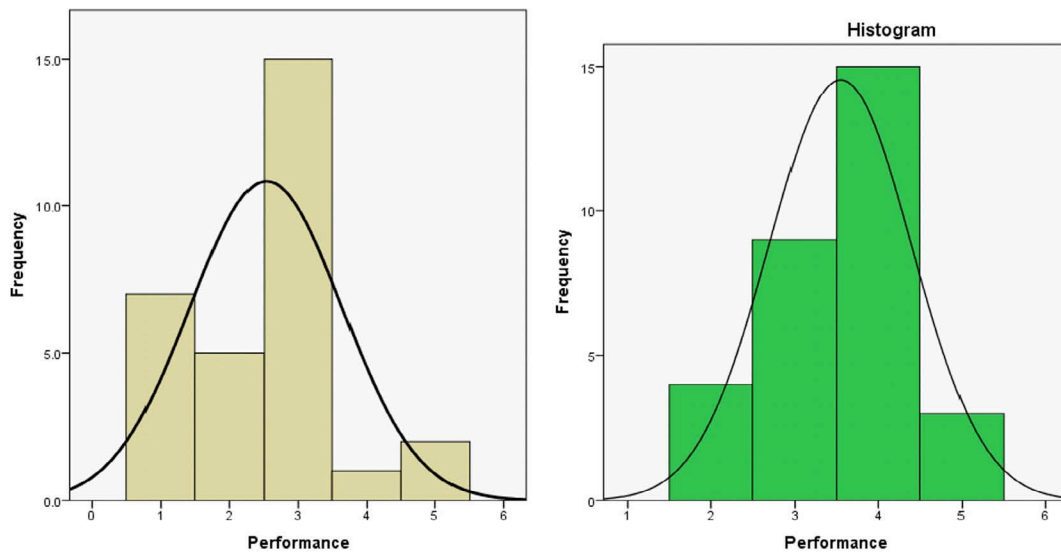


Fig. 6. Performance of the interface (left NeuroSky, right Emotiv).

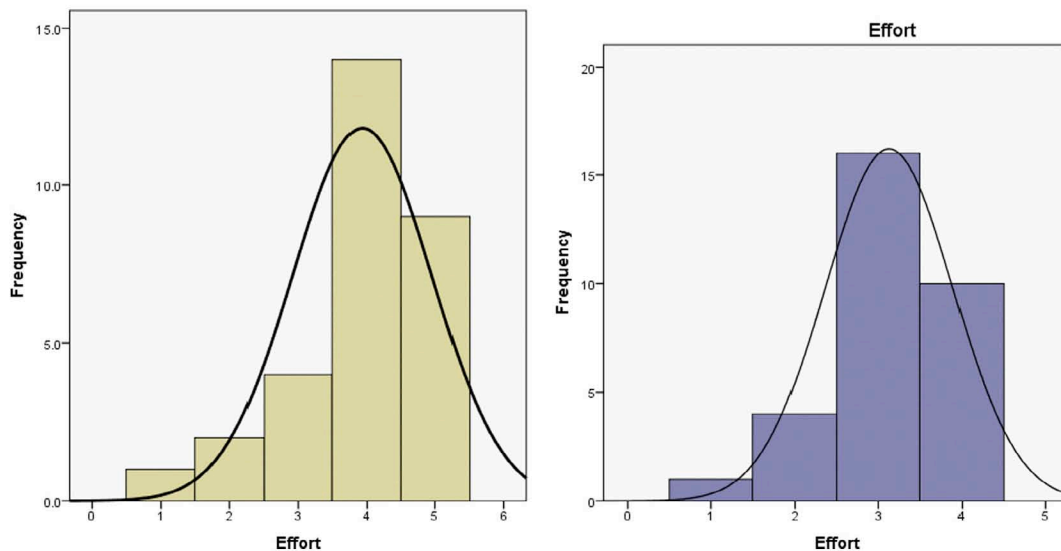


Fig. 7. Effort put by the player (left NeuroSky, right Emotiv).

not as accurate as alternative interaction devices (i.e. keyboard, mouse, joystick, etc.). NeuroSky users found it easier to use the device but much harder to adapt to it. They felt it was hard to 'train' their brains to concentrate and meditate at their first attempt. However, after some self-training it is possible to improve performance considerably. This varies according to the user and more studies need to be performed on this issue. On the contrary, the Emotiv device was easier to perform the training procedure and control the avatar more accurately, but takes a lot more time due to the increased complexity. Also, setup (which includes creating a user profile) is much more complicated compared to the NeuroSky.

In terms of the quantitative analysis, the Emotiv device proved to be more effective for controlling the avatar into the serious game and for learning on how to use the interaction device, whereas the NeuroSky device performed better in terms of satisfaction of the player. There were significant differences between the two devices for the four aspects of the questionnaire reported. Controlling the character using the Emotiv was considered less effort as expected due to the increased sensitivity of the Emotiv compared to the NeuroSky. The performance was also better for the Emotiv for similar reasons. The NeuroSky was considered more satisfying perhaps owing to the immediate use of it due to the lack of setup as compared to the Emotiv. Learnability was also considered better for Emotiv likely due to the same reasons as Performance and Effort.

Moving towards to out-of-the-box BCI Games, training and adaptation is required for both the user and the machine for a successful interaction. It is an iterative learning process in which two entities (humans and computers) are adapting in a closed loop with the serious game for optimising the following: (a) interaction experience from the human side and (b) predefined desired result on the computer side. We can construct this as three different parts: (a) the user, (b) the machine where the data acquisition and translation is taking place, and (c) game/virtual environment where actions are visualised and adapted to the internal elements of the game based on the inputs, closing the loop. From one side, while the user learns to operate the interface the game is adapted based on user performance and proficiency, on the other hand the translated electrophysiological data must be used in a meaningful way, minimising random actions that could trigger false feedback back to the user. Over time, the skill is honed; hence, the feedback must reflect the user's task in an appropriate way. As it was mentioned before, all current brain controlled games are proof of concepts and they do not satisfy the requirements to be treated equally as any other games. In order to be able to create a serious game controlled with such a unique interface, we have to create mechanisms that will ensure a successful interaction. This can be achieved through game design principles modified for brain-control, adaptive gameplay based on user performance, and mechanisms for engaging the participants to the process.

As a future step, further prototype developments could also include an analysis into how certain audio tracks can stimulate concentration/attention and inherently affect game-play. Additional recommendations comprise of the incorporation of more sensors and maybe eye tracking technology to enhance movement accuracy. More sophisticated non-invasive BCI devices equipped with more electrodes and sensors will be also used followed by another comparative study. This will allow for a further analysis based on the combination of different readings.

## Acknowledgements

The authors would like to thank the Human Computer Interaction Laboratory at Masaryk University, the WMG at the Warwick University as well as the Serious Games Institute (SGI) at Coventry

University members for their support and inspiration. Two videos that demonstrate both systems in operation can be found online at: <<http://www.youtube.com/watch?v=L6t4j5yu7k>> and <[http://www.youtube.com/watch?v=5Y\\_cJGGoO4Y](http://www.youtube.com/watch?v=5Y_cJGGoO4Y)>.

## References

- [1] H. Gürkök, G. Hakvoort, M. Poel, Modality switching and performance in a thought and speech controlled computer game, in: Proceedings of the 13th International Conference on Multimodal Interfaces, ACM Press, 2011, pp. 41–48.
- [2] A. Lécuyer, F. Lotte, R.B. Reilly, R. Leeb, M. Hirose, M. Slater, Brain-computer interfaces, virtual reality and video games, *IEEE Comput.* 41 (10) (2008) 66–72.
- [3] A. Nijholt, B. Reuderink, D. Plass-Oude Bos, Turning shortcomings into challenges: brain-computer interfaces for games, *Entertainment Comput.* 1 (2) (2009) 85–94.
- [4] D.P.O. Bos, H. Gürkök, B. Reuderink, M. Poel, Improving BCI performance after classification, in: Proceedings of the 14th ACM International Conference on Multimodal Interaction, ACM Press, 2012, pp. 587–594.
- [5] J.R. Wolpaw, N. Birbaumer, D.J. McFarland, G. Pfurtscheller, T.M. Vaughan, Brain-computer interfaces for communication and control, *Clin. Neurophysiol.* 113 (6) (2002) 767–791.
- [6] T.M. Cover, J.A. Thomas, *Joy Elements of Information Theory*, second ed., Springer, Berlin, 2006.
- [7] D.P.O. Bos, B. Reuderink, et al., Human-computer interaction for BCI games: usability and user experience, in: Proceedings of the 2010 Int'l Conference on Cyberworlds, Springer, 2010, pp. 277–281.
- [8] A. Nijholt, D. Tan, Brain-computer interfacing for intelligent systems, *IEEE Intell. Syst. IEEE Comput. Soc.* 23 (3) (2008) 72–79.
- [9] F. Lotte, Brain-computer interfaces for 3D games: hype or hope?, in: Proceedings of the 6th Int'l Conference on Foundations of Digital Games, ACM Press, 2011, pp. 325–327.
- [10] P. Coulton, C.G. Wylie, W. Bamford, Brain interaction for mobile games, in: Proceedings of the 15th Int'l Academic MindTrek Conference. Envisioning Future Media Environments, ACM Press, 2011, pp. 37–44.
- [11] J.J. Vidal, Real-time detection of brain events in EEG, *Proc. IEEE* 65 (5) (1977) 633–641.
- [12] L.F. Nicolas-Alonso, J. Gomez-Gil, Brain computer interfaces, a review, *Sensors* 12 (2) (2012) 1211–1279.
- [13] L. Bi, X.A. Fan, Y. Liu, EEG-based brain-controlled mobile robots: a survey, *IEEE Trans. Hum. Mach. Syst. IEEE Comput. Soc.* 43 (2) (2013) 161–176.
- [14] J.A. Pineda, D.S. Silverman, A. Vankov, J. Hestenes, Learning to control brain rhythms: making a brain-computer interface possible, *IEEE Trans. Neural Syst. Rehabil. Eng.* 11 (2) (2003) 181–184.
- [15] R. Krepki, B. Blankertz, G. Curio, K. Müller, The Berlin brain-computer interface (BBCI) – towards a new communication channel for online control in gaming applications, *Multimedia Tools Appl.* 33 (1) (2007) 73–90.
- [16] J.D. Bayliss, Use of the evoked potential P3 component for control in a virtual apartment, *IEEE Trans. Neural Syst. Rehabil. Eng.* 11 (2) (2003) 113–116.
- [17] J.D. Bayliss, S.A. Inverso, A. Tentler, Changing the P300 brain computer interface, *Cyberpsychol. Behav.* 7 (6) (2004) 694–704.
- [18] A. Finke, A. Lenhardt, H. Ritter, The MindGame: a P300-based brain-computer interface game, *Neural Netw.* 22 (9) (2009) 1329–1333.
- [19] E.C. Lalor, S.P. Kelly, C. Finucane, R. Burke, R. Smith, R.B. Reilly, G. McDarby, Steady-state VEP-based brain-computer interface control in an immersive 3D gaming environment, *EURASIP J. Appl. Sign. Process.* (2005) 3156–3164.
- [20] P. Martinez, H. Bakardjian, A. Cichocki, Fully online multicommand brain-computer interface with visual neurofeedback using SSVEP paradigm, in: Computational Intelligence and Neuroscience, Hindawi Publishing Corporation, 2007 (Article ID: 94561, 9 pages).
- [21] M.M. Jackson, R. Mappus, E. Barba, S. Hussein, G. Venkatesh, C. Shastry, A. Israeli (Eds.), 2009. Continuous Control Paradigms for Direct Brain Interfaces, Human-Computer Interaction. Novel Interaction Methods and Techniques, Lecture Notes Comput. Sci., vols. 5611, Springer-Verlag, pp. 588–595.
- [22] H. Hwang, K. Kwon, C.H. Im, Neurofeedback-based motor imagery training for brain-computer interface (BCI), *Neurosci. Methods Elsevier* 179 (1) (2009) 150–156.
- [23] M. Ko, K. Bae, G. Oh, T. Ryu, A study on new gameplay based on brain computer interface, in: Proceedings of DiGRA 2009, Brunel University, London, UK, 2009.
- [24] M.S. Yoh, J. Kwon, S. Kim, NeuroWander: a BCI game in the form of interactive fairy tale, in: Proceedings of the 12th Int'l Conference Adjunct Papers on Ubiquitous Computing, ACM Press, 2010, pp. 389–390.
- [25] B. Reuderink, A. Nijholt, M. Poel, Affective Pacman: a frustrating game for brain-computer interface experiments, in: Proceedings of INETAIN 2009, Springer, 2009, pp. 221–227.
- [26] C. Kapeller, C. Hintermüller, C. Guger, Augmented control of an avatar using an SSVEP based BCI, in: Proceedings of the 3rd Augmented Human Int'l Conference, Article No. 27, ACM Press, 2012.
- [27] K. O'Hara, A. Sellen, R. Harper, Embodiment in brain-computer interaction, in: Proceedings of the Int'l Conference on Human Factors in Computing Systems (CHI'11), ACM Press, Vancouver, Canada, 2011, pp. 353–362.
- [28] Rome Reborn, Available at: <<http://www.romereborn.virginia.edu/>>, 2013 (accessed 17.03.2013).



- [29] D. Panzoli, A. Qureshi, et al., Levels of interaction (LoI): a model for scaffolding learner engagement in an immersive environment, intelligent tutoring systems, *Lect. Notes Comput. Sci.* 6095 (2010) (2010) 393–395.
- [30] D. Panzoli, C. Peters, et al., A level of interaction framework for exploratory learning with characters in virtual environments, in: *Intelligent Computer Graphics*, in: *Studies in Computational Intelligence*, vol. 321, Springer-Verlag, Berlin, Heidelberg, 2010, pp. 123–143.
- [31] NeuroSky's eSense™ Meters and Detection of Mental State, Available at: <[http://company.neurosky.com/files/neurosky\\_esense\\_whitepaper.pdf](http://company.neurosky.com/files/neurosky_esense_whitepaper.pdf)>, 2013 (accessed 17.03.2013).
- [32] A. Vourvopoulos, F. Liarokapis, Robot navigation using brain-computer interfaces, in: *Proceedings of the 11th Int'l Conference on Ubiquitous Computing and Communications (IUCC-2012)*, IEEE Comput. Soc., Liverpool, UK, 2012, pp. 1785–1792.
- [33] Emotiv EPOC Software Development Kit, Available at: <<http://www.emotiv.com/store/hardware/299/>>, 2013. (accessed 17.03.2013).
- [34] A. Vourvopoulos, F. Liarokapis, P. Petridis, Brain-controlled serious games for cultural heritage, in: *Proceedings of the 18th Int'l Conference on Virtual Systems and Multimedia Virtual Systems in the Information Society*, IEEE Comput. Soc., Milan, Italy, 2012, pp. 291–298.
- [35] M. Ninaus, S.E. Kober, et al., Neurophysiological methods for monitoring brain activity in serious games and virtual environments: a review, *Int. J. Technol. Enhanced Learning* 6 (1) (2014) 78–103.
- [36] R. Berta, F. Bellotti, et al., Electroencephalogram and physiological signal analysis for assessing flow in games, *IEEE Trans. Comput. Intell. AI Games* 5 (2) (2013) 164–175.

## **8.12 Paper #12**

Sylaiou, S, Liarokapis, F., Kotsakis, K., Patias, P. Virtual museums, a survey and some issues for consideration, *Journal of Cultural Heritage*, Elsevier, 10(4): 520-528, 2009.

Contribution (30%): Collaboration on the collection of the material and write-up of the paper.

Review

# Virtual museums, a survey and some issues for consideration

Sylaiou Styliani<sup>a,\*</sup>, Liarokapis Fotis<sup>b</sup>, Kotsakis Kostas<sup>a,c</sup>, Patias Petros<sup>a,d</sup>

<sup>a</sup> *Inter-departmental Postgraduate Program of School of Technology 'Protection, Conservation and Restoration of Cultural Monuments', Aristotle University of Thessaloniki, Thessaloniki 54124, Greece*

<sup>b</sup> *Interactive Worlds Applied Research Group, Coventry University, Coventry, CV1 5FB, United Kingdom*

<sup>c</sup> *Department of History and Archaeology, Aristotle University of Thessaloniki, Thessaloniki, Greece*

<sup>d</sup> *Department of Rural & Surveying Engineering, Aristotle University of Thessaloniki, Thessaloniki, Greece*

Received 12 November 2007; accepted 23 March 2009

---

## Abstract

Museums are interested in the digitizing of their collections not only for the sake of preserving the cultural heritage, but to also make the information content accessible to the wider public in a manner that is attractive. Emerging technologies, such as VR, AR and Web3D are widely used to create virtual museum exhibitions both in a museum environment through informative kiosks and on the World Wide Web. This paper surveys the field, and while it explores the various kinds of virtual museums in existence, it discusses the advantages and limitation involved with a presentation of old and new methods and of the tools used for their creation.

© 2009 Elsevier Masson SAS. All rights reserved.

*Keywords:* Virtual museums; E-Heritage; Cultural informatics; Virtual reality; Augmented reality; Haptics

---

## 1. Introduction

Silverstone states that 'museums are in many respects like other contemporary media. They entertain and inform; they tell stories and construct arguments; they aim to please and to educate; they define, consciously or unconsciously; effectively or ineffectively, an agenda; they translate the otherwise unfamiliar and inaccessible into the familiar and accessible' [1, p. 162]. An extensive research work [2,3] and a survey of the European museum sector [4] have shown that information technologies such as the World Wide Web (WWW) enhanced by three-dimensional visualization tools can provide valuable help to achieve the aims mentioned above. Furthermore, their use by a wide range of cultural institutions, such as museums, has become easier due to an ever-increasing development of interactive techniques and of new information technology software and hardware, accompanied by a decrease in cost. Information technologies provide solutions to issues of space limitation, of the considerable exhibitions cost and of curator's concerns concerning the fragility of some museum artefacts. Conferences

such as the ICHIM Conferences on Hypermedia and Interactivity in Museums<sup>1</sup> started in 1991 and Museums and the Web,<sup>2</sup> established in 1997, highlight the importance of introducing new technologies in museums. The utility and the potential benefits for museums of emerging technologies such as Virtual Reality (VR) [5–7], Augmented Reality (AR) [8–10] and Web technologies [11,12] have been well documented by a number of researchers [13].

In the 1980s, museums influenced by the New Museology and began to change the way they conveyed the context information of the exhibits to the wider public. There was a shift in the museology concept towards considering that the context of a cultural artefact was more important than the item itself [14–17]. By means of innovative methods and tools and by taking advantage of the WWW potential as an information source, virtual museums were created. They have made the content and context of museum collections more accessible and attractive to the wide public and have enriched the museum experience. There is no official figure yet for the number of virtual museums presently existing worldwide but we know that there are thou-

---

\* Corresponding author. Tel.: +30 2310 996407; fax: +30 2310 994207.  
E-mail address: sylaiou@photo.topo.auth.gr (S. Styliani).

<sup>1</sup> Available at: <http://www.archimuse.com/conferences/ichim.html>.

<sup>2</sup> Available at: <http://www.archimuse.com/conferences/mw.html>.

sands of them and that their number is rapidly on the increase [18]. This article will present the results of a survey on the current state-of-the-art in virtual museums. The purpose behind this is threefold:

- (a) to review the various types and forms that a virtual museum can have and the characteristics of these;
- (b) to present an analysis of their advantages and to highlight their potential;
- (c) to present an overview of emerging technologies used by virtual museums.

## 2. Types of virtual museums

The idea of the virtual museum was first introduced by André Malraux in 1947. He put forward the concept of an imaginary museum (*le musée imaginaire*) [19], a museum without walls, location or spatial boundaries, like a virtual museum, with its content and information surrounding the objects, might be made accessible across the planet. A virtual museum is:

*“a collection of digitally recorded images, sound files, text documents and other data of historical, scientific, or cultural interest that are accessed through electronic media”* [20].

With no standard definition prevailing for the term ‘virtual museum’, the definition adopted for the purpose of this article describes it as:

*“(.. .) a logically related collection of digital objects composed in a variety of media, and, because of its capacity to provide connectedness and various points of access, it lends itself to transcending traditional methods of communicating and interacting with the visitors being flexible toward their needs and interests; it has no real place or space, its objects and the related information can be disseminated all over the world”* [21].

Another less rigid definition states that a virtual museum can be a digital collection that is presented either over the Web, or to an intranet, either via a personal computer (PC), an informative kiosk, a personal digital assistant (PDA), or even to a CD-ROM as an extension of a physical museum, or that it can be completely imaginary. Furthermore, the abstract term virtual museum can take various forms depending on the application scenario and end-user. It can be a 3D reconstruction of the physical museum [22]. Alternatively, it can be a completely imaginary environment, in the form of various rooms, in which the cultural artifacts are placed [23].

According to ICOM [24], there are three categories of virtual museums on the Internet that are developed as extensions of physical museums: the brochure museum, the content museum and the learning museum. The brochure museum aims at informing future visitors about the museum and is mainly used as a marketing tool, with basic information such as location, opening hours and sometimes a calendar of events etc. [25,26], in order to create motivation to visit the walled museum. The content museum is a website created with the purpose of making information about the museum collections available. It can

be identified to a database containing detailed information about the museum collections, with the content presented in an object-oriented way. The learning museum is a website, which offers different points of access to its virtual visitors, depending on their age, background and knowledge. The information is presented in a context-oriented, rather than object-oriented way. Moreover, the site is educationally enhanced and linked to additional information intended to motivate the virtual visitor to learn more about a subject of particular interest to them and to visit the site again. The goal of the learning museum is to make the virtual visitor come back and to make him/her establish a personal relationship with the online collection.

## 3. Emerging tools and technologies used by virtual museums

Technological advances that have emerged as areas of crucial interest are making it possible to use sophisticated tools to provide customized interfaces for the generation of virtual museums, to design a virtual museum exhibition in a number of ways [27,58] and to get used as conveyors of information for knowledge construction, acquisition and integration. New types of interfaces, interaction techniques and tracking devices are developing at a rapid pace and can be integrated into multimodal interactive VR and AR interfaces [9]. The first studies in the field were mainly focused on static presentations of texts and photos concerning a museum. Later on, the exhibits tended to be more dynamic and interactive rather than static in nature and authoritative [28,27], thus creating an approach which was closer to reality and enhancing the experience for virtual visitors. Usually, the structure of most virtual exhibitions is defined by the structure of exhibition spaces [11] that consist of two types of elements: the Virtual Galleries and the Cultural Objects. Exhibits are the principal means through which museums communicate their mission objectives and they can be static or interactive. According to research the key features of an online interactive exhibit are:

- (a) multiplicity of contexts for the user to connect with the exhibit in a seamless manner;
- (b) good instructional design;
- (c) pro-active learning contexts;
- (d) good balance between learning and leisure;
- (e) no text-heavy pages to interfere with the learning experience [29].

In this section, a brief overview of the most characteristic methods and tools currently used for the generation of virtual museum exhibitions and their exhibits are presented.

### 3.1. Imaging technology

Virtual museums need high-resolution images in order to provide as much information as possible about the virtual exhibits. However, the level-of-detail (LOD) is dependent on the resolution of the digital images and high-resolution conventional images produce very large files that are difficult to manage and

to transmit across networks because of their dependence on bandwidth availability (slow Internet connections). A strategy adopted to confront this problem is the image servers that use a “Russian doll” imaging architecture and give the user scalability and interactivity opportunities, because multiple resolutions of an image are stored in a single file and make it possible to progressively transmit an image. FlashPix and then JPEG2000 are the two image formats that introduced a new concept for imaging architecture [30]. Metadata storing is also allowed. This image format is used by various museums such [31–34]. Some of the FlashPix features are adopted by the JPEG2000 image format that also has the potential of progressive image transmission and scalability and some new features that fill the gaps for the inclusion of metadata and the protection of the content [35] of earlier standards for encoding digital media. The advantages of the image format have been extensively investigated in research work [36,37] and the JPEG2000 format has been adopted by cultural institutions [38–40].

### 3.2. Web3D exhibitions

Internet technologies have the tremendous potential of offering virtual visitors ubiquitous access via the WWW to a virtual museum environment. Additionally, the increased efficiency of Internet connections (i.e. ADSL) makes it possible to transmit significant media files relating to the artefacts of virtual museum exhibitions. The most popular technology for the WWW visualisation includes Web3D which offers tools such as VRML and X3D, which can be used for the creation of an interactive virtual museum. The Web3D consortium [41] contains open standards for real-time 3D communication and the most important standards include: VRML97 and X3D and are presented below. Many museum applications based on VRML have been developed for the web [12,42]. As from 4 April 1997, VRML97 has stood for Virtual Reality Modeling Language. Technically speaking, VRML is neither VR, nor a modelling language, but a 3D interchange format which defines most of the commonly used semantics found in today’s 3D applications such as hierarchical transformations, light sources, viewpoints, geometry, animation, fog, material properties, and texture mapping. Another definition states that VRML serves as a simple, multi-platform language for publishing 3D Web pages as well as for providing the necessary technology to integrate three dimensions, two dimensions, text, and multimedia into a coherent model. “When these media types are combined with scripting languages and Internet capabilities, an entirely new genre of interactive applications is possible” [43]. This is due to the fact that some information is best experienced in three-dimensional form, such as the information of virtual museums [11,9]. However, VRML can be excessively labour-intensive, time consuming and expensive. QuickTime VR (QTVR) and panoramas that allow animation and provide dynamic and continuous 360° views might represent an alternative solution for museums such as in [44]. As with VRML, the image allows panning and high-quality zooming. Furthermore, hotspots that connect the QTVR and panoramas with other files can be added [45]. In contrast, X3D is an Open Standards XML-enabled 3D

file format offering real-time communication of 3D data across all applications and network applications. Although, X3D is sometimes considered as an Application Programming Interface (API) or a file format for geometry interchange, its main characteristic is that it combines both geometry and runtime behavioral descriptions into a single file alone. Moreover, X3D is considered to be the next revision of the VRML97 ISO specification, incorporating the latest advances in commercial graphics hardware features, as well as improvements based on years of feedback from the VRML97 development community. For a virtual museum, making possible the presentation of virtual exhibitions, the visualization usually consists of dynamic Web pages embedded with 3D VRML models [9]. This can be enhanced with other multimedia information (i.e. movie clips, sound) and used remotely over web protocols (i.e. HTTP). A more 3D graphics format, is COLLABorative Design Activity (COLLADA) [46] which defines an open standard XML schema for exchanging digital assets among various graphics software applications that might otherwise store their assets in incompatible formats. One of the main advantages of COLLADA is that it includes more advanced physics functionality such as collision detection and friction (which Web3D does not support).

Moreover, more powerful technologies that have been used in museum environments include OpenSceneGraph (OSG) [47] and a variety of 3D game engines [48,49]. OSG is an open source multi-platform high performance 3D graphics toolkit, used by museums [50,51] to generate more powerful VR applications, especially in terms of immersion and interactivity since it supports text, video, audio and 3D scenes into a single 3D environment. On the other hand, 3D games engines are also very powerful and they provide superior visualization and physics support. Serious games is a new concept and allows for collaborative use of 3D spaces which are used for learning and educational purposes in a number of educational domains. The main strengths of serious gaming applications could be generalised as being in the areas of communication, visual expression of information, collaboration mechanisms, interactivity and entertainment. Both technologies (OSG and 3D game engines) compared to VRML and X3D can provide very realistic and immersive museum environments but they have two main drawbacks. First, they require advanced programming skills in order to design and implement custom applications. Secondly, they do not have support for mobile devices such as PDAs and 3G phones.

### 3.3. Virtual reality exhibitions

VR is a simulation of a real or imaginary environment generated in 3D by digital technologies that is experienced visually and provides the illusion of reality. Over the past few years, modeling software has become affordable and the cost of building virtual environments has fallen considerably, thus fuelling new application domains such as virtual heritage. For example, low cost and highly interactive VR experiences for museum visitors can be created on the basis of standard hardware components (a relatively low cost PC with cheap graphics accelerator, a touch screen and a sensor device, e.g. a inertia cube), some applica-

tion software and suitable browser plug-ins. VR applications can be used by distributed groups of large numbers of players, and are immersive and interactive. In a VR environment participants get immersed into a completely artificial world but there are various types of VR systems, which provide different levels of immersion and interaction. Heim believes that weak VR can be characterized by the appearance of a 3D environment on a 2D screen [52].

In contrast, strong VR is the total sensory immersion, which includes immersion displays, tracking and sensing technologies. Common visualization displays include head-mounted displays and 3D polarizing stereoscopic glasses while inertia and magnetic trackers are the most popular positional and orientation devices. As far as sensing is considered, 3D mouse and gloves can be used to create a feeling of control of an actual space. An example of a high immersion VR environment is Kivotos, a VR environment that uses the CAVE<sup>®</sup> system, in a room of 3 meters by 3 meters, where the walls and the floor act as projection screens and in which visitors take off on a journey thanks to stereoscopic 3D glasses [53]. As mentioned earlier, virtual exhibitions can be visualized in the Web browser in the form of 3D galleries, but they can also be used as a stand-alone interface (i.e. not within the web browser). In addition, a number of commercial VR software tools and libraries exist, such as Cortona [54], which can be used to generate fast and effectively virtual museum environments. However, the cost of creating and storing the content (i.e. 3D galleries) is considerably high for the medium and small sized museums that represent the majority of cultural heritage institutions. An overview of the tools and methods available to visitors visualizing a virtual museum has been already carried out [55].

### 3.4. Augmented reality exhibitions

In addition to the VR exhibitions, museum visitors can enjoy an enhanced experience by visualizing, interacting and navigating into museum collections (i.e. artifacts), or even by creating museum galleries in an AR environment. The virtual visitors can position virtual artifacts anywhere in the real environment by using either sophisticated software methods (i.e. computer vision techniques) or specialized tracking devices (i.e. InertiaCube). Although the AR exhibition is harder to achieve, it offers more advantages to museum visitors as compared to Web3D and VR exhibitions. Specifically, in an AR museum exhibition, virtual information (usually 3D objects but it can also be any type of multimedia information, such as textual or pictorial information) is overlaid upon video frames captured by a camera, giving users an impression that the virtual cultural artifacts actually exist in the real environment. Through human–computer interaction techniques users can examine thoroughly the virtual artifacts through tactile manipulation of fiducials (i.e. markers) or sensor devices (i.e. pinch-gloves). This ‘augmentation’ of the real-world environment can lead to an intuitive access to the museum information and enhance the impact of the museum exhibition on virtual visitors.

One of the earliest examples of an interactive virtual exhibition is an automated tour guide system that uses AR techniques

[56]. It can superimpose meaningful audio on the real world on the basis of the location of the user, offering the advantage of enriching visitors’ experiences. Also, the Meta-Museum guide system [57] is based on AR and artificial intelligence technologies and provides a communication environment between the real world and cyberspace to maximize the utilization of a museum’s archives and knowledge base. Furthermore, AR has been experimentally applied to make it possible to visualise incomplete or broken real objects as they were in their original state by superimposition of the missing parts [10]. Finally, the ARCO system [23,11] provides customised tools for virtual museum environments, ranging from the digitisation of museum collections to the tangible visualization of both museum galleries and artifacts. ARCO developed tangible interfaces that allow museum visitors to visualise virtual museums in Web3D, VR and AR environments sequentially. A major benefit of an AR-based interface resides in the fact that carefully designed applications can themselves provide novel and intuitive interaction without the need for expensive input devices.

### 3.5. Mixed reality exhibitions

Finally, mixed reality (MR) relies on a combination of VR, AR and the real environment. According to Milgram and Kishino’s virtuality-continuum, real world and virtual world, objects are presented together on a single display [58] with visual representation of real and virtual space [59]. An example of the use of MR techniques in a museum environment is the Situating Hybrid Assemblies in Public Environments (SHAPE) project [60] that uses hybrid reality technology to enhance users’ social experience and learning in museum and other exhibition environments, with regard to cultural artifacts and to their related contexts. It proposes the use of a sophisticated device called the periscope (it is now called the Augurscope), which is a portable mixed reality interface, inside museum environments to support visitors interaction and visualisation of artifacts.

### 3.6. Haptics

‘Haptics, from the Greek word ‘haptain’, involves the modality of touch and the sensation of shape and texture which an observer feels when exploring a virtual object’ [61]. Haptics makes it possible to achieve the extension of visual displays to render them more realistic, useful and engaging for visitors. One of the most characteristic museum applications using haptics is at the University of Southern California’s Interactive Art Museum [62]. In this case, the PHANToM device was used within a museum allowing visitors to touch and feel virtual artifacts [63] PHANToM is a desk-grounded robot that allows simulation of single fingertip contact with virtual objects through a pointing device (i.e. stylus). In addition, its actuators communicate forces back to the user’s fingertips as it detects collisions with virtual objects, simulating the sense of touch. Another application is the ‘Museum of Pure Form’ a VR system where users can interact, through the senses of touch and sight, with digital models of 3D art forms and sculptures. Its aim was to change the way normal users perceive sculptures, statues or,

more generally, any type of 3D artwork [64]. Two different presentations of this application were developed including a system placed inside several museums and art galleries around Europe as well as a system placed inside a CAVE<sup>TM</sup> environment [27].

### 3.7. Use of handheld devices in museums

Handheld devices represent a wide range, including cell phones, personal-digital assistants (PDAs) and tabloids. Improvements during the past few years in optics, processing power and ergonomics have initiated a number of museum-based applications. A prototype application is the City co-visiting system which combines VR, hypermedia technology, handheld devices and ultrasound tracking technology to allow three visitors, one on-site and two remote [65]. A location-aware PDA is used for the on-site visitor to display the ongoing positions of all three visitors on a map of the gallery while the two off-site visitors use two different environments: a web-only environment and a VE. The application also supports web-based multimedia information for the off-site visitors that are dynamically presented upon movement across the map. The San Francisco Museum of Modern Art (SFMOMA) has also presented work from its permanent collection in iPAQ handhelds [66]. Furthermore, the giCentre at City University is exploring LBS through the use of mobile computing including the use of third-generation (3G) phones and PDAs [67]. Users can interact with the virtual artifacts using either the menu interface or the stylus. In addition, using external sensors (i.e. inertia cube, accelerometers and digital compass) museum visitors can perceive virtual information about the artifacts in relation to their location inside the museum.

## 4. Real and virtual museum

According to the definition of the International Council of Museums (ICOM) about museums [68]: “A museum is a non-profit making, permanent institution in the service of society and of its development, and open to the public, which acquires, conserves, researches, communicates and exhibits, for purposes of study, education and enjoyment, material evidence of people and their environment.” Virtual museum enjoy the same functions of acquisition, storage, documentation, research, exhibition and communication as the ‘brick and mortar’ museums as set out by the above definition. They can, in addition, act in a complementary and auxiliary manner. A virtual museum website can provide worldwide publicity. Research has revealed that 70% of people visiting a museum website would subsequently be more likely to go and visit the ‘real’ museum [69]. Museum curators can digitally preserve the artifacts of their collections. The effective safeguarding of cultural artifacts can be achieved through the use of technological advances, by means of the comparison of different images across time to monitor their conservation. Furthermore, they provide the means to create digital representations of cultural artifacts and database technologies with which multimedia information about the virtual museum artifacts can be stored and retrieved whenever is needed. The digitized information can be re-used in a variety of

ways, for different purposes and probably even by other cultural institutions.

Additionally, virtual museums allow museum curators to experiment with various arrangements of 3D objects inside the gallery, to test different designs before deciding on the presentation style of a temporary exhibition. They create and disseminate to the wider public virtual models of cultural artifacts that combine archaeological accuracy and reliability with aesthetic pleasure. Finally, they visualize the digital representation of the cultural objects via VR and AR interfaces, so as to make available to the wider audience more realistic and appealing virtual museum exhibitions that can be interactively and easily explored. In addition to this, they can overcome limitations of space in respect of the number of objects accessible in the real museum [70].

The WWW is widely used by museums for putting their collections online [71], not only because it is very popular (especially among young people), but also because it is in the hands of museum curators a powerful communication tool that can deliver in a fast, user-friendly and low-cost information about the museum to potential virtual visitors and provides museum curators with a great variety of opportunities in terms of museum data dissemination. As it has been already mentioned, virtual museums, through innovative technologies, provide unrestricted round-the-clock access to their visitors through the WWW. Virtual museums can provide access from any place and to anyone, including people with special needs (visual, acoustic, speech and motor disabilities and learning difficulties). The UN Convention on the Rights of Persons with Disabilities [72], the Americans with Disabilities Act of 1990 (ADA) [73] and the Disability Discrimination Act (DDA) in the UK state that disabled people have equal rights of ‘access to goods, facilities and services’ [74]. It is therefore the responsibility of Cultural institutions, such as museums to find ways of providing access to the exhibitions to people with disabilities. Digital museums take into account the need emphasized by the Resource Disability Action Plan and formed by the Council of Museums, Archives and Libraries for efficient ways of using new technologies which allow the access to museum exhibitions to all end-user groups including virtual access to disabled people [75] using AR interfaces designed to operate on off-the shelf computer systems [9].

The cultural artifacts that are exhibited in the physical environment of a museum are usually shown in display cases, where only a limited amount of information about them is available. In virtual museum exhibitions, museum artifacts can be digitized and visualized into a virtual interactive environment. A virtual exhibit can contain information that a physical exhibit in a museum showcase cannot. Thus, museum curators are given the opportunity to offer a more rewarding experience thanks to rich multimedia context information data about the objects, in comparison to artifacts that are locked in a museum glass case with a simple description on a card. In these virtual exhibitions, users may explore exhibits in an interactive and more flexible way. Virtual museum exhibitions provide the experience of allowing virtual visitors to observe and examine an object from all angles. AR exhibitions can also involve physical interfaces (i.e. marker-cards), which are used as the link between real

and virtual worlds. Physical interfaces allow museum visitors to pick up and manipulate virtual cultural objects and examine them within the display system in their hands (i.e. flat screen) [9]. Additionally, a virtual museum gives the user the control of the virtual tour, because it may provide 3D views of a museum and a floor plan. Virtual visitors can orient themselves; know in which room of the virtual exhibition the exhibits are found and to which group of the exhibits an object belongs. The exhibits themselves can convey their meaning, when they are examined in conjunction with the other exhibits of the room and through a narrative that connects the objects and their context and ‘brings to life the potential dynamism of objects and their stories’ [76].

The communities targeted by virtual museums are the museum curators and the end-users. The second category can be divided into three subcategories: the specialists, the students and the tourists [77,78]. Virtual museum exhibitions can contain a great amount and depth of information, meant to broaden perspectives, satisfy needs and encourage a deeper understanding of virtual visitors of any of the above profiles. They can fulfil the need for ‘basic and distinguishing information’ of simple tourists [79] and they do not need any additional help to deciphering the concepts and the ideas behind museum objects [80, p. 210]. Virtual museums are also capable of providing information to a degree of detail that is sufficient for various kinds of visitors [72] while it may assist the specialised research needs including the comparative study requirements of specialists and students, by providing access not only to one, but to multiple museum collections. Furthermore, creative websites may attract audiences that ‘would not normally use libraries or museums’ [81] and do not have prior knowledge of or interest in the subject of the museum exhibition [82]. The visitors of virtual museum exhibitions are not passive nor do they lack opportunities to develop their critical skills. A virtual museum can provide visitors with the freedom to explore, to exercise autonomy and to be active participants as they create their own virtual tour and paths. Additionally, the digital tools provided are used as cognitive technologies that help the virtual visitors transcend the limitations of the human mind, such as memory or problem solving limitations [83] and construct their own knowledge. A representative example of the above is the ability provided to virtual museum visitors for creating a personal online exhibition of digitized material, a ‘gallery’ that corresponds to their interests and they can share it with others [33]. In a virtual museum environment, there are more learning opportunities via educational games than in a physical museum [84,85], as cited by [86]. Most of the virtual museums have been designed by taking into account the constructivist principles of learning through construction and learning through play [87,88] and they involve interaction, experiencing and learning at the same time. In a virtual museum environment, the visitor is not an observer but s/he interacts with the learning objects and s/he constructs her/himself the knowledge. Museum visitors use and interact with the virtual museum environment via a constructive dialogue that provide them with access to thematic information and explanations about the museum objects’ context with the level of information and the amount of detail they prefer [89]. Learning is an active process and the end-users are engaged in

hands-on involvement in an engaging experience that enhances the understanding, fosters fruitful learning interactions, awakens and keeps the interest alive and enriches aesthetic sensitivities. Most of the time, virtual visitors do not want to ‘learn something’ but rather to engage in an ‘experience of learning’ or ‘learning for fun’ that can be ‘important and enjoyable in its own right’ [90].

## 5. Problems and implications confronted

New technologies provide new possibilities and impose new restrictions [91]. Despite significant advantages, a virtual museum also presents drawbacks. The forms these take will now be examined. ‘VR’ (an oxymoron) cannot have the complexity of the real objects. Virtual museum comes from Greek *dynaton* (gr. δυνατόν=possible) and it means “that in potential” (Aristotle, *Analitici primi*) and exists in potential form and not in reality [92]. The problem is that advanced graphic systems that are used for computer reconstructions adopted by virtual museums may sometimes be too realistic. They are based on partial evidence, but they suggest an impression of good knowledge of the past. Sometimes advanced graphic systems present the ‘image’ as true, giving the sense of misleading accuracy [93,94]. When the reconstructed item has a lot of missing elements then – obviously – scientists must use their imagination or rely on ethno-historical information on how similar cases might have looked like, in order to reconstruct it. However, in these cases, the result will not be an explanation of the past, but a personal and subjective way of seeing it. A good ‘image’ can give the impression to the viewer that museologists know more than they actually do. Some products of computer reconstructions can be considered as scientifically accurate, because they seem to be accurate. The term “user” is used for virtual museum visitors, because, in order to retrieve information on virtual exhibits, computer skills are required [86]. This means that the computer illiterate are automatically excluded and a lot of visitors encounter difficulties with understanding the use of plug-ins and other applications that need to be downloaded from the Internet and installed in order to retrieve information from sophisticated virtual museum exhibitions.

The idea of the ambiguity between reality and virtuality can be first traced in the Metaphor of the Cave in The Republic of Plato, where people take as real a fact that is an illusion [95]. Prisoners that have been chained and held immobile can only see at a wall in front of them. Behind them, there is a fire and between them and the fire there is a walkway with shadows of moving things and creatures. So, they consider the shadows and the echoes as the only ‘reality’ and the reflections of objects more important than the objects themselves.

When it comes to building virtual reconstructions, even if there is a degree of accuracy, the one-sided view of the reconstructed site is still wrong. Computer reconstructions that offer only one aspect of the subject they examine and do not provide any alternative reconstructions, contradict the fact that there are many ways to examine the Past. In virtual reconstructions there is only one aspect of the subject that has been reconstructed and no alternative reconstructions have been created. Some



high-quality and sophisticated virtual museums involve collaborations between museologists and computer experts. In such cases, communication problems often arise between those with theoretical knowledge in museology and those with practical knowledge of computers. In most circumstances, the software itself used by virtual museums is not accessible to museologists and computer scientists stand between them and the data. In some cases, it is probable that the Past is both misinterpreted and misrepresented. The visualization results are impressive, thus fulfilling a primary goal, more specifically general public consumption, but without, in turn, serving the museum goals. Virtual museums may provide users with fragmented museum-related information that often bear no obvious information with each other or refer to a useful context. In addition to this, some virtual museums suffer from the lack of clearly identified purposes. Their design must be carried out according to their *raison d'être* and the information provided must be organized in order to construct a narrative [96]. A virtual museum has to define its target community/ies, its aims, its content and how this will be structured and delivered. Throughout all the creation phases of the virtual museum, evaluation studies that involve real users must be undertaken, in order to identify the parts of the program that need further improvement [97].

## 6. Conclusions

In this paper, the various types of virtual museums in the light of a range of classifications have been discussed. With the use of imaging technology, Web3D, VR, AR, MR, haptics and handheld devices as PDAs, museums can exploit all possibilities of the new media, analyze and answer in various ways to visitors' needs, enable an intuitive interaction with the displayed content and provide an entertaining and educational experience. The benefits of virtual museums are noteworthy as far as museum curators are concerned and in terms of documentation, conservation, research and exhibition. The virtual museums have the potential to both preserve and disseminate the cultural information in an effectively and at a low-cost through innovative methods and tools. They are an engaging medium with great appeal to a variety of groups of visitors and can promote the 'real sites' by providing information about museum exhibitions and offer an enhanced display of museum artifacts through emerging technologies. Various groups of end-users such as tourists, students and specialists can take advantage of them and satisfy their learning and entertainment needs. The visit of virtual museums can be an enjoyable and productive experience that draws the user into involvement and participation and help the promotion of real museums [98].

The virtual museums enrich the museum experience by allowing an intuitive interaction with the virtual museum artifacts. A comparison between real and virtual museums indicates that there still are important issues for virtual museums to solve. Good collaboration must be ensured between cultural heritage specialists (museum curators, historians, archaeologists, etc.) and information science specialists to achieve optimal results and in order to avoid dependence on market-produced software and to promote open-source software that may be produced with

the aid of cultural heritage specialists. Virtual museums cannot and do not intend to replace the walled museums. They can be characterised as 'digital reflections' of physical museums that do not exist *per se*, but act complementarily to become an extension of physical museums exhibition halls and the ubiquitous vehicle of the ideas, concepts and 'messages' of the real museum. Their primary aim is (or should be) to investigate and propose models for the exploration of the real purpose and conceptual orientation of a museum.

## References

- [1] R. Silverstone, The medium is the museum, in: R. Miles, L. Zavala (Eds.), *Towards the Museum of the Future*, Routledge, London/New York, 1994, pp. 161–176.
- [2] J. Jones, M. Christal, *The Future of Virtual Museums: On-Line, Immersive, 3D Environments*, Created Realities Group, 2002.
- [3] G. Scali, M. Segbert, B. Morganti, Multimedia applications for innovation in cultural heritage, in: *Proceedings of 68th IFLA Council and General Conference*, August 2002, Glasgow, U.K., 2002.
- [4] ORION Report on Scientific/Technological Trends and Platforms, available at: <http://www.orion-net.org>.
- [5] D. Pletinckx, D. Callebaut, A. Killebrew, N. Silberman, Virtual-reality heritage presentation at Ename, *IEEE Multimedia* 7 (2) (2000) 45–48.
- [6] M. Roussou, Immersive interactive virtual reality in the museum, in: *Proceedings of TILe*, June 2001, London, U.K., 2001.
- [7] R. Wolciechowski, K. Walczak, M. White, W. Cellary, Building Virtual and Augmented Reality Museum Exhibitions, in: *Proceedings of the 9th Int. Conference on 3D Web Technology*, California, USA, April 2004, ACM SIGGRAPH, 2004, pp. 135–144.
- [8] A. Brogni, C.A. Avizzano, C. Evangelista, M. Bergamasco, Technological approach for cultural heritage: augmented reality, in: *Proceedings of the RO-MAN 99 Conference*, Pisa, Italy, September 1999, 1999, pp. 206–212.
- [9] F. Liarakapis, S. Sylaïou, A. Basu, N. Mourkoussis, M. White, P.F. Lister, An interactive visualisation interface for virtual museums, in: K. Cain, Y. Chrysanthou, F. Niccolucci, N. Silberman (Eds.), *Proceedings of the VAST 2004 Conference*, Belgium, EPOCH Publication, Belgium, 2004, pp. 47–56.
- [10] F. Liarakapis, M. White, Augmented reality techniques for museum environments, *The Mediterranean Journal of Computers and Networks* 1 (2) (2005) 90–96.
- [11] M. White, N. Mourkoussis, J. Darcy, P. Petridis, F. Liarakapis, P.F. Lister, K. Walczak, R. Wolciechowski, W. Cellary, J. Chmielewski, M. Stawniak, W. Wiza, M. Patel, J. Stevenson, J. Manley, F. Giorgini, P. Sayd, F. Gaspard, ARCO—An Architecture for digitization, management and presentation of virtual exhibitions, in: *Proceedings of the CGI'2004 Conference*, Hersonisos, Crete, June 2004, Los Alamitos, California: IEEE Computer Society, 2004, pp. 622–625.
- [12] P.A.S. Sinclair, K. Martinez, D.E. Millard, M.J. Weal, Augmented reality as an interface to adaptive hypermedia systems. New review of hypermedia and multimedia, *Special Issue on Hypermedia beyond the Desktop* 9 (1) (2003) 117–136.
- [13] P. Patias, Y. Chrysanthou, S. Sylaïou, H. Georgiadis, S. Stylianidis, The development of an e-museum for contemporary arts, in: *Proceedings of the VSMM Conference on Virtual Systems and Multimedia dedicated to Cultural Heritage 2008*, 20–25 October, Nicosia, Cyprus, 2008.
- [14] S.M. Pearce, Thinking about things. approaches to the study of artifacts, *Museum Journal* (1986) 198–201.
- [15] W.E. Washburn, Collecting information, not objects, *Museum News* 62 (3) (1984) 5–15.
- [16] G. McDonald, S. Alsford, The museum as information utility, *Museum Management and Curatorship* 10 (1991) 305–311.
- [17] S. Alsford, Museums as hypermedia: Interacting on a museum-wide scale, in: D. Bearman (Ed.), *Proceedings of the ICHIM '91 Conference*, Pittsburgh, Pennsylvania, USA, October 1991, 1991, pp. 7–16.

- [18] Information today, December 2005, pp. 31–34, available at: <http://www.infotoday.com>.
- [19] A. Malraux, *La Musée imaginaire*, Gallimard, Paris, 1996 [orig. 1947].
- [20] Encyclopaedia Britannica online, available at: <http://www.britannica.com/eb/article-9000232>.
- [21] W. Schweibenz, The virtual museum: new perspectives for museums to present objects and information using the Internet as a knowledge base and communication system, in: H. Zimmermann, H. Schramm (Eds.), Proceedings of the 6th ISI Conference, Prague, November 1998, Konstanz, UKV, 1991, pp. 185–200.
- [22] 3D Van Gogh, Museum Virtual Tour', available at: <http://www3.vangoghmuseum.nl/vgm/index.jsp?page=49335&lang=en>.
- [23] ARCO, ARCO (Augmented Representation of Cultural Objects) Consortium. Available at: <http://www.arco-web.org>.
- [24] ICOM News, no. 3, 2004, available at: [http://icom.museum/pdf/E\\_news2004/p3\\_2004-3.pdf](http://icom.museum/pdf/E_news2004/p3_2004-3.pdf).
- [25] L. Teather, A museum is a museum. Or is it?: Exploring museology and the web, in: D. Bearman, J. Trant (Eds.), Proceedings of the Conference Museums and the Web, 1998, Pittsburgh, 1998.
- [26] M. McDonald, The Museum and the Web: Three Case Studies, available at: <http://xroads.virginia.edu/~MA05/macdonald/museums/method.html>.
- [27] M. Bergamasco, A. Frisoli, F. Barbagli, Haptics Technologies and Cultural Heritage Applications, in: S. Kawada (Ed.), IEEE Proceedings of the CA Conference 2002, Geneva, Switzerland, June 2002, IEEE Computer Society Press, 2002, pp. 25–32.
- [28] S. Worden, Thinking critically about virtual museums, in: D. Bearman, J. Trant (Eds.), Proceedings of the Conference Museums and the Web, 1997, Pittsburgh, 1997, pp. 93–109.
- [29] L.T.W. Hin, R. Subramaniam, A.K. Aggarwal, Virtual Science Centers: A new genre of learning in web-based promotion of science education, in: Proceedings of the 36th Annual HICSS'03 Conference, IEEE Computer Society, 2003, pp. 156–165.
- [30] O. Georgoula, P. Patias, Visualization tools using FlashPix image format, in: A. Gruen, S. Murai, J. Niederoest, F. Remondino (Eds.), Proceedings of the International Archives of Photogrammetry, Remote Sensing and Spatial Information Sciences, vol. XXXIV, PART 5/W10, February 2003, 2003.
- [31] San Francisco Fine Arts Museums, <http://www.thinker.org/>.
- [32] N. Talagala, S. Asami, D. Patterson, B. Futernick, D. Hart, The Berkeley-San Francisco Fine Arts image database, in: B. Kobler, P.C. Hariharan (Eds.), Proceedings of the 15th IEEE Symposium on Mass Storage Systems, Maryland, USA, March 1998, 1998, available at: [http://romulus.gsfc.nasa.gov/msst/conf1998/B5\\_06/TALAGA.pdf](http://romulus.gsfc.nasa.gov/msst/conf1998/B5_06/TALAGA.pdf).
- [33] Metropolitan Museum of Art in New York, available at: <http://www.metmuseum.org/>.
- [34] D. Marshak, J. Paul Getty Museum Re-Architects Technology to Enhance Visitors Experience, available at: [http://www.sun.com/products-n-solutions/edu/casestudy/pdf/getty\\_museum.pdf](http://www.sun.com/products-n-solutions/edu/casestudy/pdf/getty_museum.pdf).
- [35] Virtual Display Case, Making Museum Image Assets Safely Visible, 3rd ed., available at: [http://www.chin.gc.ca/English/Intellectual\\_Property/Virtual\\_Display\\_Case/](http://www.chin.gc.ca/English/Intellectual_Property/Virtual_Display_Case/).
- [36] A.N. Skodras, C.A. Christopoulos, T. Ebrahimi, The JPEG2000 still image compression standard, IEEE Signal Processing Magazine 18 (5) (2001) 36–58.
- [37] S. Sylaiou, P. Patias, O. Georgoula, L. Sechidis, Digital image formats suitable for museum publications, in: Proceedings of the 2nd International Museology Conference in Technology for the Cultural Heritage, Lesvos, Greece, June 2004, 2004.
- [38] Charles Olson's Melville Project digital library, available at: [http://charlesolson.uconn.edu/Works\\_in\\_the\\_Collection/Melville\\_Project/index.htm](http://charlesolson.uconn.edu/Works_in_the_Collection/Melville_Project/index.htm).
- [39] National Archives of Japan, available at <http://jpimg.digital.archives.go.jp/kouseisai/word/abc.html>.
- [40] Digital Collections of the University of South Carolina Libraries, available at: <http://www.sc.edu/library/digital/>.
- [41] Web3D consortium, available at: <http://www.web3d.org> (accessed at 6/08/2007).
- [42] S. Goodall, P.H. Lewis, K. Martinez, P. Sinclair, M. Addis, C. Lahanier, J. Stevenson, Knowledge-Based Exploration of Multimedia Museum Collections, in: Proceedings of the EWIMT Conference, November 2004, London, U.K., 2004.
- [43] VRML, The Annotated VRML 97 Reference, available at: <http://accad.osu.edu/~pgerstma/class/vn/resources/info/AnnotatedVrmlRef/ch1.htm>.
- [44] Benaki Museum, available at: <http://www.benaki.gr>.
- [45] Rembrandt House Museum, available at: [http://www.rembrandthuis.nl/cms\\_pages/index\\_main.html](http://www.rembrandthuis.nl/cms_pages/index_main.html).
- [46] COLLADA – Digital Asset Schema Release 1.5.0, [http://www.khronos.org/files/collada\\_spec\\_1.5.pdf](http://www.khronos.org/files/collada_spec_1.5.pdf).
- [47] OpenSceneGraph, 2009. Available at: <http://www.openscenegraph.org/projects/osg>.
- [48] QuakeDev, 2009. Available at: <http://www.quakdev.com/>.
- [49] Second Life, 2009. Available at: <http://secondlife.com/>.
- [50] J. Looser, R. Grasset, H. Seichter, M. Billinghurst, 2006. OSGART – A pragmatic approach to MR, In ISMAR 2006.
- [51] L. Calori, C. Camporesi, M. Forte, A. Guidazzoli, S. Pescarin, Openheritage: integrated approach to web 3D publication of virtual landscape, in: Proceedings of the ISPRS Working Group V/4 Workshop, 3D-ARCH 2005: Virtual Reconstruction and Visualization of Complex Architectures, Mestre-Venice, Italy, 22–24 August, 2005.
- [52] M. Heim, *The Metaphysics of Virtual Reality*, Oxford University Press, Oxford, 1993.
- [53] Foundation of the Hellenic World, available at: <http://www.ime.gr>.
- [54] Cortona, VRML Client – Web3D Products, available at: <http://www.parallelgraphics.com/products/cortona/>.
- [55] Y.M. Kwon, J.E. Hwang, T.S. Lee, M.J. Lee, J.K. Suhl, S.W. Ryul, Toward the synchronized experiences between real and virtual museum, in: Proceedings of Conference of APAN, January 2003, Japan, 2003.
- [56] B.B. Bederson, Audio augmented reality: a prototype automated tour guide, in: R. Mack, J. Miller, I. Katz, L. Marks (Eds.), Proceedings of the ACM Conference on CHI'95, Denver Colorado, USA, May 1995, ACM Press, New York, 2003.
- [57] K. Mase, R. Kadobayashi, R. Nakatsu, Meta-Museum: A Supportive Augmented-Reality Environment for Knowledge Sharing, in: Proceedings of the Conference VSMM '96, Japan, September 1996, IEEE Computer Society Press, 1996, pp. 107–110.
- [58] P. Milgram, F. Kishino, A Taxonomy of Mixed Reality Visual Displays, IEICE Transactions on Information and Systems, Special issue on Networked Reality, E77-D (12), (1994) 1321–1329.
- [59] C.E. Hughes, C.B. Stapleton, D.E. Hughes, E. Smith, Mixed reality in education, entertainment and training: An interdisciplinary approach, IEEE Computer Graphics and Applications 26 (6) (2005) 24–30.
- [60] T. Hall, L. Ciolfi, M. Fraser, S. Benford, J. Bowers, C. Greenhalgh, S. Hellstrom, S. Izadi, H. Schnadelbach, The visitor as virtual archaeologist: using mixed reality technology to enhance education and social interaction in the museum, in: S. Spencer (Ed.), Proceedings of the VAST 2001 Conference, Greece, November 2001, ACM Press, New York, 2001.
- [61] B. Baird, Using haptics and sound in a virtual gallery, in: M.A. Srinivasan (Ed.), Proceedings of the Fifth Annual PHANTOM Users Group Workshop, October 2000, Aspen, Colorado, USA, 2000.
- [62] S. Brewster, The impact of haptic 'Touching' technology on cultural applications, in: J. Hemsley (Ed.), Proceedings of the EVA 2001 Conference, Glasgow, UK, July 2001, Academic Press, Vasari UK, 2001, pp. 1–14, s28.
- [63] M. McLaughlin, G. Sukhatme, J. Hespanha, C. Sharabi, A. Ortega, G. Medioni, The haptic museum, in: V. Cappellini, J. Hemsley (Eds.), Proceedings of the EVA 2000 Conference, Florence, Italy, March 2000, Pitagora Editrice Bologna, 2000.
- [64] M. Bergamasco, Le musée de formes purés, in: M. Bergamasco (Ed.), Proceedings of the EVA 2000 Conference, Proceedings of the 8th IEEE International Workshop on Robot and Human Interaction, RoMan '99, Pisa, Italy, September 1999, IEEE Computer Society, 1999, pp. 27–29.
- [65] A. Galani, M. Chalmers, B. Brown, I. McColl, C. Randell, A. Steed, Developing a mixed reality co-visiting experience for local and remote museum companions, in: J. Jacko, C. Stephanidis (Eds.), Proceedings of the 10th Conference of HCI, Crete, Greece, June 2003, Lawrence Erlbaum Associates, 2003, pp. 1143–1147.

- [66] San Francisco Museum of Modern Art, 2001, 'Points of Departure', available at: <http://www.sfmoma.org/press/pressroom.asp?arch=y&id=117&do=events>.
- [67] S. Sauer, S. Göbel, Focus your young visitors: kids innovation – fundamental changes in digital edutainment, in: M. Bergamasco (Ed.), Proceedings of the Conference Museums and the Web 2003, Charlotte, USA, Toronto, 2003, pp. 131–141.
- [68] Development of the Museum Definition according to ICOM Statutes (1946–2001), available at: [http://icom.museum/hist\\_def\\_eng.html](http://icom.museum/hist_def_eng.html).
- [69] R.J. Loomis, S.M. Elias, M. Wells, 2003. Website availability and visitor motivation: An evaluation study for the Colorado Digitization Project. Unpublished Report. Fort Collins, CO: Colorado State University, available at: [http://www.cdpheritage.org/resource/reports/loomis\\_report.pdf](http://www.cdpheritage.org/resource/reports/loomis_report.pdf).
- [70] S. Sylaiou, F. Liarokapis, L. Sechidis, P. Patias, O. Georgoula, Virtual museums, the first results of a survey on methods and tools, in: Proceedings of the CIPA and the ISPRS Conference, Torino, Italy, 2005, pp. 1138–1143.
- [71] Museums in the USA, available at: <http://www.museumca.org/usa/alpha.html>.
- [72] UN Convention on the Rights of Persons with Disabilities, <http://www.ohchr.org/english/law/disabilities-convention.htm>.
- [73] Americans with Disabilities Act of 1990 (ADA), <http://www.usdoj.gov/crt/ada/adahom1.htm>.
- [74] Disability Discrimination Act 1995, available at: <http://www.disability.gov.uk/dda/>.
- [75] Resource Disability Action Plan, available at: <http://www.mla.gov.uk/documents/dap.pdf>.
- [76] Research on 'Quality' in Online Experiences for Museum Users, available at: [http://www.chin.gc.ca/English/Digital\\_Content/Research\\_Quality/about\\_vmc.html](http://www.chin.gc.ca/English/Digital_Content/Research_Quality/about_vmc.html).
- [77] S. Filippini-Fantoni, Museums with a personal touch, in: J. Hemsley, V. Cappellini, G. Stanke (Eds.), Proceedings of EVA 2003 Conference, University College London, July 2003, Vasari, UK, s25, 2003, pp. 1–10.
- [78] J.P. Bowen, S. Filippini-Fantoni, Personalization and the web from a museum perspective, in: D. Bearman, J. Trant (Eds.), Proceedings of the Conference Museums and the Web 2004, Arlington, Virginia, USA, April 2004, 2004, pp. 63–78.
- [79] F. Paternò, C. Mancini, Effective levels of adaptation to different types of users in interactive museum systems, *Journal of the American Society for Information Science* 51 (1) (2000) 5–13.
- [80] E. Hooper-Greenhill, *Museums and the Shaping of Knowledge*, Routledge, London, 1992.
- [81] M. McDonald, The Museum and The Web, Comparing the Virtual and the Physical Visits, available at: <http://xroads.virginia.edu/~ma05/macdonald/museums/virtual.pdf>.
- [82] M. Economou, The evaluation of museum multimedia applications: lessons from research, *Museum Management and Curatorship* 17 (2) (1998) 173–187.
- [83] R.D. Pea, Beyond amplification: Using the computer to reorganize mental functioning, *Educational Psychologist* 20 (4) (1985) 167–182.
- [84] J. Davallon, Une écriture éphémère: l'exposition face au multimedia, *Degrés* (92–93) (1998) 25–26.
- [85] M. Mokre, New technologies and established institutions, in: *How Museum Present Themselves in the World Wide Web*, Technisches Museum Wien, Austria, 1998.
- [86] R. Bernier, The uses of virtual museums: the French viewpoint, in: D. Bearman, J. Trant (Eds.), Proceedings of the Conference Museums and the Web 2002, Boston, USA, April 2002, 2002, available at: <http://www.archimuse.com/mw2002/papers/bernier/bernier.html>.
- [87] G. Hein, Constructivist Learning Theory, in: Proceedings of Developing Museum Exhibitions for Lifelong Learning Conference, ICOM/CECA, Israel, 1991, 1991, available at: <http://www.exploratorium.edu/IFI/resources/constructivistlearning.html>.
- [88] J.H. Falk, L.D. Dierking, *Learning from Museums: Visitor Experiences and the Making of Meaning*, Altamira Press, Walnut Creek, CA, 2000.
- [89] F. Liarokapis, S. Sylaiou, D. Mountain, Personalizing Virtual and Augmented Reality for Cultural Heritage Indoor and Outdoor Experiences, in: Proceedings of the 9th International Symposium on Virtual Reality, Archaeology and Cultural Heritage (VAST 08), Eurographics, Braga, Portugal, 2–5 Dec, 55–62, (2008), 2008.
- [90] J. Packer, Learning for fun: The unique contribution of educational leisure experiences, *Curator: The Museum Journal* 49 (3) (2006) 329–344.
- [91] S. Sylaiou, P. Patias, Virtual reconstructions in archaeology and some issues for consideration, *IMEROS, Journal for Culture and Technology* (4) (2004) 180–191.
- [92] M. Forte, About virtual archaeology: disorders, cognitive interactions and virtuality, in: J. Barcelò, M. Forte, D.H. Sanders (Eds.), *Virtual Reality in Archaeology*, BAR International Series 843, Archaeopress, Oxford, 2000, pp. 247–259.
- [93] P. Miller, J. Richards, The good, the bad, and the downright misleading: archaeological adoption of computer visualization, in: J. Hugget, N. Ryan (Eds.), Proceedings of the CAA 1995 Conference, Oxford, U.K., Tempus Reparatum, BAR International Series 600, 1995, pp. 19–22.
- [94] N.S. Ryan, Computer based visualisation of the past: technical 'realism' and historical credibility, *Imaging the Past, British Museum Occasional Paper* 114 (1996) 95–108.
- [95] Plato: Republic. Barnes & Noble Books, 2004.
- [96] J. Bonnett, New technologies, new formalisms for historians: The 3D virtual buildings, *Literary and Linguistic Computing* 19 (3) (2004) 273–287.
- [97] S. Sylaiou, M. Economou, A. Karoulis, L.M. White, The evaluation of ARCO: a lesson in curatorial competence and intuition with new technology *ACM Computers in Entertainment*, 6(2), ACM Press, New York, 2008.
- [98] R. Jackson, M. Bazley, D. Patten, M. King, Using the web to change the relation between a museum and its users, in: D. Bearman, J. Trant (Eds.), Proceedings of the Conference Museums and the Web 1998, Toronto, Canada, April 1998, Archives and Museum Informatics, Pittsburgh, 1998.

### **8.13 Paper #13**

Liarokapis, F., Anderson, E. Using Augmented Reality as a Medium to Assist Teaching in Higher Education, Proc. of the 31<sup>st</sup> Annual Conference of the European Association for Computer Graphics (Eurographics 2010), Education Program, Norrkoping, Sweden, 4-7 May, 9-16, 2010.

Contribution (90%): Implementation of the AR interface and collection of all the experimental data. Write-up of most of the paper.

# Using Augmented Reality as a Medium to Assist Teaching in Higher Education

Fotis Liarokapis<sup>1</sup> and Eike Falk Anderson<sup>1</sup>

<sup>1</sup>Interactive Worlds Applied Research Group, Coventry University, United Kingdom

---

## Abstract

*In this paper we describe the use of a high-level augmented reality (AR) interface for the construction of collaborative educational applications that can be used in practice to enhance current teaching methods. A combination of multimedia information including spatial three-dimensional models, images, textual information, video, animations and sound, can be superimposed in a student-friendly manner into the learning environment. In several case studies different learning scenarios have been carefully designed based on human-computer interaction principles so that meaningful virtual information is presented in an interactive and compelling way. Collaboration between the participants is achieved through use of a tangible AR interface that uses marker cards as well as an immersive AR environment which is based on software user interfaces (UIs) and hardware devices. The interactive AR interface has been piloted in the classroom at two UK universities in departments of Informatics and Information Science.*

---

## 1. Introduction

Until the emergence of new technological innovations the most common teaching method has been the direct communication between students and teachers usually taking place in the classroom. Although current teaching methods work successfully, most higher education institutions are interested in introducing more productive methods for improving the learning experience and increasing the level of understanding of the students. The emergence of new technological innovations in computing technologies has provided the potential for improving them. For instance, the web-based Virtual Learning Environments (VLEs) that many universities have adopted for aiding the teaching process are characteristic for this. A recent study has shown that virtual learning applications can provide the tools to allow users to learn in a quick and happy mode by playing in virtual environments [PCY\*06].

In particular, the introduction of Information and Communication Technologies (ICT) not only in schools but also in higher education institutions has been welcomed by students and educators worldwide. Most multimedia applications available for higher education purposes utilise teach-

ing material in a number of formats including text, images, video, animations and sound. These tools usually build upon traditional teaching methods, making the subject matter more interesting and challenging for both the students and the lecturer. Consequently, any future systems and techniques must take into consideration the current trends and needs of the higher education sector which is adopting these new technologies fast.

Many universities are eager in exploiting new visualisation methods to improve the current teaching models and one of the most promising technologies that currently exist is augmented reality (AR). In technical terms, AR is an amalgamation of computer graphics, vision and multimedia, which enhance the user's perception of the real world through the addition of virtual information [ABB\*01]. In AR the real environment must be harmonised (and synchronised) with the virtual in position and context to provide an understandable and meaningful view.

To improve student retention and participation in computing disciplines, Carter suggests that learning should be "more fun" [Car06], therefore one of the primary aims of this research is to increase the level of understanding of

the students in complex learning issues through an effective use of engaging audio-visual augmentation, which should be “attractive and highly motivating to today’s generation of media-conscious students” [MLC04], which is an educational goal in itself.

This paper is structured as follows. A discussion of work related to the use of AR in teaching is followed by the specification of the requirements for applying AR technology to learning in higher education. The following section introduces a theoretical framework for enhancing the learning cycle and style using AR. This is followed by a presentation of approaches for effectively designing teaching material in the context of AR scenarios for higher education. Furthermore, multimedia augmentation techniques are presented for enhancing traditional teaching methods. Several case studies exploring collaborative learning through the use of human computer interaction techniques are presented, followed by a preliminary evaluation together with our conclusions.

## 2. Education with AR in a Classroom Setting

Modern classrooms are frequently enhanced through the addition of new technologies, such as multitouch and computer gaming technologies [MDD09], and AR is one of these new technologies. Research has shown that learning does occur in virtual environments [Har06], and one of the earliest works in this area, applying AR to an educational context, is the ‘Classroom of the Future’ [Coo01], which conceptualises how it could be possible to enhance interaction between instructor and students by employing AR technologies. Another example is the higher education AR learning system for mathematics and geometry education Construct3D [KS02], which allows teachers and students to interact through various interactive scenarios in a collaborative environment. An alternative experimental education application demonstrates the AR enhanced teaching of undergraduate geography students about earth-sun [SH02]. In a similar AR application educators use AR to explain to students how specific parts of a computer could work in practice [FM03].

Other researchers have compared the use of AR and physical models in chemistry education [Che06], where results showed that some students liked to manipulate AR by rotating the markers to see different orientations of the virtual objects whereas others preferred to interact with physical models to get a feeling of physical contact. In another study, two modules, ‘Solar System’ and ‘Plant System’, were developed for mixed reality (MR) in the classroom, “providing support for classroom teaching and self-learning” [LCMLT07]. This study was influenced directly by its perceived usefulness, and indirectly through perceived ease of use and social influence, and preliminary results seemed to indicate the participants’ intention to use MR for learning. Elsewhere, researchers have studied the integration of physical objects that are computationally-augmented to support and encourage face-to-face interaction between disabled stu-

dents and virtual objects [APK\*09]. Here, initial results have indicated the importance of inclusion in novel technology-enhanced learning approaches for science education.

Collaborative AR allows participants to work in teams, experiment and interact with the superimposed virtual information in a natural way. In such environments, multiple users may access a shared space populated with digital information and thus maximizing the transfer of knowledge [Kau03]. A good example is the use of a real physical book for the development of a visually augmented reality book, however with drawbacks that include the affordances of the book resulting in book-like interaction by the users, which creates challenges in terms of technology deployment [GDB08]. Another collaborative approach for teachers and trainees is an AR system that simulates a web-based training and teaching environment for distance education and training. AR may be used successfully to provide assistance to the user necessary to carry out difficult procedures [SL03] or understand complex problems. Demonstration in lecture and seminar rooms is one of the most effective means of transferring knowledge to large groups of people [LMW\*04].

The main advantage of AR over more traditional teaching methods is that learners can actually ‘see’ and ‘listen to’ supplementary digital information. Additionally, students can intuitively manipulate the virtual information, allowing them to repeat a specific part of the augmentation as many times as they want. One of the main aims of this research is to contribute in resolving the perceptual discontinuities [DNT01] initiated by scattered sources of information during the learning process. To better understand these discontinuities we have developed a prototype AR learning system focused on higher education with a particular interest in computing courses. The basic idea is an AR table-top learning environment that integrates the real teaching environment (i.e. lecture theatre) with virtual learning scenarios in a student-friendly and engaging manner. Tangible interfaces are the medium used to allow students examine and experiment with the virtual teaching material in a natural manner.

One of the most important aspects of learning is effective interaction and collaboration. The concept of collaboration appears to be widely accepted as a manner of encouraging learning to take place in the classroom. The educational goal for which collaboration is intended plays a significant role on the nature of collaboration [Brn98]. In cases where learning is the primary educational goal it is essential to investigate how students become capable of collaboration [BBTJ97]. It is highly believed that AR support for collaboration can be improved if there is awareness for the level of collaboration required. Past research has shown that virtual and augmented reality learning environments can provide assistance to learners to perform cooperative and collaborative studies [KS02, POVP04, LMW\*04].

### 3. Learning in Augmented Reality

Apart from the subject content itself, pedagogical and psychological issues also need to be considered when designing a higher education learning system. The shared presence of virtual environments can enhance the opportunities for effective educational applications [HM97]. As an extension to VR systems, AR systems can be extremely effective in providing information to a user that deals with multiple tasks at the same time [KHS\*00]. Previous studies [BHBS99] have shown that they can improve performance times when the training is conducted on real objects. Taking into consideration the fact that individual persons have different learning styles and different ways of communication it is very important that we design a system following a user-focused approach. Educators not only need to recognise a unique learning style but also recognise this correctly for the successful development of effective learning and teaching strategies. The potential benefits of AR applied to higher education include:

- multi-modal visualisation of difficult theoretical concepts
- practical exploration of the theory through tangible examples
- natural interaction with multimedia representations of teaching material
- effective collaboration and discussion amongst the participants

Multi-modal visualisation allows students to switch interactively between different visualisation media. Tangible AR interfaces can help students to explore the multi-dimensional augmentation of the teaching material in various levels of detail. Students could navigate through the augmented lecture data and therefore concentrate and study in detail any part of the lecture. It is also important to consider the technological issues when introducing AR into teaching and learning processes [LMW\*04]. Ideally an educational AR system fulfils at least the following requirements:

- be simple and robust
- provide the learner with clear and concise information
- enable the educator to input information in a simple and effective manner
- enable easy interaction between learners and educators
- make complex procedures transparent to the learners and educators
- be cost effective and easily extensible.

A carefully planned and managed presentation of AR technology is also needed to avoid the teaching and learning environment being solely focused on technological issues rather than educational ones. This, together with the use of technological improvements, can direct higher education institutions to improve their teaching and learning processes.

### 4. An Affordable Augmented Reality System

Different learning styles can be applied from a single classroom session to a whole degree programme [HJ00]. Traditional methods of presenting teaching material can satisfy student demands only to a certain degree. For example, when teaching computer graphics it is really difficult for students to understand the concept of 3D if only traditional methods (i.e. PowerPoint presentations, websites, etc) are used. A cost effective AR audio-visual presentation of virtual multimedia content can be utilised to exploit the potential benefits of using alternative technologies to improve current teaching methods. Using virtual multimedia content, students can see real-life 3D examples of the principles they are studying as well as interact with them in a natural way.

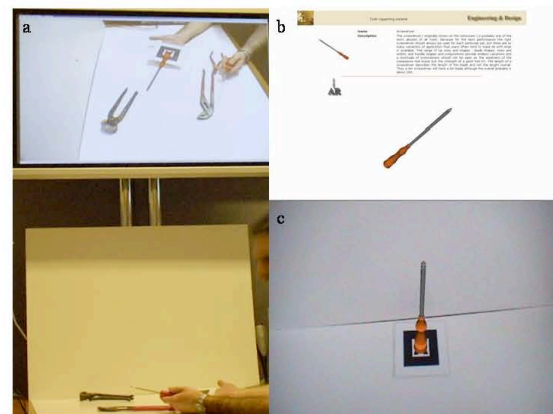


Figure 1: Operation of our AR system.

The majority of educational AR applications operate in indoor environments [FS99, SH02] and this is where this research work is focused. Registration between the real and virtual information is one of the most important issues of AR can be achieved either with the use of sensor devices or via computer vision techniques. Although, both techniques have proven advantages and disadvantages it seems that vision systems work much better in indoor environments while sensor-based systems are preferred for urban outdoor environments. In this paper, a vision-based AR learning solution has been adopted since the target environment is the classroom and it is much easier to control environmental parameters such as the lighting conditions. In addition, the cost of most of the technologies used in vision-based AR systems can be afforded by most higher education institutions – the cost of the system used here is a standard computer with a web camera (Figure 1).

### 5. Design of Teaching Material

To experimentally prove the feasibility of the higher education application, different teaching scenarios from different subject areas have been investigated and implemented. All

scenarios are specifically engaged with the improvement of learning and teaching techniques in the fields of Informatics at the University of Sussex and Information Science at City University. The presented scenarios have the capabilities of providing a rewarding learning experience that is otherwise difficult to obtain. Although the system suggests ways of enhancing teaching methods currently applied in computing courses it has been designed in such a way that it can be easily adapted and applied to other educational domains. The design of the teaching material has to be formulated in a way that satisfies the AR requirements of learning in higher education proposed above. As a consequence, the end result of the teaching material must transfer practical, theoretical, functional and constructive knowledge to the students. The teacher's expertise and the adaptation capabilities of the students serve as a cornerstone in deciding which methodology and scenarios should be followed. Depending on the nature of the subject areas (i.e. science or arts) the educator will have to design the course accordingly. In this research, courses that belong to the computing area have been used as case-studies. The development of the teaching material itself is an off-line process and consists of the following three interrelated parts each of those are explained in detail below:

- a set of distinctive marker cards – this provides the link between the real and the digital information so it needs to be easily distinctive
- digital information – the digital information including pictures, 3D models, textual descriptions, video animations and auditory information
- educational tutorials – a number of predefined learning scenarios which combine theory and practice at the same time

### 5.1. Design of Distinct Marker Cards

The precise definition of the teaching material is of great importance in any learning and teaching environment. In addition, the way the material is presented to students plays a significant role. The purpose of using marker cards for presenting the teaching material was selected because it provides a tangible interface between the real and the synthetic information, while it gives an indication to students what each marker contains at the same time. Our system presents the virtual multimedia information on marker cards of various shapes and sizes [Lia05]. The teaching material, course notes, diagrams and video animations were decomposed into appropriate components and for each a single marker card was created. This set of meaningful marker cards was carefully designed to create a complete representation of the theoretical elements of the subject matter during teaching. Students are given a selection of marker cards associated with the teaching material so they can have a visual choice before the information is presented. Each marker contains a link to the learning components (i.e. 3D objects and scenes, textual information, pictures and diagrams, video animations

and spatial sound), allowing students to experiment with different combinations. The use of marker cards allows learners to pick up and examine the superimposed information in an intuitive manner. The lecturer needs only to devise the learning strategy, which is to create the most appropriate learning scenarios. That is, the sequence in which they need to instruct the students to observe the markers and hence display the learning material in the AR environment. This can be achieved using a number of different ways. The lecturer can either manipulate the marker cards himself on a big table so that all students can enjoy an augmented presentation or leave the students to organise for themselves the sequence used in the teaching session. It may be that a combination of both methods could be more beneficiary for the students because they will be able to spend more time on the material they prefer or for which they need more assistance. This encourages students to become more interested in the subject because the teaching information is presented in an appealing way compared to more traditional methods.

### 5.2. Digital Information

The digitisation of the teaching material must have an educational flavour, as otherwise the virtual information will not have any effect when presented to students. To keep the complexity of the system low, only a few video animations were created and more emphasis was placed on the collection and generation of 3D objects, 3D sound, images and descriptive textual information. Most of the 3D models used were digitised using a technique known as Image-based Modelling (IBM). Although IBM techniques are not ideal for producing highly realistic models with detailed geometry, they have the advantage of rapidly generating 3D content. Alternative 3D content creation methods include laser scanning or manual modelling. The collection and digitisation of the static images was the most time-consuming task, as the information had to be organised in categories, followed by a thorough selection process to meet the requirements of the AR learning visualisation. Parts of relevant textbooks (i.e. from Multimedia and VR) were originally scanned and modified to make them suitable for classroom presentation. For storage, the TGA file format was selected as it provides an alpha channel that can be used to make images appear transparent. The textual descriptions were created in a text editor and stored in ASCII format. For generating the sound material, a standard microphone and an audio processing tool were employed.

### 5.3. Tutorial Generation

Aspen and Helm [AH04] proposed that a mixture of related approaches can be of substantial benefit to learners, resulting in effective commitment in a number of situations. The teaching material that we used was adapted from the official course material provided to the students in the classroom and digitised appropriately, aiming to address a wide range



of students, including visually impaired students. The digitised information has been categorised into different types. Each one aims at introducing different sections of the teaching material to the students. The tutorials that have been implemented up to this point can be categorised into three parts: theoretical; practical; and assessment tutorials;

- Theoretical tutorials – only the most important aspects of the theory are described through visual and auditory means of augmentation. The theoretical tutorials are the easiest to design technologically as no interaction with the user is required. The instructor simply needs to present the appropriate information to the students. However, the toughest challenge is the representation of the information in an appropriate manner for understanding and learning. In some cases, the information that has been transformed from 2D to 3D can confuse the user instead of helping, which is caused by a phenomenon known as information overload.
- Practical tutorials – based on the theory, students have to use a specific set of marker cards to describe a simple but complete process (i.e. building a computer using hardware devices). The practical tutorials encourage students to explore and interact with audio-visual information. Students have to use natural means (the specific set of marker cards) to explore a complex process or system, such as understanding how a computer works and replacing a component of a computer (i.e. processor), or solving a 3D puzzle that consists of – in our experiments – six pieces (Figure 4).
- Assessment tutorials – 3D graphical representations of theoretical and practical issues are assessed in a semi-automatic manner. The assessment tutorials aim to provide a less stressful way to evaluate the learning ability of the students. In the simplest scenario, students are provided with an augmentation of a 3D object or scene and a related question is asked. In more complex scenarios, AR quizzes can present a sequence of questions and have the potential of creating an enjoyable method for assessing students. It is important to note that all of our tutorials were designed on an experimental basis and only the theoretical and practical tutorials were demonstrated to higher-education students. The assessment tutorials have not yet been applied in real teaching.

In the next section, two real-life examples of how AR learning scenarios can be applied in ‘Informatics’ and ‘Information Science’ are presented.

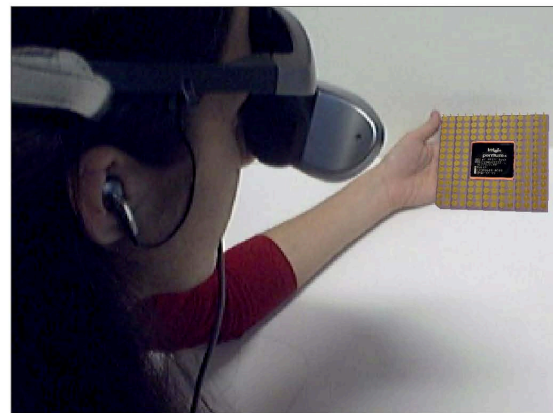
## 6. Case Studies

### 6.1. AR in Informatics Teaching

First of all, it is important to ascertain whether augmented reality technology can be effective in aiding education and learning. This part of the pilot study focused on this issue having the following hypothesis: Augmented reality can

be effectively combined with traditional methods to help students understand complex concepts and operations during practical laboratories. Based on studies that suggest that multimedia augmentation provides a more effective way of presenting teaching material compared to traditional methods [LPLW02], an interactive scenario was designed and demonstrated to two courses. The first was a foundation-level course called ‘Introduction to Information Technology’, and the other one was a postgraduate course, called ‘Virtual Environments’. The audience in ‘Introduction to Information Technology’ course consisted of 15 foundation students while ‘Virtual Environments’ had a cohort of 5 postgraduate students. In both courses, qualitative responses were collected from the students based on the ‘thinking aloud’ technique [DFAB04].

#### 6.1.1. Tutorials



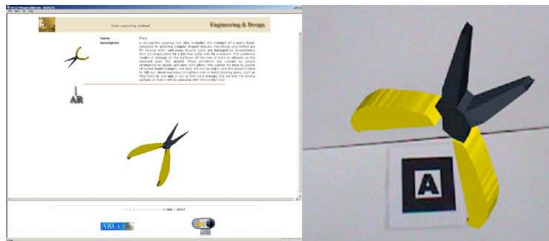
**Figure 2:** Student examining a 3D representation of an Intel CPU.

The first step in the sequence of the lecture material described the basic hardware components of a computer system and only the most significant components were presented including the motherboard, RAM, CD-ROM, DVD-ROM, hard disk, fan etc. To present a more engaging way of learning, 3D representations of the hardware were combined with human-computer interaction techniques. Students were able to examine the virtual information naturally. They could either use a head-mounted display (HMD) or a standard display monitor to view the virtual information. Each type of visualisation display has advantages and disadvantages since a HMD-based augmentation provides more immersion but may cause nausea and motion sickness and increases the overall cost of the application. On the other hand, monitor-based augmentation is much cheaper and it is more suitable for collaboration but provides less immersion.

The application of AR in learning can be demonstrated through the use of a practical tutorial called the virtual computer. The aim of this tutorial was to combine traditional

methods (i.e. PowerPoint presentation) with AR technologies to illustrate how a real computer looks like in reality. Students and lecturer gathered around a table on which the interior of a computer was overlaid. Based on the course's syllabus, augmented guidance was provided to the students using meaningful textual annotations in two ways: as references or as instructions. References (also called labels) were used to augment the most characteristic hardware devices whereas instructions were provided to act as an additional electronic guide for students (Figure 2).

To emphasise the 3D representations of the computer's parts the lecturer could isolate only the 3D models so that students were more focused and able to perform basic interactions on them such as rotations, translations and scaling operations.



**Figure 3:** Web3D-based VR presentation with metadata and AR-based presentation.

For the 'Virtual Environments' course, a second tutorial was designed and presented which allowed users to interact with 3D Web content (Web3D) using virtual and augmented reality. This allowed the exploration of the potential benefits of Web3D and AR technologies in education and learning. Multimedia content could be viewed locally or over the Internet, as well as in a tablet AR environment (Figure 3).

To better manipulate the 3D model, sensor devices, such as 3D mouse were integrated within the system. The advantage of using a 3D mouse is that it usually provides button menus together with a puck allowing six degrees of freedom interaction with 3D information, which makes for a more efficient interface than the keyboard.

### 6.1.2. Evaluation

The two educational scenarios were designed to investigate the effectiveness and presentation of multimedia augmented reality. The operation of the system was presented to the laboratory and students in both courses were asked to comment on the effectiveness of the system and whether it should be used as an additional tool for teaching the courses. Two different configurations were used, one for collaborative learning using standard monitor displays and another one using HMDs.

As far as the user feedback is concerned, all students agreed the presented technology is very promising and

should be applied in the classroom. However, the feedback received for individual aspects (i.e. visualisation and interaction) varied within the different groups of students (undergraduate and postgraduate). Specifically, the undergraduate students were impressed with the ease of use, the flexibility and the capabilities of the learning interface. In terms of visualisation, they preferred the monitor-based augmentation compared to the HMD-based augmentation which they found distracting and difficult to use. They commented that the monitor-based augmentation can help to increase collaborative learning and enhanced interaction and engagement with the subject matter. On the other hand, the postgraduate students found it extremely useful to be able to 'see' and 'interact' with related multimedia information in three-dimensions using the HMD-based augmentation. Especially in the case of computer components where they could examine a number of 3D representations concerning not only the specific hardware components but also learn by the virtual multimedia scenarios how a computer works. One particular student pointed out that the use of AR technology is the best means of teaching the 'Virtual Environments' course since it makes it easier to understand the underlying theories and concepts for three-dimensions.

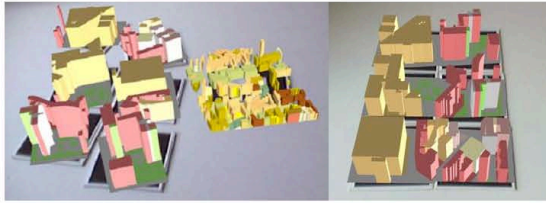
In addition, all students agreed that the physical interaction of multimedia information provides an exciting means of collaboration between the lecturer and the students. However, almost all students criticised the fact that only a few scenarios had been implemented.

## 6.2. AR for Teaching Multimedia

A second pilot was conducted in a postgraduate course called 'Multimedia'. As before, the experimental demonstration of the AR technology was qualitatively evaluated. The course cohort consisted of 10 students including 5 male and 5 female. Since the use of AR technology as a learning tool was determined to be helpful the second stage of the research was to explore some aspects of it in more detail. The hypothesis consisted of three parts: augmentation of audio-visual information is more effective than other multimedia mediums and it can aid the learning process; interaction in AR is more effective than other means and will speed up the learning process; and current learning can be enhanced through the use of AR.

### 6.2.1. Tutorials

Collaboration in classroom situations is largely determined by the students' roles [KS02] so the lecturer is responsible for controlling the sequence of the demonstration using the AR interface environment. The AR Puzzle application designed for this study was used to present geographical information in an attractive manner that would make learners more interested in geography. As an example scenario the topology of the university's campus was presented in an interactive 3D puzzle. At the beginning of the game, students



**Figure 4:** Collaborative AR learning game: unsolved AR puzzle pieces with the complete solution (left) and solved AR puzzle (right).

were provided with a set of marker cards. Each card corresponded to a different 3D component of the campus. Students were able to examine the components of the puzzle interactively in a natural way. Furthermore, they could experiment with different topology combinations by placing the pieces of the puzzle next to each other (Figure 4).

Students could collaborate in natural ways by swapping the cards around and utilising all types of communication that humans use when collaborating in real life (i.e. verbal, gesture, etc). This is one of the most significant advantages of the AR system over other educational and learning technologies. It offers students a selection of different interaction techniques ranging from natural interactions to software and hardware interactions.

### 6.2.2. Evaluation

The feedback received from the participants can be categorised into three types including: visualisation experience; interaction and movement; and usefulness in learning. In respect to the visualisation experience, most students agreed that the enhancement of AR was responsive, helpful with user-friendly interface. The visualisation experience as a whole was very useful and can be ranked as very good since the 3D models were very realistic. Some students argued that the tangible interface is very effective because it is possible to examine 3D objects from any angle in a natural manner. Three students mentioned that 3D perception in a classroom is much better than 2D because it helps to visualise the atmosphere in a better way and looks closer to reality.

As far as interaction and movement is concerned, students were impressed with the tangibility of the AR technology and the ability to naturally manipulate the learning material (i.e. navigating into a new place in 3D to wayfind from one position to another). Most of them mentioned that interaction seems to be very easy a very enjoyable and one said that it is “the most interactive interface that I have experienced in the classroom”. One stated that it is much better to interact with 3D objects in AR rather than interacting using the mouse and keyboard. Another student said that it is better than traditional methods since it is closer to reality.

The general feedback received was that AR is a useful

tool for learning not only in multimedia but also in other courses of computer science and informatics. Most students were impressed with the capabilities of AR and liked using it for exploration and learning. In particular one said “by far the most interesting lecture I have ever had and made me want to explore more and more”. Another liked the fact that they can zoom-in and explore particular features. One student mentioned that it would be good to have the whole course using AR but this would probably require a lot of time in the generation of tutorials. Finally, two students said that it is complicated and need some time to adapt to it so they proposed to do some training before using it.

## 7. Conclusions

In this paper we have presented an interactive, low-cost AR educational environment. The innovation of the system is that it can provide students with a multimedia augmentation of teaching material in a compelling and engaging way. Students can collaborate with the learning environment using either a physical interface (i.e. marker cards) or hardware devices (i.e. keyboard, mouse, 3D mouse) and software user interfaces (i.e. graphical user interfaces). The technology has been trialled in three different academic departments of different UK universities. Our initial evaluation suggests that AR technology is a promising and stimulating tool for learning and that it can be effective when used in parallel with traditional methods.

The current focus of the research is on implementing more robust educational scenarios and evaluating the operation of the system as well as the learning scenarios through extensive user studies. One of the greatest advantages of the presented system is that it can be used as a platform for creating other learning applications in other higher education fields such as computer graphics (as mentioned in section 4 of this paper), biology, chemistry, physics, archaeology, geology, geography etc. Although AR has been around for a while now, only recently have researchers started designing and implementing experimental applications including entertainment, medicine, construction, collaborative design, military, archaeology and many others. However, many issues related to technology remain to be improved as well as wide-ranging user studies within universities must be completed before AR learning environments can become a standard component of higher education.

## 8. Acknowledgements

We would like to acknowledge two websites, <http://www.3dcafe.com> and <http://www.ocnus.com/models/>, from which some of the used VRML objects were downloaded. Additionally, special thanks need to be extended to Panagiotis Petridis for creating part of the virtual multimedia content and Maria Cheiladaki for her assistance.

## References

- [ABB\*01] AZUMA R., BAILLOT Y., BEHRINGER R., FEINER S., JULIER S., MACINTYRE B.: Recent advances in augmented reality. *Computers & Graphics* 21, 6 (2001), 34–47. 1
- [AH04] ASPEN L., HELM P.: Making the connection in a blended learning environment. *Educational Media International* 41, 3 (2004), 245–252. 4
- [APK\*09] ARVANITIS T. N., PETROU A., KNIGHT J. F., SAVAS S., SOTIRIOU S., GARGALAKOS M., GIALOURI E.: Human factors and qualitative pedagogical evaluation of a mobile augmented reality system for science education used by learners with physical disabilities. *Personal Ubiquitous Comput.* 13, 3 (2009), 243–250. 2
- [BBTJ97] BURTON M., BRNA P., TREASURE-JONES T.: *Splitting the collaborative atom: How to support learning about collaboration*. IOS Press, 1997. 2
- [BHBS99] BOUD A. C., HANIFF D. J., BABER C., STEINER S. J.: Virtual reality and augmented reality as a training tool for assembly tasks. In *Proceedings of the International Conference on Information and Visualization* (1999), pp. 32–36. 3
- [Brn98] BRNA P.: Models of collaboration. In *Proceedings of the Workshop on Informatics in Education, XVIII Congresso Nacional da Sociedade Brasileira de Computacao Rumoa Sociedade do Conhecimento in Belo Horizonte* (1998). 2
- [Car06] CARTER L.: Why students with an apparent aptitude for computer science don't choose to major in computer science. *ACM SIGCSE Bulletin* 38, 1 (2006), 27–31. 1
- [Che06] CHEN Y.-C.: A study of comparing the use of augmented reality and physical models in chemistry education. In *VRCA '06: Proceedings of the 2006 ACM international conference on Virtual reality continuum and its applications* (New York, NY, USA, 2006), ACM, pp. 369–372. 2
- [Coo01] COOPERSTOCK J. R.: The classroom of the future: Enhancing education through augmented reality. In *Proceedings of the International Conference on Human-Computer Interaction* (2001). 2
- [DFAB04] DIX A. J., FINLAY J., ABOARD G. D., BEALE R.: *Human-Computer Interaction, Third Edition*. Pearson: Prentice Hall Europe, 2004. 5
- [DNT01] DUBOIS E., NIGAY L., TROCCAZ J.: Consistency in augmented reality systems. In *Proceedings of the 8th Conference on Engineering for Human Computer Interaction* (2001). 2
- [FM03] FERNANDES B., MIRANDA J. C.: Learning how computer works with augmented reality. In *Proceedings of the 2nd International Conference on Multimedia and Information & Communication Technologies in Education* (2003). 2
- [FS99] FUHRMANN A., SCHMALSTIEG D.: *Concept and Implementation of a Collaborative Workspace for Augmented Reality*. Tech. Rep. TR-186-2-99-04, 1999. 3
- [GDB08] GRASSET R., DÜNSER A., BILLINGHURST M.: Entertainment with a mixed reality book: a visually augmented illustrative childrens' book. In *ACE '08: Proceedings of the 2008 International Conference on Advances in Computer Entertainment Technology* (New York, NY, USA, 2008), ACM, pp. 292–295. 2
- [Har06] HARRINGTON M. C. R.: Situational learning in real and virtual space: lessons learned and future directions. In *SIGGRAPH '06: ACM SIGGRAPH 2006 Educators program* (New York, NY, USA, 2006), ACM, p. 48. 2
- [HJ00] HEALEY M., JENKINS A.: Kolb's experiential learning theory and its application in geography in higher education. *Journal of Geography* 99 (2000), 185–195. 3
- [HM97] HUGHES C., MOSHELL J. M.: Shared virtual worlds for education: The explorenet experiment. *ACM Multimedia* 5, 2 (1997), 145–154. 3
- [Kau03] KAUFMANN H.: Collaborative augmented reality in education, position paper for keynote speech at imagina 2003 conference. In *Proceedings of Imagina03 CDROM* (2003). 2
- [KHS\*00] KALAWSKY R. S., HILL K., STEDMON A. W., COOK C. A., YOUNG A.: Experimental research into human cognitive, processing in an augmented reality environment for embedded training systems. *Virtual Reality* 5, 1 (2000), 39–46. 3
- [KS02] KAUFMANN H., SCHMALSTIEG D.: Mathematics and geometry education with collaborative augmented reality. In *SIGGRAPH 2002 Conference Abstracts and Applications: SIGGRAPH 2002 Educators Program* (2002), pp. 37–41. 2, 6
- [LCMLT07] LIU W., CHEOK A. D., MEI-LING C. L., THENG Y.-L.: Mixed reality classroom: learning from entertainment. In *DIMEA '07: Proceedings of the 2nd international conference on Digital interactive media in entertainment and arts* (New York, NY, USA, 2007), ACM, pp. 65–72. 2
- [Lia05] LIAROKAPIS F.: *Augmented Reality Interfaces – Architectures for Visualising and Interacting with Virtual Information*. PhD thesis, University of Sussex, 2005. 4
- [LMW\*04] LIAROKAPIS F., MOURKOSSIS N., WHITE M., DARCY J., SIFNIOTIS M., PETRIDIS P., BASU A., LISTER P.: Web3d and augmented reality to support engineering education. *World Transactions on Engineering and Technology Education* 3, 1 (2004), 11–14. 2, 3
- [LPLW02] LIAROKAPIS F., PETRIDIS P., LISTER P. F., WHITE M.: Multimedia augmented reality interface for e-learning (marie). *World Transactions on Engineering and Technology Education, UTCEE 1, 2* (2002), 173–176. 5
- [MDD09] MUTO W., DOBIES J., DIEFENBACH P.: Applications of Multitouch and Gaming Technology for the Classroom. In *EG 2009 - Education Papers* (2009), pp. 1–6. 2
- [MLC04] MOSKAL B., LURIE D., COOPER S.: Evaluating the effectiveness of a new instructional approach. *ACM SIGCSE Bulletin* 36, 1 (2004), 75–79. 2
- [PCY\*06] PAN Z., CHEOK A. D., YANG H., ZHU J., SHI J.: Virtual reality and mixed reality for virtual learning environments. *Computers & Graphics* 30, 1 (2006), 20–28. 1
- [POVP04] PRADA R., OTERO N., VALA A., PAIVA A.: Belife: Teaching greenhouse management using an agent based simulator. In *Agent Based Simulation Workshop - ABS 2004* (2004). 2
- [SH02] SHELTON B. E., HEDLEY N. R.: Using augmented reality for teaching earth-sun relationships to undergraduate geography students. In *Proceedings of the First IEEE International Augmented Reality Toolkit Workshop* (2002). 2, 3
- [SL03] SCHWALD B., LAVAL B.: An augmented reality system for training and assistance to maintenance in the industrial context. *Journal of WSCG* (2003). 2

### **8.14 Paper #14**

Anderson, E.F., Peters, C., Halloran, J., Every, P., Shuttleworth, J., Liarokapis, F., Lane, R., Richards, M. In *at the Deep End: An Activity-Led Introduction to First Year Creative Computing*, Computer Graphics Forum, Wiley-Blackwell, 31(6): 1852-1866, September, 2012.

Contribution (10%): Collaboration on the teaching methods and write-up of the paper.



# In at the Deep End: An Activity-Led Introduction to First Year Creative Computing

E. F. Anderson, C. E. Peters, J. Halloran, P. Every, J. Shuttleworth, F. Liarokapis, R. Lane and M. Richards

Interactive Worlds ARG, Coventry University, UK  
eikea@siggraph.org

---

## Abstract

*Misconceptions about the nature of the computing disciplines pose a serious problem to university faculties that offer computing degrees, as students enrolling on their programmes may come to realise that their expectations are not met by reality. This frequently results in the students' early disengagement from the subject of their degrees which in turn can lead to excessive 'wastage', that is, reduced retention. In this paper, we report on our academic group's attempts within creative computing degrees at a UK university to counter these problems through the introduction of a 6 week long project that newly enrolled students embark on at the very beginning of their studies. This group project, involving the creation of a 3D etch-a-sketch-like computer graphics application with a hardware interface, provides a breadth-first, activity-led introduction to the students' chosen academic discipline, aiming to increase student engagement while providing a stimulating learning experience with the overall goal to increase retention. We present the methods and results of two iterations of these projects in the 2009/2010 and 2010/2011 academic years, and conclude that the approach worked well for these cohorts, with students expressing increased interest in their chosen discipline, in addition to noticeable improvements in retention following the first year of the students' studies.*

**ACM CCS:** [Computers and Education]: K.3.2 Computer and Information Science Education—Computer science education

---

## 1. Introduction

When applying for a university degree in the computing disciplines [ACM06], relatively few potential students have a fully accurate conception of what their chosen degree may entail. Many may believe computing or computer science to be an extension of the use of office suites that they are familiar with from ICT (information and communication technologies) courses at school, confusing the degree programmes with basic computer literacy [BM05]. This problem appears to be exacerbated by current teaching practices in schools [Roy10]. In other words, school or college learners may not be aware of the differences between ICT, based on the use of computing technology, and Computer Science with its emphasis on problem solving and the production of solutions which often involve programming. As a result, many students

are disappointed when they enrol at university and, to their dismay, discover their mistake. This is reflected in the observable decline of retention in computing programmes, and to remedy this, it has been suggested to modify degrees to become 'more fun' and to offer 'multidisciplinary and cross-disciplinary programs' [Car06] that will keep students interested in the subject. Unfortunately, retention problems are not restricted to traditional computing courses, but also extend to some of the multidisciplinary and cross-disciplinary degree programmes, such as creative computing degrees. Creative computing degrees are those degree programmes that expose students to the use of computing outside of the traditional desktop computing context. They include computing for the creative industries (see <http://www.skillsset.org>) and also explore the creative use of computing itself, for example, in wireless sensor networks, embedded in consumer devices or

as collections of services that augment our physical and digital environments. In these degrees, there is the potential to find a completely new set of misconceptions, where potential students confuse programmes such as multimedia computing, for example, with more vocational training courses for content creation software packages or web design. These applicants often demonstrate very strong expectations that their courses will predominantly feature artistic and creative content production topics, usually at the expense of more low-level technical topics such as mathematics, computer architectures or programming. Furthermore, the complexity of undergraduate computing degree programmes tends to be greatly underestimated. Once students become aware of this, they often disengage from the subject matter, often resulting in assessment failure or in the worst cases, withdrawal from their degree programmes. Consequently, retaining computing students remains a serious problem, one possible solution for which is to deepen the students' engagement with the subject.

Following the adoption of a new pedagogic model by the Faculty of Engineering and Computing (EC), Coventry University (UK), the solution of the Creative Computing subject group to address this problem has been the development of an integrative, interdisciplinary learning experience, providing new students with a breadth-first introduction to their chosen academic discipline. Newly enrolled students embark on a subject-spanning group project dubbed the 'Six Week Challenge' (see <http://vimeo.com/neophyte/pressplay>), that encompasses the first 6 weeks of their first year at university, replacing the regular teaching schedule and combining various aspects of the courses that make up the first year of the creative computing degree programmes. This project, which is not formally assessed, aspires to confront students with a challenging and ambitious task requiring them to take on a proactive role in problem-solving and to use their own initiative if they want their 'product' to succeed. They are encouraged to 'learn by doing', assuming responsibility for their student experience in the process, aiming to engage them closely with the subject matter of their degree programme while improving cohort cohesion, engagement and retention.

In this paper, we first present (Section 2) the background details of Activity-Led Learning (ALL), which has been adopted as the educational methodology in the 6 week group project. In Section 3 we describe how the group project of building a 3D etch-a-sketch-like computer graphics application engages students with activities integrating software and hardware development with usability evaluation, viral marketing techniques and academic writing. We present the results of an evaluation of the methods, based on student surveys, in Section 4. and discuss implications of the results in Section 5. We conclude in Section 6 with insights gained and issues for further consideration.

## 2. Activity-Led Learning (ALL)

One of the goals of higher education is to prepare students for life by enabling them to become independent learners. Independent learning does not come easy to students who have adapted to becoming passive participants in the learning process, where they are presented with all of the required learning material, a learning style that many of them acquired during their secondary education. The students of this 'Plug&Play Generation' [AM06, AM07] are sometimes described as suffering from shorter attention spans and impatience with the expectation to achieve quick and effortless results. However, 'active involvement in learning helps the student to develop the skills of self-learning while at the same time contributing to a deeper, longer lasting knowledge of the theoretical material' [MK02]. This is a key reason why our faculty has adopted ALL [WM08, IJP\*08, PJB\*10], a student-centred approach that has its roots in problem-based learning (PBL) [SBM04]. PBL is a constructivist instructional method [SD95] that provides a 'complex mixture of a general teaching philosophy, learning objectives and goals' [VB93].

### 2.1. Advantages

The problems that students are required to solve in PBL are usually much broader and more extensive than the relatively small, self-contained and well-defined exercises used in more traditional teaching sequences [BFG\*00]. Furthermore, in PBL and similar approaches, such as ALL, educators take on the role of facilitator, guiding the students' learning and monitoring their progress [HS04], which some studies on the subject have concluded may be superior in some aspects over more traditional teaching methods [VB93]. Such activity based educational approaches are supposed to work especially well in group projects, as they take advantage of group members' distributed expertise by allowing the whole group to tackle problems that would normally be too difficult for individuals [HS04], including other students in mutually supporting roles, as well as tutors and faculty [AM93].

PBL has gained some acceptance as an effective approach within a variety of disciplines in higher education environments [YG96, Fe196, BFG\*00]. This may be attributed to it providing an environment where the student is immersed, receiving guidance and support from fellow students and where the learning process is functional [Per92].

ALL and PBL not only lend themselves to the teaching of computing [BFG\*00] and computer graphics [MGJ06], but the use of computer graphics itself offers the possibility of defining interesting PBL scenarios whilst also enabling collaborative or mediated learning activities that could lead to better learning [Tud92]. Learning occurs through multiple interactions within the learning environment

[SD95, Cam96] and thus a potential added benefit of using computer graphics in combination with PBL scenarios is that learners engage with these using different senses, helping them to fully immerse themselves in the learning situation [Csi90] which could be expected to result in learning gains [CGSG04].

## 2.2. Pitfalls

This type of student-centred education is not without problems, however. It has been criticised due to the amount of guidance given to students [KSC06], relying on the use of ‘scaffolding’, that is, close guidance of the learner’s discovery, which others consider a simple improvement of a fundamentally ineffective approach [SKC07]. Finding an adequate balance for the amount of guidance given to students is one of the challenges of this type of educational approach [BBA09], as students might become too dependent on the provision of guidance, defeating one of the main aims of this type of approach, that is, to create independent problem solvers. It has been suggested that one precondition for the success of activity centred instruction is that participants need to already be highly motivated, well educated and possessing some degree of base competency in the subject area before engaging in activities [Mer07] and that the success of PBL approaches may depend upon the ability of students to work together to identify and analyse problems and to generate solutions [Cam96].

The use of scaffolding is not universally seen as a negative, and it has been suggested that the idea of PBL implies a ‘minimally guided’ form of education is wrong [HSDC07]. Our experience has been that to be useful, it is far from minimally guided, but also that this does not imply that the students are encouraged to become dependent on constant guidance from staff. The staff time requirements, however, are significant in comparison to traditional teaching.

## 2.3. Creative computing at Coventry University

The Creative Computing subject group of Coventry University’s EC faculty delivers degree courses which aim to produce graduates and computing professionals capable of working in environments where art and technology meet. Our courses have a strong Computer Science core, balanced with studies in design theory, game development, programming, graphics and content creation, pervasive and sensing technologies, usability and video and sound production. The teaching team strives to develop a strong interdisciplinary environment integrating content from these distinct domains.

Computing curriculum recommendations state that ‘the breadth of the discipline should be taught early in the curriculum’ [Tuc96]. This is realised in a breadth-first computing curriculum, where students are exposed to the computing

domain through a broad introduction to the major areas of Computer Science [VW00], allowing them to gain a more comprehensive understanding and appreciation of the discipline. They are able to gain ‘a holistic view of a topic before they learn about more complicated details’ [DG06] that empowers them. Important concepts are touched upon early on to provide students with the basis for a much larger range of activities than would be possible in more traditional/conservative teaching sequences. This is because students experience the tasks that they embark upon in the wider context of the computing discipline, rather than as isolated subject matter. While to many students this may seem intimidating at first, it nevertheless tends to result in much deeper understanding.

In line with a faculty driven move towards more activity-led teaching and learning, the Creative Computing subject group has developed a 6 week group project. The project aims to immerse first year students in an engaging activity designed to address some of their apprehensions, while introducing, in microcosm, the entire spread of topics in the first year curriculum. The design of the project is described in more detail next, in Section 3

## 3. A Six Week Challenge—Learning by Doing

First piloted at the start of the 2009/2010 academic year [SEA\*10] (see also <http://vimeo.com/neophyte/pressplay>), the activity for our creative computing degrees, including Multimedia Computing and Games Technology pathways, integrates software and hardware development with usability evaluation, viral marketing techniques and academic writing. In its refined second iteration at the onset of the 2010/2011 academic year, the software development aspect focussed on computer graphics, resulting in the students’ creation of a computer graphics application with a physical hardware interface. Our creative computing degree programmes are heavily reliant on modern multimedia concepts and technologies. ‘Multimedia—while embracing computer graphics—describes the foray of other disciplines into the digital realm’ [Gon00] and through their projects our students not only ‘learn computer graphics’, but also ‘learn through computer graphics’, effectively making our students’ learning experience a hybrid of both aspects of teaching computer graphics in context [CC09].

The purpose of a Six Week Challenge is to allow students to evaluate the flavour of the course they are about to embark upon, addressing a number of issues in the orientation of new students whilst promoting high levels of engagement, which aim at both deep learning and increased retention. The activities were intended to be challenging and engaging without requiring assessment to monitor progress or encourage participation. Next, we describe our rationale for finding a suitable challenge (Section 3.1) and the details related to running one (Section 3.2).



### 3.1. Finding a suitable challenge

To meet our goal of engaging the students with the creative computing discipline we had to face our own challenge of finding a suitable set of integrative activities for students. In the development of such activities it is important that they are meaningful to the student [Cun99], and appropriate for the intended student group, which in our case are absolute beginners embarking on their first steps in higher education. The activities designed for the 6 week group project would have to be related to the degree programmes of the students, complex enough to appear challenging, yet achievable within the set time-frame. At the same time the problem that ‘students . . . expect to see immediate (and spectacular) results, often before they have learned enough to achieve anything remotely spectacular’ [AM07] needs to be addressed by enabling the students to achieve results that appear ‘spectacular’. We first delivered a Six Week Challenge in the 2009/2010 academic year, and did so again in the 2010/2011 academic year. The student cohorts, staff numbers and tasks set for both years were as follows:

- The 2009/2010 cohort consisted of 56 students, with 6 faculty members and one graduate intern involved, of whom only 4 faculty members were actively delivering content. Students were tasked with developing a hardware controlled media player (see [SEA\*10] for more details).
- The 2010/2011 cohort consisted of 54 students supported by 6 faculty and 2 teaching assistants. The students were tasked with the development of a graphics application based on the popular *Etch A Sketch*<sup>®</sup> drawing toy by the Ohio Art Company (<http://www.etch-a-sketch.com>), the computer implementation of which would not only involve graphics, but would also provide an interesting exercise in user interface design and evaluation [Bux86]. To provide students with an additional challenge, we extended the basic concept of a 2D drawing toy to the third dimension: a 3D etch-a-sketch-like graphics application with turnable knobs as inputs for drawing on the three axes.

In both the 2009/2010 and 2010/2011 challenges, *Processing* [RF06] (<http://www.processing.org>) was chosen as the development environment for the task. It is a Java-derivative language for computer arts creation, which lends itself well to introductory programming and computer graphics education [PBTF09] and also interfaces with the Arduino micro-controller [Sto09] (<http://www.arduino.cc/>) that we chose for the development of the hardware interface. The Arduino is an Open Hardware design that has been successfully employed as an educational tool [FW10], which allows the easy creation of input devices for computers. The kits we used were ideal for our purposes as they did not require any soldering, allowing the hardware to be simply slotted together.

**Table 1:** A Six Week Challenge consists of six sub-challenges, or themes. Each theme adds a new element to the overall project and can be completed by students within a week.

Week	Theme	Section
1	2D graphics programming	3.2.1.
2	3D graphics programming	3.2.1.
3	Hardware design & interfacing	3.2.2.
4	Usability evaluation	3.2.3.
5	Viral marketing	3.2.4.1
6	Academic communication & reporting	3.2.4.2

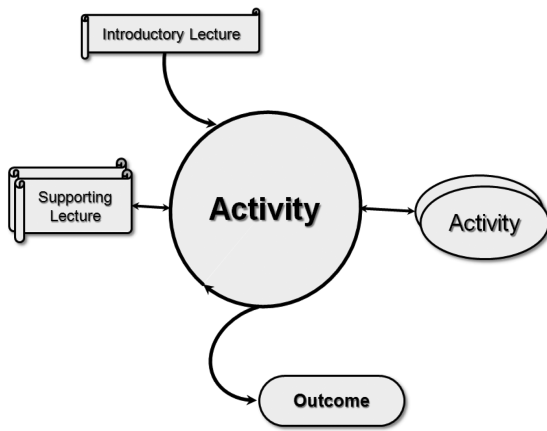
Our careful selection and presentation of topics was aimed to provide students with the opportunity to quickly evaluate the flavour of the course they were about to embark upon, addressing a number of issues in the orientation of new students and attempting to promote high levels of engagement, deep learning and increased retention.

### 3.2. Running the challenge

Since the Six Week Challenge is a group project, the 2010/2011 cohort of 54 students was split up into groups of 6 to 7 students. For the duration of the project normal delivery of teaching was suspended entirely whilst the teaching team worked collaboratively with the student groups to develop their products.

The task of creating the 3D etch-a-sketch-like graphics application with a dedicated hardware interface was broken down into six sub-challenges, or themes (Table 1), that each added new elements to the overall project and that each could be completed within 1 week, including:

- graphics programming and software control, consisting of 2D and 3D graphics programming in the Processing language and the mapping of manipulation functions to keyboard controls (see Section 3.2.1).
- hardware interface, concerning the construction of a hardware interface with the Arduino micro-controller for the 3D etch-a-sketch-like application (see Section 3.2.2).
- usability evaluation, to consider usability aspects of the controller—this gives the students their first experience of what it means for software not just to be correctly implemented, but also acceptable to users. This topic, which relates to human-computer interaction (HCI) and usability, is particularly important on the degrees for which this programme was developed (Section 3.2.3).
- dissemination (Section 3.2.4), consisting of a viral marketing campaign (Section 3.2.4.1) and academic communication (Section 3.2.4.2).



**Figure 1:** *The Activity-Led Instruction cycle. The main activity is introduced through an introductory lecture to subject-specific aspects of the students' task, which they then solve independently; this may lead to further activities and additional lectures that are based on students' needs/demands.*

We employed an activity-led instruction cycle [AP09] (Figure 1) in which students were first, in a Monday morning briefing at the start of each week, introduced to the sub-challenges. Important subject related information was covered in a short introductory lecture, followed by a variety of guided learning activities that focussed on the challenge for the week which the students participated in. The students were then left to work out how to solve each of the sub-challenges, being allowed to organise their remaining time as they saw fit. Teaching staff were available throughout the week to provide encouragement and additional guidance when requested and, depending on the student groups' progress, to run additional sessions to cover subject areas that the students discovered while working on their projects, with these support sessions timetabled from Tuesday to Thursday. No group structure was enforced, although in many groups individuals began to take on obvious roles, such as leader, tester or presenter.

A special 'show and tell' session consisting of a gathering of all of the students and lecturers involved in the project was organised for the end of every week (Section 3.2.4.3). This was an opportunity for students to demonstrate their work to the whole cohort and to members of the faculty.

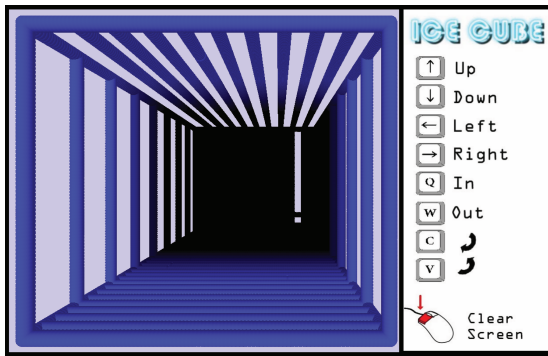
Overall, this mode of delivery allows students to actively influence the direction of their learning, as they are given some level of control of the delivery of subject-specific information, that is, while students receive an introductory lecture to subject-specific aspects in support of their activities, any additional teaching sessions (lectures and/or tutorials) are dependent on the students' needs and/or demands.

### 3.2.1. Graphics programming and software control

The first set of tasks for the student groups concentrated on the development of the graphics application. This required each team to develop, using the 'new to them' Processing language, know-how in the creation of the graphical elements needed to create the etch-a-sketch-like application. It consisted of:

- implementation of the drawing environment itself, commencing with the drawing of simple 2D points, and progressing to lines, squares and more complicated 2D shapes. A primary goal here was the understanding of how objects could be created for display on the screen, particularly their specification using vertices, edges and faces. Some experimentation took place with simple 3D objects.
- placing newly created objects within the drawing environment in 2D and 3D. Developing an understanding of basic affine transformations in 2D and 3D (i.e. translation, scaling and rotation). In many cases, this led to animation attempts that required an exploration of aspects related to the composition and redisplay of scenes, such as single- and double-buffering.
- definition of a changeable camera/view. The need for knowledge about the camera naturally arose from incidents where objects unexpectedly disappeared from view for some groups, either due to being placed outside of the viewing area of the window or outside of a poorly defined view frustum during 3D experimentation. Some groups also wished to be able to move the camera around in a manner similar to popular first person shooter games, motivating them to learn more about camera parameters.
- user interaction, to allow the program to process input from the keyboard and mouse. This involved a basic understanding of event handling and the event processing loop and was initially based on predefined keyboard input (i.e. controls that allow a user to limit movement to X, Y and Z), while students grasped the relation between the event loop, user input processing and scene redisplay for animation. Basic mouse control was also introduced.
- an appropriate graphical user interface design, building on topics learned during user interaction, but going somewhat further to consider the ease of use for the user and performance issues.

All of the student groups achieved at least a basic implementation of the features and demonstrated prototypes capable of drawing to the screen in 2D and 3D, and allowing the screen to be cleared subsequently. Most groups exceeded the basic requirements (see for example, Figure 2) and included diverse additional features. Many of these related to the selection of different drawing colours from a predefined palette, either by manual selection or, in some cases, automatic schemes that accounted for the drawing depth by



**Figure 2:** One of the student-created 3D drawing applications. Keyboard controls allow the turtle to be moved in three dimensions and enable the screen to be cleared.

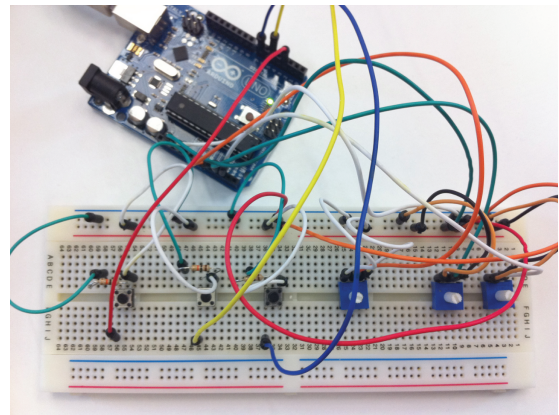
changing some of the colour characteristics. A number of implementations also featured the use of 2D shapes as brushes with which to draw.

As students experimented with shapes and drawing in 3D, important questions arose. For example, technical issues relating to camera set-up, object and scene rotation, and 3D object positioning using transformations, all arose naturally as the task was feature-driven. Furthermore, in cases where groups redisplayed the scene each frame, they also required a means for storing and updating previously drawn lines or shapes so that a full sketch could be displayed each frame. This represented an interesting and challenging problem for the students, who investigated a number of data structures and methods to do this. In this way, students discovered for themselves the need to understand these concepts, which might otherwise have seemed obscure or unimportant.

Students were also encouraged to investigate different interaction schemes, for example, by mapping different keys onto controls and considering mouse movement. In particular, they were tasked with attempting to control the application using the minimum number of keys possible and to create novel mouse-keyboard methods for control. This added an extra challenge beyond the more obvious 1:1 mapping between keys and functions, and additionally helped raise important issues for consideration during the modelling of the physical controller (Section 3.2.2) and in the usability evaluation (Section 3.2.3). A number of groups succeeded in enabling more advanced interaction control by combining both the mouse and the keyboard. This proved to be very useful when the groups were subsequently asked to design user interaction tasks for usability studies.

### 3.2.2. Hardware interface

Once the basic graphics application was developed, teams were asked to integrate their application with a dedicated



**Figure 3:** Example of a student group's hardware interface. Three potentiometers (right side of board) allow the user to draw in three dimensions.

hardware interface and then to evaluate the hardware prototypes. For the hardware task, students used the Arduino prototyping platform, which allows users to quickly construct devices ranging from simple flashing lights to autonomous spy-planes and hand-held consoles. Online resources were provided to the student groups, including eBooks and hardware tutorials. In addition, students were given instructions on how to create a blinking LED device using resistors and potentiometers. Resistors were used to protect the circuit and the potentiometer to control the speed of a LED Scanning light effect. At the end of the task, all of the groups had created circuit diagrams for their etch-a-sketch-like applications using the 'Fritzing' application [KWC09], and many students created solutions with three potentiometers for controlling the drawing (see for example, Figure 3), similar to the Digital Airbrush by Batagelj *et al.* [BMTM09]. Some of the most important keyboard functions that were assigned to hardware buttons included: change of colour, drawing speed change, background colour change, clearing the screen, restoring the screen, precision mode, camera movement, zoom in/out, and the provision of a help screen. Some groups also decided to provide a combination of two or more button pushes to perform a particular action, solving the problem of having too many key assigned features.

### 3.2.3. Usability evaluation

The usability component of the Six Week Challenge involved students in designing a simple usability study for their etch-a-sketch. Usability is perhaps the key issue in HCI. HCI is concerned with how systems are used in practice, and usability is about how to design systems so that using them is easy, effective and enjoyable [RSP11]. Thus, students focussed on one or two key tasks for their etch-a-sketch, running the study on four or five users, collecting data, and analysing it

to develop an informed view on whether or not the interface to their graphics application was usable in terms of the tasks tested.

There are alternative approaches to usability in HCI. Broadly, studies can be 'ethnographic', or 'lab-based'. Lab-based studies [DFAB03] feature pre-defined tasks carried out by users in controlled settings where there is informed consent. Aspects of performance like completion times and error rates are measured. This approach is useful for establishing common issues across a range of users, and is particularly appropriate in the context of system development, for evaluating prototypes. In contrast, ethnographic studies [Cra03] are of naturally occurring use of real systems in authentic contexts of use (i.e. the real world), unscripted and uncontrolled. This generates descriptive rather than numeric data, and is good for looking at particular cases of use in depth. It is particularly appropriate where the system is a product rather than a prototype.

The lab-based approach was what was experienced by students during the Six Week Challenge. The students engaged in the design of a usability study in which the setting, tasks and measurements were all pre-defined and kept uniform across different users to allow comparability of results.

In designing their usability tests, groups were first instructed to define the core tasks required of users by the *etch-a-sketch*. To do this, it was suggested a simple task analysis [Hor10] was created, showing the steps and any substeps required to complete the task. This required students to think about scoping: what should the realistic limits of the tested task be? How long should it take? What should count as its beginning and end, and what is the necessary sequence of actions? Following this, students turned to metrics: what aspects of the users' performances could and should be counted? This relates to a quantitative approach to data, where numbers are the basis for claims about usability. Students made sensible suggestions: for example, number of errors made, and time taken overall. This naturally led into the need for 'baseline' measures, that is, benchmark performances with which to compare user performance, and how these should be established [Nie93]. After addressing this issue, students were asked to prepare observational instruments (paper forms) they would use to record data, and to explain to tutors in advance how they would carry out data analysis, which led into consideration of individual and mean scores, variance and representation, for example, by bar charts. Crucially, students needed to be able to explain how they would make usability claims on the basis of their data. Most groups realised that the numerical scores they got from users needed to approach or equal baselines. In that case, it could be claimed that, in terms of tasks tested, their design was usable. Conversely, students were asked to consider what they could say about design revision if the numbers were further away from baselines, that is, it was more difficult to claim usability. This issue links usability studies to

technology design and is crucial to start negotiating early on in the study of human-computer interaction.

These methods and techniques, although elementary, are crucial to usability studies [BTT05], but can be hard to teach. The most difficult issue is that students, while they may be able to perform aspects of the practical work, are frequently not so clear on how to design it or why they are doing it in the first place. In the context of ALL, one goal of the usability week was to start to inculcate a scientific approach, where claims about usability are evidence-based, and the process is explicit, repeatable and replicable. This was eased by the fact that the groups had a vested interest in showing the usability of their designs. This helped leverage understanding of these principles: in other words, it was important for groups to show that their claims were not just their own subjective opinion, but evidence-based according to scientific practice, in such a way that they would gain credibility. This is a crucial hurdle for students to clear, and the motivation provided by the Six Week Challenge undoubtedly helped (although developing a scientific attitude is not immediate). That there was general appreciation of this was clear from the end-of-week group-to-peer presentations made at the end of the usability week.

Having developed their usability tests, students had to run them. This means engaging with users in systematic ways. In particular, instructions needed to be developed and kept consistent across users. Students had to learn not to interrupt or make hints to users, and crucially to keep their own behaviour discreet and uniform across users to control for any researcher effect. This resulted in tests being run in ways that began to approach professional practice. Many students worked out that in addition to the metrics they were using, they could add in other qualitative observational data, for example, questions users asked, things they said, facial expressions they made and so on. This spontaneous activity was the beginning of the important process of gathering both quantitative and qualitative data and looking for the complementarities between these, particularly how qualitative data can help explain numbers: for example, where time was slow, did the user ask a lot of questions? If so, this might indicate confusion, which helps explain slow times.

The main difficulty in teaching HCI, including usability, is that it is highly conceptual and often abstract. Typically it is taught by asking students to run studies on interfaces they may not have a personal interest in. The Six Week Challenge meant that students had a strong motivation to show their designs were usable. Personal investment in the work helped leverage engagement in many issues which can be a challenge to teach, in particular the forming of a research question for a usability study, the collection and analysis of different types of data, realistic and relevant scoping of user tasks, and the correct setting up and running of user sessions. The embedding of advanced usability material within the Six Week Challenge increased its accessibility: there was impressive work within a short period. Our activity-led

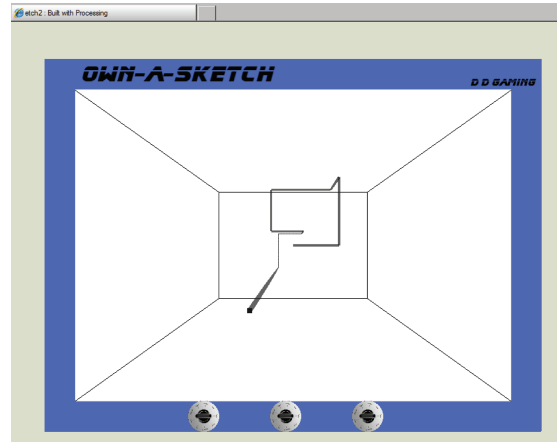
approach in general can be claimed to ease the transition from pre-degree to degree education, particularly helping to ameliorate the feelings of dismay and difficulty we identified in the introduction (Section 1).

### 3.2.4. Dissemination

An important further aim beyond developing students' technical abilities and team work was to develop their awareness of the importance of dissemination and how dissemination should be tailored to both target audience and goal. Students disseminated their work both internally and externally through group demonstrations, a viral marketing campaign and academic communication methods. The aims of dissemination were to inform the work of other groups, to provide them with the experience of presenting work to different external target-groups, highlighting the necessity of differing dissemination methods based on the target audience (e.g. academics, consumers), and to think about ways in which quantitative and qualitative feedback could be collected. In addition to internal demonstration, students also had to disseminate their work externally in two ways: through a viral marketing campaign (Section 3.2.4.1) and in the form of an academic manuscript (Section 3.2.4.2).

**3.2.4.1. Viral marketing campaign** This challenge involved student groups generating publicity for their products by creating web-pages for presenting their programs and gathering usage statistics, as well as an online viral advert linked to their product to tempt back users to their groups' product homepages. The stipulation of numbers was an important inclusion as students would need to solve the problem of digital verification and customer tracking. A suggested operational strategy for the week was to upload their source code to an open source repository, and to upload their executable to an online storage site or a hosted product website. Students were encouraged to create a video or other promotional device and to disseminate this through social networks. For visitor tracking we demonstrated the use of Google analytics software [Cli10] and tracking code.

This task allows the students to work in media that most of them are familiar with already: blogs, on-line videos, social networking sites and so on. Rather than simply allow them to demonstrate their familiarity and facility with these media, however, the marketing task asks them to think more critically about what they can achieve through them, how they might be applied in their studies or careers and ensures at least a basic level of skill in the minority of students who, before coming to university, have not had any experience in this area. The students, who we might describe as 'digital natives' [Pre01], do not always have a great deal of skill in transferring their skills [KJCG08] or realise how they might be of use in their studies or careers. The goal of this media week component of the challenge was to get the students to think about the context under which their future productive



**Figure 4:** Example of a student group's graphics application embedded in their website. Control knobs on the interface allow the user to create a sketch interactively in three dimensions.

activities may take place, and how to shape products and messages for a particular audience. Whilst presented in a light-hearted fashion, the media week provided opportunities for discussions about the nature of digital goods, ethics and piracy, copyright, open source and creative commons solutions to intellectual property rights problems.

We found that many students published their work by placing interactive demonstrations of their graphics applications on their web-pages (Figure 4) and loading pre-recorded videos on YouTube [BG09]. The Processing system provided the necessary facilities for allowing students to do this themselves, as it allows interactive graphics programs to be embedded in websites. Most of the student groups successfully completed the website integration of their applications and interfaces, while some groups chose to provide downloadable executables of their applications instead. Unsurprisingly, very few of the students exhibited any difficulty with the technical components of the week's challenge: producing simple web pages, embedding JavaScript tracking code, uploading video content and accessing analytic data. The students performed particularly well during this week, being able to share their existing knowledge of how to resource web activity for free and they welcomed the opportunity to proudly demonstrate their achievements to their friends on social networking sites. The graphical nature of the work seemed particularly amenable to such sites, as a means for attracting interest from peers and potential employers, and also serving as a starting point for the creation of a graphics programming portfolio. In fact, in hind-sight, we probably set the 'number of viewers' stipulation too low, as between them each group could possess over one thousand contacts on social networking sites. The more interesting learning outcomes of the week occurred in the conversations that the tasks entailed. Some

students worried about how to protect their products from piracy (even though they were free) and then had to consider this in the light of the fact that the tools with which they had made them were free also. Students were encouraged to read about copyright and creative commons solutions to the problem of intellectual property. Similarly, the mechanics of viral marketing were a topic for discussion during the week, leading students to examine what makes an individual share a link with their friends on the internet and which content was most likely to trigger exponential sharing. This also helped raise an awareness that the dissemination method must account for the target audience, which may also include potential employers.

This week's activities also served to raise the question of feedback, by looking at ways in which qualitative and quantitative data could be collected. This involved accounting for simple metrics, such as tracking the number and types of comments and views that their work attracted. The issue of feedback is sometimes underestimated from the students' point of view. Graphics work published on the web may be a very useful way for attracting comments from more skilled graphics practitioners from around the world, as a way for students to obtain broader formative feedback on their portfolio work from a diverse audience.

At the end of the week presentation all of the groups had met their viewing targets, a few were able to share customer's comments' and one group had even 'monetized' their website and were deriving an income stream.

**3.2.4.2. Academic communication** The academic writing and research component deals directly with the process of critically evaluating students' own work and the work of others, reading academic texts, synthesising arguments and presenting information; skills that will be used and developed throughout any degree course, yet are not necessarily obviously critical to students beginning a technical degree.

The task involved preparing a short paper (3–4 pages of collaborative academic writing) providing background information to their projects and stressing the relevance of this research to their product. Each student group was presented with a different research question. Many of these were related to the graphics techniques they used and that they were tasked with describing in their short papers. For this the groups had to:

- engage with a number of academic texts, providing a basic understanding of academic writing (language and style), some of which [LR88, Lar09], originating from the computer graphics community, were provided to them;
- adopt appropriate strategies for finding and evaluating relevant textual sources [Gri09], including the use of citation databases;

- learn to organise information in a logical manner, suitable for presentation in written form, as well as for oral presentation [Ger04].

The introductory lecture for the academic communication week provided students with an overview of academic writing, that is, the academic writing style and the structure of academic texts, which students were exposed to in a light-hearted manner [Sch96], as well as considerations of good academic conduct, including issues of proper citing of sources. Students were then introduced to literature search strategies, as well as the LaTeX document preparation system [Lam86] to ease them into the practice of preparing consistently formatted documents. Students were then directed towards the compilation of a comprehensive reading list of academic articles that appeared relevant to their set research questions, providing the basis for their short review/survey paper. Throughout this activity, students were repeatedly briefed on the principles of academic honesty to prevent problems like plagiarism.

The resulting short papers showed an unexpected level of maturity, rarely seen in students in their first year at university. The students also developed a much greater appreciation for the academic writing style, contrasting it to the much more informal communication forms they were familiar with before (Section 3.2.4.1).

**3.2.4.3. Group demonstrations** Over the course of the Six Week Challenge, a special 'show and tell' session consisting of a gathering of all of the students and lecturers involved in the project was organised for the end of every week, so that students could demonstrate the week's results to the other groups of their cohort, as well as to members of the faculty. This was primarily a student-driven activity: while lecturers had the opportunity to provide feedback on the work of the students, the student demonstration sessions focused on students commenting on the work of others. Most importantly, it allowed groups to demonstrate any innovative features that they had implemented over the course of the previous week. We believe that the fostering of this type of constructive competition between groups was a major contributing factor in motivating them to seek new and interesting features to be demonstrated the following week.

#### 4. Evaluation

The previous sections all have an evaluative aspect, in indicating the gains accruing from the Six Week Challenge for the teaching of the discipline represented in each week. This suggests that ALL has definite advantages over more traditional teaching methods. In terms of overall evaluation, a range of anonymous surveys were carried out, and the Six Week Challenge was also externally evaluated, concluding that the Six Week Challenge 'potentially represents one of the most

interesting developments in PjBL across the UK' [Gra10] (Graham refers to PBL as PjBL—Project Based Learning). The external expert 'was particularly impressed by the extent of the students' awareness and understanding of the active learning approach that had been adopted. Hearing them reflect on their own learning, it was clear that this awareness was an important element of their development through the 6-week activity' [Gra09].

In a faculty-wide student survey of the activities offered by the 11 subject groups in the EC faculty, conducted at the end of the 2010/2011 group project, our group's project was found to have received the overall best feedback from students [WM11].

This survey asked students some key questions concerning the relevance and importance of their learning to their futures; how far ALL challenges are achievable; if students felt part of a learning community; and whether the workload was right. All these questions met with high average scores of the order of 4 out of 5 (on a 5 point Likert-type scale), indicating high satisfaction. Thus, it appears that, despite the potential disorientation that computer science students can face at degree level, discussed in the introduction (see Section 1), students generally felt what they learned was relevant and important, reported a sense of belonging, and believed the workload was feasible. Students were asked more specifically about their learning: whether or not the ALL experience had developed their subject knowledge; how far the teaching staff encouraged them to learn effectively; and if there were sufficient opportunities to learn from others. Responses to such questions are important to gauge, to see whether the passing of initiative and direction to students that ALL implies results in any difficulties compared to traditional alternatives. The average response scores for these questions were all of the order of 4.3 out of 5, which again indicates high satisfaction.

To complement the questions about learning, questions were also asked about teaching: the extent to which students were satisfied with how they were being taught; how far tutors were available informally and whether the group size and teaching environment were right. Again, these are important questions to ask, especially concerning the more agile, ad-hoc tutor responsiveness required throughout an ALL process, and whether this works compared with the more formal traditional alternatives. Again, the responses are of the order of 4 of 5 and over, indicating high satisfaction.

These scores are gratifying and indicate that students were happy with the teaching and learning that took place during the Six Week Challenge. Importantly, there seems to have been a sense of engagement and involvement which could help mitigate attrition rates, which, as we saw in Section 1, are a problem in degree-level computing education. Graphics is a tough area of computer science, but the Six Week Challenge

**Table 2:** Student responses to the prompt 'Taking part in the six week activity has helped improve my ...'. Results are based on the responses of 56 students from the 2009/2010 cohort and displayed as percentages, where SD, Strongly Disagree; D, Disagree; N, Neutral; A, Agree; SA, Strongly Agree.

Improvement	SD(%)	D(%)	N(%)	A(%)	SA(%)
Problem solving	0	5	17	71	7
Team-working	0	7	8	46	39
Communication	0	0	15	56	29
Time-management	0	7	27	29	37
Self confidence	0	2	34	49	15
Analytical & critical abilities	0	2	35	56	7

indicates that if an ALL approach is taken, graphics plus linked relevant disciplines can be effectively taught with high satisfaction at this level.

The results of the EC faculty survey are very similar to a further anonymous survey that we conducted of the students of the 2009/2010 cohort in our subject group, for which the students' responses were also highly positive. We have been particularly concerned to track how students reflect on their own learning during the 6 week period, particularly in the absence of traditional lectures and tutorials (Table 2). Asked if they would recommend this type of learning to other students, 98% of our 2009/2010 cohort agreed that they would.

'The Six Week Challenge began as difficult and uncertain but the results showed our potential. This was a triumph' (Student feedback on the 2009/2010 activity).

## 5. Discussion

Our student-centred, activity-led introduction to creative computing through the development of a simple, yet intriguing interactive computer graphics application, appears to have achieved its aims. Over the course of the six weeks, we observed the transformation in our students from 'nervous and unsure' to 'confident and proud' as they became increasing capable communicators. The group presentations at the end of each week especially were an arena where the groups competed in terms of the features and capabilities of their product. Indeed, we believe that this competitive atmosphere was crucial to driving student effort and engagement, allowing us to forego assessment as a means of motivation.

We have found that:

- by introducing students to all components of their course in a concentrated short term exercise, they are better able

to assess quickly what the coming 3 years will involve in terms of content and approach.

- by working in small groups alongside, and supported by, the teaching team, students are rapidly introduced to our academic community. This is further enhanced by social activities which help to develop a strong sense of cohort identity.
- by focussing on activity and production, students are introduced to the practical nature of their subject and, by example, realise that their learning will be active, rather than passive, and that the production of technically sound artefacts will be a predominant feature of their course.

One reason for the success may be the novelty effect of our approach, which Vernon and Blake believed to be a possible factor of the success of PBL, as ‘participating in something new and different . . . may create positive attitudes by psychological mechanisms that are unrelated to the theory, content, or learning objectives’ [VB93]. However, a review conducted by the EC faculty of the six week project designed by our group has led to our project being characterised as a ‘true “high impact” activity’ [WM11] as described by the US National Survey of Student Engagement [Nat07], which could explain the success that this project seems to have had with the participating students.

### 5.1. High points

We have experienced a number of other positive outcomes. Our first year students have retained a significant degree of group cohesion throughout the year, organising social events and often speaking with one voice on issues that affect them. The early use of group-based activities, which can provide a social support structure that helps to retain students who might otherwise consider leaving their degree programme, is likely to be one factor that has influenced this apparent success. Furthermore, many students have retained some of the good habits they learned in the six week group project, particularly in academic writing, and assessments submitted by the students so far appear to be of a better quality than previously observed. Furthermore, the students appear more amenable to challenging material and tasks than in previous years. Finally, the introduction of the Six Week Challenge has coincided with a significant improvement in first year student retention. In the 2010/2011 academic year we have suffered no early withdrawals and at the end of this academic year we expect year 1 retention to be over 90%.

### 5.2. Comparison of the 2009/2010 and 2010/2011 challenges

Although the 2009/2010 media player challenge involved some notable computer graphics aspects, particularly relating to the visualisation of the audio component, the graphics

component was not core to the functionality of the player and the implementation of 3D graphical effects was voluntary. In contrast, the nature of the 2010/2011 3D etch-a-sketch meant that 3D graphics programming formed a mandatory core of the challenge, necessitating the use of linear algebra and transformations to define and animate interactive scenes. Computer graphics and mathematics lie at the core of the students’ degree programmes; it is essential for the games technology course and important for multimedia computing. The 2010/2011 3D etch-a-sketch challenge therefore seemed more relevant to the students’ degrees, providing a better practical introduction to the use of mathematics and programming for defining 3D scenes and interactive animations.

### 5.3. Issues for further consideration

The Six Week Challenge is highly resource intensive both in terms of staffing, accommodation and technology. One important observation that should be taken seriously is that despite our approach’s expectation that the students should demonstrate initiative and solve the set challenges on their own, this does not imply reduced responsibility or workload on behalf of faculty involved in preparing the challenges and developing teaching materials for the sessions that are led by an instructor. In the 2010/2011 Six Week Challenge, the project involved 6 academics and 2 teaching assistants working with a cohort of 54 students and it is unclear how well this activity would ‘scale up’ for larger cohorts, especially as the instructors need to closely monitor the students’ progress to ensure that the learning goals are met. As the student groups have freedom in the way in which they approach any task, their solution may very well miss a specific aspect of vital importance to the outcome of their activities and instructors must watch for these ‘wrong turns’ and if the need should arise, make the students aware of potential problems with their chosen approach. One of the more demanding aspects of the Six Week Challenge for the support staff (besides the physical requirements of extensive ad-hoc student support) was ensuring that each member of the student groups was participating as much as possible, and it was not uncommon to find some students trying to avoid doing parts of the tasks they did not enjoy by taking a back seat. Generally this could be rectified by engaging these students and trying to get them to think about the problem faced by the group and to provide input. This monitoring was not implemented as a formal process or assessment, but as part of the close relationship developed between groups and their personal tutor. Additionally due to the problem-based, self discovery structure of the Six Week Challenge, support staff would often find the demand for guidance from the students would fluctuate throughout the week depending on the overall complexity of the task. One issue that did become apparent during the programming element of the project was that we found that within the groups a minority of the students had previous



experience with programming, resulting in these students tending to take on the majority of the workload in this area. This often caused a divide in the group and would further isolate the students who were new to computer science. The main solution to this was for staff to encourage the students to share knowledge with the group and for the less technically experienced members to understand that even small contributions to technical aspects, coupled with the experience of seeing a software project develop, was beneficial. Despite these efforts, some students did still become disillusioned during this activity. This could be addressed by running optional programming orientation sessions for students who are completely new to computer science.

Finally, we have found that student expectations are significantly higher at the end of the Six Week Challenge, in terms of pace and direction of their degree programme. Management of these expectations can be problematic as the students return to more traditional classroom formats. The delivery of the latter has also been affected by the Six Week Challenge, as a side effect of the suspension of regular teaching activities for the duration of the project has been the need to redesign courses which started after the project which now have to run within a shorter time frame. During each academic year, students are asked to complete a feedback questionnaire on each module of study. When the data for this year is available, we intend to examine how the attitudes of students might have changed in comparison to previous years. Combined with a study of the performance of the year group, we hope to have a much better understanding of the longer-term effects of this kind of introductory programme.

## 6. Conclusions

Our mode of delivery has very much followed the concept of activity-led instruction, which in this context refers to the instruction of students on how to embrace the ALL process. At the introduction for every sub-challenge (Section 3.2), exemplar-based activity sessions were organised with the primary purpose of familiarising students with the process, rather than the task's content *per se*. Students were thus provided with a concrete, real-world example of the processes involved in addressing the challenges, eventually turning them into pro-active problem solvers who were not 'afraid' to face new problem domains. In this respect the weekly 'show and tell' sessions were also highly useful, as the competition they instilled between the different student groups prompted many students to independently investigate different techniques, which they then disseminated among their peers—effectively students took on the role of instructors. The evaluation of the students' experience during the Six Week Challenge suggests that students have reflected on themselves and their learning and the reasons for which they enrolled at university, which in itself is a positive outcome of the six week group project.

## Acknowledgements

The authors would like to thank the creative computing students who achieved much more than we had anticipated. Furthermore we wish to thank Sarah Wilson-Medhurst, who conducted the faculty-wide survey for the evaluation of the six week group project in the 11 subject groups of the EC faculty. The screenshots shown in this article show the projects of two of the student groups. Figure 2 shows the program developed by the student group *Clumsy Penguin Entertainment* (Sareena Hussain, Sennel Ionus, Charnjeet Kaur, Jaipreet Panesar, Shaun Richardson, Anthony Rickhuss and Sarah Wardle). Figure 4 shows the program developed by the group *DDG* (Ahsan Ahmed, Sean Bhadrinath, William Brady, Thomas Bridger, Constantin Cercel and Ian Evans).

## References

- [ACM06] ACM – ASSOCIATION FOR COMPUTING MACHINERY, INC: Computing disciplines & majors. ACM Computing Careers Website: <http://computingcareers.acm.org/> (2006). Accessed 18 March 2012.
- [AM93] ALBANESE M., MITCHELL S.: Problem based learning: A review of literature on its outcomes and implementation issues. *Academy of Medicine* 68, 1 (1993), 52–81.
- [AM06] ANDERSON E., MCLOUGHLIN L.: Do robots dream of virtual sheep: Rediscovering the karel the robot paradigm for the plug&play generation. In *Proceedings of the Fourth Game Design and Technology Workshop and Conference (GDTW 2006)* (Liverpool, UK, 2006), pp. 92–96.
- [AM07] ANDERSON E., MCLOUGHLIN L.: Critters in the classroom: A 3D computer-game-like tool for teaching programming to computer animation students. In *Proceedings of the ACM SIGGRAPH 2007 Educators Program* (San Diego, CA, 2007).
- [AP09] ANDERSON E. F., PETERS C. E.: On the provision of a comprehensive computer graphics education in the context of computer games: An activity-led instruction approach. In *Proceedings of the Eurographics 2009 - Education Papers* (Munich, Germany, 2009), G. Domik and R. Scateni (Eds.), Eurographics Association, pp. 7–14.
- [BBA09] BRUNSTEIN A., BETTS S., ANDERSON J.: Practice enables successful learning under minimal guidance. *Journal of Educational Psychology* 101, 4 (2009), 790–802.
- [BFG\*00] BARG M., FEKETE A., GREENING T., HOLLANDS O., KAY J., KINGSTON J., CRAWFORD K.: Problem-based learning for foundation computer science courses. *Computer Science Education* 10, 2 (2000), 109–128.
- [BG09] BURGESS J., GREEN J.: *YouTube: Online Video and Participatory Culture*. Polity press, Cambridge, UK, 2009.

- [BM05] BEAUBOUF T., MASON J.: Why the high attrition rate for computer science students: Some thoughts and observations. *ACM SIGCSE Bulletin* 37, 2 (2005), 103–106.
- [BMTM09] BATAGELJ B., MAROVT J., TROHA M., MAHNIC D.: Digital airbrush. In *Proceedings of the 51st International Symposium ELMAR-2009* (Zadar, Croatia, 2009), pp. 305–308.
- [BTT05] BENYON D., TURNER P., TURNER S.: *Designing Interactive Systems*. Addison Wesley, Harlow, UK, 2005.
- [Bux86] BUXTON W.: There's more to interaction than meets the eye: Some issues in manual input. In *User Centered System Design: New Perspectives on Human-Computer Interaction*, D. Norman, S. Draper (Eds.). Lawrence Erlbaum Associates, Hillsdale, NJ, 1986, pp. 319–337.
- [Cam96] CAMP G.: Problem-based learning: A paradigm shift or a passing fad? *Medical Education Online* 1, 2 (1996).
- [Car06] CARTER L.: Why students with an apparent aptitude for computer science don't choose to major in computer science. *ACM SIGCSE Bulletin* 38, 1 (2006), 27–31.
- [CC09] CASE C., CUNNINGHAM S.: Teaching computer graphics in context—CGE 09 workshop report. Available from: <http://education.siggraph.org/conferences/eurographics/eurographics-2009-computer-graphics-education-09-workshop/teaching-computer-graphics-in-context/> (2009).
- [CGSG04] CRAIG S., GRAESSER A., SULLINS J., GHOLSON B.: Affect and learning: An exploratory look into the role of affect in learning. *Journal of Educational Media* 29 (2004), 241–250.
- [Cli10] CLIFTON B.: *Advanced Web Metrics with Google Analytics* (2nd edition). SYBEX Inc., Alameda, CA, 2010.
- [Cra03] CRABTREE A.: *Designing Collaborative Systems: A Practical Guide to Ethnography*. Springer-Verlag, Secaucus, NJ, 2003.
- [Csi90] CSIKSZENTMIHALYI M.: *Flow: The Psychology of Optimal Experience*. Harper and Row, NY, 1990.
- [Cun99] CUNNINGHAM S.: Re-inventing the introductory computer graphics course: Providing tools for a wider audience. In *GVE '99: Proceedings of the Graphics and Visualization Education Workshop* (Coimbra, Portugal, 1999), pp. 45–50.
- [DFAB03] DIX A., FINLAY J. E., ABOARD G. D., BEALE R.: *Human-Computer Interaction* (2nd edition) Prentice-Hall, Upper Saddle River, NJ, 2003.
- [DG06] DOMIK G., GOETZ F.: A breadth-first approach for teaching computer graphics. In *Proceedings of the EG Education Papers* (Vienna, Austria, 2006), pp. 1–5.
- [Fel96] FELTON J.: Problem-based learning as a training modality in the occupational medicine curriculum. *Occupational Medicine-Oxford* 46, 1 (1996), 5–11.
- [FW10] FURMAN B., WERTZ E.: A first course in computer programming for mechanical engineers. In *Proceedings of the IEEE/ASME International Conference on Mechatronics and Embedded Systems and Applications (MESA)* (Qingdao, Shandong, 2010), pp. 70–75.
- [Ger04] GERODIMOS R.: How to present at conferences: A guide for graduate students. *PSA Graduate Network Newsletter (October 2004)* (2004), 13–16.
- [Gon00] GONZALEZ R.: Disciplining multimedia. *Multimedia, IEEE* 7, 3 (2000), 72–78.
- [Gra09] GRAHAM R.: Personal communication, 2009.
- [Gra10] GRAHAM R.: UK approaches to engineering project-based learning. White Paper sponsored by the Bernard M, Gordon MIT Engineering Leadership Program, MIT, Boston, MA, 2010.
- [Gri09] GRISWOLD W.: How to read an engineering research paper. Available at: <http://cseweb.ucsd.edu/users/wgg/CSE210/howtoread.html> (2009). Accessed 18 March 2012.
- [Hor10] HORNSBY P.: Hierarchical task analysis. UX Matters. Available at: <http://www.uxmatters.com/mt/archives/2010/02/hierarchical-task-analysis.php> (2010). Accessed 18 March 2012.
- [HS04] HMELO-SILVER C.: Problem-based learning: What and how do students learn? *Educational Psychology Review* 16, 3 (2004), 235–266.
- [HSDC07] HMELO-SILVER C., DUNCAN R., CHINN C.: Scaffolding and achievement in problem-based and inquiry learning: A response to Kirschner, Sweller, and Clark (2006). *Educational Psychologist* 42, 2 (2007), 99–107.
- [IJP\*08] IQBAL R., JAMES A., PAYNE L., ODETAYO M., AROCHENA H.: Moving to activity-led-learning in computer science. In *Proceedings of iPED 2008* (Coventry, UK, 2008).
- [KJCG08] KENNEDY G. E., JUDD T. S., CHURCHWARD A., GRAY K.: First year students' experiences with technology: Are they really digital natives? *Australasian Journal of Educational Technology* 24, 1 (2008), 108–122.

- [KSC06] KIRSCHNER P. A., SWELLER J., CLARK R. E.: Why minimal guidance during instruction does not work: An analysis of the failure of constructivist, discovery, problem-based, experiential, and inquiry-based teaching. *Educational Psychologist* 41, 2 (2006), 75–86.
- [KWC09] KNÖRIG A., WETTACH R., COHEN J.: Fritzing: a tool for advancing electronic prototyping for designers. In *Proceedings of the 3rd International Conference on Tangible and Embedded Interaction* (Cambridge, UK, 2009), pp. 351–358.
- [Lam86] LAMPORT L.: *Latex: a document preparation system*. Addison-Wesley Longman Publishing Co., Inc., Boston, MA, 1986.
- [Lar09] LARAMEE R. S.: How to write a visualization research paper: The art and mechanics. In *Proceedings of the EG 2009 - Education Papers* (Munich, Germany, 2009), pp. 59–66.
- [LR88] LEVIN R., REDELL D. D., : An evaluation of the ninth sosp submissions or how (and how not) to write a good systems paper. *SIGGRAPH Computer Graphics* 22 (1988), 264–266.
- [Mer07] MERRILL M.: A task-centered instructional strategy. *Journal of Research on Technology in Education* 40, 1 (2007), 33–50.
- [MGJ06] MARTÍ E., GIL D., JULIÀ C.: A PBL experience in the teaching of computer graphics. *Computer Graphics Forum* 25, 1 (2006), 95–103.
- [MK02] MCCOWAN J., KNAPPER C.: An integrated and comprehensive approach to engineering curricula, part one: Objectives and general approach. *International Journal of Engineering Education* 18, 6 (2002), 633–637.
- [Nat07] NATIONAL SURVEY OF STUDENT ENGAGEMENT: *Experiences that matter: Enhancing student learning and success*. NSSE Annual Report 2007, 2007.
- [Nie93] NIELSEN J.: *Usability Engineering*. Morgan Kaufman, San Francisco, CA, 1993.
- [PBTF09] PELLICER J. L., BLANES J. S., TORMOS P. M., FRAU D. C.: Using processing.org in an Introductory Computer Graphics Course. In *Proceedings of the Eurographics 2009 - Education Papers* (Munich, Germany, 2009), pp. 23–28.
- [Per92] PERELMAN L.: *School's out: hyperlearning, the new technology, and the end of education*. William Morro, NY, 1992.
- [PJB\*10] POOLE N., JINKS R., BATE S., OLIVER M., BLAND C.: An activity led learning experience for first year electronic engineers. In *Proceedings of the 2010 Engineering Education (EE2010) Conference* (Birmingham, UK, 2010).
- [Pre01] PRENSKY M.: Digital natives, digital immigrants part 1. *On the Horizon* 9 (2001), 1–6.
- [RF06] REAS C., FRY B.: Processing: Programming for the media arts. *AI&Society* 20 (2006), 526–538.
- [Roy10] ROYAL SOCIETY: Current ICT and computer science in schools: Damaging to UK's future economic prospects? Available at: <http://royalsociety.org/Current-ICT-and-Computer-Science-in-schools/> (2010). Accessed 18 March 2012.
- [RSP11] ROGERS Y., SHARP H., PREECE J.: *Interaction Design: Beyond Human - Computer Interaction* (3rd edition) Wiley Publishing, Chichester, UK, 2011.
- [SBM04] SAVIN-BADEN M., MAJOR C.: *Foundations of Problem Based Learning*. Open University Press, Buckingham, UK, 2004.
- [Sch96] SCHULMAN E.: How to write a scientific paper. *Annals of Improbable Research* 2, 5 (1996), 8–9.
- [SD95] SAVERY J., DUFFY T.: Problem based learning: An instructional model and its constructivist framework. *Educational Technology* 35, 5 (1995), 31–38.
- [SEA\*10] SHUTTLEWORTH J., EVERY P., ANDERSON E., HALLORAN J., PETERS C., LIAROKAPIS F.: Press play: An experiment in creative computing using a novel pedagogic approach. *Anglo Higher* 2, 1 (2010), 23–24.
- [SKC07] SWELLER J., KIRSCHNER P., CLARK R.: Why minimally guided teaching techniques do not work: A reply to commentaries. *Educational Psychologist* 42, 2 (2007), 115–121.
- [Sto09] STORNI C.: The ambivalence of engaging technology: Artifacts as products and processes. In *Proceedings of the NORDIC Design Research Conference* (Oslo, Norway, 2009).
- [Tuc96] TUCKER A. B., : Strategic directions in computer science education. *ACM Computing Surveys* 28 (1996), 836–845.
- [Tud92] TUDGE J.: Processes and consequences of peer collaboration: A vygotskian analysis. *Child development* 63 (1992), 1364–1379.
- [VB93] VERNON D., BLAKE R.: Does problem-based learning work? A meta-analysis of evaluative research. *Academic Medicine* 68, 7 (1993), 550–563.
- [VW00] VANDERBERG S., WOLLOWSKI W.: Introducing computer science using a breadth-first approach and functional

- programming. In *Proceedings of SIGCSE 2000* (Austin, TX, 2000).
- [WM08] WILSON-MEDHURST S.: Towards sustainable activity led learning innovations in teaching, learning and assessment. In *Proceedings of the 2008 Engineering Education (EE2008) Conference* (Loughborough, UK, 2008).
- [WM11] WILSON-MEDHURST S.: Key findings from 2010/11 first year UG first integrative ALL experience. Unpublished Report, Faculty of Engineering and Computing, Coventry University, 2011.
- [YG96] YATES W., GERDES T.: Problem-based learning in consultation psychiatry. *General Hospital Psychiatry* 18, 3 (1996), 139–144.

### **8.15 Paper #15**

Anderson, E.F., McLoughlin, L., Liarokapis, F., Peters, C., Petridis, P., de Freitas, S. Developing serious games for cultural heritage: a state-of-the-art review, *Virtual Reality*, Springer, 14(4): 255-275, 2010.

Contribution (20%): Write-up of the serious games, virtual and augmented reality sections of the paper. Also co-written the introduction and conclusions.

# Developing serious games for cultural heritage: a state-of-the-art review

Eike Falk Anderson · Leigh McLoughlin ·  
Fotis Liarokapis · Christopher Peters ·  
Panagiotis Petridis · Sara de Freitas

Received: 14 September 2009 / Accepted: 12 October 2010 / Published online: 16 November 2010  
© Springer-Verlag London Limited 2010

**Abstract** Although the widespread use of gaming for leisure purposes has been well documented, the use of games to support cultural heritage purposes, such as historical teaching and learning, or for enhancing museum visits, has been less well considered. The state-of-the-art in serious game technology is identical to that of the state-of-the-art in entertainment games technology. As a result, the field of serious heritage games concerns itself with recent advances in computer games, real-time computer graphics, virtual and augmented reality and artificial intelligence. On the other hand, the main strengths of serious gaming applications may be generalised as being in the areas of communication, visual expression of information, collaboration mechanisms, interactivity and entertainment. In this report, we will focus on the state-of-the-art with respect to the theories, methods and technologies used in serious heritage games. We provide an overview of existing literature of relevance to the domain, discuss the strengths and weaknesses of the described methods and point out unsolved problems and challenges. In addition, several case studies illustrating the application of methods and technologies used in cultural heritage are presented.

**Keywords** Cultural heritage · Serious games · Computer games technology

## 1 Introduction

Computer games with complex virtual worlds for entertainment are enjoying widespread use, and in recent years we have witnessed the introduction of serious games, including the use of games to support cultural heritage purposes, such as historical teaching and learning, or for enhancing museum visits. At the same time, game development has been fuelled by dramatic advances in computer graphics hardware—in turn driven by the success of video games—which have led to a rise in the quality of real-time computer graphics and increased realism in computer games. The successes of games that cross over into educational gaming—or serious gaming, such as the popular *Civilization* (although “abstract and ahistorical” (Apperley 2006)) and *Total War* series of entertainment games, as well as games and virtual worlds that are specifically developed for educational purposes, such as *Revolution* (Francis 2006) and the *Virtual Egyptian Temple* (Jacobson and Holden 2005), all of which exist within a cultural heritage context, reveal the potential of these technologies to engage and motivate beyond leisure time activities.

The popularity of video games, especially among younger people, makes them an ideal medium for educational purposes (Malone and Lepper 1987). As a result, there has been a trend towards the development of more complex, serious games, which are informed by both pedagogical and game-like, fun elements. The term ‘serious games’ describes a relatively new concept, computer games that are not limited to the aim of providing entertainment, that allow for collaborative use of 3D spaces that

---

E. F. Anderson (✉) · F. Liarokapis · C. Peters  
Interactive Worlds Applied Research Group (iWARG),  
Coventry University, Coventry, UK  
e-mail: eikea@siggraph.org

L. McLoughlin  
The National Centre for Computer Animation (NCCA),  
Bournemouth University, Bournemouth, UK

P. Petridis · S. de Freitas  
Serious Games Institute (SGI), Coventry University,  
Coventry, UK

are used for learning and educational purposes in a number of application domains. Typical examples are game engines and online virtual environments that have been used to design and implement games for non-leisure purposes, e.g. in military and health training (Macedonia 2002; Zyda 2005), as well as cultural heritage (Fig. 1).

This report explores the wider research area of interactive games and related applications with a cultural heritage context and the technologies used for their creation. Modern games technologies (and related optimisations (Chalmers and Debattista 2009) allow the real-time interactive visualisation/simulation of realistic virtual heritage scenarios, such as reconstructions of ancient sites and monuments, while using relatively basic consumer machines. Our aim is to provide an overview of the methods and techniques used in entertainment games that can potentially be deployed in cultural heritage contexts, as demonstrated by particular games and applications, thus making cultural heritage much more accessible.

Serious games can exist in the form of mobile applications, simple Web-based solutions, more complex ‘mashup’ applications (e.g. combinations of social software applications) or in the shape of ‘grown-up’ computer games, employing modern games technologies to create virtual worlds for interactive experiences that may include socially based interactions, as well as mixed reality games that combine real and virtual interactions, all of which can be used in cultural heritage applications. This state-of-the-art report focuses on the serious games technologies that can be found in modern computer games.

The report is divided into two main sections:

- The first of these is concerned with the area of cultural heritage and serious games, which integrate the core technologies of computer games with principled pedagogical methodologies. This is explored in a range of characteristic case studies, which include entertainment games that can be used for non-leisure purposes as well

as virtual museums and educationally focused and designed cultural heritage projects.

- The second part investigates those computer games technologies that are potentially useful for the creation of cultural heritage games, such as real-time rendering techniques, mixed reality technologies and subdomains of (game) artificial intelligence. This literature review includes discussions of strengths and weaknesses of the most prominent methods, indicating potential uses for cultural heritage serious games and illustrating challenges in their application.

## 2 The state-of-the-art in serious games

The state-of-the-art in Serious Game technology is identical to the state-of-the-art in Entertainment Games technology. Both types of computer game share the same infrastructure, or as Zyda notes, “applying games and simulations technology to non-entertainment domains results in serious games” (Zyda 2005). The main strengths of serious gaming applications may be generalised as being in the areas of communication, visual expression of information, collaboration mechanisms, interactivity and entertainment.

Over the past decade, there have been tremendous advances in entertainment computing technology, and “today’s games are exponentially more powerful and sophisticated than those of just three or four years ago” (Sawyer 2002), which in turn is leading to very high consumer expectations. Real-time computer graphics can achieve near-photorealism and virtual game worlds are usually populated with considerable amounts of high quality content, creating a rich user experience. In this respect, Zyda (2005) argues that while pedagogy is an implicit component of a serious game, it should be secondary to entertainment, meaning that a serious game that is not ‘fun’ to play would be useless, independent of its pedagogical content or value. This view is not shared by all, and there exist design methodologies for the development of games incorporating pedagogic elements, such as the four-dimensional framework (de Freitas and Oliver 2006), which outlines the centrality of four elements that can be used as design and evaluation criteria for the creation of serious games. In any case, there is a need for the game developers and instructional designers to work together to develop engaging and motivating serious games for the future.

### 2.1 Online virtual environments

There is a great range of different online virtual world applications—at least 80 virtual world applications existed



**Fig. 1** ‘Roma Nova’—experiencing ‘Rome Reborn’ as a game

in 2008 with another 100 planned for 2009. The field is extensive, not just in terms of potential use for education and training but also in terms of actual usage and uptake by users, which is amply illustrated by the online platform Second Life (Linden Labs), which currently has 13 million registered accounts worldwide. The use of Second Life for supporting seminar activities, lectures and other educational purposes has been documented in a number of recent reports and a wide range of examples of Second Life use by UK universities has been documented (Kirriemuir 2008). Online virtual worlds provide excellent capabilities for creating effective distance and online learning opportunities through the provision of unique support for distributed groups (online chat, the use of avatars, document sharing, etc.). This benefit has so far been most exploited in business where these tools have been used to support distributed or location-independent working groups or communities (Jones 2005). Online virtual worlds in this way facilitate the development of new collaborative models for bringing together subject matter experts and tutors from around the world, and in terms of learning communities are opening up opportunities for learning in international cohorts where students from more than one country or location can learn in mixed reality contexts including classroom and non-classroom based groups (<https://lg3d-wonderland.dev.java.net>). Online virtual worlds also notably offer real opportunities for training, rehearsing and role playing.

## 2.2 Application to cultural heritage: case studies

This section provides an overview of some of the most characteristic case studies in cultural heritage. In particular, the case studies have been categorised into three types of computer-game-like applications including: prototypes and demonstrators, virtual museums and commercial historical games.

### 2.2.1 Prototypes and demonstrators

The use of visualisation and virtual reconstruction of ancient historical sites is not new, and a number of projects have used this approach to study crowd modelling (Arnold et al. 2008; Maim et al. 2007). Several projects are using virtual reconstructions in order to train and educate their users. Many of these systems have, however, never been released to the wider public, and have only been used for academic studies. In the following section, the most significant and promising of these are presented.

**2.2.1.1 Roma Nova** The Rome Reborn project is the world's largest digitisation project and has been running for 15 years. The main aims of the project are to produce a

high-resolution version of Rome at 320 AD (Fig. 2), a lower resolution model for creating a 'mashup' application with 'Google Earth' (<http://earth.google.com/rome/>), and finally the collaborative mode of the model for use with virtual world applications and aimed primarily at education (Frischer 2008).

In order to investigate the efficacy of the Rome Reborn Project for learning, exploration, re-enactment and research of cultural and architectural aspects of ancient Rome the serious game 'Roma Nova' is currently under development. In particular, the project aims at investigating the suitability of using this technology to support the archaeological exploration of historically accurate societal aspects of Rome's life, with an emphasis on political, religious and artistic expressions.

To achieve these objectives, the project will integrate four cutting-edge virtual world technologies with the Rome Reborn model, the most detailed three-dimensional model of Ancient Rome available. These technologies include:

- the Quest3D visualisation engine (Godbersen 2008)
- Instinct(maker) artificial life engine (Toulouse University) (Sanchez et al. 2004)
- ATOM Spoken Dialogue System (<http://www.agilingua.com>)
- High-resolution, motion-captured characters and objects from the period (Red Bedlam).

The use of the Instinct artificial life engine enables coherent crowd animation and therefore the population of the city of Rome with behaviour-driven virtual characters. These virtual characters with different behaviours can teach the player about different aspects of life in Rome (living conditions, politics, military) (Sanchez et al. 2004). Agilingua ATOM's dialogue management algorithm allows determining how the system will react: asking questions, making suggestions, and/or confirming an answer.



**Fig. 2** 'Rome Reborn' serious game



This project aims to develop a researchers' toolkit for allowing archaeologists to test past and current hypotheses surrounding architecture, crowd behaviour, social interactions, topography and urban planning and development, using Virtual Rome as a test-bed for reconstructions. By using such a game the researches will be able to analyse the impact of major events. For example, the use of this technique would allow researchers to analyse the impact of major events, such as grain distribution or the influx of people into the city. The experiences of residents and visitors as they pass through and interact with the ancient city can also be explored.

**2.2.1.2 Ancient Pompeii** Pompeii was a Roman city, which was destroyed and completely buried in the first recorded eruption of the volcano Mount Vesuvius in 79 AD (Plinius 79a, Plinius 79b). For this project, a model of ancient Pompeii was constructed using procedural methods (Müller et al. 2005) and subsequently populated with avatars in order to simulate life in Pompeii in real time. The main goal of this project was to simulate a crowd of virtual Romans exhibiting realistic behaviours in a reconstructed district of Pompeii (Maim et al. 2007). The virtual entities can navigate freely in several buildings in the city model and interact with their environment (Arnold et al. 2008).

**2.2.1.3 Parthenon Project** The Parthenon Project is a short computer animation that “visually reunites the Parthenon and its sculptural decorations” (Debevec 2005). The Parthenon itself is an ancient monument, completed in 437 BC, and stands in Athens, while many of its sculptural decorations reside in the collection of the British Museum, London (UK). The project goals were to create a virtual version of the Parthenon and its separated sculptural elements so that they could be reunited in a virtual representation.

The project involved capturing digital representations of the Parthenon structure and the separate sculptures, recombining them and then rendering the results. The structure was scanned using a commercial laser range scanner, while the sculptures were scanned using a custom 3D scanning system that the team developed specifically for the project (Tchou 2002). The project made heavy use of image-based lighting techniques, so that the structure could be relit under different illumination conditions within the virtual representation. A series of photographs were taken of the structure together with illumination measurements of the scene's lighting. An inverse global illumination technique was then applied to effectively ‘remove’ the lighting. The resulting “lighting-independent model” (Debevec et al. 2004) could then be relit using any lighting scheme desired (Tchou et al. 2004; Debevec et al. 2004).

Although the Parthenon Project was originally an off-line-rendered animation, it has since been converted to work in real-time (Sander and Mitchell 2006; Isidoro and Sander 2006). The original Parthenon geometry represented a large dataset consisting of 90 million polygons (after post-processing), which was reduced to 15 million for the real-time version and displayed using dynamic level-of-detail techniques. Texture data consisted of 300 MB and had to be actively managed and compressed, while 2.1 GB of compressed High-Dynamic Range (HDR) sky maps were reduced in a pre-processing step. The reduced HDR maps were used for lighting, and the extracted sun position was used to cast a shadow map.

## 2.2.2 Virtual museums

Modern interactive virtual museums using games technologies (Jones and Christal 2002; Lepouras and Vassilakis 2004) provide a means for the presentation of digital representations for cultural heritage sites (El-Hakim et al. 2006) that entertain and educate visitors (Hall et al. 2001) in a much more engaging manner than was possible only a decade ago. A recent survey paper that examines all the technologies and tools used in museums was recently published (Sylaiou et al. 2009). Here, we present several examples of this type of cultural heritage serious game, including some virtual museums that can be visited in real-world museums.

**2.2.2.1 Virtual Egyptian Temple** This game depicts a hypothetical Virtual Egyptian Temple (Jacobson and Holden 2005; Troche and Jacobson 2010), which has no real-world equivalent. The temple embodies all of the key features of a typical New Kingdom period Egyptian temple in a manner that an untrained audience can understand. This Ptolemaic temple is divided into four major areas, each one of which houses an instance of the High Priest, a pedagogical agent. Each area of this virtual environment represents a different feature from the architecture of that era.

The objective of the game ‘Gates of Horus’ (Jacobson et al. 2009) is to explore the model and gather enough information to answer the questions asked by the priest (pedagogical agent). The game engine that this system is based on is the Unreal Engine 2 (Fig. 3) (Jacobson and Lewis 2005), existing both as an Unreal Tournament 2004 game modification (Wallis 2007) for use at home, as well as in the form of a Cave Automatic Virtual Environment (CAVE Cruz-Neira et al. 1992) system in a real museum.

**2.2.2.2 The Ancient Olympic Games** The Foundation of the Hellenic World has produced a number of gaming applications associated with the Olympic Games in ancient Greece (Gaitatzes et al. 2004). For example, the ‘Olympic



**Fig. 3** New Kingdom Egyptian temple game

Pottery Puzzle’ exhibit the user must re-assemble a number of ancient vases putting together pot shards. The users are presented with a colour-coded skeleton of the vessels with the different colours showing the correct position of the pieces. They then try to select one piece at a time from a heap and place it in the correct position on the vase. Another game is the ‘Feidias Workshop’, which is a highly interactive virtual experience taking place at the construction site of the 15-m-tall golden ivory statue of Zeus, one of the seven wonders of the ancient world. The visitors enter the two-storey-high workshop and come into sight of an accurate reconstruction of an unfinished version of the famous statue of Zeus and walk among the sculptor’s tools, scaffolding, benches, materials and moulds used to construct it. They take the role of the sculptor’s assistants and actively help finish the creation of the huge statue, by using virtual tools to apply the necessary materials onto the statue, process the ivory and gold plates, apply them onto the wooden supporting core and add the finishing touches. Interaction is achieved using the navigation wand of the Virtual Reality (VR) system, onto which the various virtual tools are attached. Using these tools, the user helps finish the work on the statue, learning about the procedures, materials and techniques applied for the creation of these marvellous statues. The last example is the ‘Walk through Ancient Olympia’, where the user, apart from visiting the historical site, learns about the ancient games themselves by interacting with athletes in the ancient game of pentathlon (Fig. 4). The visitors can wonder around and visit the buildings and learn their history and their function: the Heraion, the oldest monumental building of the sanctuary dedicated to the goddess Hera, the temple of Zeus, a model of a Doric peripteral temple with magnificent sculpted decoration, the Gymnasium, which was used for the training of javelin throwers, discus throwers and runners, the Palaestra, where the wrestlers, jumpers and boxers trained, the Leonidaion, which was where the official guests stayed, the Bouleuterion, where athletes, relatives

and judges took a vow that they would uphold the rules of the Games, the Treasuries of various cities, where valuable offerings were kept, the Philippeion, which was dedicated by Philip II, king of Macedonia, after his victory in the battle of Chaeronea in 338 BC and the Stadium, where most of the events took place. Instead of just observing the games, the visitors take place in them. They can pick up the discus or the javelin and they try their abilities in throwing them towards the far end of the stadium. Excited about the interaction they ask when they will be able to interact with the wrestler one on one. A role-playing model of interaction with alternating roles was tried here with pretty good success as the visitors truly immersed in the environment, wishing they could participate in more games (Gaitatzes et al. 2004).

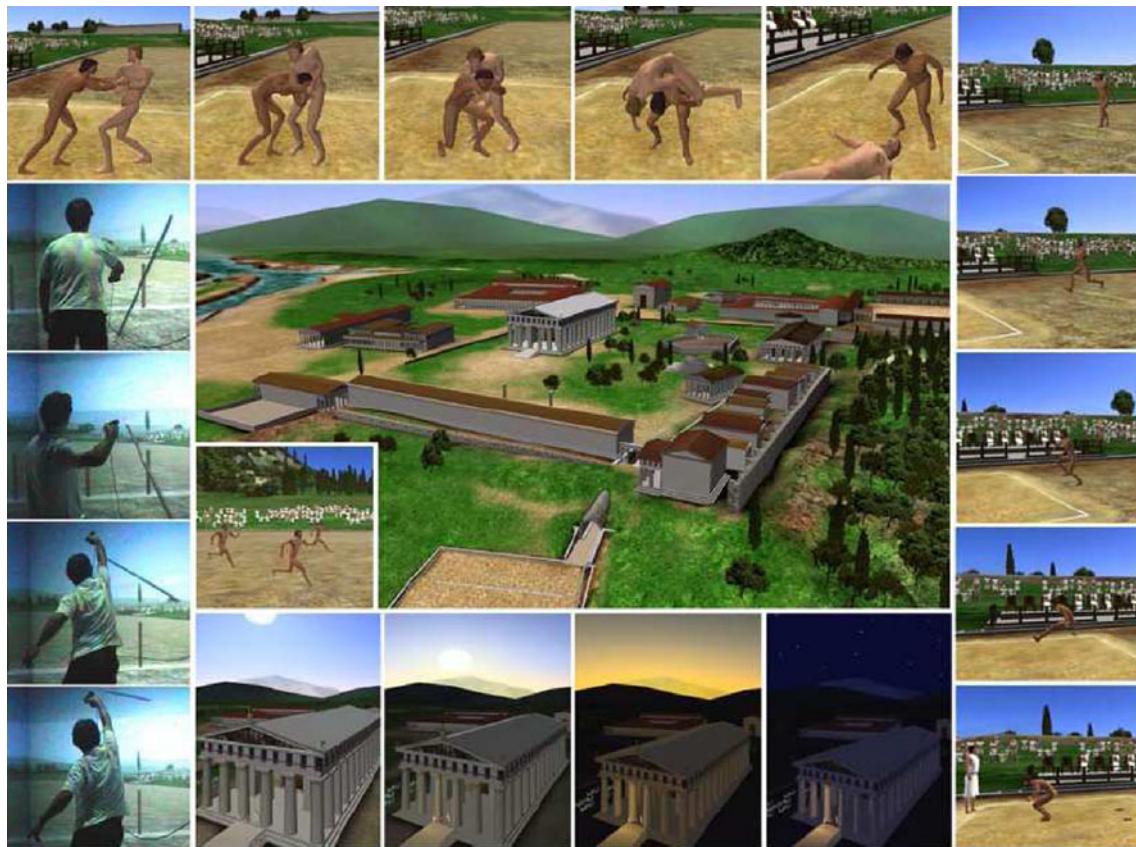
**2.2.2.3 Virtual Priory Undercroft** Located in the heart of Coventry, UK, the Priory Undercrofts are the remains of Coventry’s original Benedictine monastery, dissolved by Henry VIII. Although archaeologists revealed the architectural structure of the cathedral, the current site is not easily accessible for the public. Virtual Priory Undercroft offers a virtual exploration of the site in both online and offline configurations.

Furthermore, a first version of a serious game (Fig. 5) has been developed at Coventry University, using the Object-Oriented Graphics Rendering Engine (OGRE) (Wright and Madey 2008). The motivation is to raise the interest of children in the museum, as well as cultural heritage in general. The aim of the game is to solve a puzzle by collecting medieval objects that used to be located in and around the Priory Undercroft. Each time a new object is found, the user is prompted to answer a question related to the history of the site. A typical user-interaction might take the form of: “What did St. George slay?—Hint: It is a mythical creature. —Answer: The Dragon”, meaning that the user then has to find the Dragon.

### 2.2.3 Commercial historical games

Commercial games with a cultural heritage theme are usually of the ‘documentary game’ (Burton 2005) genre that depict real historical events (frequently wars and battles), which the human player can then partake in. These are games that were primarily created for entertainment, but their historical accuracy allows them to be used in educational settings as well.

**2.2.3.1 History Line: 1914–1918** An early representative of this type of game was History Line: 1914–1918 (Blue Byte 1992), an early turn-based strategy game depicting the events of the First World War. The game was realised using the technology of the more prominent game Battle Isle,



**Fig. 4** Walk through Ancient Olympia (Gaitatzes et al. 2004)



**Fig. 5** Priory Undercroft—a serious game

providing players with a 2D top-down view of the game world, divided into hexagons that could be occupied by military units, with the gameplay very much resembling traditional board games.

The game's historical context was introduced in a long (animated) introduction, depicting the geo-political situation of the period and the events leading up to the outbreak of war in 1914. In between battles the player is provided with additional information on concurrent events that

shaped the course of the conflict, which is illustrated with animations and newspaper clippings from the period.

**2.2.3.2 Great Battles of Rome** More recently, a similar approach was used by the History Channel's Great Battles of Rome (Slitherine Strategies 2007), another 'documentary game', which mixes interactive 3D real-time tactical simulation of actual battles with documentary information (Fig. 6), including footage originally produced for TV documentaries, which places the battles in their historical context.

**2.2.3.3 Total War** The most successful representatives of this type of historical game are the games of the Creative Assembly's Total War series, which provide a gameplay combination of turn-based strategy (for global events) and real-time tactics (for battles). Here, a historical setting is enriched with information about important events and developments that occurred during the timeframe experienced by the player. While the free-form campaigns allow the game's players to change the course of history, the games also include several independent battle-scenarios with historical background information that depict real events and allow players to partake in moments of historical significance.



**Fig. 6** Great Battles of Rome

The use of up-to-date games technology for rendering, as well as the use of highly detailed game assets that are reasonably true to the historical context, enables a fairly realistic depiction of history. As a result, games from the Total War series have been used to great effect in the visualisation of armed conflicts in historical programmes produced for TV (Waring 2007).

The latest titles in the series, ‘Empire: Total War’ (released in 2009), depicting events from the start of the eighteenth century to the middle of the nineteenth century, and ‘Napoleon: Total War’ (released in 2010), depicting European history during the Napoleonic Wars, make use of some of the latest developments in computer games technology (Fig. 7). The games’ renderer is scalable to support different types of hardware, including older systems, especially older graphics cards (supporting the programmable Shader Model 2), but the highest visual fidelity is only achieved on recent systems (Shader Model 3 graphics hardware) (Gardner 2009). If the hardware allows for this, shadows for added realism in the virtual world are



**Fig. 7** Reliving the battle of Brandywine Creek (McGuire 2006) in ‘Empire: Total War’

generated using Screen Space Ambient Occlusion (Mittring 2007; Bavoil and Sainz 2008), making use of existing depth-buffer information in rendered frames. Furthermore the virtual world of the game is provided with realistic vegetation generated by the popular middleware system SpeedTree (Interactive Data Visualization, Inc.), which “features realistic tree models and proves to be able to visualise literally thousands of trees in real-time” (Fritsch and Kada 2004). As a result, the human player is immersed in the historical setting, allowing the player to re-live history.

### 3 The technology of cultural heritage serious games

Modern interactive virtual environments are usually implemented using game engines, which provide the core technology for the creation and control of the virtual world. A game engine is an open, extendable software system on which a computer game or a similar application can be built. It provides the generic infrastructure for game creation (Zyda 2005), i.e. I/O (input/output) and resource/asset management facilities. The possible components of game engines include, but are not limited to the following: rendering engine, audio engine, physics engine, animation engine.

#### 3.1 Virtual world system infrastructure

The shape that the infrastructure for a virtual environment takes is dictated by a number of components, defined by function rather than organisation, the exact selection of which determines the tasks that the underlying engine is suitable for. A game engine does not provide data or functions that could be associated with any game or other application of the game engine (Zerbst et al. 2003). Furthermore, a game engine is not just an API (Application Programming Interface), i.e. a set of reusable components that can be transferred between different games, but also provides a glue layer that connects its component parts. It is this glue layer that sets a game engine apart from an API, making it more than the sum of its components and sub-systems.

Modern game engines constitute complex parallel systems that compete for limited computing resources (Blow 2004). They “provide superior platforms for rendering multiple views and coordinating real and simulated scenes as well as supporting multiuser interaction” (Lewis and Jacobson 2002), employing advanced graphics techniques to create virtual environments. Anderson et al. (2008) provide a discussion of several challenges and open problems regarding game engines, which include the precise definition of the role of content creation tools in the game

development process and as part of game engines, as well as the identification of links between game genres and game engine architecture, both of which play a crucial role in the process of selecting an appropriate game engine for a given project.

Frequently, the technology used for the development of virtual environments, be they games for entertainment, serious games or simulations, is limited by the development budget. Modern entertainment computer games frequently require “a multimillion-dollar budget” (Overmars 2004) that can now rival the budgets of feature film productions, a significant proportion of which will be used for asset creation (such as 3D models and animations). Some of these costs can be reduced through the use of procedural modelling techniques for the generation of assets, including terrain (Noghani et al. 2010), vegetation (Lintermann and Deussen 1999) or whole urban environments (Vanegas et al. 2009). Game developers are usually faced with the choice of developing a proprietary infrastructure, i.e. their own game engine, or to use an existing engine for their virtual world application. Commercially developed game engines are usually expensive, and while there are affordable solutions, such as the Torque game engine which is favoured by independent developers and which has been successfully used in cultural heritage applications (Leavy et al. 2007; Mateevitsi et al. 2008), these generally provide fewer features, thus potentially limiting their usefulness. If one of the project’s requirements is the use of highly realistic graphics with a high degree of visual fidelity, this usually requires a recent high-end game engine, the most successful of which usually come at a very high licensing fee.

There are alternatives, however, as several older commercially developed engines have been released under Open Source licences, such as the Quake 3 engine (id Tech 3) (Smith and Trenholme 2008; Wright and Madey 2008), making them easily accessible, and while they do not provide the features found in more recently published games, they nevertheless match the feature sets of the cheaper commercial engines. Furthermore, there exist open source game engines such as the Nebula Device (Rémond and Mallard 2003), or engine components, such as OGRE (Rémond and Mallard 2003; Wright and Madey 2008) or ODE (Open Dynamics Engine) (Macagon and Wünsche 2003), which are either commercially developed or close to commercial quality, making them a viable platform for the development of virtual worlds, although they may lack the content creation tools that are frequently packaged with larger commercial engines.

Finally, there is the possibility of taking an existing game and modifying it for one’s own purposes, which many recent games allow users to do (Wallis 2007; Smith and Trenholme 2008). This has the benefit of small

up-front costs, as the only requirement is the purchase of a copy of the relevant game, combined with access to high-spec modern game engines, as well as the content development tools that they contain. Examples for this are the use of the game Civilization III for the cultural heritage game The History Game Canada (<http://historycanadagame.com>) or the use of the Unreal Engine 2 (Smith and Trenholme 2008) for the development of an affordable CAVE (Jacobson and Lewis 2005), which has been used successfully in cultural heritage applications (Jacobson and Holden 2005).

### 3.2 Virtual world user interfaces

There are different types of interface that allow users to interact with virtual worlds. These fall into several different categories, such as VR and Augmented Reality (AR), several of which are especially useful for cultural heritage applications, and which are presented in this section.

#### 3.2.1 Mixed reality technologies

In 1994, (Milgram and Kishino 1994) tried to depict the relationship between VR and AR. To illustrate this, he introduced two new terms called Mixed Reality (MR), which is a type of VR but has a wider concept than AR, (Tamura et al. 2001) and Augmented Virtuality (AV). On the left-hand side of the Reality-Virtuality continuum, there is the representation of the real world and on the right-hand side there is the ultimate synthetic environment. MR stretches out in-between these environments, and it can be divided into two sub-categories: AR and AV (Milgram and Kishino 1994). AR expands towards the real world, and thus it is less synthetic than AV which expands towards virtual environments. To address the problem from another perspective, a further distinction has been made. This refers to all the objects that form an AR environment: real objects and virtual objects. Real objects are these, which always exist no matter what the external conditions may be. On the other hand, a virtual object depends on external factors but mimics objects of reality. Some of the most interesting characteristics that distinguish virtual objects, which include holograms and mirror images, and real objects are illustrated below (Milgram and Kishino 1994).

The most obvious difference is that a virtual object can only be viewed through a display device after it has been generated and simulated. Real-world objects that exist in essence, on the contrary, can be viewed directly and/or through a synthetic device. Another factor is the quality of viewed images that are generated using state-of-the-art technologies. Virtual information cannot be sampled directly but must be synthesised, therefore, depending on the chosen resolution, displayed objects may appear real,

but their appearance does not guarantee that the objects are real. Virtual and real information may be distinguished depending on the luminosity of the location that it appears in. Images of real-world objects receive lighting information from the position at which they appear to be located while virtual objects do not necessarily, unless the virtual scene is lit exactly like the real-world location in which objects appear to be displayed. This is true for directly viewed real-world objects, as well as displayed images of indirectly viewed objects.

### 3.2.2 Virtual reality

Ivan Sutherland originally introduced the first Virtual Reality (VR) system in the 1960s (Sutherland 1965). Nowadays VR is moving from the research laboratories to the working environment by replacing ergonomically limited HMDs (Head-Mounted Displays) with projective displays (such as the well known CAVE and Responsive Workbench) as well as online VR communities. In a typical VR system the user's natural sensory information is completely replaced with digital information. The user's experience of a computer-simulated environment is called immersion. As a result, VR systems can completely immerse a user inside a synthetic environment by blocking all the signals of the real world. In addition, a VR simulated world does not always have to obey all laws of nature. In immersive VR systems, the most common problems of VR systems are of emotional and psychological nature including motion sickness, nausea, and other symptoms, which are created by the high degree of immersiveness of the users.

Moreover, internet technologies have the tremendous potential of offering virtual visitors ubiquitous access via the World Wide Web (WWW) to online virtual environments. Additionally, the increased efficiency of Internet connections (i.e. ADSL/broadband) makes it possible to transmit significant media files relating to the artefacts of virtual museum exhibitions. The most popular technology for WWW visualisation includes Web3D which offers tools such as the Virtual Reality Modeling Language (VRML—<http://www.web3d.org/x3d/vrml/>) and its successor X3D (<http://www.web3d.org/x3d/>), which can be used for the creation of an interactive virtual museum. Many cultural heritage applications based on VRML have been developed for the Web (Gatermann 2000; Paquet et al. 2001; Sinclair and Martinez 2001). Another 3D graphics format, is COLLABorative Design Activity (COLLADA – <https://collada.org>) which defines an open standard XML schema (<http://www.w3.org/XML/Schema>) for exchanging digital assets among various graphics software applications that might otherwise store their assets in incompatible formats. One of the main advantages of COLLADA is that

it includes more advanced physics functionality such as collision detection and friction (which Web3D does not support).

In addition to these, there are more powerful technologies that have been used in museum environments, which include the OpenSceneGraph (OSG) high performance 3D graphics toolkit (<http://www.openscenegraph.org/projects/osg>) and a variety of 3D game engines. OSG is a freely available (open source) multi-platform toolkit, used by museums (Calori et al. 2005; Looser et al. 2006) to generate more powerful VR applications, especially in terms of immersion and interactivity since it supports the integration of text, video, audio and 3D scenes into a single 3D environment. An alternative to OpenSceneGraph, is OpenSG, which is an open-source scene graph system used to create real-time VR applications (<http://www.opensg.org/>) On the other hand, 3D game engines are also very powerful and they provide superior visualisation and physics support. Both technologies (OSG and 3D game engines), compared to VRML and X3D, can provide very realistic and immersive museum environments but they have two main drawbacks. First, they require advanced programming skills in order to design and implement custom applications. Secondly, they do not have support for mobile devices such as PDAs and third-generation mobile phones.

### 3.2.3 Augmented reality

The concept of AR is the opposite of the closed world of virtual spaces (Tamura et al. 1999) since users can perceive both virtual and real information. Compared to VR systems, most AR systems use more complex software approaches, usually including some form of computer vision techniques (Forsyth and Ponce 2002) for sensing the real world. The basic theoretical principle is to superimpose digital information directly into a user's sensory perception (Feiner 2002), rather than replacing it with a completely synthetic environment as VR systems do. An interesting point is that both technologies—AR and VR—may process and display the same digital information and that they often make use of identical dedicated hardware. Although AR systems are influenced by the same factors, the amount of influence is much less than in VR since only a portion of the environment is virtual. However, there is still a lot of research to be done in AR (Azuma 1997; Azuma et al. 2001; Livingston 2005) to measure accurately its effects on humans.

The requirements related to the development of AR applications in the cultural heritage field have been well documented (Brogni et al. 1999; Liarokapis et al. 2008; Sylaiou et al. 2009). An interactive concept is the Meta-Museum visualised guide system based on AR, which tries to establish scenarios and provide a communication

environment between the real world and cyberspace (Mase et al. 1996). Another AR system that could be used as an automated tour guide in museums is the automated tour guide, which superimposes audio in the world based on the location of the user (Bederson 1995). There are many ways in which archaeological sources can be used to provide a mobile AR system. Some of the wide range of related applications includes the initial collection of data to the eventual dissemination of information (Ryan 2000). MARVINS is an AR assembly, initially designed for mobile applications and can provide orientation and navigation possibilities in areas, such as science museums, art museums and other historical or cultural sites. Augmented information like video, audio and text is relayed from a server via the transmitter-receiver to a head-mounted display (Sanwal et al. 2000).

In addition, a number of EU projects have been undertaken in the field of virtual heritage. The SHAPE project (Hall et al. 2001) combined AR and archaeology to enhance the interaction of persons in public places like galleries and museums by educating visitors about artefacts and their history. The 3DMURALE project (Cosmas et al. 2001) developed 3D multimedia tools to record, reconstruct, encode and visualise archaeological ruins in virtual reality using as a test case the ancient city of Sagalassos in Turkey. The Ename 974 project (Pletinckx et al. 2000) developed a non-intrusive interpretation system to convert archaeological sites into open-air museums, called TimeScope-1 based on 3D computer technology originally developed by IBM, called TimeFrame. ARCHEOGUIDE (Stricker et al. 2001) provides an interactive AR guide for the visualisation of archaeological sites based on mobile computing, networking and 3D visualisation providing the users with a multi-modal interaction user interface. A similar project is LIFEPLUS (Papagiannakis et al. 2002), which explores the potential of AR so that users can experience a high degree of realistic interactive immersion by allowing the rendering of realistic 3D simulations of virtual flora and fauna (humans, animals and plants) in real-time.

AR technologies can be combined with existing game engine subsystems to create AR game engines (Lugrin and Cavazza 2010) for the development of AR games. AR has been applied successfully to gaming in cultural heritage. One of the earliest examples is the Virtual Showcase (Bimber et al. 2001) which is an AR display device that has the same form factor as a real showcase traditionally used for museum exhibits and can be used for gaming. The potentials of AR interfaces in museum environments and other cultural heritage institutions (Liarokapis 2007) as well as outdoor heritage sites (Vlahakis et al. 2002) have been also briefly explored for potential educational applications. A more specific gaming example are the MAGIC

and TROC systems (Renevier et al. 2004) which were based on a study of the tasks of archaeological fieldwork, interviews and observations in Alexandria. This takes the form of a mobile game in which the players discover archaeological objects while moving.

Another cultural heritage AR application is the serious game SUA that was part of the BIDAIATZERA project (Linaza et al. 2007). This project takes the form of a play which recreates the 1813 battle between the English and the French in San Sebastian. Researchers developed an interactive system based on AR and VR technologies for recreational and educational applications with tourist, cultural and socio-economical contents, the prototype for which was presented at the Museo del Monte Urgull in San Sebastian.

### 3.3 Advanced rendering techniques

One of the most important elements of the creation of interactive virtual environments is the visual representation of these environments. Although serious games have design goals that are different from those of pure entertainment video games, they can still make use of the wide variety of graphical features and effects that have been developed in recent years. The state-of-the-art in this subject area is broad and, at times, it can be difficult to specify exactly where the ‘cutting edge’ of the development of an effect lies. A number of the techniques that are currently in use were originally developed for offline applications and have only recently become adopted for use in real-time applications through improvements in efficiency or hardware. Here, the ‘state-of-the-art’ for real-time lags several years behind that for offline—good examples of this would be raytracing or global illumination, which we shall briefly examine. A number of effects, however, are developed specifically for immediate deployment on current hardware and can make use of specific hardware features—these are often written by hardware providers themselves to demonstrate their use or, of course, by game developers. Other real-time graphical features and effects can be considered to follow a development cycle, where initially they are proven in concept demonstrations or prototypes, but are too computationally expensive to implement in a full application or game. Over time these techniques may then be progressively optimised for speed, or held back until the development of faster hardware allows their use in computer games.

The primary reason for the proliferation of real-time graphics effects has been due to advances in low-cost graphics hardware that can be used in standard PCs or games consoles. Modern graphics processing units (GPUs) are extremely powerful parallel processors and the graphics pipeline is becoming increasingly flexible. Through the use

of programmable shaders, which are small programs that define and direct part of the rendering process, a wide variety of graphical effects are now possible for inclusion in games and virtual environments, while there also exist a range of effects that are currently possible but still too expensive for practical use beyond anything but the display of simple scenes.

The graphics pipeline used by modern graphics hardware renders geometry using rasterisation, where an object is drawn as triangles which undergo viewing transformations before they are converted directly into pixels. In contrast, ray-tracing generates a pixel by firing a corresponding ray into the scene and sampling whatever it may hit. While the former is generally faster, especially using the hardware acceleration on modern graphics cards, it is easier to achieve effects such as reflections using ray-tracing. Although the flexibility of modern GPUs can allow ray-tracing (Purcell et al. 2002) in real-time (Horn et al. 2007; Shirley 2006), as well as fast ray-tracing now becoming possible on processors used in games consoles (Benthin et al. 2006), rasterisation is currently still the standard technique for computer games.

Although the modern graphics pipeline is designed and optimised to rasterise polygonal geometry, it should be noted that other types of geometry exist. Surfaces may be defined using a mathematical representation, while volumes may be defined using ‘3D textures’ of voxels or, again, using a mathematical formula (Engel et al. 2006). The visualisation of volumetric ‘objects’, which are usually semi-opaque, is a common problem that includes features such as smoke, fog and clouds. A wide variety of options exist for rendering volumes (Engel et al. 2006; Cerezo et al. 2005), although these are generally very computationally expensive and it is common to emulate a volumetric effect using simpler methods. This often involves drawing one or more rectangular polygons to which a four-channel texture has been applied (where the fourth, alpha, channel represents transparency)—for example a cloud element or wisp of smoke. These may be aligned to always face the viewer as billboards (Akenine-Möller et al. 2008), a common game technique with a variety of uses (Watt and Policarpo 2005), or a series of these may be used to slice through a full volume at regular intervals. An alternative method for rendering full volumes is ray-marching, where a volume is sampled at regular intervals along a viewing ray, which can now be implemented in a shader (Crassin et al. 2009), or on processors that are now being used in games consoles (Kim and Jaja 2009).

It is sometimes required to render virtual worlds, or objects within worlds, that are so complex or detailed that they cannot fit into the graphics memory, or even the main memory, of the computer—this can be especially true when dealing with volume data. Assuming that the hardware

cannot be further upgraded, a number of options exist for such rendering problems. If the scene consists of many complex objects at varying distances, it may be possible to adopt a level-of-detail approach (Engel et al. 2008) and use less complex geometry, or even impostors (Akenine-Möller et al. 2008), to approximate distant objects (Sander and Mitchell 2006). Alternatively, if only a small subsection of the world or object is in sight at any one time, it may be possible to hold only these visible parts in memory and ‘stream’ replace them as new parts come into view, which is usually achieved by applying some form of spatial partitioning (Crassin et al. 2009). This streaming approach can also be applied to textures that are too large to fit into graphics memory (Mittring and Crytek 2008). If too much is visible at one time for this to be possible, a cluster of computers may be used, where the entire scene is often too large for a single computer to hold in memory but is able to be distributed among the cluster with the computers’ individual renders being accumulated and composited together (Humphreys et al. 2002) or each computer controlling part of a multi-screen tile display (Yin et al. 2006).

### 3.3.1 *Post-processing effects*

One important category of graphical effect stems from the ability to render to an off-screen buffer, or even to multiple buffers simultaneously, which can then be used to form a feedback loop. A polygon may then be drawn (either to additional buffers or to the visible framebuffer) with the previously rendered texture(s) made available to the shader. This shader can then perform a variety of ‘post-processing’ effects.

Modern engines frequently include a selection of such effects (Feis 2007), which can include more traditional image processing, such as colour transformations (Burkersroda 2005; Bjorke 2004), glow (James and O’Rorke 2004), or edge-enhancement (Nienhaus and Döllner 2003), as well as techniques that require additional scene information such as depth of field (Gillham 2007; Zhou et al. 2007), motion blur (Rosado 2008) and others which will be mentioned in specific sections later.

The extreme of this type of technique is deferred shading, where the entire lighting calculations are performed as a ‘post-process’. Here, the scene geometry is rendered into a set of intermediate buffers, collectively called the G-buffer, and the final shading process is performed in image-space using the data from those buffers (Koonce 2008).

### 3.3.2 *Transparency, reflection and refraction*

The modern real-time graphics pipeline does not deal with the visual representation of transparency, reflection or



refraction and their emulation must be dealt with using special cases or tricks. Traditionally, transparency has been emulated using alpha blending (Akenine-Möller et al. 2008), a compositing technique where a ‘transparent pixel’ is combined with the framebuffer according to its fourth colour component (alpha). The primary difficulty with this technique is that the results are order dependent, which requires the scene geometry to be sorted by depth before it is drawn and transparency can also present issues when using deferred shading (Filion and McNaughton 2008). A number of order-independent transparency techniques have been developed, however, such as depth-peeling (Everitt 2001; Nagy and Klein 2003).

Mirrored background reflections may be achieved using an environment map (Blinn and Newell 1976; Watt and Policarpo 2005), which can be a simple but effective method of reflecting a static scene. If the scene is more dynamic, but relatively fast to render, reflections on a flat surface may be achieved by drawing the reflective surface as transparent and mirroring the entire scene geometry about the reflection surface, drawing the mirrored geometry behind it (Fig. 8) or, for more complex scenes, using reduced geometry methods such as impostors (Tatarchuk and Isidoro 2006). Alternatively, six cameras can be used to produce a dynamic environment map (Blythe 2006). Alternative methods have also been developed to address the lack of parallax, i.e. apparent motion offsets due to objects at different distances, which are missing in a fixed environment map (Yu et al. 2005).

Perhaps surprisingly on first note, simple refraction effects can be achieved using very similar techniques to those used for reflection. The only differences are that the sample ray direction points inside the object and that it is bent due to the difference in refractive indices of the two



**Fig. 8** Achieving a mirror effect by rendering the geometry twice (Anderson and McLoughlin 2007)

materials, in accordance with Snell’s Law (Akenine-Möller et al. 2008). Thus, environment mapping can be used for simple refractions in a static scene, which may be expanded to include chromatic dispersion (Fernando and Kilgard 2003). In some cases, refraction may also be achieved as a post-processing effect (Wyman 2007).

### 3.3.3 Surface detail

The simplest method of adding apparent detail to a surface, without requiring additional geometry, is texture mapping. The advent of pixel shaders means that textures can now be used in more diverse ways to emulate surface detail (Rost 2006; Watt and Policarpo 2005; Akenine-Möller et al. 2008).

A variety of techniques exist for adding apparent high-resolution bump detail to a low-resolution mesh. In normal mapping (Blinn 1978) the texture map stores surface normals, which can then be used for lighting calculations. Parallax mapping (Kaneko et al. 2001) uses a surface height map and the camera direction to determine an offset for texture lookups. Relief texture mapping (Oliveira et al. 2000; Watt and Policarpo 2005) is a related technique which performs a more robust ray-tracing of the height map and can provide better quality results at the cost of performance.

### 3.3.4 Lighting

The old fixed-function graphics pipeline supported a per-vertex Gouraud lighting model [OpenGL ARB], but programmable shaders now allow the developer to implement their own lighting model (Rost 2006; Hoffman 2006). In general, though, the fixed-function lighting equation is split into: a diffuse component, where direct lighting is assumed to be scattered by micro-facets on the surface; a specular component, which appears as a highlight and is dependent on the angle between the viewer and the light; and an ambient component, which is an indirect ‘background’ lighting component due to light that has bounced off other objects in the scene (Akenine-Möller et al. 2008).

**3.3.4.1 Shadows** Although the graphics pipeline did not originally support shadows, it does now provide hardware acceleration for texture samples of a basic shadow map (Akenine-Möller et al. 2008; Engel et al. 2008). However, this basic method suffers from aliasing issues, is typically low resolution and can only result in hard shadow edges. Except in certain conditions, the majority of shadows in the real world exhibit a soft penumbra, so there is a desire within computer graphics to achieve efficient soft shadows, for which a large number of solutions have been developed (Hasenfratz et al. 2003; Bavoil 2008). Shadowing complex

objects such as volumes can also present issues, many of which have also been addressed (Lokovic and Veach 2000; Hadwiger et al. 2006; Ropinski et al. 2008).

**3.3.4.2 High-Dynamic Range Lighting** HDR Lighting is a technique that has become very popular in modern games (Sherrod 2006; Engel et al. 2008). It stems from the fact that real world luminance has a very high dynamic range, which means that bright surface patches are several orders of magnitude brighter than dark surface patches—for example, the sun at noon “may be 100 million times brighter than starlight” (Reinhard et al. 2006). In general, this means that the 8-bit integers traditionally used in each component of the RGB triplet of pixels in the framebuffer, are woefully inadequate for representing real luminance ranges. Thankfully, modern hardware now allows a greater precision in data types, so that calculations may be performed in 16 or even 32-bit floating-point format, although it should be noted that a performance penalty usually occurs when using more precise formats.

One of the most striking visual effects associated with HDR lighting is bloom, where extremely bright patches appear to glow. Practically, this is usually applied as a post-process effect in a similar way to a glow effect, where bright patches are drawn into a separate buffer which is blurred and then combined with the original image (Kawase 2004; Kawase 2003). This can also be applied to low-dynamic range images, to make them appear HDR (Sousa 2005).

Modern displays still use the traditional 8-bit per colour component format (with a few exceptions (Seetzen et al. 2004)), so the HDR floating point results must be converted, which is the process of tonemapping (Reinhard et al. 2006). Some tonemapping methods allow the specification of a brightness, or exposure value as taken from a physical camera analogy. In an environment where the brightness is likely to change dramatically this exposure should be automatically adjusted—much like a real camera does today. Various methods are available to achieve this, such as by downsampling the entire image to obtain the average brightness (Kawase 2004), or by asynchronous queries to build a basic histogram of the brightness level to determine the required exposure (McTaggart et al. 2006; Sheuermann and Hensley 2007).

**3.3.4.3 Indirect lighting: global illumination** Incident light on a surface can originate either directly from a light source, or indirectly from light reflected by another surface. Global illumination techniques account for both of these sources of light, although in such methods it is the indirect lighting component that is usually of most interest and the most difficult to achieve. The main difficulty is that in order to render a surface patch, the light that is reflected by

all other surface patches in the scene must be known. This interdependence can be costly to compute, especially for dynamic scenes, and although indirect lighting accounts for a high proportion of real world illumination, the computational cost of simulating its effects has resulted in very limited use within real-time applications (Dutr et al. 2003).

The simplest inclusion of indirect lighting is through pre-computed and baked texture maps, which can store anything from direct shadows or ambient occlusion results to those from radiosity or photon mapping (Mittring 2007). However, this technique is only viable for completely static objects within a static scene. Another simple global illumination technique, which is commonly associated with HDR lighting, is image-based lighting (Reinhard et al. 2006). Here, an environment map stores both direct and indirect illumination as a simple HDR image, which is then used to light objects in the scene. The image may be captured from a real-world location, drawn by an artist as an art asset or generated in a pre-processing stage by sampling the virtual environment. Multiple samples can then be used to light a dynamic character as it moves through the (static) environment (Mitchell et al. 2006). Although the results can be very effective, image-based lighting cannot deal with fully dynamic scenes without having to recompute the environment maps, which may be costly.

Fully dynamic global illumination techniques generally work on reduced or abstracted geometry, such as using discs to approximate the geometry around each vertex for ambient occlusion (Shanmugam and Arikan 2007; Hoberock and Jia 2008). It is also possible to perform some operations as a post-process, such as ambient occlusion (Mittring 2007) and even approximations for single-bounce indirect lighting (Ritschel et al. 2009). The general-purpose use of the GPU has also allowed for radiosity at near real-time for very small scenes (Coombe and Harris 2005) and fast, but not yet real-time, photon mapping (Purcell et al. 2003). The latter technique can also be used to simulate caustics, which are bright patches due to convergent rays from a refractive object, in real-time on the GPU (Krüger et al. 2006), although other techniques for specifically rendering caustics are also possible (Wand and Straßer 2003), including as an image-space post-process effect (Wyman 2007), or by applying the ‘Caustic Cones’ that utilise an intensity map generated from real photographic images (Kider et al. 2009).

### 3.4 Artificial intelligence

Another important aspect of the creation of populated virtual environments as used in cultural heritage applications is the creation of intelligent behaviour for the inhabitants of the virtual world, which is achieved using artificial intelligence (AI) techniques.

It is important to understand that when we refer to the AI of virtual entities in virtual environments, that which we refer to is not truly AI—at least not in the conventional sense (McCarthy 2007) of the term. The techniques applied to virtual worlds, such as computer games, are usually a mixture of AI related methods whose main concern is the creation of a believable illusion of intelligence (Scott 2002), i.e. the behaviour of virtual entities only needs to be believable to convey the presence of intelligence and to immerse the human participant in the virtual world.

The main requirement for creating the illusion of intelligence is perception management, i.e. the organisation and evaluation of incoming data from the AI entity's environment. This perception management mostly takes the form of acting upon sensor information but also includes communication between or coordination of AI entities in environments which are inhabited by multiple entities which may have to act co-operatively. The tasks which need to be solved in most modern virtual world applications such as computer games and to which the intelligent actions of the AI entities are usually restricted to (by convention rather than technology) are (Anderson 2003):

- decision making
- path finding (planning)
- steering (motion control)

The exact range of problems that AI entities within a computer game have to solve depends on the context in which they exist and the virtual environment in which the game takes place. Combs and Ardoit (2004) state that a popular method for the implementation of game AI is the use of an 'environment-based programming style', i.e. the creation of the virtual game world followed by the association of AI code with the game world and the entities that exist in it. This means that the AI entity intelligence is built around and is intrinsically linked to the virtual game environment. This type of entity intelligence can be created using 'traditional' methods for 'decision making', 'path finding' and 'steering'.

Of the three common AI tasks named above, 'decision making' most strongly implies the use of intelligence. Finite state machines (FSMs) are the most commonly used technique for implementing decision making in games (Fu and Houlette 2004). They arrange the behaviour of an AI entity in logical states—defining one state per possible behaviour—of which only one, the entity's behaviour at that point in time, is active at any one time. In game FSMs each state is usually associated with a specific behaviour and an entity's actions are often implemented by linking behaviours with pre-defined animation cycles for the AI entity that allow it to enact the selected behaviour (Orkin 2006). It is relatively simple to program a very stable FSM

that may not be very sophisticated but that "will get the job done". The main drawback of FSMs is that they can become very complex and hard to maintain, while on the other hand the behaviour resulting from a too simple FSM can easily become predictable. To overcome this problem sometimes hierarchical FSMs are used that break up complex states into a set of smaller ones that can be combined, allowing the creation of larger and more complex FSMs.

In recent years, there has been a move towards performing decision making using goal-directed techniques to enable the creation of nondeterministic behaviour. Dybsand describes this as a technique in which an AI entity "will execute a series of actions ... that attempt to accomplish a specific objective or goal" (Dybsand 2004). In its simplest form, goal-orientation can be implemented by determining a goal with an embedded action sequence for a given AI entity. This action sequence, the entity's plan, will then be executed by the entity to satisfy the goal (Orkin 2004a). Solutions that allow for more diverse behaviour can improve this by selecting an appropriate plan from a pre-computed 'plan library' (Evans 2001) instead of using a built-in plan. More complex solutions use plans that are computed dynamically, i.e. 'on the fly', as is the case with Goal-Oriented Action Planning (GOAP) (Orkin 2004a). In GOAP the sequence of actions that the system needs to perform to reach its end-state or goal is generated in real-time by using a planning heuristic on a set of known values which need to exist within the AI entity's domain knowledge. To achieve this in his implementation of GOAP, Orkin (2004b) separates the actions and goals, implicitly integrating preconditions and effects that define the planner's search space, placing the decision making process into the domain of the planner. This can be further improved through augmenting the representation of the search space by associating costs with actions that can satisfy goals, effectively turning the AI entity's knowledge base into a weighted graph. This then allows the use of path planning algorithms that find the shortest path within a graph as the planning algorithm for the entity's high-level behaviour (Orkin 2006). This has the additional benefit of greater code re-use as the planning method for high-level decision making, as well as path planning is the same and can therefore be executed by the same code module (Orkin 2004b) if the representations of the search space are kept identical. The most popular path planning algorithm used in modern computer games is the A\* (A-Star) algorithm (Stout 2000; Matthews 2002; Nareyek 2004), a generalisation of Dijkstra's algorithm (1959). A\* is optimal, i.e. proven to find the optimal path in a weighted graph if an optimal solution exists (Dechter and Pearl 1985), which guarantees that AI entities will find the least costly path if such a solution exists within the search space.

Challenges in game AI that are relevant to serious games include the construction of intelligent interfaces (Livingstone and Charles 2004), such as tutoring systems or virtual guides, and particularly real-time strategy game AI, part of which is concerned with the modelling of great numbers of virtual entities in large scale virtual environments. Challenges there include spatial and temporal reasoning (Buro 2004), which can be addressed through the use of potential fields (Hagelbäck and Johansson 2008).

### 3.4.1 Crowd simulation

The AI techniques described in the previous section are important tools with which more complex systems can be constructed. A domain of great potential relevance to cultural heritage that is derived from such techniques is the simulation of crowds of humanoid characters. If one wishes to reconstruct and visualise places and events from the past, a crowd of real-time virtual characters, if appropriately attired and behaving, can add new depths of immersion and realism to ancient building reconstructions. These characters can feature merely as a backdrop (Ciechomski et al. 2004) to add life to a reconstruction, or can assume the centre stage in more active roles, for example, as virtual tour guides to direct the spectator (DeLeon 1999). Indeed, the type of crowd or character behaviour to be simulated varies greatly with respect to the type of scenario that needs to be modelled. In this vein, (Ulicny and Thalmann 2002) model crowd behaviour of worshippers in a virtual mosque, while (Maim et al. 2007) and (Ryder et al. 2005) focus on the creation of more general pedestrian crowd behaviours, the former for populating a virtual reconstruction of a city resembling ancient Rome.

More general crowd synthesis and evaluation techniques are also directly applicable to crowd simulation in cultural heritage. A variety of different approaches have been taken, most notably the use of social force models (Helbing and Molnar 1995), path planning (Lamarche and Donikian 2004), behavioural models incorporating perception and learning (Shao and Terzopoulos 2005) sociological effects (Musse and Thalmann 1997) and hybrid models (Pelechano et al. 2007).

The study of real world corpus has also been used as a basis for synthesising crowd behaviour in approaches that do not entail the definition of explicit behaviour models. Lerner et al. (2007) manually track pedestrians from an input video containing real world behaviour examples. They use this data to construct a database of pedestrian trajectories for different situations. At runtime, the database is queried for similar situations matching those of the simulated pedestrians: the closest matching example from the database is selected as the resulting trajectory for each pedestrian and the process is repeated.

Lee et al. (2007) simulate behaviours based on aerial-view video recordings of crowds in controlled environments. A mixture of manual annotation and semi-automated tracking provides information from video about individuals' trajectories. These are provided as inputs to an agent movement model that can create crowd behaviours of a similar nature to those observed in the original video.

Human perception of the animation of crowds and characters has been increasingly recognised as an important factor in achieving more realistic simulations. Research has been conducted regarding the perception of animation and motion of individuals (Reitsma and Pollard 2003; McDonnell et al. 2007), groups (Ennis et al. 2010a; McDonnell et al. 2009a) and crowds (Peters et al. 2008; Ennis et al. 2010b). For example, (Peters et al. 2008) examined the perceptual plausibility of pedestrian orientations and found that participants were able to consistently distinguish between those virtual scenes where the character orientations matched the orientations of the humans in the corresponding real scenes and those where the character orientations were artificially generated, according to a number of different rule types. The results of such perceptual studies can be linked to synthesis, in order to create more credible animations (McDonnell et al. 2009b).

A key factor of differentiation between crowd control methods concerns where knowledge is stored in the system. One approach is to endow knowledge separately to individual characters, an extreme example of which would create autonomous agents that have their own artificial perceptions, reasoning, memories, etc. with respect to the environment, as in (Lamarche and Donikian 2004). Another method is to place knowledge into the environment itself, to create a shared or partially shared database accessible to characters. According to this smart object methodology (Peters et al. 2003), graphical objects are tagged with behavioural information and may inform, guide or even control characters. Such an approach is applicable also to crowd simulation in urban environments. For example, navigation aids, placed inside the environment description, may be added by the designer during the construction process. These have been referred to as annotations (Doyle and Hayes-Roth 1998). The resulting environment description (Farenc et al. 1999; Thomas and Donikian 2000; Peters and O'Sullivan 2009) contains additional geometric, semantic and spatial partitioning information for informing pedestrian behaviour, thus transferring a degree of the behavioural intelligence into the environment. In (Hostetler 2002), for example, skeletal splines are defined that are aligned with walkways. These splines, called ribbons, provide explicit information for groups to use, such as the two major directions of travel on the walkway. In addition to environment annotation and mark-up, interfaces for managing the definition of crowd

scenarios have also been investigated. Crowdbush (Ulicny et al. 2004) provides an intuitive way for designers to add crowds of characters into an environment using tools analogous to those found in standard 2D painting packages. It allows designers to paint crowds and apply attributes and characteristics using a range of different tools in real-time, obtaining immediate feedback about the results.

### 3.4.2 Annotated entities and environments

A fairly recent method for enabling virtual entities to interact with one another as well as their surroundings is the use of annotated worlds. The mechanism for this, which we refer to using the term ‘Annotated Entities’, has been described using various names, such as ‘Smart Terrain’ (Cass 2002), ‘Smart Objects’ (Peters et al. 2003; Orkin 2006) and ‘Annotated Environment’ (Doyle 2002), all of which are generally interchangeable and mostly used with very similar meanings, although slight differences in their exact interpretation sometimes remain. A common aspect to all of the implementations that utilise this mechanism is the indirect approach to the creation of believable intelligent entities.

The idea of annotated environments is a computer application of the theory of affordance (or affordance theory) (Cornwell et al. 2003) that was originally developed in the fields of psychology and visual perception. Affordance theory states that the makeup and shape of objects contains suggestions about their usage. Affordance itself is an abstract concept, the implementation of which is greatly simplified by annotations that work like labels containing instructions which provide an explicit interpretation of affordances. Transferred into the context of a virtual world, this means that objects in the environment contain all of the information that an AI controlled entity will need to be able to use them, effectively making the environment ‘smart’.

A beneficial side effect of this use of ‘annotated’ objects (Doyle 1999) is that the complexity of the entities is neutral to the extent of the domain knowledge that is available for their use, i.e. the virtual entities themselves can not only be kept relatively simple, but they do not need to be changed at all to be able to make use of additional knowledge. This allows for the rapid development of game scenarios (Cornwell et al. 2003) and if all annotated objects use the same interface to provide knowledge to the world’s entities then there is no limit to the scalability of the system, i.e. the abilities of AI controlled entities can practically be extended indefinitely (Orkin 2002) despite a very low impact on the system’s overall performance. Furthermore, this method provides an efficient solution to the ‘anchoring problem’ (Coradeschi and Saffiotti 1999) of matching

sensor data to the symbolic representation of the virtual entity’s knowledge as objects in the world themselves have the knowledge as to how other virtual entities can interact with them.

Annotations have been employed in several different types of applications in order to achieve different effects. They have proven popular for the animation of virtual actors in computer animation production, where they facilitate animation selection (Lee et al. 2006), i.e. the choice of appropriate animation sequences that fit the environment. Other uses of annotations include the storage of tactical information in the environment for war games and military simulations (Darken 2007), which is implemented as sensory annotations to direct the virtual entities’ perception of their environment. Probably the most common form of annotations found in real-time simulated virtual environments affects behaviour selection, usually in combination with animation selection (Orkin 2006), i.e. the virtual entity’s behaviour and its visual representation (animation) are directed by the annotated objects that it uses.

Virtual entities that inhabit these annotated worlds can be built utilising rule-based system based on simple FSMs in combination with a knowledge interface based on a trigger system that allows the entities to ‘use’ knowledge (instructions) for handling the annotated objects. The interaction protocol employed to facilitate the communication between entity and ‘smart’ object needs to enable the object to ‘advertise’ its features to the entities and then allow them to request from the object relevant instructions (annotations) on its usage (Macedonia 2000). The success of this technique is demonstrated by the best-selling computer game *The Sims*, where ‘Smart Objects’ were used for behaviour selection to great effect. Forbus and Wright (2001) state that in *The Sims* all game entities, objects as well as virtual characters, are implemented as scripts that are executed in their own threads within a multitasking virtual machine. A similar approach, based on a scripting language that can represent the behaviours of virtual entities, as well as the objects that they can interact with, has been presented more recently by Anderson (2008). These scripting-language based approaches are most likely to provide solutions for the creation of large scale virtual environments, such as the serious game component of the *Rome Reborn* project. This is the automatic generation of AI content (Nareyek 2007), which in combination with techniques such as procedural modelling of urban environments (Vanegas et al. 2009), will require the integration of the creation of complex annotations with the procedural generation of virtual worlds, automating the anchoring of virtual entities into their environment.

## 4 Conclusions

The success of computer games, fuelled among other factors by the great realism that can be attained using modern consumer hardware, and the key techniques of games technology that have resulted from this, have given way to new types of games, including serious games, and related application areas, such as virtual worlds, mixed reality, augmented reality and virtual reality. All of these types of application utilise core games technologies (e.g. 3D environments) as well as novel techniques derived from computer graphics, human computer interaction, computer vision and artificial intelligence, such as crowd modelling. Together these technologies have given rise to new sets of research questions, often following technologically driven approaches to increasing levels of fidelity, usability and interactivity.

Our aim has been to use this state-of-the-art report to demonstrate the potential of games technology for cultural heritage applications and serious games, to outline key problems and to indicate areas of technology where solutions for remaining challenges may be found. To illustrate that first we presented some characteristic case studies illustrating the application of methods and technologies used in cultural heritage. Next, we provided an overview of existing literature of relevance to the domain, discussed the strengths and weaknesses of the described methods and pointed out unsolved problems and challenges. It is our firm belief that we are only at the beginning of the evolution of games technology and that there will be further improvements in the quality and sophistication of computer games, giving rise to serious heritage games of greater complexity and fidelity than is now achievable.

**Acknowledgments** The authors would like to thank the following: The Herbert Art Gallery & Museum (Coventry, UK), Simon Bakkevig, and Lukasz Bogaj. This report includes imagery generated using the Virtual Egyptian Temple, which is a product of PublicVR (<http://publicvr.org>).

## References

- Akenine-Möller T, Haines E, Hoffman N (2008) Real-time rendering 3rd edn. A. K. Peters, Natick
- Anderson EF (2003) Playing smart—artificial intelligence in computer games. In: Proceedings of zfxCON03 conference on game development
- Anderson EF (2008) Scripted smarts in an intelligent virtual environment: behaviour definition using a simple entity annotation language. In: Future Play '08: Proceedings of the 2008 conference on future play, pp 185–188
- Anderson EF, McLoughlin L (2007) Critters in the classroom: a 3d computer-game-like tool for teaching programming to computer animation students. In: SIGGRAPH '07: ACM SIGGRAPH 2007 educators program, p 7
- Anderson EF, Engel S, McLoughlin L, Comminos P (2008) The case for research in game engine architecture. In: Future Play '08: Proceedings of the 2008 conference on future play, pp 228–231
- Apperley TH (2006) Virtual unaustralia: Videogames and australia's colonial history. In: UNAUSTRALIA 2006: Proceedings of the cultural studies association of Australasia's annual conference
- Arnold D, Day A, Glauert J, Haegler S, Jennings V, Kevelham B, Laycock R, Magnenat-Thalmann N, Mam J, Maupu D, Papagiannakis G, Thalmann D, Yersin B, Rodriguez-Echavarria K (2008) Tools for populating cultural heritage environments with interactive virtual humans. In: Open digital cultural heritage systems, EPOCH final event Rome
- Azuma R (1997) A survey of augmented reality. Presence: teleoperators and virtual environments 6(4):355–385
- Azuma R, Baillyot Y, Behringer R, Feiner S, Julier S, MacIntyre B (2001) Recent advances in augmented reality. IEEE Comput Graph Appl 21(6):34–47
- Bavoil L (2008) Advanced soft shadow mapping techniques. Presentation at the game developers conference 2008
- Bavoil L, Sainz M (2008) Screen space ambient occlusion. NVIDIA developer information: <http://developers.nvidia.com>
- Bederson BB (1995) Audio augmented reality: a prototype automated tour guide. In: CHI '95: Conference companion on human factors in computing systems, pp 210–211
- Benthin C, Wald I, Scherbaum M, Friedrich H (2006) Ray tracing on the cell processor, pp 15–23
- Bimber O, Frhlich B, Schmalstieg D, Encarnao LM (2001) The virtual showcase. IEEE Comput Graph Appl 21(6):48–55
- Bjorke K (2004) Color controls. In: Fernando R (ed) GPU gems, Pearson Education, pp 363–373
- Blinn JF (1978) Simulation of wrinkled surfaces. SIGGRAPH Comput Graph 12(3):286–292
- Blinn JF, Newell ME (1976) Texture and reflection in computer generated images. Commun ACM 19(10):542–547
- Blow J (2004) Game development harder than you think. ACM Queue 1(10):28–37
- Blythe D (2006) The direct3d 10 system. ACM Trans Graph 25(3):724–734
- Brogni B, Avizzano C, Evangelista C, Bergamasco M (1999) Technological approach for cultural heritage: augmented reality. In: RO-MAN '99: Proceedings of the 8th IEEE international workshop on robot and human interaction, pp 206–212
- Burkersroda R (2005) Colour grading. In: Engel W (eds) Shader X3: advanced rendering with DirectX and OpenGL. Charles River Media, Hingham, pp 357–362
- Buro M (2004) Call for ai research in rts games. In: Proceedings of the AAAI-04 workshop on challenges in game AI, pp 139–142
- Burton J (2005) News-game journalism: history, current use and possible futures. Aust J Emerg Technol Soc 3(2):87–99
- Calori L, Camporesi C, Forte M, Guidazzoli A, Pescarin S (2005) Openheritage: integrated approach to web 3d publication of virtual landscape. In: Proceedings of the ISPRS working group V/4 workshop 3D-ARCH 2005: virtual reconstruction and visualization of complex architectures
- Cass S (2002) Mind games. IEEE Spectrum 39(12):40–44
- Cerezo E, Perez-Cazorla F, Pueyo X, Seron F, Sillion F (2005) A survey on participating media rendering techniques. The Visual Comput 21(5):303–328
- Chalmers A, Debattista K (2009) Level of realism for serious games. In: VS-Games 2009: Proceedings of the IEEE virtual worlds for serious applications first international conference, pp 225–232
- Ciechowski PDH, Ulicny B, Cetre R, Thalmann D (2004) A case study of a virtual audience in a reconstruction of an ancient roman odeon in aphrodisias. In: The 5th international symposium on virtual reality, archaeology and cultural heritage, VAST (2004)

- Combs N, Ardoit J (2004) Declarative versus imperative paradigms in Games AI. Available from: <http://www.red3d.com/cwr/games/>
- Coombe G, Harris M (2005) Global illumination using progressive refinement radiosity. In: Pharr M (ed) GPU gems 2, Pearson Education, pp 635–647
- Coradeschi S, Saffiotti A (1999) Symbolic object descriptions to sensor data. Problem statement. Linköping Electronic Articles in Computer and Information Science 4(9)
- Cornwell J, O'Brien K, Silverman B, Toth J (2003) Affordance theory for improving the rapid generation, composability, and reusability of synthetic agents and objects. In: BRIMS 2003: Proceedings of the twelfth conference on behavior representations in modeling and simulation
- Cosmas J, Itegiaki T, Green D, Grabczewski E, Weimer F, Van Gool L, Zalesny A, Vanrintel D, Leberl F, Grabner M, Schindler K, Karner K, Gervautz M, Hynst S, Waelkens M, Pollefeys M, DeGeest R, Sablatnig R, Kampel M (2001) 3d murale: a multimedia system for archaeology. In: VAST '01: Proceedings of the 2001 conference on virtual reality, archeology, and cultural heritage, pp 297–306
- Crassin C, Neyret F, Lefebvre S, Eisemann E (2009) Gigavoxels: ray-guided streaming for efficient and detailed voxel rendering. In: I3D '09: Proceedings of the 2009 symposium on interactive 3D graphics and games, pp 15–22
- Cruz-Neira C, Sandin DJ, DeFanti TA, Kenyon RV, Hart JC (1992) The cave: audio visual experience automatic virtual environment. *Commun ACM* 35(6):64–72
- Darken CJ (2007) Level Annotation and Test by Autonomous Exploration. In: AIIDE 2007: Proceedings of the third artificial intelligence and interactive digital entertainment conference
- Debevec P (2005) Making “The Parthenon”. 6th international symposium on virtual reality, archaeology, and cultural heritage
- Debevec P, Tchou C, Gardner A, Hawkins T, Poullis C, Stumpfel J, Jones A, Yun N, Einarsson P, Lundgren T, Fajardo M, Martinez P (2004) Estimating surface reflectance properties of a complex scene under captured natural illumination. Tech. rep., University of Southern California, Institute for Creative Technologies
- Dechter R, Pearl J (1985) Generalised best-first search strategies and the optimality of A\*. *J ACM* 32(3):505–536
- DeLeon VJ (1999) Vrnd: notre-dame cathedral: a globally accessible multi-user real time virtual reconstruction. In: Proceedings of virtual systems and multimedia
- Dijkstra EW (1959) A note on two problems in connexion with graphs. *Numerische Mathematik* 1:269–271
- Doyle P (1999) Virtual intelligence from artificial reality: building stupid agents in smart environments. In: AAAI '99 spring symposium on artificial intelligence and computer games
- Doyle P (2002) Believability through context. In: AAMAS '02: Proceedings of the first international joint conference on autonomous agents and multiagent systems, pp 342–349
- Doyle P, Hayes-Roth B (1998) Agents in annotated worlds. In: AGENTS '98: Proceedings of the second international conference on autonomous agents, pp 173–180
- Dutr P, Bekaert P, Bala K (2003) Advanced global illumination. A. K. Peters, Natick
- Dybsand E (2004) Goal-directed behaviour using composite tasks. In: AI game programming wisdom 2, Charles River Media, pp 237–245
- El-Hakim S, MacDonald G, Lapointe JF, Gonzo L, Jemtrud M (2006) On the digital reconstruction and interactive presentation of heritage sites through time. In: International symposium on virtual reality, archaeology and intelligent cultural heritage, pp 243–250
- Engel K, Hadwiger M, Kniss JM, Rezk-Salama C, Weiskopf D (2006) Real-time volume graphics. A. K. Peters, Wellesley
- Engel W, Hoxley J, Kornmann R, Suni N, Zink J (2008) Programming vertex, geometry, and pixel shaders. Online book available at: <http://wiki.gamedev.net/>
- Ennis C, McDonnell R, O'Sullivan C (2010a) Seeing is believing: body motion dominates in multisensory conversations. *ACM Trans Graph* 29(4):1–9
- Ennis C, Peters C, O'Sullivan C (2010b) Perceptual effects of scene context and viewpoint for virtual pedestrian crowds. *ACM Trans Appl Percept* (in press)
- Evans R (2001) AI in computer games: the use of AI techniques in Black & White. Seminar notes, available from: <http://www.dcs.qmul.ac.uk/research/logic/seminars/abstract/EvansR01.html>
- Everitt C (2001) Interactive order-independent transparency. NVIDIA Whitepaper
- Farenc N, Boulic R, Thalmann D (1999) An informed environment dedicated to the simulation of virtual humans in urban context. *Comput Graph Forum* 18(3):309–318
- Feiner S (2002) Augmented reality: a new way of seeing. *Sci Am* 286(4):48–55
- Feis A (2007) Postprocessing effects in design. In: Engel W (ed) Shader X5: advanced rendering techniques. Charles River Media, pp 463–470
- Fernando R, Kilgard MJ (2003) The Cg tutorial. Addison Wesley
- Filion D, McNaughton R (2008) Effects & techniques. In: SIGGRAPH '08: ACM SIGGRAPH 2008 classes, pp 133–164
- Forbus KD, Wright W (2001) Some notes on programming objects in The Sims™. Class notes, available from: <http://qrg.northwestern.edu/papers/papers.html>
- Forsyth DA, Ponce J (2002) Computer vision: a modern approach. Prentice Hall, Upper Saddle River
- Francis R (2006) Revolution: learning about history through situated role play in a virtual environment. In: Proceedings of the American educational research association conference
- de Freitas S, Oliver M (2006) How can exploratory learning with games and simulations within the curriculum be most effectively evaluated?. *Comput Educ* 46:249–264
- Frischer B (2008) The rome reborn project. How technology is helping us to study history. OpEd, November 10, University of Virginia
- Fritsch D, Kada M (2004) Visualisation using game engines. ISPRS commission 5, pp 621–625
- Fu D, Houlette R (2004) The ultimate guide to FSMs in games. In: AI game programming Wisdom 2. Charles River Media, pp 283–302
- Gaitatzes A, Christopoulos D, Papaioannou G (2004) The ancient olympic games: being part of the experience. In: VAST 2004: The 5th international symposium on virtual reality, archaeology and cultural heritage, pp 19–28
- Gardner R (2009) Empire total war—graphics work shop. Available from (the official) Total War blog: <http://blogs.sega.com/totalwar/2009/03/05/empire-total-war-graphics-work-shop/>
- Gatermann H (2000) From vrml to augmented reality via panorama-integration and eai-java. In: SIGRaDi'2000—Construindo (n)o espacio digital (constructing the digital Space), pp 254–256
- Gillham D (2007) Real-time depth-of-field implemented with a postprocessing-only technique. In: Engel W (ed) Shader X5: advanced rendering techniques. Charles River Media, pp 163–175
- Godbersen H (2008) Virtual environments for anyone. *IEEE Multimedia* 15(3):90–95
- Hadwiger M, Kratz A, Sigg C, Bühler K (2006) Gpu-accelerated deep shadow maps for direct volume rendering. In: GH '06: Proceedings of the 21st ACM SIGGRAPH/EUROGRAPHICS symposium on graphics hardware, pp 49–52
- Hagelbäck J, Johansson SJ (2008) The rise of potential fields in real time strategy bots. In: AIIDE 08: Proceedings of the fourth

- artificial intelligence and interactive digital entertainment conference, pp 42–47
- Hall T, Ciolfi L, Bannon L, Fraser M, Benford S, Bowers J, Greenhalgh C, Hellström SO, Izadi S, Schnädelbach H, Flintham M (2001) The visitor as virtual archaeologist: explorations in mixed reality technology to enhance educational and social interaction in the museum. In: VAST '01: Proceedings of the 2001 conference on virtual reality, archeology, and cultural heritage, pp 91–96
- Hasenfratz JM, Lapiere M, Holzschuch N, Sillion F (2003) A survey of real-time soft shadows algorithms
- Helbing D, Molnar P (1995) Social force model for pedestrian dynamics. *Phys Rev E* 51(5):4282–4286
- Hoferock J, Jia Y (2008) High-quality ambient occlusion. In: Nguyen H (ed) GPU gems 3. Pearson Education, pp 257–274
- Hoffman N (2006) Physically based reflectance for games
- Horn DR, Sugerman J, Houston M, Hanrahan P (2007) Interactive k-d tree gpu raytracing. In: I3D '07: Proceedings of the 2007 symposium on interactive 3D graphics and games, pp 167–174
- Hostetler TR (2002) Controlling steering behavior for small groups of pedestrians in virtual urban environments. PhD thesis, The University of Iowa
- Humphreys G, Houston M, Ng R, Frank R, Ahern S, Kirchner PD, Klosowski JT (2002) Chromium: a stream-processing framework for interactive rendering on clusters. *ACM Trans Graph* 21(3):693–702
- Isidoro JR, Sander PV (2006) Animated skybox rendering and lighting techniques. In: SIGGRAPH '06: ACM SIGGRAPH 2006 courses, pp 19–22
- Jacobson J, Holden L (2005) The Virtual Egyptian Temple. In: ED-MEDIA: Proceedings of the world conference on educational media. Hypermedia & Telecommunications
- Jacobson J, Lewis M (2005) Game engine virtual reality with CaveUT. *IEEE Comput* 38(4):79–82
- Jacobson J, Handron K, Holden L (2009) Narrative and content combine in a learning game for virtual heritage. In: Computer applications to archaeology 2009
- James G, O'Rorke J (2004) Real-time glow. In: Fernando R (ed) GPU gems. Pearson Education, pp 343–362
- Jones C (2005) Who are you? theorising from the experience of working through an avatar. *E-Learning* 2(4):414–425
- Jones G, Christal M (2002) The future of virtual museums: On-line, immersive, 3d environments. Created realities group
- Kaneko T, Takahei T, Inami M, Kawakami N, Yanagida Y, Maeda T, Tachi S (2001) Detailed shape representation with parallax mapping. In: Proceedings of ICAT 2001, pp 205–208
- Kawase M (2003) Frame buffer postprocessing effects in double-s.t.e.a.l (wreakless). Presentation at the game developers conference 2003
- Kawase M (2004) Practical implementation of high dynamic range rendering. Presentation at the game developers conference 2004
- Kider JT, Fletcher RL, Yu N, Holod R, Chalmers A, Badler NI (2009) Recreating early islamic glass lamp lighting. In: VAST09: The 10th international symposium on virtual reality, archaeology and intelligent cultural heritage, pp 33–40
- Kim J, Jaja J (2009) Streaming model based volume ray casting implementation for cell broadband engine. *Sci Program* 17(1–2): 173–184
- Kirriemuir J (2008) Measuring the impact of second life for educational purposes. Eduserv foundation, Available from: [http://www.eduserv.org.uk/foundation/sl/uksnapshot052008](http://www.eduserv.org.uk/foundation/sl/uknapshot052008)
- Koonce R (2008) Deferred shading in Tabula Rasa. In: Nguyen H (ed) GPU Gems 3. Pearson Education, pp 429–457
- Krüger J, Bürger K, Westermann R (2006) Interactive screen-space accurate photon tracing on GPUs. In: Rendering Techniques (Eurographics symposium on rendering–EGSR), pp 319–329
- Lamarche F, Donikian S (2004) Crowd of virtual humans: a new approach for real time navigation in complex and structured environments. *Comput Graph Forum* 23(3):509–518
- Leavy B, Wyeld T, Hills J, Barker C, Gard S (2007) The ethics of indigenous storytelling: using the torque game engine to support australian aboriginal cultural heritage. In: Proceedings of the DiGRA 2007 conference, pp 24–28
- Lee KH, Choi MG, Lee J (2006) Motion patches: building blocks for virtual environments annotated with motion data. In: SIGGRAPH '06: ACM SIGGRAPH 2006 Papers, pp 898–906
- Lee KH, Choi MG, Hong Q, Lee J (2007) Group behavior from video: a data-driven approach to crowd simulation. In: SCA '07: Proceedings of the 2007 ACM SIGGRAPH/Eurographics symposium on computer animation, pp 109–118
- Lepouras G, Vassilakis C (2004) Virtual museums for all: employing game technology for edutainment. *Virtual Real* 8(2):96–106
- Lerner A, Chrysanthou Y, Dani L (2007) Crowds by example. *Comput Graph Forum* 26(3):655–664
- Lewis M, Jacobson J (2002) Game engines in scientific research. *Commun ACM* 45(1):27–31
- Liarokapis F (2007) An augmented reality interface for visualising and interacting with virtual content. *Virtual Real* 11(1):23–43
- Liarokapis F, Sylaiou S, Mountain D (2008) Personalizing virtual and augmented reality for cultural heritage indoor and outdoor experiences. In: VAST08: the 9th international symposium on virtual reality, archaeology and intelligent cultural heritage, pp 55–62
- Linaza MT, Cobos Y, Mentxaka J, Campos MK, Penalba M (2007) Interactive augmented experiences for cultural historical events. In: VAST07: the 8th international symposium on virtual reality, archaeology and intelligent cultural heritage, pp 23–30
- Lintermann B, Deussen O (1999) Interactive modeling of plants. *IEEE Comput Graph Appl* 19(1):56–65
- Livingston MA (2005) Evaluating human factors in augmented reality systems. *IEEE Comput Graph Appl* 25(6):6–9
- Livingstone D, Charles D (2004) Intelligent interfaces for digital games. In: Proceedings of the AAAI-04 workshop on challenges in game AI, pp 6–10
- Lokovic T, Veach E (2000) Deep shadow maps. In: SIGGRAPH '00: Proceedings of the 27th annual conference on computer graphics and interactive techniques, pp 385–392
- Looser J, Grasset R, Seichter H, Billinghurst M (2006) Osgart—a pragmatic approach to mr. In: ISMAR 06: 5th IEEE and ACM international symposium on mixed and augmented reality
- Lugrin J, Cavazza M (2010) Towards ar game engines. In: SEARIS 2010–3rd workshop on software engineering and architecture of realtime interactive systems
- Macagon V, Wünsche B (2003) Efficient collision detection for skeletally animated models in interactive environments. In: Proceedings of IVCNZ '03, pp 378–383
- Macedonia M (2000) Using technology and innovation to simulate daily life. *IEEE Comput* 33(4):110–112
- Macedonia M (2002) Games soldiers play. *IEEE Spectrum* 39(3): 32–37
- Maim J, Haegler S, Yersin B, Mueller P, Thalmann D, Van Gool L (2007) Populating ancient pompeii with crowds of virtual romans. In: VAST07: the 8th international symposium on virtual reality, archaeology and intelligent cultural heritage, pp 109–116
- Malone TW, Lepper MR (1987) Making learning fun: A taxonomy of intrinsic motivations for learning. In: Snow RE, Farr MJ (eds) aptitude, learning and instruction: III. Conative and affective process analyses, Erlbaum, pp 223–253
- Mase K, Kadobayashi R, Nakatsu R (1996) Meta-museum: a supportive augmented-reality environment for knowledge sharing. In: ATR workshop on social agents: humans and machines, pp 107–110



- Matevitsi V, Sfakianos M, Lepouras G, Vassilakis C (2008) A game-engine based virtual museum authoring and presentation system. In: DIMEA '08: Proceedings of the 3rd international conference on digital interactive media in entertainment and arts, pp 451–457
- Matthews J (2002) Basic A\* pathfinding made simple. In: AI game programming wisdom, Charles River Media, pp 105–113
- McCarthy J (2007) What is Artificial Intelligence. Available from: <http://www-formal.stanford.edu/jmc/whatisai/whatisai.html>
- McDonnell R, Newell F, O'Sullivan C (2007) Smooth movers: perceptually guided human motion simulation. In: SCA '07: Proceedings of the 2007 ACM SIGGRAPH/Eurographics symposium on computer animation, pp 259–269
- McDonnell R, Ennis C, Dobbyn S, O'Sullivan C (2009a) Talking bodies: Sensitivity to desynchronization of conversations. *ACM Trans Appl Percept* 6(4):21
- McDonnell R, Larkin M, Hernández B, Rudomin I, O'Sullivan C (2009b) Eye-catching crowds: saliency based selective variation. *ACM Trans Graph* 28(3):1–10
- McGuire TJ (2006) The Philadelphia Campaign: volume one: Brandywine and the fall of Philadelphia. Stackpole Books, Washington
- McTaggart G, Green C, Mitchell J (2006) High dynamic range rendering in valve's source engine. In: SIGGRAPH '06: ACM SIGGRAPH 2006 Courses, p 7
- Milgram P, Kishino F (1994) A taxonomy of mixed reality visual displays. *IEICE Trans Inf Syst* E77-D(12):1321–1329
- Mitchell J, McTaggart G, Green C (2006) Shading in valve's source engine. In: SIGGRAPH '06: ACM SIGGRAPH 2006 Courses, pp 129–142
- Mittring M (2007) Finding next gen: Cryengine 2. In: SIGGRAPH '07: ACM SIGGRAPH 2007 courses, pp 97–121
- Mittring M, Crytek GmbH (2008) Advanced virtual texture topics. In: SIGGRAPH '08: ACM SIGGRAPH 2008 classes, pp 23–51
- Müller P, Vereenoghe T, Ulmer A, Van Gool L (2005) Automatic reconstruction of roman housing architecture. In: Recording, modeling and visualization of cultural heritage, pp 287–298
- Musse SR, Thalmann D (1997) A model of human crowd behavior: group inter-relationship and collision detection analysis. In: Computer animation and simulation '97, pp 39–52
- Nagy Z, Klein R (2003) Depth-peeling for texture-based volume rendering. In: PG '03: Proceedings of the 11th Pacific conference on computer graphics and applications, p 429
- Nareyek A (2004) Ai in computer games. *ACM Queue* 1(10):58–65
- Nareyek A (2007) Game ai is dead. long live game ai!. *IEEE Intell Syst* 22(1):9–11
- Nienhaus M, Döllner J (2003) Edge-enhancement—an algorithm for real-time non-photorealistic rendering. International Winter School of computer graphics. *J WSCG* 11(2):346–353
- Noghani J, Liarokapis F, Anderson EF (2010) Randomly generated 3d environments for serious games. In: VS-GAMES 2010: Proceedings of the 2nd international conference on games and virtual worlds for serious applications, pp 3–10
- Oliveira MM, Bishop G, McAllister D (2000) Relief texture mapping. In: SIGGRAPH '00: Proceedings of the 27th annual conference on computer graphics and interactive techniques, pp 359–368
- OpenGL Architecture Review Board, Shreiner D, Woo M, Neider J, Davis T (2007) OpenGL programming guide, 6th edn. Addison-Wesley, New York
- Orkin J (2002) 12 Tips from the trenches. In: AI game programming wisdom. Charles River Media, Hingham, pp 29–35
- Orkin J (2004a) Applying goal-oriented action planning to games. In: AI game programming wisdom 2. Charles River Media, Hingham, pp 217–228
- Orkin J (2004b) Symbolic representation of game world state: toward real-time planning in games. In: Proceedings of the AAAI-04 workshop on challenges in game AI, pp 26–30
- Orkin J (2006) Three states and a plan: the A.I. of F.E.A.R. In: Proceedings of the 2006 game developers conference
- Overmars M (2004) Teaching computer science through game design. *IEEE Comput* 37(4):81–83
- Papagiannakis G, Ponder M, Molet T, Kshirsagar S, Cordier F, Magnenat-Thalmann M, Thalmann D (2002) LIFEPLUS: revival of life in ancient Pompeii. In: Proceedings of the 8th international conference on virtual systems and multimedia (VSMM '02)
- Paquet E, El-Hakim S, Beraldin A, Peters S (2001) The virtual museum: virtualisation of real historical environments and artefacts and three-dimensional shape-based searching. In: VAA'01: Proceedings of the international symposium on virtual and augmented architecture, pp 182–193
- Pelechano N, Allbeck JM, Badler NI (2007) Controlling individual agents in high-density crowd simulation. In: SCA '07: Proceedings of the 2007 ACM SIGGRAPH/Eurographics symposium on computer animation, pp 99–108
- Peters C, O'Sullivan C (2009) Metroped: A tool for supporting crowds of pedestrian ai's in urban environments. In: Proceedings of the AISB 2009 convention: AI and games symposium, pp 64–69
- Peters C, Dobbyn S, Mac Namee B, O'Sullivan C (2003) Smart Objects for Attentive Agents. In: Proceedings of the international conference in central Europe on computer graphics, Visualization and computer vision
- Peters C, Ennis C, McDonnell R, O'Sullivan C (2008) Crowds in context: evaluating the perceptual plausibility of pedestrian orientations. In: Eurographics 2008—Short Papers, pp 33–36
- Pletinckx D, Callebaut D, Killebrew AE, Silberman NA (2000) Virtual-reality heritage presentation at ename. *IEEE MultiMedia* 7(2):45–48
- Plinius Caecilius Secundus G (79a) Epistulae vi.16. The Latin Library: <http://www.thelatinlibrary.com/pliny.ep6.html>
- Plinius Caecilius Secundus G (79b) Epistulae vi.20. The Latin Library: <http://www.thelatinlibrary.com/pliny.ep6.html>
- Purcell TJ, Buck I, Mark WR, Hanrahan P (2002) Ray tracing on programmable graphics hardware. *ACM Trans Graph* 21(3):703–712
- Purcell TJ, Donner C, Cammarano M, Jensen HW, Hanrahan P (2003) Photon mapping on programmable graphics hardware. In: HWWS '03: Proceedings of the ACM SIGGRAPH/EUROGRAPHICS conference on graphics hardware, pp 41–50
- Reinhard E, Ward G, Pattanaik S, Debevec P (2006) High dynamic range imaging: acquisition, display and image-based lighting. Morgan Kaufmann
- Reitsma PSA, Pollard NS (2003) Perceptual metrics for character animation: sensitivity to errors in ballistic motion. *ACM Trans Graph* 22(3):537–542
- Rémond M, Mallard T (2003) Rei: an online video gaming platform. In: Proceedings of the 9th international Erlang/OTP User Conference
- Renevier P, Nigay L, Bouchet J, Pasqualetti L (2004) Generic interaction techniques for mobile collaborative mixed systems. In: CADUI 2004: Proceedings of the fifth international conference on computer-aided design of user interfaces, pp 307–320
- Ritschel T, Grosch T, Seidel HP (2009) Approximating dynamic global illumination in image space. In: I3D '09: Proceedings of the 2009 symposium on interactive 3D graphics and games, pp 75–82
- Ropinski T, Kasten J, Hinrichs KH (2008) Efficient shadows for gpu-based volume raycasting. In: Proceedings of the 16th international conference in central Europe on computer graphics, visualization and computer vision (WSCG 2008), pp 17–24
- Rosado G (2008) Motion blur as a post-processing effect. In: Nguyen H (ed) GPU gems 3, Pearson Education, pp 575–581

- Rost RJ (2006) OpenGL shading language. 2nd edn. Addison-Wesley, Upper Saddle River
- Ryan N (2000) Back to reality: augmented reality from field survey to tourist guide. In: Virtual archaeology between Scientific Research and Territorial Marketing, proceedings of the VAST Euroconference
- Ryder G, Flack P, Day A (2005) A framework for real-time virtual crowds in cultural heritage environments. In: M Mudge NR, R S (eds) Vast 2005, short papers proceedings, pp 108–113
- Sanchez S, Balet O, Luga H, Duthen Y (2004) Vibes, bringing autonomy to virtual characters. In: Proceedings of the third IEEE international symposium and school on advance distributed systems, pp 19–30
- Sander PV, Mitchell JL (2006) Out-of-core rendering of large meshes with progressive buffers. In: ACM SIGGRAPH 2006: Proceedings of the conference on SIGGRAPH 2006 course notes, pp 1–18
- Sanwal R, Chakaveh S, Fostirpoulos K, Santo H (2000) Marvins–mobile augmented reality visual navigational system. Eur Res Consort Informatics Math (ERCIM News) 40:39–40
- Sawyer B (2002) Serious games: improving public policy through game-based learning and simulation. Whitepaper for the woodrow wilson international center for scholars
- Scheuermann T, Hensley J (2007) Efficient histogram generation using scattering on gpus. In: I3D '07: Proceedings of the 2007 symposium on interactive 3D graphics and games, pp 33–37
- Scott B (2002) The illusion of intelligence. In: AI game programming wisdom. Charles River Media, Hingham, pp 16–20
- Seetzen H, Heidrich W, Stuerzlinger W, Ward G, Whitehead L, Trentacoste M, Ghosh A, Vorozcovs A (2004) High dynamic range display systems. vol 23, pp 760–768
- Shanmugam P, Arikian O (2007) Hardware accelerated ambient occlusion techniques on gpus. In: I3D '07: Proceedings of the 2007 symposium on interactive 3D graphics and games, pp 73–80
- Shao W, Terzopoulos D (2005) Autonomous pedestrians. In: SCA '05: Proceedings of the 2005 ACM SIGGRAPH/Eurographics symposium on Computer animation, pp 19–28
- Sherrod A (2006) High dynamic range rendering using opengl frame buffer objects. In: Game programming gems 6. Charles River Media, Hingham, pp 529–536
- Shirley P (2006) State of the art in interactive ray tracing. In: ACM SIGGRAPH 2006 courses
- Sinclair P, Martinez K (2001) Adaptive hypermedia in augmented reality. In: Proceedings of the 3rd workshop on adaptive hypertext and hypermedia systems, ACM hypertext 2001 conference
- Smith S, Trenholme D (2008) Computer game engines for developing first-person virtual environments. Virtual Real 12(3):181–187
- Sousa T (2005) Adaptive glare. In: Engel W (eds) Shader X3: advanced rendering with directX and OpenGL. Charles River Media, Hingham, pp 349–355
- Stout B (2000) The basics of A\* for path planning. In: Game programming gems, Charles River Media, Hingham, pp 254–263
- Stricker D, Daehne P, Seibert F, Christou I, Almeida L, Carlucci R, Ioannidis N (2001) Design and development issues for AR-CHEOGUIDE: an augmented reality based cultural heritage on-site guide. In: icav3d'01: Proceedings of the international conference on augmented, virtual environments and three-dimensional imaging, pp 1–5
- Sutherland IE (1965) The Ultimate Display. In: Proceedings of the IFIP congress, vol 2. pp 506–508
- Sylaiou S, Liarokapis F, Kotsakis K, Patias P (2009) Virtual museums, a survey on methods and tools. J Cult Herit 10(4): 520–528
- Tamura H, Yamamoto H, Katayama A (1999) Steps toward seamless mixed reality. In: Ohta Y, Tamura H (eds) Mixed reality: merging real and virtual worlds. Ohmsha Ltd/Springer, Tokyo, pp 59–79
- Tamura H, Yamamoto H, Katayama A (2001) Mixed reality: future dreams seen at the border between real and virtual worlds. IEEE Comput Graph Appl 21(6):64–70
- Tatarchuk N, Isidoro J (2006) Artist-directable real-time rain rendering in city environments. In: Eurographics workshop on natural phenomena
- Tchou C (2002) Image-based models: geometry and reflectance acquisition systems. Master's thesis, University of California, Berkeley
- Tchou C, Stumpf J, Einarsson P, Fajardo M, Debevec P (2004) unlighting the parthenon. In: SIGGRAPH '04: ACM SIGGRAPH 2004 Sketches, p 80
- Thomas G, Donikian S (2000) Virtual humans animation in informed urban environments. In: Computer animation 2000, pp 112–119
- Troche J, Jacobson J (2010) An exemplar of ptolemaic egyptian temples. In: CAA 2010 the 38th conference on computer applications and quantitative methods in archaeology
- Ulicny B, Thalmann D (2002) Crowd simulation for virtual heritage. In: Proceedings of first international workshop on 3D virtual heritage, pp 28–32
- Ulicny B, de Heras Ciecchowski P, Thalmann D (2004) Crowdbrush: interactive authoring of real-time crowd scenes. In: SCA '04: Proceedings of the 2004 ACM SIGGRAPH/Eurographics symposium on computer animation, pp 243–252
- Vanegas CA, Aliaga DG, Wonka P, Müller P, Waddell P, Watson B (2009) Modeling the appearance and behavior of urban spaces. In: Eurographics 2009–State of the Art Reports, pp 1–16
- Vlahakis V, Ioannidis N, Karigiannis J, Tsotros M, Gounaris M, Stricker D, Gleue T, Daehne P, Almeida L (2002) Archeoguide: an augmented reality guide for archaeological sites. IEEE Comput Graph Appl 22(5):52–60
- Wallis A (2007) Is modding useful?. In: Game career guide 2007, CMP Media, pp 25–28
- Wand M, Straßer W (2003) Real-time caustics. In: Brunet P, Fellner D (eds) Comput Graph Forum, vol 22. p 3
- Waring P (2007) Representation of ancient warfare in modern video games. Master's thesis, School of Arts, Histories and Cultures, University of Manchester
- Watt A, Policarpo F (2005) Advanced game development with programmable graphics hardware. A. K. Peters, Natick
- Wright T, Madey G (2008) A survey of collaborative virtual environment technologies. Tech Rep 2008–11, University of Notre Dame, USA
- Wyman C (2007) Interactive refractions and caustics using image-space techniques. In: Engel W (eds) Shader X5: advanced rendering techniques. Charles River Media, Hingham, pp 359–371
- Yin P, Jiang X, Shi J, Zhou R (2006) Multi-screen tiled displayed, parallel rendering system for a large terrain dataset. IJVR 5(4):47–54
- Yu J, Yang J, McMillan L (2005) Real-time reflection mapping with parallax. In: I3D '05: Proceedings of the 2005 symposium on interactive 3D graphics and games, pp 133–138
- Zerbst S, Düvel O, Anderson E (2003) 3D-Spieleprogrammierung. Markt + Technik
- Zhou T, Chen JX, Pullen M (2007) Accurate depth of field simulation in real time. Comput Graph Forum 26(1):655–664
- Zyda M (2005) From visual simulation to virtual reality to games. IEEE Comput 38(9):25–32

### **8.16 Paper #16**

de Freitas, S., Rebolledo-Mendez, G., Liarokapis, F., Magoulas, G., Poulouvassilis, A. Learning as immersive experiences: Using the four-dimensional framework for designing and evaluating immersive learning experiences in a virtual world, *British Journal of Educational Technology*, Blackwell Publishing, 41(1): 69-85, 2010.

Contribution (20%): Collaboration on the design and evaluation of the serious game as well as the write-up of the paper.

## *Learning as immersive experiences: Using the four-dimensional framework for designing and evaluating immersive learning experiences in a virtual world*

**Sara de Freitas, Genaro Rebolledo-Mendez, Fotis Liarokapis, George Magoulas and Alexandra Poulouvassilis**

*Sara de Freitas, Genaro Rebolledo-Mendez and Fotis Liarokapis are all researchers in the field of serious games and virtual worlds. George and Alex are researchers with a specialism in data management and integration technologies. Sara de Freitas, Genaro Rebolledo-Mendez and Fotis Liarokapis are from Serious Games Institute, Coventry University. George Magoulas and Alexandra Poulouvassilis are from London Knowledge Lab, Birkbeck, University of London.*

### **Abstract**

Traditional approaches to learning have often focused upon knowledge transfer strategies that have centred on textually-based engagements with learners, and dialogic methods of interaction with tutors. The use of virtual worlds, with text-based, voice-based and a feeling of 'presence' naturally is allowing for more complex social interactions and designed learning experiences and role plays, as well as encouraging learner empowerment through increased interactivity. To unpick these complex social interactions and more interactive designed experiences, this paper considers the use of virtual worlds in relation to structured learning activities for college and lifelong learners. This consideration necessarily has implications upon learning theories adopted and practices taken up, with real implications for tutors and learners alike. Alongside this is the notion of learning as an ongoing set of processes mediated via social interactions and experiential learning circumstances within designed virtual and hybrid spaces. This implies the need for new methodologies for evaluating the efficacy, benefits and challenges of learning in these new ways. Towards this aim, this paper proposes an evaluation methodology for supporting the development of specified learning activities in virtual worlds, based upon inductive methods and augmented by the four-dimensional framework reported in a previous study.

The study undertaken aimed to test the efficacy of the proposed evaluation methodology and framework, and to evaluate the broader uses of a virtual world for supporting lifelong learners specifically in their educational choices and career decisions. The paper presents the findings of the study and considers that virtual worlds are reorganising significantly how we relate to the

design and delivery of learning. This is opening up a transition in learning predicated upon the notion of learning design through the lens of 'immersive learning experiences' rather than sets of knowledge to be transferred between tutor and learner. The challenges that remain for tutors rest with the design and delivery of these activities and experiences. The approach advocated here builds upon an incremental testing and evaluation of virtual world learning experiences.

### **Background**

The widespread reporting of Second Life (SL)—a social virtual world—has helped to highlight the more general use of immersive worlds for supporting a variety of human activities and interactions, presenting a wealth of new opportunities and challenges for enriching how we learn (eg, Boulos, Hetherington & Wheeler, 2007; Prasolova-Førland, Sourin & Sourina, 2006), as well as how we work and play. In this way, SL, in common with other virtual world applications, has opened up the potential for users and learners, teachers and trainers, policy makers and decision makers to easily collaborate together in immersive three-dimensional (3D) environments regardless of distance in real time. At the heart of the immersive experiences is the presence of the learner or user as an 'avatar' in the virtual space. This avatar represents the embodiment of the user in the virtual space and facilitates a greater sense of control within the immersive environments, allowing users to more readily engage with the experiences as they unfold in real time (Gazzard, 2009).

The more general use of virtual environments over the last few years has been facilitated greatly through Web-based technologies and applications, as well as increasing broadband connectivity and computer graphics capabilities. Together, these allow a range of options in the context of education and training, not least sharing documents and files, holding meetings and events, networking and hosting virtual seminars, lectures and conferences, running research experiments, providing forums for sharing research findings and meeting international colleagues (eg, de Freitas, 2008). Such applications also have an even greater potential for integrating different technologies by supporting social software applications (eg, Facebook, Flickr and Wikipedia), presenting e-learning materials and content, and offering learners' games and rich social interactions. In addition, custom online virtual platforms originating mainly from Universities and research institutes have also been developed mainly for educational and learning purposes (eg, Liarokapis, Petridis, Lister & White, 2002; Liarokapis *et al*, 2004). These are more experimental prototypes and usually use dedicated hardware devices such as advanced visualisation (head-mounted displays, stereoscopic displays), interaction (3D mouse, orientation and position sensors) as well as haptics (gloves). However, usually the costs involved in these types of configurations are still very high, compared to the alternatives presented above.

This flexibility of usage alongside potential global reach for users has led to a sudden and wide growth in the emergence of virtual world applications: in work preparing this paper, 80 virtual world applications were identified with another 100 planned by the end of 2009 (de Freitas, 2008). While not all of these virtual worlds have applicability for learning, and many are aimed at young children (eg, Club Penguin), the extent of the field, not just in terms of potential use for education and training, but actual usage and uptake by users is extensive. For example, SL, a social open world, currently has 13 million registered accounts (as of March 2008). This paper however is focused upon how virtual worlds can be better understood and used specifically in the context of education and training, and here the use of SL for supporting seminar activities and lectures and other educational purposes has been documented in a number of recent studies and reports (eg, Dickey, 2005; Hut, 2007; Jennings and Collins 2008; for a list of examples of SL use by UK universities, see Kirriemuir 2008). Both the broad emergence and the applicability of immersive spaces for undertaking learning have led to wide interest from learning practitioners in finding out more about how they may be best deployed in the class and seminar room.

However, the breadth of applications of virtual worlds, and their relatively swift emergence, has made this a challenging area for researchers and tutors (Hendaoui, Limayem & Thompson, 2008). The area is fragmented due to the nature of its cross-disciplinary appeal and the literature is dispersed around a range of disciplines. Suitably then, this study, undertaken as part of the Joint Information Systems Committee (JISC)-funded MyPlan project (see <http://www.lkl.ac.uk/research/myplan>), led by the London Knowledge Lab, University of London, set out to explore in a cross-disciplinary way how virtual worlds might be most effectively evaluated in relation to designed learning activities, and whether this evaluation methodology could be used as part of the design process and feedback into an iterative design of activities that could then be replicated by other researchers and learning practitioners.

Underpinning this cross-disciplinary approach to the emerging field of serious games and virtual worlds, the authors in previous work have been attempting to reconceptualise ideas around learning, in particular away from more traditional approaches and towards a notion of learning as more centred upon experience and exploration. To understand this we are considering the role of multimodal interfaces (eg, 3D interfaces) and perceptual modelling (cognitive-based approaches), in that our interactions with the environment and our social interactions with others are adopting an approach towards constructing learning experiences as a process of 'choreography' rather than based around data recall strategies (de Freitas & Neumann, 2009). This approach reorganises how we produce and develop learning activities, with a greater emphasis upon learner control, greater engagement, learner-generated content and peer-supported communities, which jointly may increase learning gains. Work outlining an 'exploratory learning model' to support this experience-based and open-ended approach to learning in training contexts is outlined elsewhere (Jarvis & de Freitas, 2009a), and this paper aims to present the outcomes from a study undertaken to

Four Dimensional Framework	
<b>Learner Specifics</b>	<b>Pedagogy</b>
Profile	Associative
Role	Cognitive
Competencies	Social/Situative
<b>Representation</b>	<b>Context</b>
Fidelity	Environment
Interactivity	Access to learning
Immersion	Supporting resources

*Figure 1: The four-dimensional framework*  
 Source: Sara de Freitas, 2008

evaluate the efficacy of using SL as a platform for supporting lifelong learners. In particular, the study was testing the ‘four dimensional framework’ developed in previous studies (de Freitas & Oliver, 2006).

### **Methodology**

Literature searches have found few other evaluative frameworks for exploring the uses and designs of learning activities in virtual worlds, and these are generally training-centred (eg, Fu, Jensen & Hinkelman, 2008). Therefore, this evaluation study adopted an inductive methodology, which requires researchers to construct theories and explanations based upon observations conducted using educational research approaches, including the use of survey data and observations (Gill & Johnson, 1997). A similar approach has been adopted in the Serious Games—Engaging Training Solutions project co-funded by the UK Technology Strategy Board, Selex Systems and TruSim (a division of Blitz Games), but this focused upon measuring the efficacy of game-based learning rather than virtual world learning activities (Jarvis & de Freitas, 2009a). The methodology was selected to address some of the wider issues of efficacy as well as highlighting some of the main issues and challenges arising from this approach to learning and support.

In addition to the inductive method, the study combined the use of the ‘four dimensional framework’ to provide a more structured approach to the synthesis and analysis of the research findings. The four-dimensional framework has been proposed in previous studies and papers, (eg, de Freitas & Oliver, 2006). The framework emerged from user studies with tutors and learners around the selection and use of game-based learning. But it has since been used to support the game design and development process (Jarvis & de Freitas, 2009a). In this study, we applied its use for supporting other immersive experiences—in virtual worlds. The framework proposes four dimensions: the learner, the pedagogic models used, the representation used and the context within which learning takes place (see Figure 1).

The first dimension involves a process of profiling and modelling the learner and their requirements. This profile ensures a close match between the learning activities and the required outcomes. The emphasis upon the learner highlights the importance of the interaction between the learner and their environment. For example, more naturalistic interactions may provide less of a gap in learning transfer. Information and communication technology (ICT) capabilities may affect the way that the learner interacts with the experience, and their abilities to become immersed in the activities in the first place. Feedback to the learner is an important aspect of reflection upon learning and may be central to the most effective learning experience—or individual perception of effectiveness (eg, Jarvis & de Freitas, 2009b).

The second dimension analyses the pedagogic perspective of the learning activities, and includes a consideration of the kinds of learning and teaching models adopted alongside the methods for supporting the learning processes. This may include the use of associative models based upon task-centred approaches of learning and consistent with training methodology (eg, Gagné, 1965), and constructivist models of learning that involve building upon existing knowledge on the part of the learner (eg, Vygotsky, 1978). ‘Situative’ models of learning involve more socially constructed approaches to learning (eg, Wenger’s model of communities of practice, 1998). Particular selection of learning theories may anticipate the types of learning outcomes that result. For example, it has been observed that immersive experiences based upon task-centred analysis and learning task construction result in task-centred outputs, and although effective may be limited to more training-based contexts for learning. Also, certain forms may reinforce particular approaches more readily.

The third dimension outlines the representation itself, how interactive the learning experience needs to be, what levels of fidelity are required and how immersive the experience needs to be. The link between fidelity and learning has been well explored in the work around simulations, but what constitutes interactivity and immersion are relatively under-researched areas and so present challenges for researchers designing experiments. The representational dimension includes the ‘diegesis’ or world of the experience, and may affect levels of engagement and motivation.

The final dimension of context may impact upon the place where learning is undertaken, for example, in school or informal contexts; it may also affect the disciplinary context, for example, which subject area is being studied, and whether the learning is conceptual or applied. Context may also include the supporting resources used for learning. The interactions between the learner and their context are particularly important as the learner may be present in a physical and a virtual space at the same time. These hybrid spaces are relatively unexplored in research terms, but may allow for different approaches to learning beyond those outlined here.

Each dimension has dependencies upon the others; however, jointly, the four dimensions provide a conceptual framework for exploring immersive learning and, we argue, have implications upon learning design as a whole, particularly when applied to immer-



sive learning environments. In part to test the efficacy of the framework and the methodology outlined, the study aims to explore this framework. For ease of use, the findings of the study are synthesised in relation to these four dimensions.

### **Using SL to support planning lifelong learning**

The JISC MyPlan project as a whole aimed to develop a personalised system for planning lifelong learning. The component of the study outlined in this paper aimed to explore the possibilities of using a virtual world for supporting lifelong learners in their career decisions and educational choices. In particular, we were interested to find out whether this method could support mentoring and social interactions for learners in a blended virtual context supplemented with face-to-face tutoring. The study was therefore designed as user studies with two defined groups of learners: learners studying at Birkbeck College on the IT Applications programme and learners from Hackney Community College studying on BTEC courses. The data collection methods for the study included pre- and post-activity surveys, video observations of real world and the in-world sessions, recordings and chat logs. The study was undertaken with ethical considerations and active consent from the participants.

The sessions were held in two computer labs at Birkbeck College, University of London (BBK) and Hackney Community College, London (HCC), and in SL. Learner groups from both institutions were selected for the study. The learners from Birkbeck's IT Applications programme were mature part-time learners all over 18 years of age, and were self-motivated learners. The learners from Hackney Community College were aged between 18 and 24 years of age and were studying for BTEC courses. The two groups offered significant contrast, allowing the researchers to test a range of different responses to the learning activities under exploration. The Learning Day sessions were constructed in order to allow for some degree of structured activities, and some degree of exploration on the part of the learner. The activities functioned as a method for highlighting the main issues arising from this mode of learning, and to aid with producing guidelines for tutors using the tools.

Although the intention was that each learner had access to the Internet, some learners' sessions at HCC were shared since not enough computers were available. User groups consisted of 7 learners at BBK and 14 at HCC. A tutor with experience of SL guided the sessions, which lasted between 2 and 3 hours. At the beginning and end of both sessions individual learners were asked to answer an online survey.

Although factors outside our control altered the sessions (see below), they aimed to take the following structure: an introduction to the session, where the tutor introduces the session, explaining the timetable and answering any questions from the learners. This is followed by an introduction to SL, where the tutor takes learners through an induction into SL, including the creation of an avatar, movement around the virtual world, and text chat functions. This is followed by sessions using a blended approach with face-to-face and virtual components. This includes the tutor and learners visiting the Universities & Colleges Admissions Service (UCAS) SL island, a session with a UCAS



Figure 2: Meeting in-world in Second Life for virtual tour  
Source: David Burden

advisor in SL, a visit to the Serious Games Institute in SL and a short session with David Burden, an expert in SL who discusses the merits of using SL. See Figure 2, where David Burden takes the participants on a virtual tour. The group then visits the IBM island where they walk around and converse with an expert. To complete the session, the tutor holds a debrief meeting with the group, including a discussion about their experience and completion of survey.

#### *Modelling the learner and their learning experiences*

As outlined above, the learner dimension provides a modelling of the needs and requirements of the learner and learner group, including their ICT capabilities. These cohorts were therefore surveyed. A total of 18 learners answered the pre-activity survey, 7 (38.89%) were BBK learners and 11 (61.11%) were from HCC. The average for self-rated ICT skills (using a scale from 1–5, where 1 = not very good and 5 = excellent) was 3.94, where BBK learners' skill was rated as 3.57 and HCC learners' skill was 4.18. The high self-rating for ICT skills, and in particular HCC learners self-rated their ICT skills considerably higher than BBK learners, may be attributed to the difference in age groups or the greater familiarity of younger learners with new technologies. Notably, though our previous user studies had also found a higher estimation of technological capabilities from mature learners (de Freitas, Harrison, Magoulas, Mee, Mohamad & Oliver, 2006).

Also the capabilities of the learners in using related games technologies were surveyed. It was found that in the user groups polled, 66.67% of learners do play video games (28.57% from BBK and 90.91% from HCC), of which 70% from HCC play every day. Video games are played once a week by 50% of BBK learners and 10% of HCC learners; two to five times a week by 20% of HCC learners; and once a month by 50% of BBK learners who play video games. The learners (when asked to select one or more options from the survey) who play video games answered that online games were the most

*Table 1: A comparison of how well liked each aspect of the session was by each user group*

<i>Aspect of session</i>	<i>BBK learners</i>	<i>HCC learners</i>
The face-to-face sessions	3	2.5
Using the SL application	3.14	2.66
Creating avatars	2.2	3.14
Moving in the virtual space	2.42	2.75
The visit to the UCAS island	2.83	3.125
The SGI presentations	3	3.125
The visit to IBM's island	2.85	3
Meeting the experts	3.14	2.87
Reacting with your fellow learners in-world	3.5	3.14

BBK, Birkbeck College, University of London; HCC, Hackney Community College, London; SL, Second Life; SGI, Serious Games Institute; UCAS, Universities & Colleges Admissions Service.

popular (40% from HCC and 100% from BBK), followed by PC (30% from HCC) and console (30% from HCC and 50% from BBK). Other forms of video games that are played are mobile games by 20% of HCC learners and virtual games by 10% of learners.

HCC learners are heavy gamers, which may explain the fact that only a few (18.18%) had seen or experienced a virtual world. This is at least partly attributable to the comparatively higher numbers of users using multiplayer online games when compared with virtual worlds. Surprisingly perhaps, only 22.22% of the sample had used virtual worlds before. Broken down by institution, 28.57% of BBK learners and 18.18% of HCC learners had used this type of application before. All of the learners who had used virtual worlds previously had chosen SL, and none had used a different platform, such as Olive. All of the learners had used SL only once.

A total of 16 learners answered the post-activity survey (two learners from HCC left after the session without having completed the post-activity survey). All seven learners from BBK and nine from HCC completed this survey. When asked how much they had enjoyed the SL session (using a scale from 1–5, where 1 = didn't enjoy the session and 5 = really enjoyed the session), BBK learners averaged 3.14 while HCC learners averaged 3.22. The survey asked learners about how much they had enjoyed the different aspects of sessions. The findings of the survey, including Likert numbers, are included in Table 1 below.

More generally, the survey synthesis found that 43.75% of the sample (42.88% from BBK and 44.44% from HCC) would recommend the use of SL to their friends. However, when asked whether SL sessions helped them to reflect upon their educational choices and career decisions, only 12.5% of the sample answered positively (14.29% from BBK and 11.11% from HCC). Nevertheless, when asked whether they would like to use SL or another virtual world as part of an educational environment for international collaboration with learners globally, the majority of the sample (81.25%) answered affirmatively (100% from BBK and 66.67% from HCC). This indicates that there were problems

with the method we used for structuring the learning activities, for example providing more time for feedback and reflection may have been advantageous.

### *Learning models and theories*

The pedagogic dimension of the study design rested largely upon a posited constructivist model where knowledge construction on the part of the learner was inferred. It was expected that the learners' experiences would build upon previous experiences, in particular using previous experience of similar formats of learning and previous knowledge of career decisions and educational choices. However, this area of the study design perhaps presupposed too much prior knowledge on the part of the learner, and some learners found it difficult to engage with the virtual world. A more structured pedagogic model and more structured activities in-world may have been more effective and warrants further testing. Existing constructivist theories of learning are being supplemented by new ones, currently being piloted (eg, the exploratory learning model of de Freitas and Neumann 2009) and the use of social virtual worlds such as SL favours social interactions. Therefore, although a more constructivist approach was favoured for the study, the findings seemed to point to greater strengths for supporting social learning. One college learner noted that: '[it] brings all people from every aspect of the world together and learn about each other [sic]'. A greater focus upon social interactions and pedagogic models designed to support more socially focused activities may be a better approach for future design. The strengths of the social virtual world need to be better reflected in learning design strategies.

The emphasis and strength of the system for supporting social interactions was supported by the comment from one learner around the use of voice capability: 'we couldn't use voice on this trial, but I'm sure this would help quite a bit.' However, in some studies tutors have expressed a preference for using text interactions due to the ease of turn taking when managing groups of learners in-world. Other studies with SL have demonstrated similar findings to this study. In particular, the study undertaken by Dr Diane Carr observing the use of SL with Masters learners at the Institute of Education, UK as outlined on the Learning in Social Worlds project blog (Carr, 2008), demonstrated some similarities, such as problems with using text chat, disorientation and ambiguity, the need to spend time getting used to the interface and the complexity around structuring experiences that are useful for supporting learning. Carr summarises this:

A great deal of 'structuring' was going on during the sessions—the tutors' frantically [sic] use of Instant Messenger, for instance, that was not visible to the learners. Also, there were 2 or 3 tutors at each session, taking on different roles in relation to content and class management (Carr, 2008).

The study also pointed to strengths of using SL in terms of enhancing social interactions, which is useful for distance and online learners, adding a greater sense of 'presence' than traditional virtual learning environments such as Blackboard, with which the use of SL was compared (rather than face-to-face learning). Carr's study also found

that some individual learners were unable to adapt to the use of virtual worlds. Our observation was that those unfamiliar with text chat had a particular disconnect from the application: as both the 3D interface and text chat were unfamiliar to them, they felt excluded from the session. Learners' capabilities with using the interfaces therefore do need to be considered in advance of using the technologies, and additional induction training may be needed in these cases, or alternative learning strategies (eg, Web-based) may be offered.

#### *Usability, interactivity and accessibility*

In the course of this study, the main area of consideration with relevance to the representational dimension focused on the usability of SL. There were clearly issues with the technology, not least because there were significant problems with connectivity and local development work being undertaken at Linden Lab that day that affected the access to the system, and had a negative impact upon the study findings. These technical issues had a clear impact upon the transfer of the learning experience.

The comments from learners underlined these technical issues. In the area of usability of the system, learners commented that 'movement was a bit sluggish, but I suppose that's more to do with the Internet connection I think.' One of the college learners noted: 'make it so it dont [sic] glitch as much and add a few more features to the island'. The connectivity problems were significant and led to some comments from the learners, such as 'a better Internet connection would have allowed us to have a "fuller" experience. I think that would have made it better'. The issues are significant and tutors aiming to use SL would have to find coping mechanisms for these kinds of problems that occur with limited broadband, accessibility issues and regular maintenance work at Linden Lab. The newness of the technologies and the architectural issues with SL has led a group of open source developers to develop OpenSim (<http://www.opensimulator.com>), with the aim of developing a more scalable architecture and allowing the application to be hosted behind institutional firewalls, reducing considerably the technical issues experienced on the day of testing.

However, despite these difficulties at least one of the mature learners could see real benefits for those using the application with disabilities:

I work with drama/theatre and people with a disability—acquired brain injury—who are on a programme getting them back to work. I think there are some really interesting possibilities in helping to develop confidence among such clients interacting virtually before or as an adjunct to 'real' life social interaction and skills development.

The representation of the virtual world itself therefore can have a negative impact upon learning, not least because of the level of expectation on the part of the learner. There is evidence that regular gamers find the graphics of virtual worlds too low level, and can experience negative transfer as a result. Learner expectation is a factor for tutors to deal with when using immersive worlds. However, if the activities are well structured and feedback is given by tutors to the learners then there are possibilities for using the tools,

in particular, where social interactions and support may be required. The representation of the virtual world then creates an additional design tool for the tutor: once usability and accessibility issues are addressed, the tutor may explore learning through the interchange between the real and virtual representations—or hybrid spaces (and experiences). In this context, virtual worlds may be used as metaphors of learning or life experiences that can be reflected upon and interacted with in social groups.

#### *Real and virtual contexts*

There were wider contextual issues that affected the efficacy of the learning experiences, and these centred upon a lack of engagement with the virtual worlds due in part to specific learners' background and age. For example, one or two learners did have problems relating to the format. One mature learner commented, 'I am afraid that I cannot relate to the virtual world'. Another learner commented that 'I think anyone new to SL would need someone to show them how to use it, as it is not intuitive to non computer games players.' The first learner commented that 'my worry is that it would exclude people who weren't technologically sophisticated.' The first learner felt that: 'I can't relate to a virtual world and imaginary people; it makes me restless and want to be with real people.' Interestingly, this learner found it difficult to relate to the fact that avatars were all human-driven, and felt distanced from the real people due to the interface and use of avatars. This was compounded by the fact that the learner was not familiar with the process of text chat and found it alienating for communicating with others.

In addition, the study raised particular issues around accessibility and usability, including the quality of broadband connectivity and the user interface design. It is undeniable that using SL behind the institutional firewalls is a difficult and imprecise undertaking, and negative first impressions can be off-putting to the extent that some will not return. As an indication of this, Linden Lab estimate that half of all users never return after their first hour in SL (Lorica, Magoulas & the O'Reilly Radar Team, 2008). However for those that do there are interesting applications that can be investigated (de Freitas, 2008).

It is worth considering that while the learners were participating in a study situated at college and university, and as such the context of learning was strictly formal, it would be interesting to gauge the reactions if the study were undertaken in informal learning settings, at home, or in work based settings.

#### **Discussion**

While multiplayer games may have educational potential in the future, virtual worlds are generally regarded as having greater educational potential (de Freitas, 2008). Currently this is broadly because of the focus of activities. However, the method for comparing the benefits of structured activities in games over open-ended explorations of virtual worlds is an area in need of further research. Of interest here may be how to bring together the structured activities of games with the exploration and social power

of virtual worlds. The motivational capacities of game-play when brought together with the social interactions of virtual worlds may be a powerful teaching combination in the future.

The wider trends of technical convergence between games technologies and educational uses is occurring in the shape of serious games and simulations, but while simulations and games for learning are more established approaches, and have more literature to accompany them, the uses of virtual worlds for learning is still a relatively new field, and as this preliminary study has shown there is a significant learning curve when using virtual world applications to support learning, both for tutors and learners. The main impediment lies in the context and familiarity of the form. Indeed, factors such as where the virtual world is used and the past experience of users with the system are significant aspects ensuring or preventing effective use. Additionally, prior experience of gameplay may not be a positive factor, and previous game experience may in fact have a negative impact upon learning with virtual world applications, as game-players are used to much higher levels of fidelity and interactivity than are presently available in virtual worlds. With convergence, this is in the process of changing, but as the testing session revealed issues, such as firewalls and graphics, capabilities of hardware can significantly reduce the immersion of the experience and so reduce the effectiveness of the experience.

The technical issues did significantly impede the users' seamless experience and, in contrast with other studies, the least liked aspects of the interaction in SL were creating avatars and moving in-world. This was certainly due to extremely slow connections as a result of maintenance work that day at HCC and due to multiple users on the network at BBK, both of which caused slow download times. In general, the research indicates that control over avatars can be a critical aspect of allowing users to become engaged and motivated through empowerment of controlling their own representation in-world although as Carr, for example, has indicated for some learners this can be off-putting and produce a 'pain barrier' to be overcome. From our study, it was clear that the college learners felt more familiar with the process of avatar creation and that this did hold their attention: Figure 3 shows a college learner who had personalised his avatar within a few minutes of using SL, although he had no prior knowledge of SL. Younger learners are adapting to new approaches more readily and concepts such as avatars and customisation of one's avatar are integrated into their prior knowledge of online gaming.

The research team experienced significant challenges with assessing and validating the efficacy of SL for supporting educational choices and career decisions, in terms of the methods of structuring of exercises, providing the best support for the learners and also in terms of technical issues experienced by the users. While some learners were clearly visibly engaged, more work is needed to find out ways of engaging more learners with how to structure the activities, and greater support in advance of trialling is required. More rigorous frameworks and metrics would also be useful for supporting future efficacy studies. The research team would like to undertake further larger and more longitudinal studies towards that end.



Figure 3: Photo of one of the students participating in the study in *Second Life*  
Source: Sara de Freitas, 2008

Reflecting on these difficulties, only a handful of learners tested (12.5% of learners) expressed that SL helped them to reflect upon their educational choices and career decisions. This indicates that the platform is one in which the format used with users would not be appropriate for mentoring learners. In particular the technical issues such as accessibility and usability were too jarring for the learners, and got in the way of them appreciating the value of the form. Problems with SL such as connection speed, difficulty to move around, orientation, lack of signposts, and not using voice as used in a classroom setting, impeded the study. The HCC did visit the UCAS island, but needed more support with their interactions with the information there. They also thought more signposting on the island would be helpful. They enjoyed visiting the IBM island but also needed more support and guidance in-world. Due to technical issues it was not possible to provide this. However, if the activities were better structured, and the technical issues could be overcome then the format may have potential for mentoring and other socially-driven interactions and learning modes.

On the other hand, 81.25% of learners saw positive links between using SL as part of an educational environment for international collaboration with learners globally. This indicates that there are other aspects of SL that may be used in the future for supporting socially-based learning activities designed for lifelong learners. The social dimension of SL is clearly a powerful component of the format, and when the technology becomes more stable, and broadband and sufficient graphics capabilities can be guaranteed within institutions, then it could be used for role play, mentoring and for social skills acquisition.

The main lessons arising from this study demonstrated a need to evaluate the platform with a larger sample of learners. While this study is useful for defining some of the



evaluation issues, larger numbers of learners would yield a richer dataset and more scope for analysis. In addition, there is a need to consider the design criteria for more structured activities, find ways to better orientate the learners and tutors in advance of the study, and a need to utilise more concerted and experienced technical support and resources.

While it was found that the inductive methodology of data collection was effective for providing information about the use of SL (in particular, the combination of chat logs, video footage and surveys was useful for providing a more multidimensional impression of the usage of SL), the use of in-depth semistructured interviews with some of the participants would have been useful for providing a more qualitative dimension for study findings analysis. A follow-up study examining the design, development and use of virtual worlds for tertiary education with lifelong learners would be helpful for validating this evaluation methodology. Moreover, a study using greater numbers of users exploring the patterns of use of modules being taught in SL, in particular with a comparison between face-to-face learner groups, pure distance or online learners and hybrid groups of both would be desirable.

The use of immersive learning centrally implies a shift from considering and designing learning tasks to choreographing learning experiences as a whole, mediated by structured and semistructured social interactions. This has implications upon elements of how the learning day as a whole is structured in terms of the different requirements such as duration of sessions, breaks, and necessary facilities and technical support. But it also has implications upon pedagogic considerations, such as learning theories and models applied, the role of the tutor and the context of learning. This shift merits consideration of learning experiences as involving social interactions between members of the learning group, supporting exploratory individual pathways and identification of methods of tutoring that focus more upon mentoring and guiding development. Towards this end, tutors may analyse the learner group and consider their ICT skills levels, game experience and learning approaches. Also, they may consider the pedagogic approaches needed for the subject area taught, learner group and context of learning. Use of the four-dimensional framework can support this process, in terms of the selection of media used and the questions that the tutor needs to ask themselves when structuring and considering the most appropriate ways of integrating immersive learning into their plans.

Orientation is important for new users of virtual worlds to induct them into using the platform, and for maximising their engagement with virtual worlds as a whole. As this study has demonstrated, those who are familiar with gaming and who use multiplayer games regularly often find the unstructured and open-ended aspect of virtual worlds difficult to adapt to, as they are used to more structured and purposeful activities, and it can take a long while for them to adapt to these more open and exploratory social worlds. In order to support learners who are novices or regular gamers, it would be useful to hold start-up sessions with learners in advance of learning sessions to allow learners to become orientated with the user interface. For example, sessions may be

held where learners log in remotely from home, allowing sufficient time for them to become used to the interface, and minimising the technical issues. In addition to that, orientation sensors (ie, Wiimote) may be used to allow for more tangible orientation in the virtual social worlds.

### **Conclusions**

This study set out with the intention of testing a virtual world using a predeveloped evaluation methodology and approach. The approach was based upon an assumption that learning experiences need to be designed, used and tested in a multidimensional way due to the multimodal nature of the interface. To support this, the four-dimensional framework was used with the inductive method to gather data and to synthesis and analyse the findings. As a whole, the approach has worked well in this first iteration, its main strength being that the use of the evaluation methodology allowed the research team to evaluate the learning experience according to specific criteria. The presented evaluation methodology may be used as a design tool for designing learning activities in-world as well as for evaluating the efficacy of experiences, due to its set of consistent criteria. The approach does augment the existing methods for evaluation, but needs to be tested with a larger sample and in wider contexts of use to verify its efficacy across different platforms.

While the study itself was affected by technical issues that in general were off-putting for those unfamiliar with virtual worlds, still some benefits of using SL for supporting under-served learners, for engaging learners and for supporting distributed groups of learners were highlighted, due to the engaging nature of the form and to its international reach. While it is generally considered that improvements of the SL platform, and the advent of OpenSim and other new-generation virtual worlds will significantly reduce many of the technical issues experienced by the learners, it is also recognised that such tools are still relatively immature and that more work needs to be undertaken to establish their most effective uses, to produce clear guidelines and to exploit their capabilities to the highest degree.

Particular strengths of the medium were highlighted; for example, the learners were positive about using the tools for supporting international collaboration, indicating the power of the tool for supporting distributed learning communities based upon shared interests. While the study has not proved conclusively the power of the tool for mentoring, the sessions with the mentor were very effective in practice, and in the future one-to-one sessions with mentors based abroad or not co-located could be further explored. However, more context and advance study is needed to situate the activities in-world and greater time for reflection needs to be provided. Virtual worlds may also support peer collaboration and may be used, for example, for collaborative assignments in-world with practical outputs, for example, designing a marketing campaign in-world, and work centring upon social interactions would be well served in this virtual world. Also, there is real potential for supporting online learning methods by extending the benefits of audio-graphic conferencing to provide a greater sense of presence, thereby potentially reducing non-completion rates.

The potential for using a social virtual world such as SL for supporting life decisions and educational choices has been established with this study, but thorough testing of sessions, appropriate technical support, use of established and tested pedagogical principles and well-structured sessions are essential for providing enriched experiences that are properly contextualised for the learner. In particular, this immersive learning approach could work well with distance and online learners, distributed user groups or as an additional support for face-to-face learners. The use of virtual worlds may also need to be considered with respect to using a 'blend' of other media support mechanisms, such as videoconferencing and virtual learning environments, which may help to support the community-based and social collaborative strengths of immersive environments.

## References

- Boulos, M., Hetherington, L. & Wheeler, S. (2007). Second Life: an overview of the potential of 3-D virtual worlds in medical and Health education. *Health Information & Libraries Journal*, 24, 4, 233–245.
- Carr, D. (2008). *Learning to Teach in Second Life. Report for Learning from Online Worlds; Teaching in Second Life. Institute of Education/Eduserv Foundation, April 2008*. Retrieved October 13, 2008, from <http://learningfromsocialworlds.wordpress.com/learning-to-teach-in-second-life/>
- Dickey, M. D. (2005). Three-dimensional virtual worlds and distance learning: two case studies of Active Worlds as a medium for distance education. *British Journal of Educational Technology*, 36, 3, 439–451.
- de Freitas, S. (2008). *Serious virtual worlds: a scoping study*. Bristol: Joint Information Systems Committee. Retrieved April 27, 2009, from <http://www.jisc.ac.uk/publications/publications/seriousvirtualworldsreport.aspx>
- de Freitas, S. & Neumann, T. (2009). The use of 'exploratory learning' for supporting immersive learning in virtual environments. *Computers and Education*, 52, 2, 343–352.
- de Freitas, S. & Oliver, M. (2006). How can exploratory learning with games and simulations within the curriculum be most effectively evaluated? *Computers and Education*, 46, 249–264.
- de Freitas, S., Harrison, I., Magoulas, G., Mee, A., Mohamad, F., Oliver, M. et al (2006). The development of a system for supporting the lifelong learner. *British Journal of Educational Technology*. Collaborative e-support for lifelong learning, 37, 6, 867–880.
- Fu, D., Jensen, R. & Hinkelman, E. (2008). *Evaluating game technologies for training. IEEE aerospace conference*. San Mateo, CA: Stottler Henke Assoc., Inc.
- Gagné, R. M. (1965). *The conditions of learning*. New York: Holt, Rinehart & Winston.
- Gazzard, A. (2009). The avatar and the player: understanding the relationship beyond the screen. In G. Rebolledo-Mendez, F. Liarokapis & S. de Freitas (Eds), *Proceedings of the IEEE Games and Virtual Worlds for Serious Applications conference, Coventry, UK* (pp. 190–193). *Proceedings of the 1st IEEE International Conference in Games and Virtual Worlds for Serious Applications, IEEE Computer Society, Coventry, UK*, 23–24 March.
- Gill, J. & Johnson, P. (1997). *Research methods for managers* (2nd ed.). London: Paul Chapman Publishing.
- Hendaoui, A., Limayem, M. & Thompson, C. W. (2008). 3D social virtual worlds: research issues and challenges. *IEEE Internet Computing*, 12, 1, 88–92.
- Hut, P. (2007). Virtual laboratories. *Progress of Theoretical Physics*, 164, 38–53.
- Jarvis, S. & de Freitas, S. (2009a). Towards a development approach for serious games. In T. M. Connolly, M. Stansfield & E. Boyle (Eds), *Games-based learning advancements for multi-sensory human-computer interfaces: techniques and effective practices*. Hershey, PA: IGI Global.
- Jarvis, S. & de Freitas, S. (2009b). Evaluation of an Immersive Learning Programme to support Triage Training. *Proceedings of the 1st IEEE International Conference in Games and Virtual Worlds for Serious Applications, IEEE Computer Society, Coventry, UK*, 23–24 March (pp. 117–122) ISBN: 978-0-7695-3588-3.

- Jennings, N. & Collins, C. (2008). Virtual or Virtually U: educational institutions in Second Life. *International Journal of Social Sciences*, 2, 3, 180–186.
- Kirriemuir, J. (2008). *Measuring the impact of Second Life for educational purposes*. Retrieved August 4, 2008, from <http://www.eduserv.org.uk/foundation/sl/uksnapshot052008>
- Liarokapis, F., Mourkoussis, N., White, M., Darcy, J., Sifniotis, M., Petridis, P. *et al* (2004). Web3D and augmented reality to support engineering education. *World Transactions on Engineering and Technology Education*, 3, 1, 11–14.
- Liarokapis, F., Petridis, P., Lister, P. F. & White, M. (2002). Multimedia Augmented Reality Interface for E-Learning (MARIE). *World Transactions on Engineering and Technology Education*, 1, 2, 173–176.
- Lorica, B., Magoulas, R. & the O'Reilly Radar Team (2008). Virtual worlds: a business guide: 2008. O'Reilly Radar Report. <http://radar.oreilly.com/research/virtual-world-report.html>. Publisher name: O'Reilly, Web publication, date accessed 29th October 2009.
- Prasolova-Førland, E., Sourin, A. & Sourina, O. (2006). Cybercampuses: design issues and future directions. *Visual Computing*, 22, 12, 1015–1028.
- Vygotsky, L. S. (1978). M. Cole, V. John-Steiner, S. Scribner & E. Souberman (Eds), *Mind in society*, Cambridge, Massachusetts and London England: Harvard University Press.
- Wenger, E. (1998). *Communities of practice*. Cambridge: Cambridge University Press.