Journal Subline

LNCS 7544

Fotis Liarokapis
Guest Editor

Transactions on Edutainment IX

Zhigeng Pan • Adrian David Cheok • Wolfgang Müller
Editors-in-Chief

Springer

Lecture Notes in Computer Science 7544

Commenced Publication in 1973
Founding and Former Series Editors:
Gerhard Goos, Juris Hartmanis, and Jan van Leeuwen

Editorial Board

David Hutchison
Lancaster University, UK
Takeo Kanade
Carnegie Mellon University, Pittsburgh, PA, USA
Josef Kittler
University of Surrey, Guildford, UK
Jon M. Kleinberg
Cornell University, Ithaca, NY, USA
Friedemann Mattern
ETH Zurich, Switzerland
John C. Mitchell
Stanford University, CA, USA
Moni Naor
Weizmann Institute of Science, Rehovot, Israel
Oscar Nierstrasz
University of Bern, Switzerland
C. Pandu Rangan
Indian Institute of Technology, Madras, India
Bernhard Steffen
TU Dortmund University, Germany
Madhu Sudan
Microsoft Research, Cambridge, MA, USA
Demetri Terzopoulos
University of California, Los Angeles, CA, USA
Doug Tygar
University of California, Berkeley, CA, USA
Moshe Y. Vardi
Rice University, Houston, TX, USA
Gerhard Weikum
Max Planck Institute for Informatics, Saarbruecken, Germany

Zhigeng Pan Adrian David Cheok
Wolfgang Müller Fotis Liarokapis (Eds.)

Transactions on Edutainment IX

Springer

Editors-in-Chief

Zhigeng Pan
Hangzhou Normal University
Hangzhou, China
E-mail: zhigengpan@gmail.com

Adrian David Cheok
Keio University
Kanagawa, Japan
E-mail: adriancheok@kmd.keio.ac.jp

Wolfgang Müller
University of Education
Weingarten, Germany
E-mail: mueller@md-phw.de

Guest Editor

Fotis Liarokapis
Coventry University, UK
E-mail: f.liarokapis@coventry.ac.uk

ISSN 0302-9743 (LNCS) e-ISSN 1611-3349 (LNCS)
ISSN 1867-7207 (TEDUTAIN) e-ISSN 1867-7754 (TEDUTAIN)
ISBN 978-3-642-37041-0 e-ISBN 978-3-642-37042-7
DOI 10.1007/978-3-642-37042-7
Springer Heidelberg Dordrecht London New York

CR Subject Classification (1998): K.8.0, K.4.2, K.3.1, H.5.1-3, I.3.7

© Springer-Verlag Berlin Heidelberg 2013
This work is subject to copyright. All rights are reserved, whether the whole or part of the material is concerned, specifically the rights of translation, reprinting, re-use of illustrations, recitation, broadcasting, reproduction on microfilms or in any other way, and storage in data banks. Duplication of this publication or parts thereof is permitted only under the provisions of the German Copyright Law of September 9, 1965, in ist current version, and permission for use must always be obtained from Springer. Violations are liable to prosecution under the German Copyright Law.
The use of general descriptive names, registered names, trademarks, etc. in this publication does not imply, even in the absence of a specific statement, that such names are exempt from the relevant protective laws and regulations and therefore free for general use.

Typesetting: Camera-ready by author, data conversion by Scientific Publishing Services, Chennai, India

Printed on acid-free paper

Springer is part of Springer Science+Business Media (www.springer.com)

Preface

In this issue, one of the focuses is on "serious games", which covers a broad range of applications from Adobe Flash-based applications to fully immersive 3D environments where users interact with large volumes of data through sophisticated digital interfaces. The shift toward immersive world applications being used to support, among others, education and training activities marks the beginning of new challenges that present considerable scope for collaborative and multi-disciplinary research solutions, as well as opportunities for innovative development. The set of papers on serious games (the "special section") is selected from the VS-Games 2011 conference, which addressed some of the significant challenges of the cross-disciplinary community that works around these serious application areas by bringing the community together to share case studies of practice, to present virtual-world infrastructure developments, as well as new frameworks, methodologies and theories, and to begin the process of developing shared cross-disciplinary outputs. This special section comprising eight papers is focused on educational aspects of serious games and virtual worlds for serious applications. Papers were selected on the basis of fundamental ideas and concepts rather than the thoroughness of techniques deployed.

The special section on serious games starts with a paper by Birchall and Gatzidis, entitled "The Periodic Table of Elements via an XNA-Powered Serious Game." The paper presents the production of a serious game, which illustrates both the potential of a cost-efficient platform itself for programmers of a medium skill set but also the possibilities of designing to completion a fully functional game that can be used as part of secondary education chemistry curriculums in order to teach the important topic of the periodic table of elements. The second paper is by Christopoulos et al., entitled "Digital Storytelling Within Virtual Environments: The Battle of Thermopylae," focuses on an interactive virtual reality application developed for the museum of Thermopylae located at the site of the original battle, near the city of Lamia in Greece. Storytelling techniques and principles of modern video games were used to disseminate historical knowledge about the battle and the associated legends.

The third paper is by Backlund et al., entitled "Games on Prescription! Evaluation of the Elinor Console for Home-Based Stroke Rehabilitation." It presents the feasibility of Elinor, a game-based system for stroke rehabilitation in the home. The Elinor prototype has been positively evaluated with respect to its usability, user acceptance, and motivational factors as well as its rehabilitation effect. The fourth paper is by Chilcott and Smith entitled "An Analysis of the Potential to Utilize Virtual Worlds to Enhance Edutainment and Improve the Wellbeing of the Ageing Population." This paper identifies the importance of developing and designing tailored virtual places which older members of society can utilize easily and safely within this context; with a focused application of addressing social isolation, enabling wellbeing and improving edutainment.

The fifth paper is by Vosinakis et al., entitled "Course Lectures as Problem-Based Learning Interventions in Virtual Worlds," and investigates the suitability of Virtual Worlds (VWs) as a platform for hosting PBL (Problem-Based Learning) activities and explores their affordances in terms of collaboration support and learning effectiveness. Results reveal several encouraging findings about PBL and collaboration mediated by VWs, and lead to a series of recommendations. The sixth paper is by Mathieu et al., entitled "Virtual Customers in a Multiagent Training Application." The paper argues that a relevant way to train a salesperson on their daily activities (e.g., customer relationship management, store management, and stock control) consists in immersing them in a 3D environment populated with realistic virtual customers.

The seventh paper is by Paliokas et al., entitled "Game-Based Early Programming Education: The More You Play, the More You Learn." The paper examines various widely known mini-languages with an emphasis on LOGO implementations and follows a combinational route to take advantage of both Game-Based Learning (GBL) and the use of mini-languages in the design of a new LOGO-like environment. The final paper of this special section is by Hulusic and Rizvic, entitled "Story Guided Virtual Environments in Educational Applications". The paper demonstrates a new concept of using story-guided virtual environments for cultural heritage virtual reconstruction, with live virtual guiders in an interactive Flash format.

Besides the papers on the topic of serious games in the special section, we have another 13 papers selected from differents sources such as DMDCM 2010, EGVR 2011, CIDE 2011, ICVRV 2011, DMDCM 2011, and ordinary submissions. They are described briefly in the following.

In the paper by Xunxiang Li, entitled "Rendering Technology of 3D Digital Chinese Ink-Wash Landscape Paintings Based on Maya," the author's work includes analyzing and simulating the techniques of traditional Chinese ink-wash landscape paintings, e.g. the ways of brush moving and ink painting, exploring the digital rendering modes with the characteristics of wrinkled-texture paintings. In the paper by Qian Li et al., entitled "Effects of RPG on Middle School Players' Intrapersonal Intelligence," the authors test the intrapersonal intelligence of 192 middle school students who play role-playing game (RPG) by questionnaire, aiming at exploring the effects of RPG on intrapersonal intelligence. The results showed that RPG has a positive effect on students' intrapersonal intelligence, and the effect is subject to factors such as age, frequency of playing games, and RPG type. The paper by Jun Dai et al., entitled "Implementation of Lushan Virtual Digital Plant Museum", introduces the application of virtual technology in the design process of the Lushan Virtual Digital Plant Museum.

In the paper by Keyang Cheng, entitled "Pedestrian Detection Based on Kernel Discriminative Sparse Representation," the author puts forward a novel framework for pedestrian detection tasks, and proposes a model with both sparse reconstruction and class discrimination components, jointly optimized during dictionary learning. The author presents an efficient pedestrian-detection system using mixing sparse features of HOG, FOG, and CSS to combine into a Kernel classifier. The paper by XiaoJuan Hu et al., entitled "The Design and Implementation of Computer-Aided Chinese Medicated Diet System," describes research ideas and implementation technologies on Chinese Medicated Diet System (CMDS). The design and implementation of CMDS are discussed as key issues including the design of persistence layer and database access. In the paper by Desheng Lv et al., entitled "Music-Driven Emotion Model Applied in Digitalized Dance Performance of Sacrificial Ceremony for Confucius," a digitalized system, which was developed to reconstruct the musical dance performance of the ceremony, includes music analysis based on an emotion model and an action movement library based on a virtual model for the choreography with the motion capture technique, and the action correlation with a music-dance match.

In the paper by Pengyu Zhu et al., entitled "Real-Time Rendering Framework in the Virtual Home Design System," the authors introduce a home design system with its great functions and framework design, including the scene management based on the Cell&Portal system, improved variance shadow mapping, and the recently popular real-time rendering framework called deferred lighting. In the paper by Aiping Wang et al., entitled "Multiple-Cue-Based Visual Object Contour Tracking with Incremental Learning," the author proposes a visual object contour tracking algorithm using a multi-cue fusion particle filter. A novel contour evolution energy is proposed which integrates an incrementally learnt model of object appearance with a parametric snake model. In the paper by Liangjie Zhang et al., entitled "Building Virtual Entertainment Environment with Tiled Display Wall and Motion Tracking,"an immersive and interactive entertainment environment that integrates tiled display wall and motion tracking techniques is introduced. A fast calibration method is proposed to achieve geometry alignment and color consistency of the tiled display wall. A robust motion tracking algorithm is designed to obtain the player's moving direction and speed from video streams captured by a Web camera.

In the paper by Shidu Dong et al., entitled "A Calibration Method for Removing the Effect of Infrared Camera Self-Radiance on the Accuracy of Temperature Measurement," a novel three-phase method is proposed, which is outlined below. First, from a set of samples, how the pixel value of each blackbody varies with the camera temperature is determined. Second, at the given camera temperature, the calibration function, describing how the pixel value varies with the temperature of blackbody, is formed. Finally, with the aid of the calibration function, the temperature of the probe object can be determined by its pixel value. In the paper by Shi Yan et al., entitled "Enjoying of Traditional Chinese Shadow Play – A cross-culture study", a cross-culture study was conducted to investigate the effect of Piying-induced emotion on heart rate and heart rate variability during Piying perception and performance. The result confirmed that Piying performance was far more effective in emotion induction than Piying perception. In the paper by Li Xiao et al., entitled "A Patch-Based Data Reorganization Method for Coupling Large-Scale Simulations and Parallel Visualization," a patch-based data reorganization method was presented for this coupling through a parallel file system. Based on the method, simulation data sets in application codes are reorganized by patch and written into many files in parallel. The last paper by Eric Klopfer et al., entitled "The Boom and Bust and Boom of Educational Games," is an overview paper on education games.

All the papers in the special section and on other topics clearly demonstrate the use of serious games and virtual worlds for edutainment applications. This is a new area and in the coming years a number of research issues will be further explored and new ideas will be developed. Sincere thanks to all the people who helped to make this special issue as well as all the authors for contributing their work and the anonymous reviewers for their constructive comments and suggestions.

January 2013

Fotis Liarokapis
Zhigeng Pan

Transactions on Edutainment

This journal subline serves as a forum for stimulating and disseminating innovative research ideas, theories, emerging technologies, empirical investigations, state-of-the-art methods, and tools in all different genres of edutainment, such as game-based learning and serious games, interactive storytelling, virtual learning environments, VR-based education, and related fields. It covers aspects from educational and game theories, human-computer interaction, computer graphics, artificial intelligence, and systems design.

Editors-in-Chief

Zhigeng Pan	Hangzhou Normal University, China
Adrian David Cheok	Keio University, Japan
Wolfgang Müller	University of Education Weingarten, Germany

Managing Editor

Yi Li	Nanjing Normal University, China

Editorial Board

Ruth Aylett	Heriot-Watt University, UK
Judith Brown	Brown Cunningham Associates, USA
Yiyu Cai	NTU Singapore
Maiga Chang	Athabasca University, Canada
Holger Diener	Fhg-IGD Rostock, Germany
Jayfus Tucker Doswell	Juxtopia Group, USA
Sara de Freitas	The Serious Games Institute, UK
Lynne Hall	University of Sunderland, UK
Masa Inakage	Keio University, Japan
Ido A Iurgel	Universidade do Minho, Portual
Kárpáti Andrea	Eötvös Loránd University, Hungary
Lars Kjelldahl	KTH, Sweden
James Lester	North Carolina State University, USA
Nicolas Mollet	IIT, Italy
Ryohei Nakatsu	NUS, Singapore
Ana Paiva	INESC-ID, Portugal
Abdennour El Rhalibi	JMU, UK

Daniel Thalmann	EPFL, Switzerland
Kok-Wai Wong	Murdoch University, Australia
Gangshan Wu	Nanjing University, China
Xiaopeng Zhang	IA-CAS, China
Stefan Goebel	ZGDV, Germany
Michitaka Hirose	University of Tokyo, Japan
Hyun Seung Yang	KAIST, Korea

Editorial Assistants

Ruwei Yun	Nanjing Normal University, China
Qiaoyun Chen	Nanjing Normal University, China

Editorial Office

Address: Ninghai Road 122, Edu-Game Research Center, School of Education Science, Nanjing Normal University, Nanjing, 210097, China
E-mail: njnu.edutainment@gmail.com; edutainment@njnu.edu.cn
Tel/Fax: 86-25-83598921

Table of Contents

Papers from VS-Games 2011

Regular Papers

The Periodic Table of Elements via an XNA-Powered Serious Game

James Birchall[1] and Christos Gatzidis[2]

[1] Save The Children, 1 St John's Lane, London, EC1M 4AR, UK
[2] Bournemouth University, School of Design, Engineering And Computing, Talbot Campus, Poole House, Poole, BH12 5BB, UK
J.Birchall@savethechildren.org.uk, cgatzidis@bournemouth.ac.uk

Abstract. This publication concentrates on the production of a serious game on Microsoft's XNA Studio, which illustrates both the potential of a cost-efficient, off-the-shelf development platform itself for programmers of a medium skillset (such as for example educators) but also the possibilities of designing to completion a fully functional game that can be used as part of secondary education chemistry curriculums in order to teach the important topic of the periodic table of elements. Additionally, a 15-subject pilot evaluation study of the game (titled Elemental) is included, which displays initial evidence that there can be educational benefits for the experiment participants exposed to the work.

Keywords: serious games, physics, XNA, education, games development, software engineering.

1 Introduction

It is widely accepted that the term "serious game" was originally used in the late 1970s, referring to learning from predominantly board and card-based games. This theme carried on with advancing technology, forming today's concept that categorises serious games as interactive media for some sort of purposeful use other than amusement. This article will focus on serious game development and use for (secondary) school education. Currently, through a variety of educational systems across different countries, pupils/students learn skills in ranging subjects across a number of years. It is therefore evident that there can occasionally be limited avenues in terms of offering information in new ways as a teacher/educator. Serious games could be the key to diversifying teaching mediums.

The problem facing developers interested in delivering content for this area is two-fold, firstly how easy are serious games to design and develop; and secondly, do they educate enough to supplement or replace certain materials? Video games are today, with the advent of free community tools, much simpler to develop and if teachers, lecturers and other educating bodies are willing to explore other new mediums, video games could indeed be an excellent candidate. This publication will be looking at how to form a serious game design for secondary educational chemistry training, its

Z. Pan et al. (Eds.): Transactions on Edutainment IX, LNCS 7544, pp. 1–28, 2013.
© Springer-Verlag Berlin Heidelberg 2013

subsequent detailed development and, finally, pilot evaluation of its usefulness as a learning tool using empirical data.

While the development of a full commercial game using a traditional production pipeline can be difficult, with the right tools, such as the Microsoft XNA Studio, as described in this article, even with minimal basic programming experience someone could rapidly develop a game in a few weeks. The educational game outlined here, called Elemental and deployed on the XBox 360 platform, is a chemistry game based around the periodic table of elements and atomic composition, a key part of the GCSE curriculum (an important UK secondary education academic qualification). The prototype of the Elemental game is evaluated using 15 participants during a 2-stage trial study. Using questions based around recent GCSE chemistry exam papers, a set of quizzes were developed to test participants, both before and post playing the game. The results suggest differences in evidence of the players widening their knowledge on the subject after the use of Elemental.

It should finally be noted that this publication is an extended version of the work presented by the authors at the IEEE VS Games 2011 (Third International Conference in Games and Virtual Worlds for Serious Applications) [1].

2 Background

The last few years have seen the emergence of the serious games movement; with games used in a number of new areas not associated purely with entertainment. Under the guise of terms such as edutainment, games have been used for purposes other than entertainment for a long time. It appears that only recently has this produced significant results; the modern serious games movement is now worth an estimated $200-$400 million in the US alone. Additionally, it is beginning to become characterised by the level of complexity and sophistication behind it, as well as the number of different directions explored. This fact, combined with advancing technology in both software and hardware and prior inhibitions about games in general lifted (for example, associations with violent and anti-social behaviour), has turned into interest from the educational sector in conducting further research in this field, including case studies. It remains true that the area of education, one of the first explored by serious games, even before this term was introduced, is the one that still holds the greatest potential.

There are a number of surveys in the area such as the ones by Vogel et al [2] and Dempsey et al [3] highlighting the many benefits in numerous application and subject areas that educational gaming can cultivate in an audience. Secondary education is an area of particular interest for the introduction of serious games. Current examples of pursuit of this avenue in the literature include the use of the medium in a variety of subject areas.

In science, Lopez-Morteo and Lopez [4] explored an approach for motivating students towards mathematics. This involved educational interactive software components presented through recreational mathematics. This work also included evaluation of the approach with three math courses for high-school students. Results were positive, indicating the methodology’s usefulness, particularly towards motivating students. Ke and Grabowski [5] also conducted research in the same area, that of mathematics, this time using 125 pupils. This involved the pre-test and post-test approach and also examination in a number of covariates and their correlation with the results. The results are once again very positive for game-playing but also for collaborative play which was also examined and was found to further enhance learning.

Another notable effort in teaching mathematics via game-based learning is Zombie Division [6],[7] (by Baker et al and Habgood respectively), a game which bears similarities to a typical adventure game and aims to teach elementary school pupils the subject of division. Employing basic combat mechanics and placing the player in the role of an ancient Greece hero, the pupil can progress by selecting the correct divisor number when confronted with a zombie enemy.

There is more research strongly indicating the help of computer games in the learning of maths curriculum such as the work by Sedighian and Sedighian [8], Rosas et al [9] and finally, more recently, Kebritchi et al [10]. It is interesting to see this subject area explored more than others by serious games development and manipulation. This can be partly attributed to the quantifiable manner mathematics learning can be observed and tested for.

Other subject areas have also been explored. In history, Egenfeldt-Nielsen [11] conducted a study as part of his PhD using 72 high-school students and also teachers in Denmark. This was achieved using a strategy game to draw conclusions about the potential benefits of video game playing in obtaining knowledge. Also in history, Squire used the very popular strategy game Civilization 3 to investigate learning via the medium [12]. In geography, and with particular attention to urban planning using SimCity 2000, Adams [13] also showcased positive results for the potential of game-based learning in this area.

Other research in geography-aimed game-based learning includes the work of Tüzün et al [14], who have employed a 3D MUVE called Quest Atlantis [15]. Quest Atlantis supports educational activities via different tasks in the guise of quests. This is set in the mythical land of Atlantis and offers a rich 3D virtual world experience but also the possibility of extensive collaborative teaching and learning. Because of the open nature of Quest Atlantis, work in distinct other curriculum areas, such as the teaching of writing has emerged, again supported by the use of this platform, such as the efforts of Barab et al [16] and Warren et al [17]. General science game-based learning has also been attempted on Quest Atlantis, and also in this case evaluated, as described by Lim et al [18].

Evidence of MUVE game-based learning has also been observed in biology teaching. Using the platform of the River City [19], a MUVE designed for participation-oriented immersive experiences, Dede et al [20] conducted an experiment which assessed the motivational effects of such an approach in teaching biology to pupils. Results were very encouraging, with biology knowledge and inquiry content showing signs of increase. Similarly, Nelson [21] performed a study using the same platform to establish the degree of guidance needed in teaching with MUVE as a tool, with insignificant results this time, despite demonstrating knowledge gain in the population sample.

In the area of chemistry and physics there are not many examples to draw knowledge from as it has not been as explored as much as other subject areas. The most important example is the work of Squire et al [22]. This focused on using a simulation game, called "Supercharged!", tackling the subject of electromagnetism, an important physics concept.

While educational gaming has a lot to offer there are concerns too. Randel et al [23] bring up the issue of bias from the educator. Also, there is the sometimes dubious social aspect of game-based learning as discussed by Stoll [24] (although it has to be mentioned that many games since the publication of this research have included increasingly complex collaborative routes to further strengthen their delivery in this area) and finally, the distraction and the long-term retaining of knowledge acquired by game-based learning, as argued by Clark [25].

There is also the tendency in research to rely on commercial products or titles for the development of the serious game that is to be tested or trialled in an educational setting. While that can offer many advantages, it can also considerably hinder the exploration of enabling the teacher or educator to single-handedly produce a flexible tool that could aid him/her in this area.

The rest of this article pursues this avenue, demonstrating the process of an implementation such as that via the discussion of a number of key design/technical decisions, problems, issues and eventual trial of a custom-made educational game called Elemental. There is a strong focus on the technical development of the prototype so that an interested party could be presented with a comprehensive build guide on the creation of an educational game.

Elemental focuses on the periodic table of elements, the well-known tabular display of the 117 chemical elements used to classify different forms of chemical behaviour. The periodic table of elements is a particularly important part of the chemistry curriculum for the UK General Certificate of Secondary Education (or GCSE in short) academic qualification awarded in this corresponding subject. Undertaking GCSEs it should be noted is a huge and key part of secondary education for students of ages ranging from 14 to 16, i.e. in secondary education, across the span of the United Kingdom (i.e. in England, Wales and also Northern Ireland).

3 Requirements and Analysis

This section of the article focuses on the design requirements of the Elemental game. Using the waterfall method of software development, the requirements are the first parts of the game to be defined. This will then provide the project a direction, allowing design to cater to these requirements. There are multiple methodologies developers can use for this, spanning from rapid design and implementation using extreme programming (XP), to the more categorical Waterfall method methodology. For the Elemental implementation, a custom version of the Waterfall method will be used. The reason for this is that it features all of the most important stages of software development and can be customised easily. It can also be made to suit developing a small game by a single developer (i.e. a teacher/educator, as is intended in this case) who wants to deploy it quickly in a classroom as a project that will be used as complimentary to the usual teaching delivery. Originally, the method developed by Royce was for large-scale software development [26]. It follows 7 standard steps, which progress from requirements to implementation.

The requirements mainly come from the learning objectives and the game design process. Learning objectives consist of the knowledge the students should take away after playing the game. Testing will use these requirements to deduce when to finish with the implementation. The first set of requirements is based around defining the artefact as a game:

- The artefact will be made using game development tools
- The artefact will mentally challenge the player
- The artefact will have a defined game play
- Game play will contain a set of common sequences used throughout the game in a repetitive manner; however the content must change throughout

Firstly, to make a game there are many specialised tools already in place for developers. Large-scale projects will use many middleware tools to develop games, most of which are too costly or difficult to use for a single person with no or little budget. For a teacher and/or educator the decision on the platform needs to be heading in the opposite direction. The Elemental project has selected the Microsoft XNA Studio, along with Visual Studio 2008. The toolkit, essentially a wrap-around of the more extensive Microsoft DirectX API, is a very popular choice for many hobbyists exploring the graphics/game programming and already has a middleware set of libraries making game development easier. XNA (which according to Microsoft stands for XNA's Not Acronymed) was first announced in 2004 and, after its 2006 release, it quickly managed to become an established brand of collective assets for game developers, for a number of platforms. The major appeal of it is the fact that it is now a free download, enabling hobbyists, students and of course educators to experiment with it in order to create content of their own choice and for a variety of platforms. Prior to XNA, for the creation of similar content, more complex APIs such as OpenGL and DirectX had to be used. Due to the complexity of these API libraries and the basic access to hardware they provided, including key areas such as managing memory, creating game content was almost prohibitive for a novice. XNA, in an effort to deviate from generic game engine frameworks that the aforementioned APIs favoured, included a custom-made, existing content pipeline for the inclusion of 3D models, textures and audio (amongst others). Also, timing and render loops were created in a Game class with Update and Draw Methods for further simplification. Finally, a framework similar to a game engine, called Application Model provided hardware abstraction to a GPU.

For educators with some limited programming experience, the XNA platform is a great choice, as it provides an entirely clean canvas for game content creation, with the added benefit of being simple enough to master in a short period of time. Even today, there are few platforms that can offer the creative freedom, undemanding technical requirements and cost-effective attributes of XNA. It could be argued that more comprehensive, somewhat more recent, GUI-based and user-friendly middleware tools/engines such as Unity, OGRE or Torque X have emerged that could potentially provide equal levels of support, particularly for a novice developer. However, at the time of the development of Elemental Unity required a license fee, Torque X still does to this day (and is essentially a graphical-based "wrap-around" of XNA to begin with), while OGRE demands a far steeper learning curve in terms of programming skills because of the vast range of features it can offer (particularly for 3D games development, which was not needed on this occasion).

The second requirement is more related to a game challenge. The game has to challenge the user mentally, very much like sports games challenge the players physically. The challenge could take many forms, such as getting the user to remember something, identify something or react to something. Using challenges like this the player should learn something to aid them in better understanding the game and becoming “better” at it. As the game will be educational/geared towards learning, it is expected that it will be something ‘educational’ that is added into those challenges. The next two requirements refer to giving the game a structure. By developing a game with a specific set of rules to which the basic gameplay is defined, one can then be put in the changing content within that. If the players were to put the game down and come back in a week, it would still play in the same way and hopefully they would remember that style.

The next set of requirements will define the content:

- Content should be understandable without any prior experience in chemistry or the periodic table of elements
- Content will include an introduction to explain the game, the basic of atomic composition and controls
- Content should be displayed in a visual format
- Content will include every element (all 117) from the periodic table of elements
- Content will include the name, chemical symbol and composition of each element in the periodic table
- Content will include radiation, showing how radiation can be called
- The game will be based around a high score system
- The game will include a timer
- Content will display atomic composition of elements

The final set of requirements is defining the game system requirements:

- The game will run on a PC with a basic graphics card and for the Microsoft XBox 360
- It will be controllable by either a Microsoft XBox controller or a keyboard/mouse combination

The game should be as accessible as possible. This is why the game is being developed for both PC and the XBox 360. These can also be developed in line with each other as XNA Studio allows for this. The final requirement is that both versions will work with an XBox 360 controller. The controller can be used on any Windows XP/Vista/Windows 7 PC and will allow for streamlining of controls too. Coming up with a specification for software such as a game does differ to developing a piece of software for other functions or purposes. The reason for this is that games are increasingly seen more as an art form and do not always need to be developed to provide a solution to something. In this case, the solution is to provide knowledge of the periodic table of elements and atomic composition to a dedicated audience using specific content.

4 Design

This section of the article focuses on the design of the game covering areas such as class and code design. The project will be focusing on developing an element generating system over a compound creation system. The reason for this is that creating elements as the main game will meet the requirements of teaching the player about particle composition better. The game is based around building each element of the periodic table. The main mechanic will be the "catch and release" system. This system will use the same buttons to catch and release the different particles. They will be released in a 360 degree radius outside the screen and follow a path through the very centre. To catch them, the user has to use the same button, however when catching it is key to note which particle it is as to where it can be caught. For it to be as realistic as possible, the neutrons and protons are caught inside the nucleus which is depicted by the small red circle in the centre of the screen. The electrons can be caught on the shells on the outside.

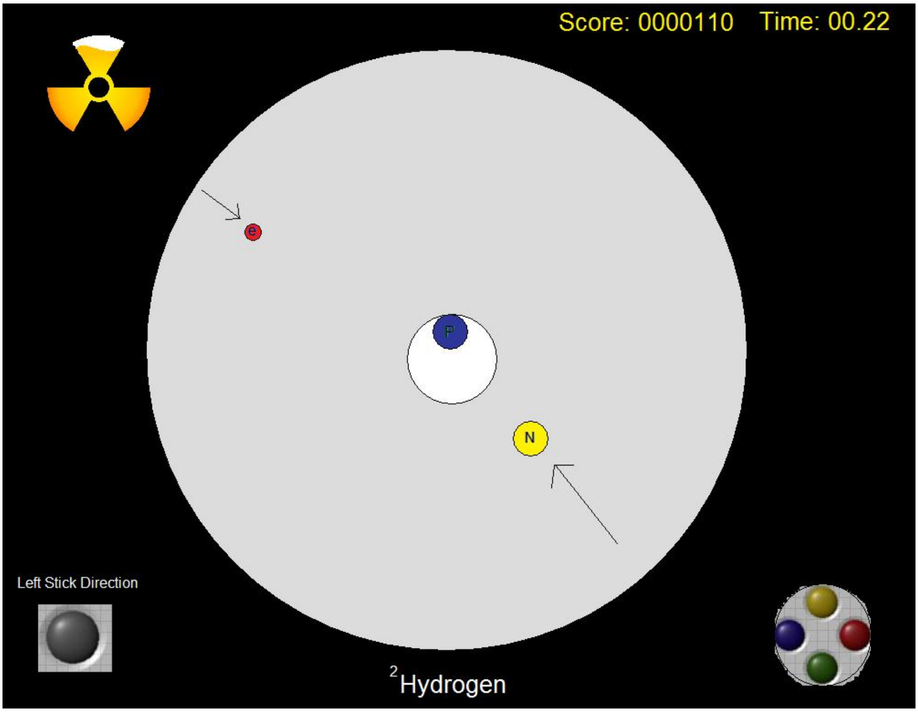


Fig. 1. Initial concept screenshot of the Elemental game

The next mechanic is the "particle gun" mechanism. The "gun" can move 360 degrees around the screen. Being able to move in both ways was also key; this was also initially added as the "charge of particles" mechanic would not work without it. The reason for this is because of how the particles work. Simulating realistic working particles of the same charge will repel, neutrons and particles of different types will cancel each other out, meaning they will disappear and will need to be re-released and caught. If the player is trying to catch a particle but misses its "catching area", it will move out of the screen in the same direction and increase radiation. Should radiation reach its limit of 100%, then the game will be 'over'. As radiation gets to a higher limit, the screen will also start to invert the colours, showing that the viewing piece is getting damaged. A scoring mechanic will increase points depending on which particle was caught. As electrons are harder, they will be worth more points. Another idea here was to also include a bonus score, should the player get a certain combination of particles.

The game will play in levels. Each level will be a different element from the periodic table. There are a few ways this could work. The first and simplest is making each element in order of atomic mass (which was the one followed in Elemental). Starting with the basic element Hydrogen and working up the periodic table was how it was finally implemented in the prototype game. Using a timer on each level will add a sense of urgency and another point of failure for the player should the time run out. Each level will add a set amount of time giving the play extra time as there will be more particles.

The ideas above need to be converted into code. As C# (the programming language XNA uses) is an object-oriented language the game will be developed with this in mind. The easiest place to start is by deciding on some classes on which to begin formulating the game. Microsoft XNA Studio was developed to be a toolset for game developers (unlike for example DirectX, which is a more encompassing graphics library). For this reason, when one sets up the first project it will setup the first class and a set of functions. The first class in this case will be the "Game" class. This class will contain the working parts of the game including the engine that runs it. The next class is the "Particle" class. This class will be used to hold particle data for manipulation by the game engine. The "Level" class will contain the settings for each level. The last class is another developed by the toolkit and is called "Program". This class calls the "Game" class and contains the "main" entry method of the game.

The games constructor class will contain code to set up outside variables. This is to make sure they contain a value and are not null. The reason for this is when using global values they are not always defined, until certain requirements are met but will be needed by many functions. By default the XNA toolkit includes a variable to setup the Graphics Device Manager. This automatically detects the setting of the machine and implements the default resolution and display buffers. Other variables defined here will be ones such as the Window name and the root directory of the project contents.

The Load Content method will be used to define the content in the game. This includes loading in the element definitions, the sprite manager, textures such as buttons and also generating the lists holding the co-ordinates for the Nucleus, the electron shells and the particle gun circle. The Update function is where the logic of the program is run. The logic will define which level the player is on, check which particles have been caught, decide whether the level is complete and the process of collision detection and radiation. The Draw function will perform the drawing of content to the graphics device, which will be the screen. This includes drawing the buttons, particles, radiation, score and timer. As the function will be too large incorporating all this, it has been split and calls other functions to draw specific parts.

These are the main functions of the Game class. However, other methods are called where necessary to do specific tasks, these include:

- Draw Buttons, this function decides which buttons need to be drawn to the screen. The buttons will change to show a depressed button like that on the controller
- Get Mouse State, which checks the mouse wheel state. For the PC version, the mouse wheel will be used to rotate the "particle gun" by incrementing or decrementing the co-ordinates
- Get Key State, this will determine which state the keys are in on the keyboard or controller. Depending on whether it is ready to catch or release then it will also call the Create Particle function
- Create Particle, this function determines which type of particle should be created and adds it into the game
- Draw Particle, called by the Draw function it goes through each particle current in the game and draws them
- Draw Score, this uses the setup font and draws the current score to the player's view
- Draw Radiation, similar to Draw Score but with the radiation limit
- Draw Timer, similar to Draw Score but with a countdown timer
- Draw Limits, this function will draw the circle depicting the nucleus boundary and the outer shells
- Check Collisions, this function will check collisions between particles at every update

The main Particle class functions include:

- Particle constructor, called when a new particle is created, all data is input at this point
- Change State, this function changes the state of the particle to depict whether it is active or not
- Update End Position, this function will change the end position for when a collision happens and both particles are the same, it reverses the position so that it looks like it has been repelled

The only function Level has is the constructor. This is used to setup the level initially when called in the Load Content method of the Game class. No change ever occurs to the levels so no other functions are needed.

Some of the functions are specialised to the game. The first piece of code that will be challenging to design is the code for getting the circle points generated. The nucleus, particle gun, all outer shells and also the particles will be using these. The particle gun co-ordinates have to be setup slightly different than the other co-ordinates, because not only will each particle need the starting point, but it will also need to get the opposite point to move towards. To get around this, one could generate a half circle instead of a full circle and put it in an array. One can then generate another set of points to mimic the opposite side and store them in another array. When lining up the array, each co-ordinate will be mapped to its opposite co-ordinate in the other array.

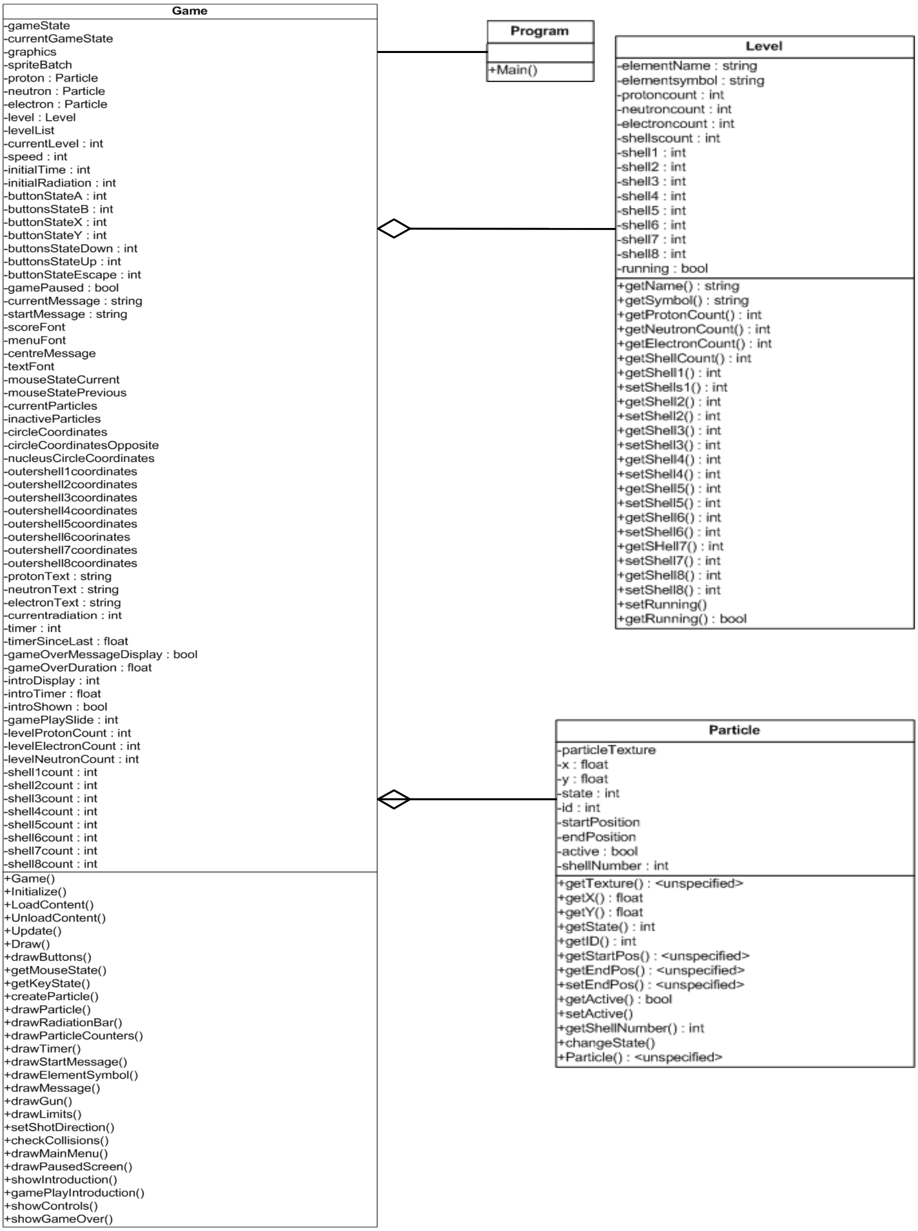


Fig. 2. Class diagram for the final version of the Elemental game

The code for generating the circle point will work as showcased in Figure 3. To develop the 2 half circles you can stop the loop after getting to PI and starting the next loop with the angle equalling PI. Processing the particles currently active in the game could cause problems as there will be multiple lists holding different types of particles.

One list will hold particles currently in motion and ready to be caught by the player. Another will be used as storage, while the particles are processed after moving to it.

1. START
2. Define X as resolution width / 2
3. Define Y as resolution height / 2
4. Define radius of co-ordinates
5. Define angle
6. WHILE angle is less than 2 * PI(Pie)
7. Define temporary x as Cosine of the angle * radius * X
8. Define temporary y as Sine of the angle * radius * Y
9. Define temporary vector with the co-ordinates X, Y
10. Add the temporary vector to the correct list
11. Increment the angle by a certain amount

Fig. 3. Pseudo-code for the generation of the circle point

The final is for particles which are now out of use after being correctly caught. Neutrons and protons are caught in the same way and share the code for being caught. Electrons however can be caught in different shells. This means having to check the position of the electron being processed against all known shell points. For each particle in the currently active list the program will need to determine how far they have travelled each update. To do this, both x and y co-ordinates of the particle are checked against the position it is moving to. The statements then adjust the co-ordinates factoring in speed. The final Elemental prototype game (and its help screen) are shown in Figures 4, 5 and 6.

Fig. 4. Control help screen for the Elemental game

Fig. 5. Final game played during the experiment

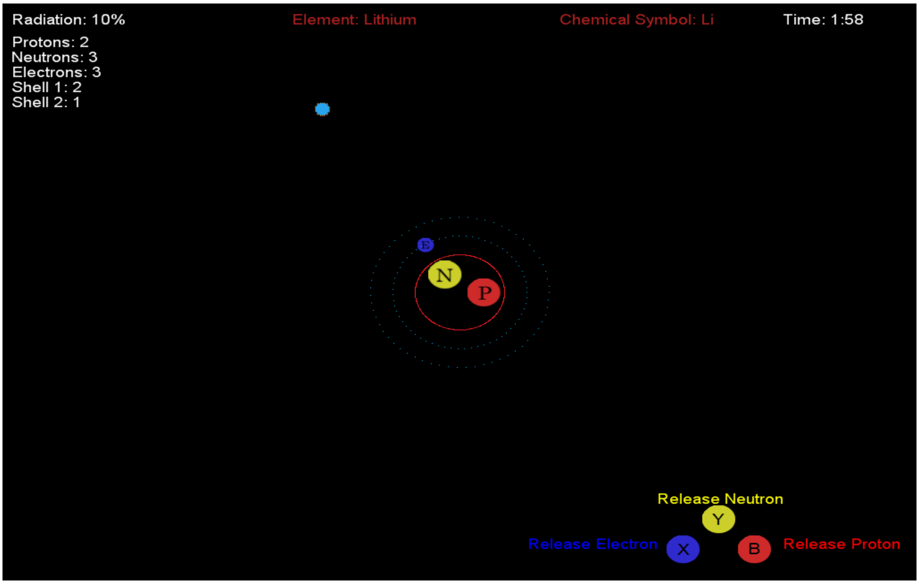


Fig. 6. Final game screenshot

5 Evaluation Study

The pilot evaluation took place in 3 stages and focused on whether the game could fit into a teaching environment and be classed as a serious game. More importantly, the facilitation of learning within the context area (periodic table of elements) needed to be examined. The first stage was to develop a suitable questionnaire. The second was the process of getting individuals to evaluate the game. The third involved the analysis of the results. Wilson describes the first part of designing a questionnaire is to have clear objectives [27].

These objectives are going to be based in 3 areas and will cover these 4 hypotheses:

- Does the game show indication of teaching the players in the field of chemistry?
- Does the game challenge players in a coherent style?
- Can the game be played by players who lack experience in the field of video games?
- Can the game be played by players who lack experience in the field of chemistry?

The first hypothesis looks into whether the game can teach a player. This is part of the criteria to make the game serious and forms the most important hypothesis of the project, essentially asking the question "can the game teach?" The second hypothesis is used to define the software as a video game. Going back to the introduction, it was stated that video games need to challenge players mentally, while offering a stylised way of playing. Without having a questionnaire, we can already interpret the game possesses this last attribute. It clearly plays in one style but at the same time does have a range of deferent content to keep players engaged.

The last two hypotheses are looking at whether the game can be played by someone with little or no knowledge of video games or chemistry. This was originally a single question but splitting this up meant that we can see which factor may make more of an impact in learning from a serious game. The next concept from Wilson is to get people to answer questions carefully and completely [27]. To do this, the questions will mainly use a 5-point Likert scale. This will give enough scope for the user to put a reasonable answer, while restricting the ambiguous nature of open answers. However, one of the flaws of using this approach is that it does restrict the participants from expressing themselves. For this reason, a comments box will be added to the end of the questionnaire for anything else the participants wanted to express but could not in their Likert-type answer. The questionnaire data collection took place in 4 stages and followed this format:

Initially participants were given a short set of demographics questions. This included age, sex and occupation. The next 4 questions focused on understanding the experience level of the participant. This was expressed as experience in video games, educational games, chemistry and the periodic table of elements. These 4 questions (and their corresponding answers) could be used to group participants into skill levels when analysing the results.

The quiz part was taken immediately after the demographic questions. There were two quizzes given to the participants which both followed a similar format. The questions themselves mainly came from past GCSE exam papers and were multiple choice (three possible answers). Sample questions included the following; "What is the chemical symbol for Lithium?", "How many electrons does Nitrogen have?", "Which element has the chemical symbol: B?", "How many protons does Neon have?" etc. The first quiz was taken and they were then shown the Elemental game, which was to be played for 15 minutes in total. Once the 15 minutes had elapsed, participants would then take the second quiz. For both quizzes, questions were to be answerable from playing the game during a 15-minute time period.

After the post-gameplay quiz participants then had the last section of the questionnaire to fill in. This section related to the game, the content, knowledge gained and how the participants now perceived serious games. A comments box at the end was for anything extra which was qualitatively-natured information that may have been of use during the analysis of the evaluation study results.

The pilot evaluations took place during a 2-week period using an XBox 360. 15 participants took part with each session lasting about 30 minutes (10 male and 5 female). The pilot study was conducted mainly in Bournemouth University, UK, therefore the majority of participants were students and also staff from the University itself. Demographics data for the experiment are shown in Figure 7.

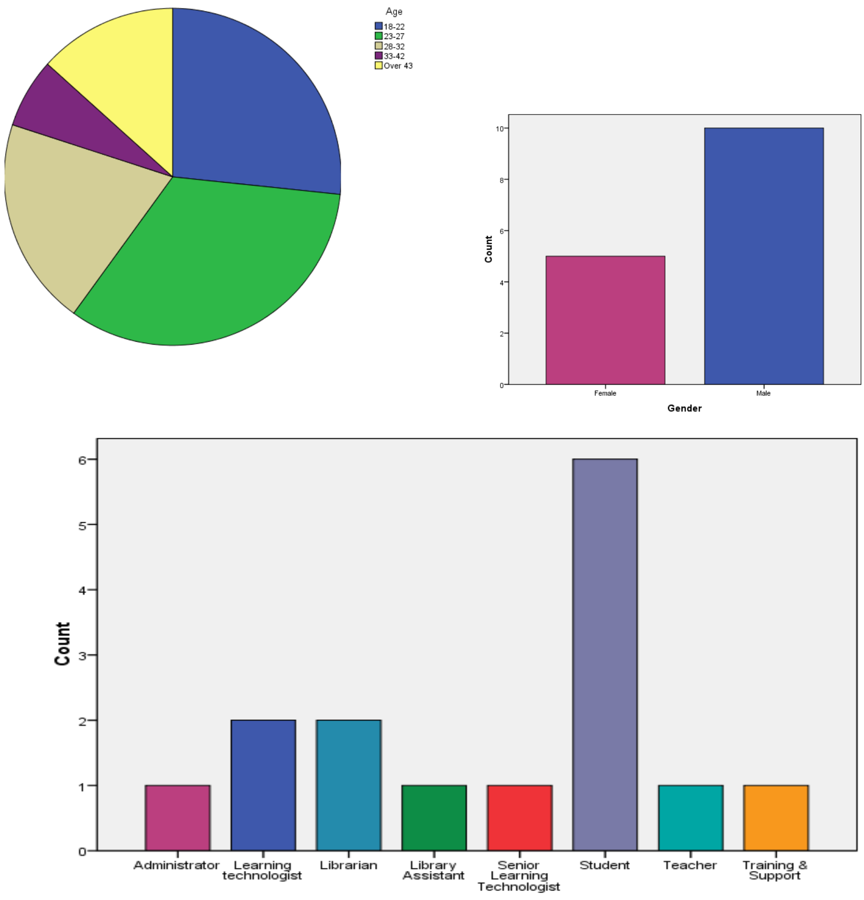


Fig. 7. Demographics data for the Elemental experiment

Experience is something that the study cannot change. Each individual will have different experience levels in different areas, so to cater for this it was decided to ask them to record it in a self-reported capacity. All collected experience results from the study are shown in Figure 8. Starting with experience in playing computer or video games, the overall data showed experience in this area to be moderate to high. It is clear that those who rate themselves as higher are in the younger categories of age. As serious games are (to this date) fairly specialised, the result from experience in educational games is expected, with the majority scoring themselves as low in experience. A similar response was given with the experience in chemistry. This was possibly due to the fact that no one who participated was involved in any work related to chemistry. The last question before the pre-game quiz was regarding the experience of the periodic table. Figure 8 shows that the majority of participants rated themselves between no experience and average experience. Surprisingly, although most participants expressed that they had low prior chemistry experience, the highest count for knowledge of the periodic table was higher than the chemistry knowledge. One reason for this may be that due to there being a high number of student participants, they may not have remembered much about chemistry in general but remember the periodic table. Another reason may be that because the periodic table of elements is only a small part of chemistry (and its curriculum), the participants generally felt they knew more in this specific area.

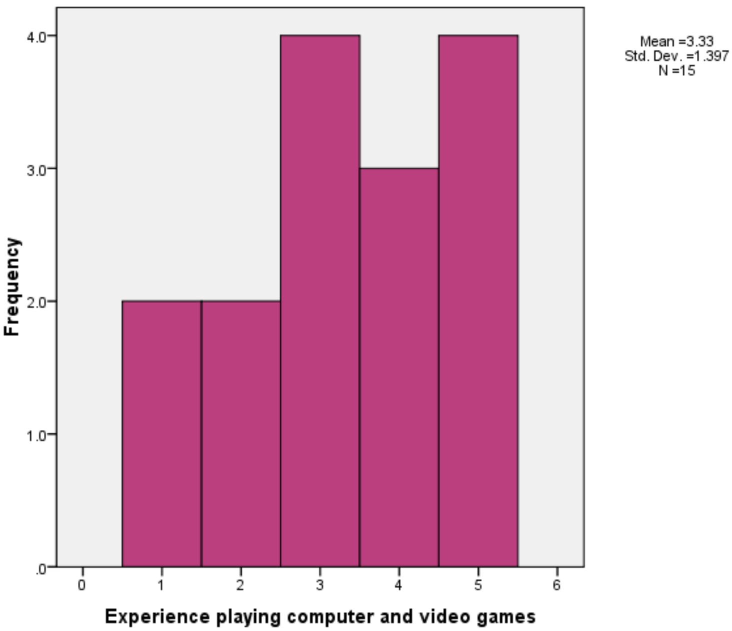
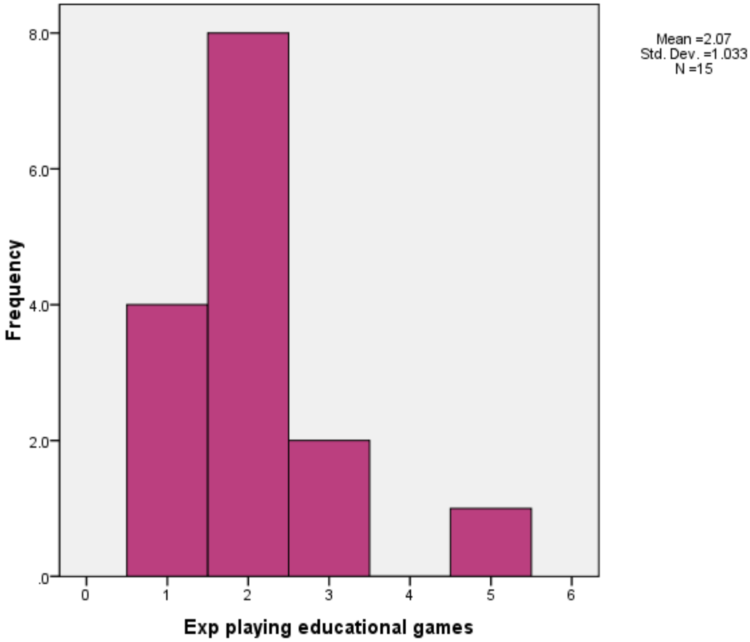


Fig. 8. Prior experience data collected



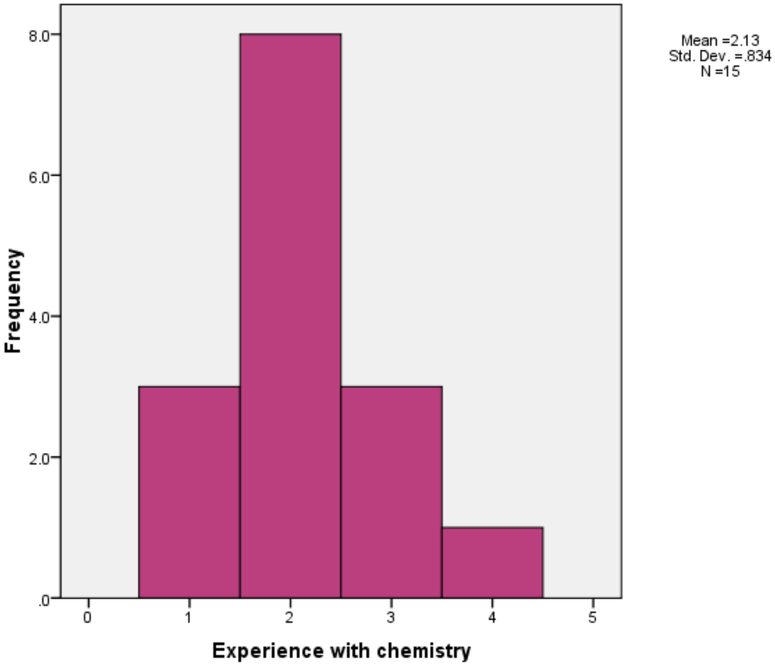


Fig 8. (*Continued*)

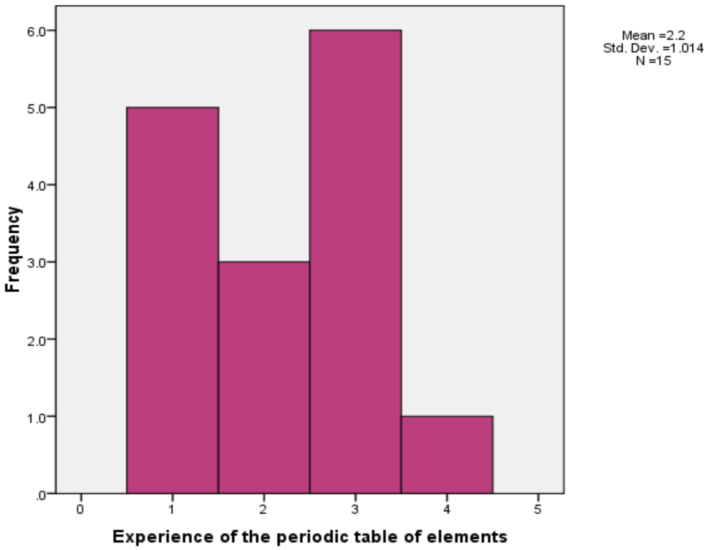


Fig 8. (*Continued*)

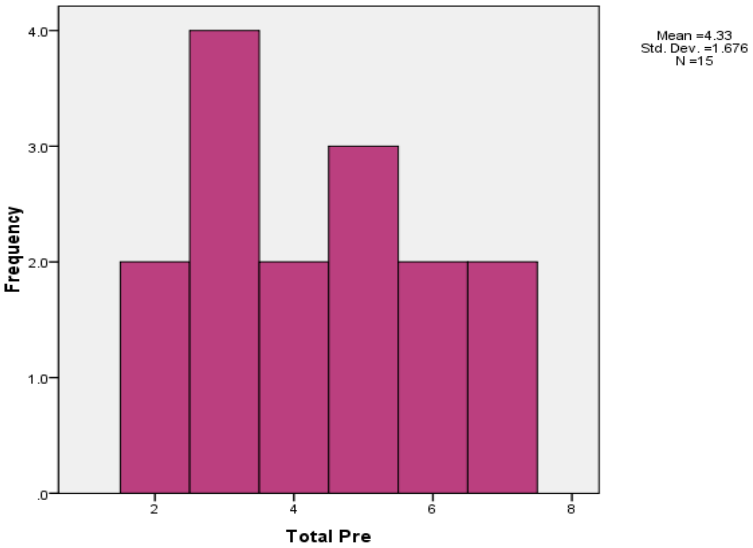


Fig. 9. Pre/post gameplay quiz results

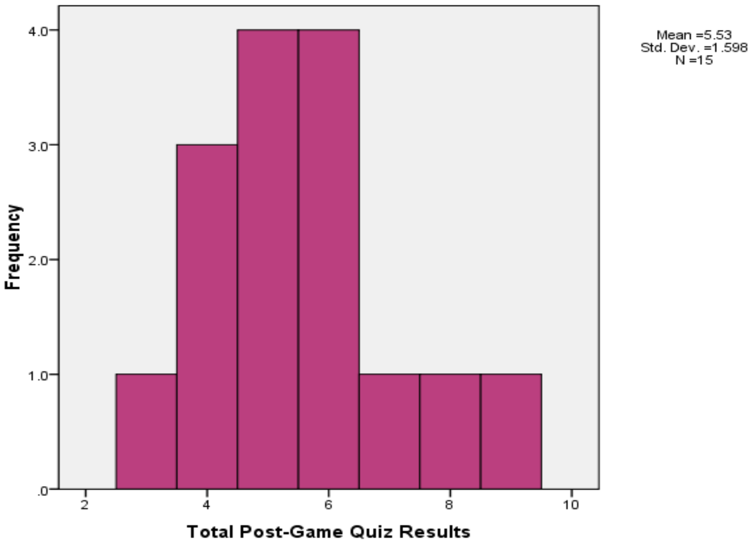


Fig. 9. (*Continued*)

Figure 9 shows the results of the pre-game quiz. The Figure shows that 8 out of 15 participants answered under half of the quiz correct. Figure 9 also shows the results of the quiz after playing the game. Initially, the results display that the highest result was higher than anything in the pre-game quiz. Also, 8 people (over half of the participants) scored between 5 and 6 (out of the total 9 questions) right. The pre-game quiz showed that over half of the participants scored a total under 5/9. Calculating the mean/average, the results show that each participant was roughly scoring 5.53 questions correct after playing the game.

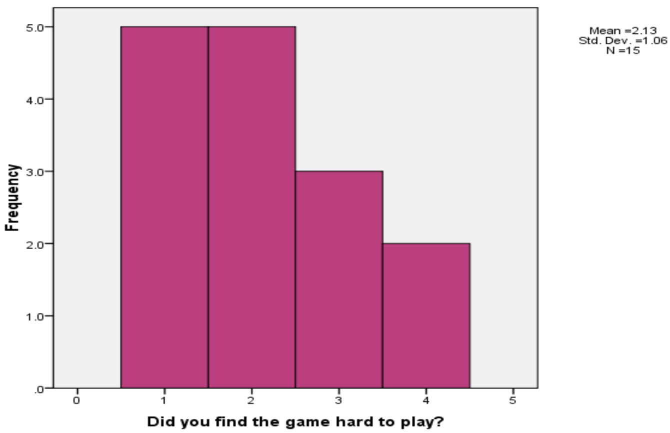


Fig. 10. Perceived difficulty-in-play results

The pre-game quiz had a mean/average result of 4.33 questions per participant. This "jump" in grading represents almost a whole extra question per participant. The results here appear to indicate that, on average at least, the participants did take in some knowledge from the game.

The questions presented after the game were used to show more about the game and how it played. Figure 10 shows the results from how difficult participants found the game. The scale was defined as following; 1 as easy to play and 5 as the hardest. The majority of players found the game to be in the first 2 categories. This means that overall participants felt the game was easy to play. Considering this is a serious game for secondary education, that is very positive, as games with hard to learn controls and mechanisms may cause the players to lose interest.

Similarly, Figure 11 shows that the majority of the participants found the controller easy to use. The controller was part of the game and would have also had a learning curve to those with less experience with it. General observations during the evaluations showed that participants differed greatly in the way they would use the controller. Some would hold it in a way able to access all buttons; others would use it on a hard surface without picking it up.

Figure 12 shows the count of people who found the introduction helpful in understanding how to play the game. It is interesting to see that the introduction was not regarded as understandable by all participants.

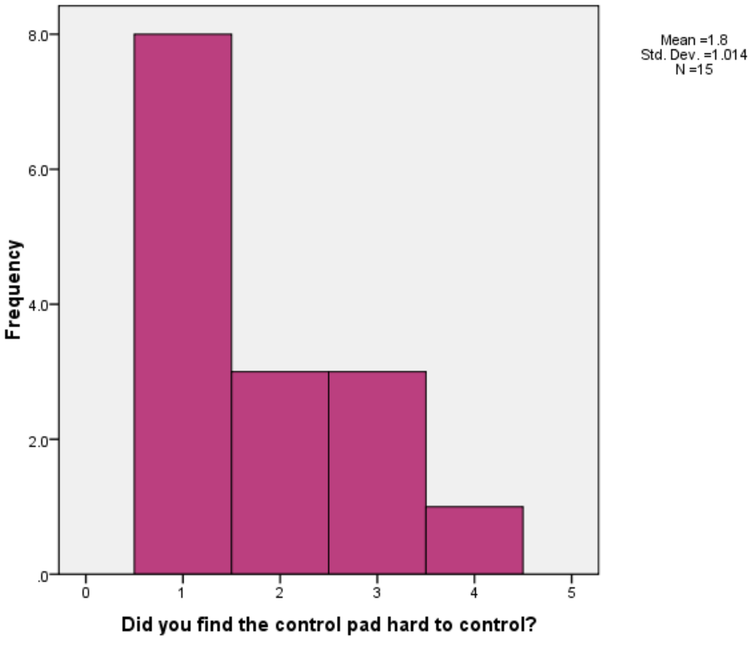


Fig. 11. Perceived difficulty in control

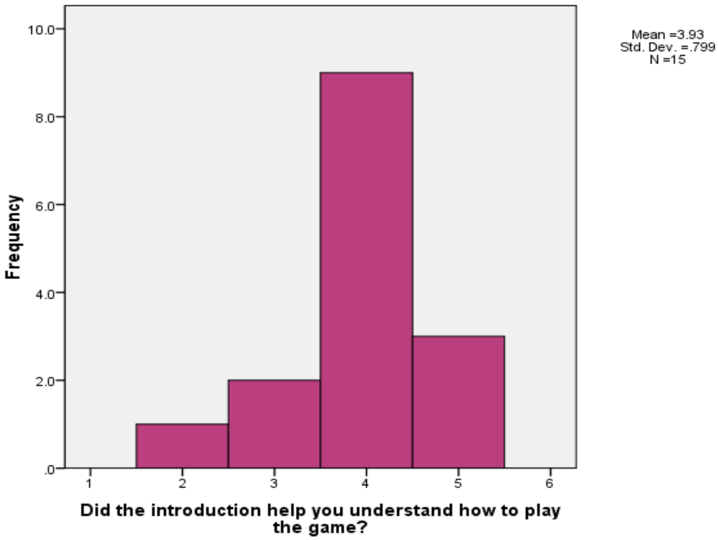


Fig. 12. Perceived difficulty in introduction-helpfulness results

The quizzes were designed using past GCSE papers. The game was also designed on a subject which had specific GCSE questions asked about it each year. The purpose of the quiz was to determine whether the game had any teaching potential. To relate the quizzes to the gameplay, the questions needed to be something the game could provide the answer for. The purpose of this question was to determine whether the participants did think and/or perceive they could get the answer from the game after playing. Figure 13 shows the relation to be perceived as high between all the participants. This means that a high majority of players felt they could answer both quizzes with enough time on the game.

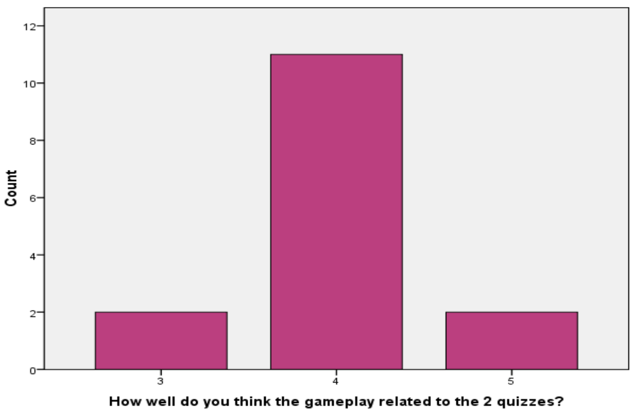


Fig. 13. Perceived game relation to periodic table of elements results

Figure 14 shows the results of whether participants thought they gained additional knowledge of the periodic table of elements. Over half of the participants rated this as 4 and 5. This means that participants generally thought they gained some knowledge about the periodic table of elements.

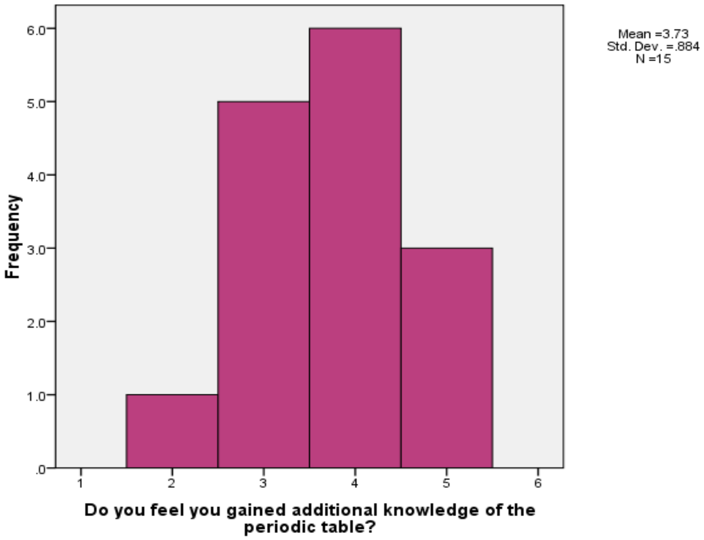


Fig. 14. Perceived knowledge generation results

Another important question asked was whether participants felt engaged while playing the game and found it to be fun. Figure 15 shows there was no rating in the bottom 2 categories showing that participants thought the game was an activity they found fun to play with and kept them engaged while playing. This question was used to see if the serious game fitted to the "video games" objectives.

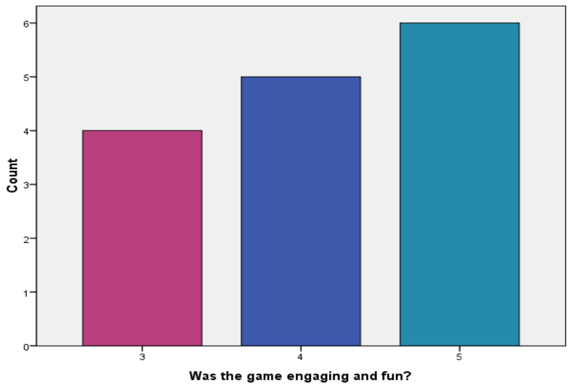


Fig. 15. Perceived fun and engagement results

The final question asked in the evaluation study was directed towards grasping the perception of the participants on serious games. The question specifically asked about how they would perceive a game in helping independent learning. The reason for this was to get a general consensus on how video games are perceived at present. Figure 16 shows the results, which are suggesting that almost every participant felt that games could be very successful in helping someone to independently learn.

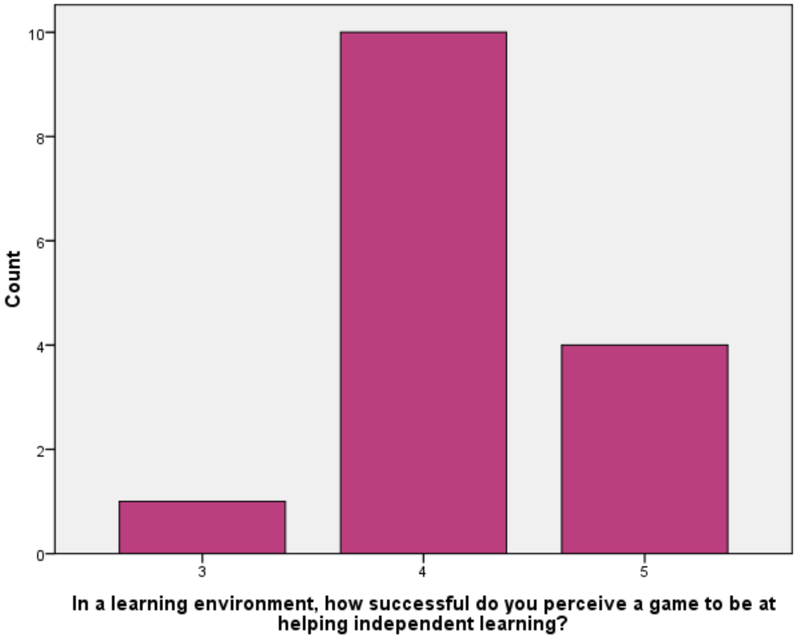


Fig. 16. Perceived learning results

We can also focus on a deeper analysis of the quiz results. This will be achieved by analysing the experience levels against the results in order to understand which ones had the biggest and smallest impacts. Figure 17 shows the pre-game quiz results with video game experience. The Figure shows a varying scale of results, comparing the participant's final result against their gaming experience.

It does demonstrate an indication that some of the less experienced gamers did worse in the tests and some of the more experienced gamers managed to do better in the post-game quiz. This could show that prior experience of gaming has made the game easier for participants, leading them to learn more about chemistry from the Elemental game.

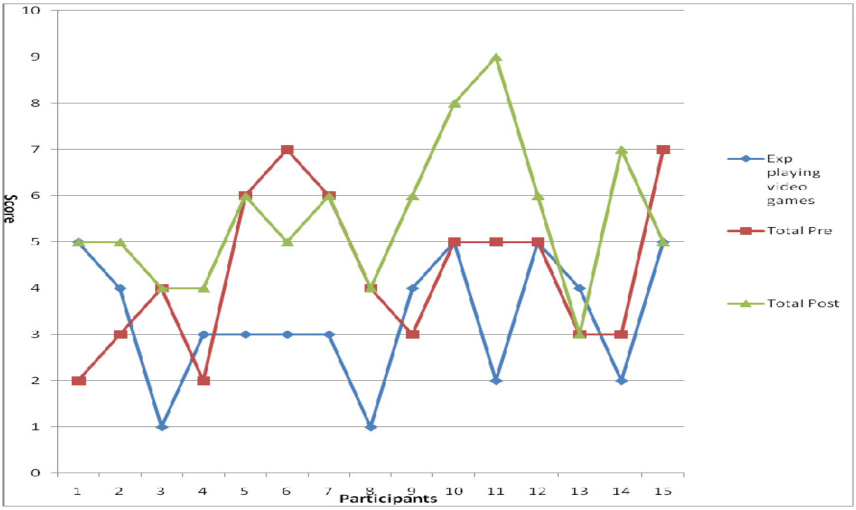


Fig. 17. Correlations between game experiences and pre/post gameplay results in the quiz

The education game experience correlation Figure 18 displays similar results, which suggest that the participants associated experience of educational games as a lesser version of experience with games. Looking at this a little bit more closely, the majority of participants had a difference of -1 to -2 between experience of games and experience of educational games.

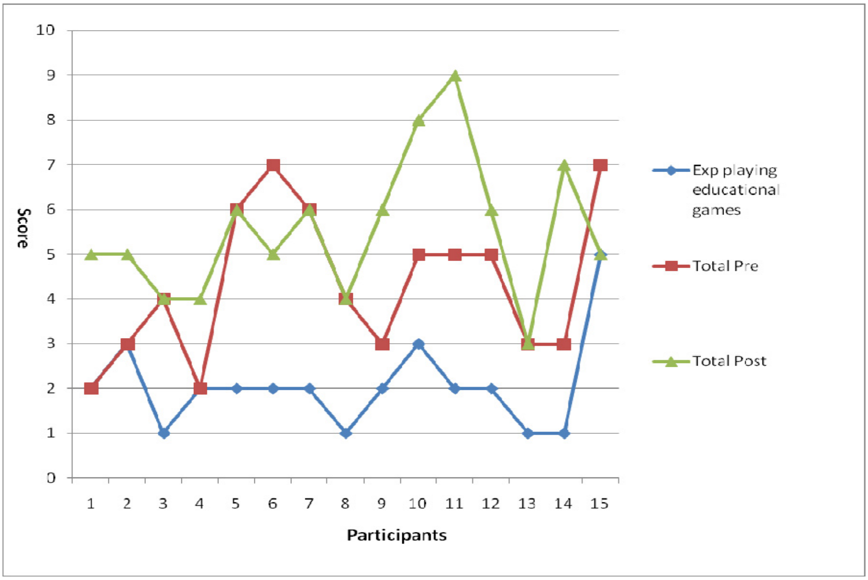


Fig. 18. Correlations between serious games experiences and pre/post gameplay results in the quiz

Figure 19 shows the correlation between experience in chemistry and the results of both quizzes. Although the effect appears marginal, those with more knowledge did generally better in both quizzes. The issue with this question is that chemistry is a broad subject and both the quizzes and the game only focus on a small portion/part of it. So while participants may have a good knowledge of chemistry, they may have had it in a different area of the subject.

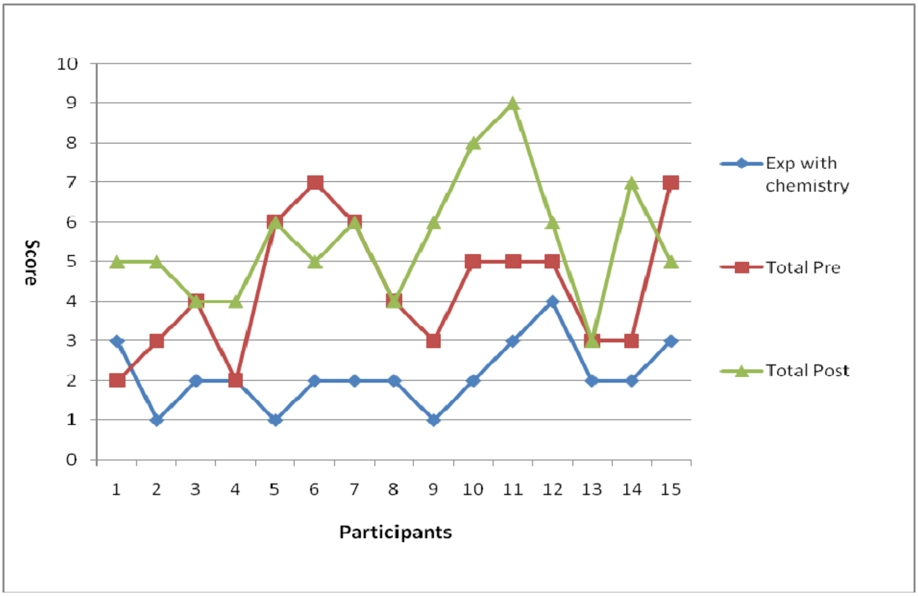


Fig. 19. Correlations between chemistry knowledge and pre/post gameplay results in the quiz

Finally, Figure 20 shows the correlation between results and experience of the periodic table. Looking at the results, it is clear that many of the peaks and troughs were mostly in the same locations. This indicates that those with less experience in the periodic table of elements generally did worse in the quizzes. However, looking at the post-game quiz results alongside it, it does suggest that they learned or stayed the same in the majority of cases. The higher experience participants seemed to do better by a wide margin in the post-game quiz.

While there are indications of the game facilitating learning of the desired content (the periodic table of elements), the population sample in the trial described is far too small to provide data that can be analysed with the appropriate statistical models (the McNemar's test would be more appropriate on this particular occasion, given the multiple-choice format of the periodic table of element questions).

It would be very difficult to gauge statistical significance of pre/post gameplay knowledge for example from this experiment. Correlations with age and gender, which could provide very useful data for further discussion, would also need a larger population sample for appropriate conduction.

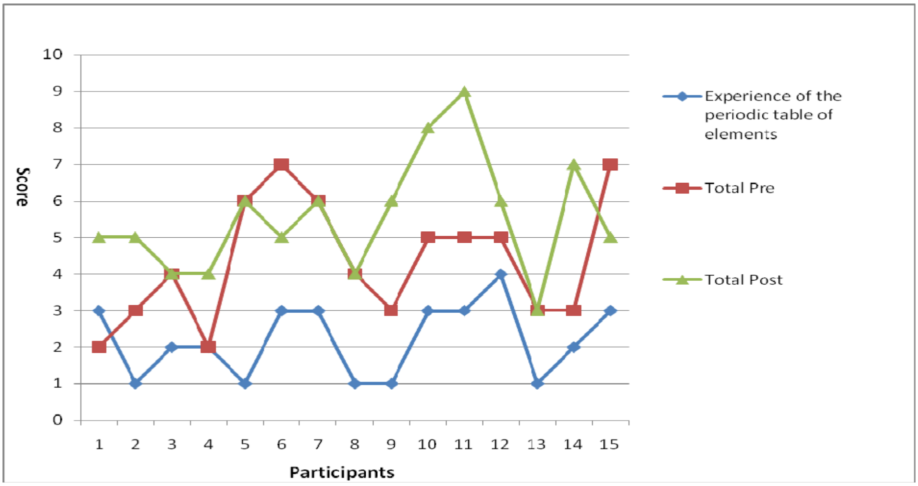


Fig. 20. Correlations between periodic table of elements knowledge and pre/post gameplay results in the quiz

6 Conclusion

This article explores the development and pilot evaluation of a chemistry educational game called Elemental, created using XNA and deployed on the PC/XBox 360 platforms. The results of the evaluation appear positive regarding the value of the game. There were 4 original hypotheses; the first was the game showing indication of teaching the players in the field of chemistry. The post-game quiz managed a better means result that the pre-game quiz. The mean/average of the first being 4.3 and the second being 5.53. On average, each participant scored better by 1 question. The second hypothesis was that the game challenged players in a coherent style. All participants addressed this with a value of 3 or more in the Likert scale used. This shows that the game had a common style, which allowed players to engage well with the game. They also had fun playing and generally this is not done with content which is too easy or too hard, therefore this hypothesis is true. The last two hypotheses were also both true with all levels of video game and chemistry experience participants being able to play the game. Some of the participants with a low experience in chemistry received a better mark in the post-game quiz, as did some of the participants with low experience in video games.

These hypotheses were all shown to be true through a pilot evaluation of the game with 15 participants. The game developed had the characteristics of other video games plus also managed to show educational usefulness, despite being developed using fairly modest techniques and inexpensive tools. The next step would be to explore the Elemental's game use with its intended audience; secondary school students.

Finally, overall recommendations from the experience of creating and also evaluating the Elemental game include the following;

- background research reveals that it is not uncommon to target, with a variety of gaming development platforms and, more importantly, gaming genres traditional areas of classroom curriculum, as many successful approaches to this are already in existence

- XNA can be an ideal tool for developing small scale game-based learning projects intended for classroom use for a variety of reasons; it is a free download, supports a number of Microsoft platforms, is supported itself by a large development community, is simpler to use than other game development programming APIs and can offer great control to a novice programmer, while at the same time providing many pre-made game development classes for use

- the requirements set out for a game created for classroom use need to "draw" as much as possible from prior experience the pupils inevitably will have of computer / video games. Building in game mechanics such as recognisable controls, an informative introduction and help system, as glossy as possible visuals, a timer feature, a high score feature etc. could potentially improve the implementation's desired learning impact

- iterative creation (there were over ten distinctly different versions of the game developed before it was exposed to a classroom) can allow for progressive improvement and feature additions to the game and also for a gradual development of the programming skills of the educator himself/herself whilst creating the implementation

- based on the empirical results, which suggest some pre/post test questionnaire differentiation (with post-test yielding a more informed mean to GSCE-level periodic table of elements questions), the approach of using a game similar to Elemental, developed using the methodology described in this paper, can be of benefit in classroom use to enhance and further support learning of traditional curriculum subjects such as chemistry.

Acknowledgements. Bournemouth University colleagues are thanked for their support, guidance and help throughout carrying out the research work which led to the Elemental game and its eventual trial evaluation.

Additionally, all the participants of the experiment described in the publication are also thanked for investing their time towards the completion of this work.

References

1. Birchall, J., Gatzidis, C.: Elemental: An insight into the development and evaluation of a secondary education chemistry game using XNA. In: Proceedings of the 3rd IEEE International Conference in Games and Virtual Worlds for Serious Applications 2011, pp. 32–39 (2011)
2. Vogel, J.J., Vogel, D.S., Cannon-Bowers, J., Bowers, C.A., Muse, K., Wright, M.: Computer gaming and interactive simulations for learning: A meta-analysis. Journal of Educational Computing Research 34(3), 229–243 (2006)

3. Dempsey, J.V., Rasmussen, K., Lucassen, B.: Instructional gaming: implications for instructional technology. In: The Annual Meeting of the Association for Educational Communications and Technology (1994)
4. Lopez-Morteo, G., Lopez, G.: Computer support for learning mathematics: A learning environment based on recreational learning objects. Computers & Education 48(4), 618–641 (2007)
5. Ke, F., Grabowski, B.: Game playing for math learning: cooperative or not? British Journal of Educational Technology 38(2), 249–259 (2007)
6. Baker, R.S.J.d., Habgood, M.P.J., Ainsworth, S.E., Corbett, A.T.: Modeling the Acquisition of Fluent Skill in Educational Action Games. In: Conati, C., McCoy, K., Paliouras, G. (eds.) UM 2007. LNCS (LNAI), vol. 4511, pp. 17–26. Springer, Heidelberg (2007)
7. Habgood, M.P.J.: The Effective Integration of Digital Games and Learning Content. PhD Thesis at the University of Nottingham (2007)
8. Sedighian, K., Sedighian, A.S.: Can Educational Computer Games Help Educators Learn About the Psychology of Learning Mathematics in Children? In: The 18th Annual Meeting of the International Group for the Psychology of Mathematics Education, Florida (1996)
9. Rosas, R., Nussbaum, M., Cumsille, P., Marianov, V., Correa, M., Flores, P., Graua, V., Lagosa, F., Lopez, X., Lopez, V., Rodriguez, P., Salinasa, M.: Beyond Nintendo: design and assessment of educational video games for first and second grade students. Computers and Education 40(1), 71–94 (2003)
10. Kebritchi, M., Hirumi, A., Bai, H.: The Effects of Modern Math Computer Games on Learners' Math Achievement and Math Course Motivation in a Public High School Setting. Research Brief. University of Central Florida (2008)
11. Egenfeldt-Nielsen, S.: Beyond Edutainment: Exploring the educational potential of computer games. PhD thesis at the IT-University of Copenhagen (2005)
12. Squire, K.: Replaying History: Learning World History through Playing Civilization III. PhD thesis at Indiana University (2004)
13. Adams, P.C.: Teaching and Learning with SimCity 2000. Journal of Geography 97(2), 47–55 (1998)
14. Tüzün, H., Yilmaz-Soylu, M., Karakus, T., Inal, Y., Kizilkaya, G.: The effects of computer games on primary school students' achievement and motivation in geography learning. Computers and Education 52(1), 68–77 (2009)
15. Quest Atlantis Project, `http://atlantis.crlt.indiana.edu/`
16. Barab, S., Warren, S., Ingram-Goble, A.: Conceptual play spaces. In: Ferdig, R. (ed.) Handbook of Research on Effective Electronic Gaming in Education, pp. 1–20. IGI Global Publications, Hershey (2008)
17. Warren, S., Barab, S.A., Dondlinger, M.J.: A MUVE Towards PBL Writing: Effects of a digital learning environment designed to improve elementary student writing. Journal of Research on Technology in Education 41(1), 121–147 (2008)
18. Lim, C., Nonis, D., Hedberg, J.: Gaming in a 3D multiuser virtual environment: Engaging students in science lessons. British Journal of Educational Technology 37(2), 211–231 (2006)
19. The River City Project, `http://muve.gse.harvard.edu/rivercityproject/index.html`
20. Dede, C., Clarke, J., Ketelhut, D., Nelson, B., Bowman, C.: Fostering motivation, learning, and transfer in multi-user virtual environments. In: Proceedings of the Annual Conference of American Educational Research Association (AERA) (2005)

21. Nelson, B.: Exploring the use of individualized, reflective guidance in an educational multi-user virtual environment. The Journal of Science Education and Technology 16(1), 83–97 (2007)
22. Squire, K., Barnett, M., Grant, J.M., Higginbotham, T.: Electromagnetism Supercharged! Learning physics with digital simulation games. In: Proceedings of the Sixth International Conference of the Learning Sciences, pp. 513–520 (2004)
23. Randel, J.M., Morris, B.A., Wetzel, C.D., Whitehill, B.V.: The effectiveness of games for educational purposes: a review of recent research. Simulation and Gaming 23(3), 261–276 (1992)
24. Stoll, C.: High tech heretic - reflections of a computer contrarian. Anchor, New York (1999)
25. Clark, D.: Computer games in education and training. In: LSDA seminar Learning by playing: can computer games and simulations support teaching and learning for post-16 learners in formal, workplace and informal learning contexts? (2003)
26. Royce, W.: Managing the development of large software systems. In: Proceedings of IEEE WESCON, pp. 328–338 (1970)
27. Wilson, C.: Designing useful and usable questionnaires. Interactions 14(3) (2007)

Digital Storytelling within Virtual Environments: "The Battle of Thermopylae"

Dimitrios Christopoulos, Pavlos Mavridis, Anthousis Andreadis, and John N. Karigiannis

Department of 3D Graphics and Virtual Reality
Foundation of the Hellenic World
Athens, Greece
christop@fhw.gr

Abstract. Until recently virtual environments and videogame applications rarely incorporated any deep cultural or educational principles. Current research in the area of virtual reality applications, clearly indentify that they are extremely motivating for learners and therefore can be employed as an innovative, more accessible framework to deliver education while at the same time entertain the public. On the other hand, recent advances in videogame applications and human computer interfaces, demonstrate that a combination of game applications with effective learning principles and intuitive human computer interfaces could potentially transform virtual environments to a significant educational tool that could significantly facilitate the learning process. This paper describes an interactive virtual reality application that we developed for the museum of Thermopylae located at the site of the original battle, near the city of Lamia in Greece. We utilized storytelling techniques and principles of modern videogames to disseminate historical knowledge about the battle and the associated legends. We present the hardware and software components comprising the proposed installation, while we elaborate over the educational techniques designed to reinforce the strength of virtual reality technology as a mean of designing educational experiences in the context of cultural heritage related information.

Keywords: virtual reality, interactive storytelling, educational virtual environments, video games, thermopylae battle simulation.

1 Introduction

Enabled by the advent of interactive digital media, interactive digital storytelling redefines the experience of narrative by allowing its audience to actively participate in the story. Moreover, the design of interactive storytelling scenarios in a virtual reality environment, exhibits social cognitive and technical challenges, that need to be addressed in order to educate and entertain the public. Creating VR environments for storytelling requires the development of complex software, mixing multiple disciplines: computer games, graphics and engineering, physics simulation, pedagogical approaches, and significant experience in the area of the education. The

Z. Pan et al. (Eds.): Transactions on Edutainment IX, LNCS 7544, pp. 29–48, 2013.
© Springer-Verlag Berlin Heidelberg 2013

recent success of computer games fueled by current games technology forms new types of applications such as serious games, influencing also related application areas such as Virtual Environments (VE). VEs have been valued as being extremely motivating for learners [1] and therefore attract significant attention from institutions of informal education aiming to deliver not only knowledge in a more accessible way to the public but also entertainment [34]. Common type of information disseminated by these institutions is historical descriptions of significant battles. In most cases only textual descriptions exist for these events without any archeological findings of relics or monuments that actually could reveal what happened. In most cases, to facilitate the presentation of this information, dioramas, maps, 2D animations or movie archives are employed. Although informative, these techniques often fail to capture the interest of the visitors and often require prior knowledge.

Thermopylae unarguably one of the most famous battles in European ancient history, repeatedly referenced in ancient, recent and contemporary culture, has become a symbol of courage against overwhelming odds. The Battle of Thermopylae was fought over the course of 3 days between an alliance of Greek city-states 7000 men strong, led by king Leonidas and his 300 Spartans, and the 250.000 strong force of the Persian Empire of Xerxes I in 480 B.C. [2] [3]. Both ancient and modern writers have used the Battle of Thermopylae as an example to demonstrate exceptional military strategy, determination and courage.

The challenge of creating an educational Virtual Reality attraction for the newly created “Innovative Center of Historical Awareness of Thermopylae” consisted of finding new ways to combine the techniques used in videogames, virtual reality and the entertainment industry for telling the story of a historic event in an educational and attractive way. A holistic vision of information was needed to avoid the pitfall focusing on only one aspect such as graphics, hardware or interface. Factors such as the design of the hardware in order to support immersion, the interaction and communication with the visitors to support acclimation, the definition of clear educational goals to create an engaging storyline are all factors which contribute to the success of an educational exhibit. Taking into account the gaming culture that has been established even for older age groups through the use of portable devices like smartphones and tablets, new opportunities arise for institutions to reach users in new ways by applying elements of play to newly created educational VR attractions.

The remainder of the paper is organized as follows: Section II provides the motivation behind our work; next Section III analyzes the importance of the hardware setup utilized for creating an engaging onsite installation, Section IV analyses the software architecture of our rendering engine and all the major software components that comprise it, Section V gives the design methodologies used and describes the application. Next, Section VI highlights the importance of our decision to employ a human guide for the educational VE while the subsequent Section VII describes the engagement model that was adopted. The paper concludes by providing a visitors survey and with Section IX that summarizes the key points of our work.

2 Motivation

Those who believe in using games for education usually start from the observation that game players regularly exhibit persistence, risk taking, attention to detail and problem solving skills, all behavior that ideally would be regularly demonstrated in school. Interactive games are the medium of attention for youth who spend on average 50 minutes playing them each day [4]. Using 3D simulation technology for educational purposes is not a new concept; in the past two decades immersive VR has attracted the attention of many researchers and educators [5] [6] [7]. However recent advances in creating videogame applications imbedded with effective learning principles [8] and the study of human computer interfaces for the entertainment industry [9] [10], suggest that a combination of successful practices in these areas could provide motivating experiences and strong tools for learning in educational VE.

Recent evolution of mobile computing, connected devices like smartphones and tablets, made gaming available anywhere and at anytime, to a wider range of age groups. Moreover people claim that gaming becomes second nature to a continuously increasing percentage of the population. As a result, people expect that games will continue to move out of the traditional entertainment environment, blending seamlessly with the real world in areas such as learning and education. The stereotype of the reclusive gamer is outdated. These days, games are everywhere, and gamers are social, tech-savvy, goal-oriented people with a real drive to improve themselves and the world around them. Recent studies [35] on age groups 15-54 confirm that users view games also as a medium for self-improvement and expressed the desire to fit them into daily life activities, gamify boring and repetitive tasks and see them applied foremost for education and learning.

Computer generated interactive experiences allow visitors to travel through space and time without stepping out of the museum building [11] and can therefore be used as a new way of communicating information to the public, making abstract and intangible knowledge to become concrete and manipulable, allowing visitors to immerse in historical events and explore them first hand. The majority of Virtual Heritage (VH) projects deal with reconstruction of architecture providing the experience of guided or first person world exploration in which visualization dominates [12]. Creating computer generated representation of towns and scenery is one thing, but what about less concrete things like processes, feelings or atmosphere. This is especially the case when recreating historical events like battles where the need to de-emphasize reconstruction is evident in order to place them in their historical context and make visitors understand larger connections.

Mixing 3D real-time games with documentary information has been attempted before using commercial games embedded with a cultural heritage theme that depict real historical events which the player can partake in [8] [13]. Using VR for representation of historical events was also attempted in [14] representing the siege of Syracuse by the Romans in 212 B.C. Our aim on that project was to cross-fertilize edutainment VR, core game technologies and novel techniques derived from the entertainment industry providing an educational experience that should put visitors in touch with what is fundamentally engaging about the historical battle of Thermopylae, help them build a scaffolding of the historical core concepts and motivate them to go deeper into the subject themselves. The documentation of the theories that guided our development and their implementation using techniques from a wide selection of fields should contribute to the challenging and outstanding task to study the right and applicable use of VR in education [15].

3 The Importance of Hardware

Better education media to assist teaching has constantly been sought by researchers, highlighting the importance of the technology used for visualization. This project was specifically designed with projection based VR systems in mind.

Although it can be exhibited on all displays, even monitors, we carefully designed the visitor experience for this educational, historical VR attraction on the assumption that the minimum system requirements should be a powerwall type VR system. This means high resolution images projected in stereo 3D by a powerful graphics PC on a 3D stereo capable screen and spatial surround sound. Of course systems with more degrees of immersion (more screens, tracking) can be used as well, but usually are out of reach for educational institutions for financial reasons. The main reason for usage of specialized VR equipment for cultural applications is to provide a technologically impressive setting which attract visitors, enable them to immerse easily into the attraction and also take into account the social aspect of a museum visit.

Providing a better learning experience means from a technological stand point, providing a setup that helps in creating motivation and presence. Technology by its self is a common motivator but also creates certain expectations leading to a conflict of expectation for visitors of historical VR worlds. On one hand visitors want their experience to be realistic and learn something of historical and scientific value, but at the same time they expect to have an exciting experience. With the continuous advances of game consoles, VR systems simply cannot keep up to remain fresh from a visual only standpoint. Therefore they must address more senses and provide an impressive setting using large screens, surround sound and 3D stereo that cannot be experienced easily at home.

3D glasses not only put the action right in front of the user but also make the projector screen disappear creating a convincing world. One of the main factors contributing to immersion is the Place Illusion [16] which is the illusion of the sense of “being there” in spite of the knowledge that you are not there. In the case of simple setups using desktop displays (monitors, simple projectors) this feeling is reported to come, if it comes at all, after much greater exposure to the medium and requires deliberate attention. The VR experiences that are provided in informal educational environments are relatively short in time and often come with waiting queues for visitors, thus it is essential to use the appropriate technology that helps the immersion and suspension of disbelieve.

Besides the learning experience, the social aspect of a museum visit is often underestimated. People tend to visit in groups to have shared experiences. Projection based VR can provide shared visual and audio experiences in which all hear the same and see form the same point of view facilitating social interaction and communication.

4 Software Design

Hardware alone cannot ensure the immersion of the audience. The Software Design plays an equally distinct and important role in the process. The in-house developed Enhanced Visualization System (EVS) was used to create the high-resolution stereoscopic output [14] [17] [18]. EVS was designed specifically for creating and running immersive Virtual Reality applications in a variety of hardware configurations.

EVS supports multi channel setups, allowing monoscopic, passive stereo, active stereo or left/right individual eye operation through a custom developed display library [19] that handles arbitrary display surfaces and viewing modes. In that way an application can be configured to run on a variety of Virtual Reality systems such as CAVE, Reality Center, Dome, Powerwall as well as on single screen desktop monitors and HMD. Different setups are easily configured through a simple and effective XML script allowing multiple configurations to be present in a single file and share some common features if necessary.

Handling of the various input devices is performed through a generic open source architecture network interface, the Virtual Reality Peripheral Network (VRPN) [20]. VRPN provides a device-independent and network-transparent interface to virtual reality peripherals. The library consists of a set of classes that enable interfacing between application programs and physical devices (trackers, buttons, dials, force feedback devices etc.) usually employed in virtual reality (VR) systems.

EVS supports cluster rendering through a custom protocol that was developed [19] and designed for the synchronization of multiple cluster units consisting of a central unit (master) and multiple subunits (slaves). The master unit synchronizes the slaves using special synch packets. In this way consistent representation of the virtual world from multiple viewpoints on multiple screens is achieved. The communication between the units is established through LAN network.

Our engine is script based, meaning that ASCII script files describe the virtual world and the user interactions. EVS also allows the programming of dynamic interaction and events utilizing the Lua [21] script language. The actual rendering module was developed on top of OpenSceneGraph [22] which is, an open source 3D graphics framework that offers high performance rendering along with a feature-rich scene graph implementation. It is written in C++ and uses OpenGL for cross-platform 2D and 3D graphics rendering. “Fig. 1” shows the architecture of the whole environment that was developed on a Linux operating system.

The pipeline for the creation of our virtual reality environment and assets was extended and incorporated apart from modelers and texture artists, historians and architects, as the main goal of the project was to attribute with accuracy the historical events that took place in 480 B.C. All the needed assets were initially cross-referenced with historical sources and designed in the correct proportions by architects. The prototypes produced were high-polygon models sculpted to depict accurately all the fine details. Since real-time poses limitations in terms of geometric complexity of the rendered objects, we couldn’t use the actual prototypes for our VR application. From each high polygon model a simplified version was created, baking all the details of the high polygon model in normal maps and applying them to a low polygon representation. Real time material shaders use the baked normal maps to

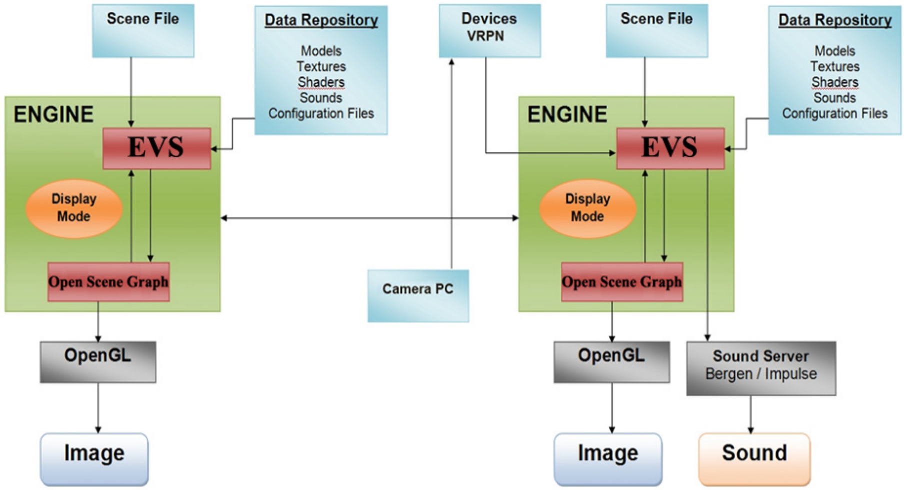


Fig. 1. The Architecture of the EVS Engine

render these low polygon models as detailed 3D surfaces during runtime, greatly enhancing their appearance.

A dedicated VR engine such as EVS gives us several advantages compared to classical Game Engines that by design addresses mainly single screen or desktop systems. The rendering demands for immersive systems that usually need stereo rendering along multiple screens become prohibitive for a single GPU/CPU computer and establishe the cluster rendering feature obligatory, favoring greatly the choice of a dedicated engine. A Cave setup for example requires five distinct projections. While in games the main purpose is to entertain, VR setups have to achieve the immersion utilizing stereo projection and advanced interaction methods using specialized input devices for tracking. On the other hand EVS had to borrow many techniques seen in modern games in terms of graphics, interaction and presentation. Young visitors are already accustomed to impressive 3D environments from playing games at home and seek this type of visual quality sometimes even more that realism.

All the shading and lighting calculations in EVS, are performed through programmable shaders. Shaders are small computer programs, written in a shading language, that are executed on individual elements (vertices or pixels) and affect the properties of this element (pixel color, vertex position). The concept of shaders was introduced with the RenderMan Shading Language (RSL)[36] and since then it is widely used in the creation of special fx for motion pictures, using offline rendering systems. The evolution of the graphics hardware in the last years permitted the introduction of shaders in real-time graphics applications, with the introduction of the OpenGL Shading Language. GLSL is a lower level language than RSL, and lighting calculations require explicit knowledge about the number of lights in the scene and their properties. EVS provides an enhanced version of the OpenGL Shading Language (GLSL) [23], called ESL (Enhanced Shading Language). ESL is a pure superset of the GLSL, meaning that any GLSL shader will work on EVS, but also exposes a set of convenience functions and uniform variables for the computation of lighting, shadowing and vertex skinning. These functions are compiled during the runtime to the equivalent GLSL code, permitting fast and easy development of shaders, and more importantly reusability of the shaders between different scenes and projects.

Several techniques have been implemented in our engine that use real-time shaders, such as Environment Mapping, Per-Pixel Lighting, Bump Mapping, and Parallax Mapping with offset limiting. Shaders are essential in our project for the realistic presentation of cloth, wood and metallic surfaces. The power of shaders was also used in rendering of the human skin. We wanted to achieve a realistic skin effect with subsurface scattering that would run in real-time. Our solution is a combination of the methods described in [37] and in [38] and achieves a realistic skin shading with subsurface reflectance and micro-surface detail.

For the critical part of shadowing in-order to achieve antialiased real-time shadows for our environments we have implemented Parallel Split Shadow Mapping (PSSMs) [39, 40]. In this technique the view frustum is split into depth layers that are parallel to the view plane. For each layer an independent shadow map is rendered. Furthermore to alleviate the shadow boundaries aliasing issues we use Percentage-closer filtering (PCF) that produces smoother results by performing multiple depth tests for each pixel.

Post-process filtering is a major part of most modern visualization engines. Real Time Ambient Occlusion, Color, Correction, Depth of Field, Motion Blur and Volumetric Light Scattering are some of the common techniques are used as post effects to increase the visual appearance of games and make them look more cinematic an immersive. For EVS we developed a parametric solution that enables us to implement all modern special effects as mentioned above. For each scene we choose the appropriate set that will enhance and dignify the feeling we want to achieve.

The gaming industry had also a strong influence in our choice of input device used for interactivity. Since visitors can actually interact with the virtual environment and participate in activities, the interaction had to be close to the one that young audience is used to with modern gaming consoles. For this reason, we used a standard gamepad with 2-analog sticks using the left one for movement and the right one for steering.

For usage in cut scenes, the engine was adapted to support high definition stereoscopic video playback by integrating the open source mplayer [24] technology. Since the target system supported stereoscopy, all produced movies consisted of two streams (one for each eye) which were converted for efficient real time playback to one side by side stereo stream using an offline procedure that utilized the open source mencoder and ImageMagick tools [24].

5 Design Goals and Description of the Project

The Battle of Thermopylae in 480 B.C. was fought between an alliance of Greek city-states led by Sparta, and the Persian Empire of Xerxes I, over the course of three days, during the second Persian invasion of Greece. This project was designed as an educational game which aims to deliver the historical context and importance of the battle, the art of warfare of the two civilizations, their cultural differences and how the Greeks were so successful defending Thermopylae for 3 days before they were betrayed leading to the heroic last stand of the Spartans.

Due to the flexibility they provide, virtual environments can compensate for the shortcomings of traditional instructional strategies, satisfying the needs of different cognitive styles and contribute to the understanding of abstract ideas making them visually concrete [25]. Games promote understanding, motivation, enjoyment and are terrific at immersing players in complex, feedback rich problem spaces. And while they are most often not sufficient in and of themselves for a course study they can help many students advance beyond the memorization of facts and procedures, attainments that are usually lost when classes stop [26]. Thus our goal besides investigating the educational potential was to be accessible, letting the visitor experience the story in an extraordinary new way, from unique points of view and motivate them to go deeper.

The project combines 3D computer generated imagery (CGI) and interactive games by dividing the experience in two parts. The first part consists of a short (10 min.) high resolution, stereoscopic, CGI movie which employs a documentary style approach. It introduces the historical context of the battle, how the Greek and Persian troops assembled, and the route they took and why they met at Thermopylae. Zooming into a map of the ancient world and onto the coast of Euboea we witness the naval battle of Artemisium which took place simultaneously with the famous land battle of Thermopylae and was vital for its success. 270 Greek battleships had to hold off at the straits of Artemisium the entire Persian fleet of 800 ships to avoid the overrunning of the land troops, stressing the importance of the naval fleet for the outcome of the invasion.

The depiction of the land battle was divided into 3 parts, one for each day, showing all the important facts. The picture alternates between action packed close ups and documentary style explanations of the terrain topology, the advances of troops and battle tactics “Fig. 2”. Cinematic bullet time visual effects stop the action to explain topics and points of interest such as the fighting style of the famous Greek hoplite phalanx, the armor of a hoplite, the battle style of the Persians and their weapons. Special attention had to be given to the correct depiction of the terrain using topographical surveys and geographical data and to the representation of weaponry by consulting historians, archeologists and experts in ancient warfare. We refrained from the Hollywood approach of using exaggerations, gore violence and bloody scenes. Although casualties are depicted after the battle there is no scene where someone is actually killed. Instead we decided to use other cinematic techniques like visual effects, fast camera movements, depictions of the battlefield after the end of each day to infer the drama and tension. After the movie the storyline transfers the action to the interactive game. It was of primary importance that the movie and game aesthetics were matched to produce a visual unified experience. Both productions featured the

Fig. 2. Stills form the CGI Movie. Top left, the battle of Artemisium. Top right, the battle of the first day starts. Bottom left, combat inside the phalanx. Bottom right, the battlefield after the first day.

same render quality, the movie was rendered using game like shading and polygon techniques, and the interactive game raised the average graphics quality using expensive high detail material shaders and animations.

The interactive game which follows the movie travels the visitors back in time to the dawn of the first day of the battle, where both the Greek and Persian camps have to be visited in order to help prepare the two main protagonists, the Greek Dienekis and Persian Aribazo for battle. These names were not selected randomly, according to historical sources high ranking officers of those names actually existed and participated at the battle. The visitors are called to traverse the camp, find the corresponding equipment and weapons and apply them one by one on the main characters. During this quest they learn about the culture, battle preparations, tactics and weapons of these ancient warriors. The tasks of finding and applying weapons or equipment are laid out in such a way that essentially all the camp has to be visited in order to complete the final goal. Every time a task is completed the user is rewarded with narration about its usage and other relevant historical information. During the search in the camp a variety of interest spots are met like places for rest, food preparation, conversation, armor repairs, and camp fires which are populated with virtual characters and objects relevant to the battle "Fig. 3". Some of these spots transmit historical information automatically when visited but their primary usage is to provoke the interest of the visitor to inspect them nearer, guiding him essentially through the story and giving the opportunity to a human guide for personalization of the experience according to the target group. We decided to have a basic level of

Fig. 3. Screenshots from the interactive game. Top left, the Spartan camp. Top right, Spartan interactivity hot-spot for the chest armor. Bottom left, Persian interactivity hot-spot for hand arms. Bottom right the Persian camp.

information transmitted automatically by the story in the game and let a museum educator decide whether more information is required. The museum educator is a human guide who operates a game pad controller, controls the pace of the storyline, gives commentary in accordance with the story written, helps during the interactive parts and adapts the experience according to the visitor level. After the completion of all the tasks in both camps, the interactive game zooms out and focuses the camera onto the rendering of an old book with which the battle outcome is narrated. As the narration progresses, the book's pages are browsed showing hand drawn pictures of what is just narrated. This technique refers to the way fairy tales are told and is used to soften the impact of an otherwise brutal ending.

Besides providing entertainment the experience supports several learning methodologies as proposed by [26]. It can be used as a content system to deliver understanding about the subject or content area, as a trigger system to create an experiential context for understanding around this topic, as a point of view system allowing visitors to take on certain identities and associated points of view by visiting both the Greeks and Persian camps comparing differences in strategy, choice and values held by the two main characters of the story and as an assessment system for visitors who have already been taught that subject to asses and consolidate their understanding. Furthermore through the usage of a museum educator which we will cover in the next section, opportunities for reflection and discussion are provided to let the visitor see the experience as part of a larger body of knowledge on that subject.

6 The Human Guide

Despite the fact that the community has used for year's museum educators as mediators for VR tours, we deliberately choose to design the experience with the participation of a museum educator as a prerequisite, since it proves to be very beneficial for the educational context. Recent studies [27] [28] [29] [30] highlight that one of the most crucial factors in terms of educational benefit, entertainment and construction of meaning for visitors remains the human factor of the guide as presenter and mediator. The studies conclude that the majority of visitors prefer to explore the application with a human guide, because it allows direct interpersonal interaction with which the museum educator can immediately adapt to the audience and solve doubts or provide more information. The experience is regarded to be more alive, spontaneous and not standardized comparatively to passive mediums like movies or planetariums.

A museum educator affects the type of experience visitors will have allowing them to focus more easily not only on the technology but also on the content. Even a simple, human guided tour of the VR world creates a space of kinesthetic illusion which is both visual spatial (3D reconstruction) and textual spatial (guided tour) [27]. Nevertheless the traditional use of a human guide seemed unidirectional, letting information and action flow only one way, from the guide to the visitor. This is not the way games work, which have to provide feedback and information to the user, inducing him to act. Slater [16] argues that a major component for the suspension of disbelieve is the Plausability illusion which is the illusion that what is apparently happening in the VE is really happening and defines that the key component for its realization is the existence of events in the virtual world over which the user has no direct control and which refer directly to her. Based on this we introduced a bidirectional flow of information and action, by handing at certain moments control over to the application, so that not all information and actions would come from the museum educator. Through the usage of in-game characters at the various interest points in the VE and mainly during the assignment or completion of a task, historical information is provided and events initiated which advance the storyline. After the playback of the initial movie both the museum educator and visitor are forced to cooperate with the interactive application in order to advance the story, providing a game like experience. The reception of information and tasks both from the museum educator and the application allows not only the direct participation in an interactive experience but also the passive observation placing the visitor in the middle of the action.

Using this combination of constructivist and guided exploration techniques as a recipe for an effective virtual experience we aimed to trigger the Kolbs experiential learning cycle which in effect is triggered when learners have also to observe rather than only directly participate in an interactive experience [7].

7 Creating Engagement

Research suggests that most VH projects should strive to have a successful blend of three components. Representation to accurately visualize or reconstruct data, experience to present the virtual environment with motivating elements that incorporate knowledge and interaction to provide the ability to gain insight by actively engaging and even modify the experience [31]. Based on these 3 components we tried to incorporate the principles of current VR, entertainment and game industry practices to support the educational and motivational aspect of the project.

7.1 Representation

The first priority of any educational institution is to maintain its scientific and educational character. Not only has the rendered environment, the characters and the weapons of the Thermopylae project to be appealing as in modern games but also historically correct. A variety of disciplines ranging from designers, artists, architects, archeologists and historians worked together to produce all our 3D assets, historical information and storyline. As described in the software section we made elaborate use of game technology and next generation graphics for rendering, trying to keep a unified appearance across the movie and interactive game. Younger visitors and children are already accustomed to the look and feel of these high detail 3D representations through the usage of computer games minimizing the time of acclimation.

Believability is not necessarily linked with perfect realism; on the contrary maximal realism can be counterproductive for students with little knowledge. Despite the use of high profile rendering, we don't pursue perfect realism but preferred to create emotional response providing personalization in character, environment, animation and sound design. Using personalization to provide a specific mood and aesthetic to the application and create something memorable that stands out, is a common practice in games design [32]. Especially during the interactive game the sound, the color ramp, shading and lighting was chosen to transfer the tension of a camp before the battle. The environment needs to be impressive but also has to make the unimportant details natural so that the visitor focuses on the important parts [33]. In general the quality of the representation influences the overall experience since as we will explain later on, it is one of the three major factors that control interest. The more beautiful the art, the more interesting and compelling the visitors will find the experience.

7.2 Experience

A successful entertainment experience should use the right combination of interest to focus attention, empathy to make us feel we are part of the story world and imagination to let the visitor fantasize alternative realities [10] [9]. Storytelling proves to be the best vehicle for that. The more compelling the story the more focused the listener will be and the more likely that true understanding will take place.

Fig. 4. A graph showing the designed visitor interest curve during the experience, as the storyline progresses. The initial spikes both in the CGI Movie and Interactive game are the "hooks" which grab the attention of the visitor, followed by smaller interest spikes used to unravel the story and keep the interest.

The story for this project was designed in such a way as to provide the desired amount of historical information and foremost maintain a good interest curve "Fig. 4". Usually the visitor enters the experience with high interest and expectation fueled by the hardware 3D capabilities of the system and the shared group experience which is about to start. The high action stereoscopic CGI movie shown at the start acts as a "hook" that grabs the attention and provides an interest spike that helps stay focused and show what is to come. Although the 10 minute duration of this movie might be considered too long it was necessary in order to be able to narrate the whole story. Through a rich arsenal of cinematic and stereoscopy techniques the tension and interest is attenuated throughout the movie playback so as to keep the interest. The switch to the interactive game fuels again the interest curve motivating the visitors to pay attention to the historical facts narrated and complete the tasks assigned to unravel the story leading to the climax and final resolution. Although the definition of interest during an experience might seem at first a little abstract, research [10] has determined that it is composed of inherent interest, representation and psychological proximity.

Inherent interest defines the motivation and interest in the actual historical event. Since the project is about one of the most famous battle in history and presented in high tech 3D it provides an incredibly rich back story that everyone can relate to. The attraction and place immediately gets everyone in the correct mind set.

Representation defines the aesthetics of the entertainment experience which, as explained earlier, were carefully designed using next generation graphics with strong personalization of the environment to serve the purpose of the story line.

Psychological proximity is what compels visitors to use the powers of empathy and imagination to put themselves into the experience. Although traditional passive media like the introductory movie shown lag in this respect, the following interactive game was planned to create an environment with events that visitors can emphasize with.

Fig. 5. The two main protagonists who during the story seek assistance from the visitors. On the left the Spartan Dienekis and on the right the Persian Aribazo. According to historical sources high ranking officers of those names actually existed and participated at the battle.

The game assigns a specific role to the visitors which are asked by the main virtual character in the story to complete specific tasks in order to help him "Fig. 5". The storyline and the additional explanations of the museum educator help bring the places and characters visited in the camp into a human and cultural context. Initially visitors are strangers to the environment and its characters, but as the events progress through the interactive tasks completed by the visitors, the interest and understanding about the culture, customs and way of living of these warriors' increases, allowing them to mentally enter the world and place themselves into the experience. To support these goal virtual characters were placed into the environment for interaction, performing actions of social and cultural importance. As an example for the Spartan camp, the warrior helmet that has to be found in one of the tasks is given by a Spartan who prepares himself for battle combing his hair and anointing his body with olive oil. This action refers to the physical and mental preparations of the ancient Spartans before battle. Besides the automatic narration explaining this, the museum educator can then enrich this information with relevant myths and stories.

7.3 Interactivity

Interactivity is considered a cornerstone in creating engagement building an overlap between perception and imagination allowing the directly manipulation of the story world. This is why games can take events of low inherent interest or representation and still be compelling, because they make up for in psychological proximity. Even though the interactive world gives the freedom to go anywhere, a well designed environment will lead the visitor through the highlights letting her think she found them herself. By allowing interactivity the creator loses control over the sequence of events and exerts only indirect control. Perception studies [33] show that we usually turn towards movement and strong colors or other markers. The camp environment of the interactive game was designed to draw the attention of the visitor guiding him to interesting places. Animated virtual characters, objects with shiny colors, camp fires, trees, smoke and rocks were strategically placed in the environment to define spots for exploration. The use of in-game characters asking for help and the use of storytelling provided clear goals and reason for participation. Being focused on a particular task not only helped to achieve high motivation but was also the primary method for influencing the visitor who by completing a series of goals essentially creates a linear storyline.

8 Visitor Survey

In order to evaluate the effectiveness of our story telling techniques at disseminating historical knowledge, in the context of our virtual reality exhibit, we gathered data through in situ observations and face-to-face questionnaires with visitors. Since the experience of the visitors contains a lot of interaction with the museum educators, especially in the interactive part, we believe that the observation of the visitor reactions during the showcase, through the questions they pose to the human guides, is essential to have a complete picture. Along with the results of the user survey, we present a qualitative analysis of the gathered data.

The survey was conducted to a group of twelve visitors, of age's nine to twelve. Data where gathered by observations during their visit and with semi-guided interviews, taken when the visit was over. We concentrated on children of that age group because the historical information presented in our work is also covered as core material in their educational school cycle. During the interviews, we verified that the subject of our exhibition was interesting to this age group. As we said earlier, the main goal of our survey was to investigate the value of the exhibit as a means to pass historical information, but also questions were posed to evaluate the amount of immersion achieved and the usability of the exhibit. Finally the visitors were asked to state if they would like the traditional means of education in schools to be augmented with the story telling methods presented in this paper. Fig. 6 presents an overview of the survey results. In the following paragraphs we will further analyze the answers to the survey questions, in order to better evaluate the results.

The first series of questions were related to the historical details presented during the CGI film and the interactive game, such as the place and the exact date of the historical event, the outcome of the battle and some questions about the combat outfit of the Greek and the Persian army. These questions were used to evaluate the amount of historical knowledge acquired by the visitors. All the children answered correctly in these series of questions, and with the exception of one child, they even remembered the exact date when the battle took place. It is noteworthy that all the children could correctly name at least three items from the combat outfit of both the

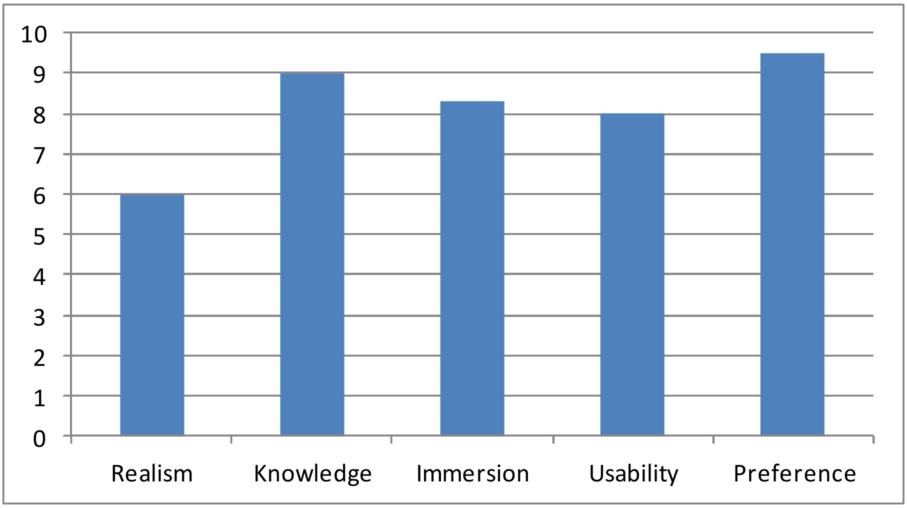


Fig. 6. An overview of the visitor survey results. The rates were deduced from the visitor answers to questions related to the realism of the graphical representation, the historical knowledge they acquired from the production, the amount of immersion, the usability of the real-time game, and finally if they prefer the story telling methods presented in this paper to be used in traditional education. We observe that while the graphical presentation was not rated as photorealistic, the amount of historical knowledge acquired by the production is impressive.

Spartan and the Persian army. Those results give us confidence that this exhibition kept the attention of the children focused during the whole period of their visit, something that we believe an educator with traditional teaching practices will have difficulty to achieve, for children of such ages. Recent surveys [29] [30] compared VR exhibits and traditional museum exhibitions dealing with the same historical topic and reported that the visitors from the guided VR tours (probably consumed by the audio and video) could not remember a single concrete historical fact from the show. It therefore seems that our approach to include linked elements of text, interactivity, 3D graphics and verbal explanations in an engaging storytelling experience overcomes this problem. Surprisingly in many cases even detailed historical facts which were explained by the museum educator were remembered. Finally, when asked, all the children stated that they would prefer some history classes in the school to be taught this way, further reinforcing the educational value of storytelling techniques in the context of education.

With a second series of questions we tried to evaluate the amount of immersion achieved during the film and the interactive game. When asked, the children stated that the virtual world was not photo-realistic, but the majority of them liked the aesthetic result as it reminded them of the video games they play at home. The most frequent answer, when asked about the role they believe they had during the game, was that of a Helot (servant), who has to help the soldiers prepare for battle. The second most frequent answer was that they were spectators. All of them stated that they wanted to help the heroes when they asked for help and although the story urges you the help both the Greek and Persian hero, some stated that as being Greeks they preferred to help the Greek hero, confirming the creation of empathy and psychological proximity during the story. Most children stated that their favorite moment was the battle as depicted in the film, and the second most frequent answer was the search for the weapons in the interactive game. Based on these answers we believe that immersion and engagement was achieved in a great extend in this production, initially by using the CGI movie as a “hook” to build interest and then keeping it by introducing interactive story driven events.

A third series of questions tried to evaluate some usability aspects of our interactive production. The children properly identified that the goal of the simulation was to find the military equipment of the two soldiers and return it to them. When asked about how easy was to identify the missing objects through the environment, they stated that although they had some difficulty to identify them, due to the low luminance of the projection screen, the fact that they were shining helped them a lot. Many remembered with enthusiasm a piece or problem that they found or solved first, stating that it was much better that they had to actually do something themselves instead of just watching a guide to do it for them. As detailed in a previous section, we designed the important objects to stand out of the environment, to draw the attention of the user and to circumvent the low contrast and low brightness of some VR projections environments and the results of this survey support our design decision. It is important to take into consideration the limitations of the underlying hardware when designing content for virtual reality exhibitions. Some also requested more interaction and be able to fight in the battle themselves, indicating that interactivity is something visitors seek in such experiences.

During the interactive game, the importance of the presence of the museum educator was apparent. The main problem in many interactive virtual reality exhibitions is that the visitor can be confused about the rules and the purpose of the simulation or get too fascinated by the experience focusing more on the technology then the content. The presence of the human guide prevented this from happening, by explaining the purpose of this exhibit, answering any questions of the visitors, and by guiding them through the virtual environment. After our observations, we believe that the presence of a museum educator is essential for the success of such exhibits.

The questions posed to the museum educator by the visitors also revealed some aspects of the story were not clearly communicated by the story, and needed further clarification. One such case, asked by the children, was about the chronological order of the events in the film and the interactive game. It was not apparent to all the visitors that we were witnessing the preparation of the two armies for the battle we saw at the film. The social interaction between the peer group and its guide also helped to satisfy the curiosity characterizing those ages by answering simple questions regarding the content, like why Spartans had red coats or why all Persians had mustaches. Those types of questions, which arise perhaps from some omissions of the digital content, are easy to answer with the presence of the museum educator.

All the information and the user feedback we gathered proved valuable in order to improve both the educational and the entertainment value our future productions.

9 Conclusion

Virtual heritage applications should start to have a holistic vision of information and not focus on only one aspect such as graphics, hardware or interface. The design of an educational experience must take into account factors such as the design of the hardware in order to support the immersion, the interaction and communication with the visitors using a museum educator or any other method to help visitor acclimation, define clear educational goals and create an engaging storyline which fosters the learning of these.

The Thermopylae project investigated on the fusion of techniques used in videogames, virtual reality and the entertainment industry for telling the story of a historic event in an educational and attractive way. Current research in the field of immersive VR, videogames, human computer interaction and educational studies were used to drive the design of the whole experience. The evaluation results appear to be quite promising, indicating that new practices in the area of games could significantly facilitate learning and overall produce more effective learning environments.

Tomorrow's gamers are interested in applying elements of play to the world around them. This creates opportunities for educational institutes to reach users in new ways. It is at the intersection of videogames and entertainment industry where virtual environments should find new ideas and opportunities for creating innovative educational experiences.

The presented work has been developed within the project "Innovative Center of Historical Awareness of Thermopiles, of Municipality of Lamia", for the creation of a virtual museum in the area of Thermopylae. It is one of the exhibits of the new museum. The project has been co-funded by European Commission and National Funds within the 3rd framework of Support.

References

1. Bricken, M.: Virtual Reality Learning Environments: Potentials and Challenges. Computer Graphics 25, 178–184 (1991)
2. History of the Greek nation volume B', Ekdotike Athenon, Athens (1971)
3. Herodotus, Rawlinson, G. (Trns): The History of Herodotus: Polymnia (2005)
4. Roberts, D.F., Foehr, U.G., Rideout, V.J., Generation, M.: Media in the Lives of 8-18 year-olds. Kaiser Family Foundation, Menlo Park (2005)
5. Lee, E.A.-L., Wong, K.W.: A Review of Using Virtual Reality for Learning. In: Pan, Z., Cheok, D.A.D., Müller, W., El Rhalibi, A. (eds.) Transactions on Edutainment I. LNCS, vol. 5080, pp. 231–241. Springer, Heidelberg (2008)
6. Gaitatzes, A.G., Christopoulos, D., Roussou, M.: Virtual Reality Interfaces for the Broad Public. In: The Proceedings of Human Computer Interaction 2001, Panhellenic Conference with International Participation, December 7-9. Patras, Greece (2001)
7. Roussou, M., Slater, M.: A Virtual Playground for the Study of the Role of Interactivity in Virtual Learning Environments. In: Proc. of PRESENCE 2005: The 8th Annual International Workshop on Presence, London, UK, September 21-23. International Society for Presence Research, pp. 245–253 (2005)

8. Anderson, E.F., McLoughlin, L., Liarokapis, F., Peters, C., Petridis, P., de Freitas, S.: Serious Games in Cultural Heritage. In: 10th VAST International Symposium on Virtual Reality, Archaeology and Cultural Heritage (VAST 2009), VAST-STAR, Short and Project Proceedings, Eurographics, Malta, pp. 22–25, 29-48 (2009) ISBN: 978-99957-807-0-8
9. Jesse, S., Shochet, J.: Designing Interactive Theme Park Rides. IEEE Computer Graphics and Applications, 11–13 (July-August 2001)
10. Shell, J.: Understanding entertainment: story and gameplay are one. In: Jacko, J.A., Sears, A. (eds.) The Human-Computer Interaction Handbook, ch. 43. Lawrence Erlbaum (2003)
11. Roussou, M.: Incorporating Immersive Projection-based Virtual Reality in Public Spaces. In: Proceedings of 3rd International Immerse Projection Technology Workshop, Stuttgart, Germany, pp. 33–39 (May 1999)
12. Paul, H.: VR as Performance for an Audience. Leonarda Electronic Almanac 13(11) (2005)
13. Waring P.: Representation of Ancient Warfare in Modern Video Games. Master's thesis, School of Arts, Histories and Cultures, University of Manchester (2007)
14. Christopoulos, D., Gaitatzes, A.: Multimodal Interfaces for Educational Virtual Environments. In: IEEE Proceedings of the 13th Panhellenic Conference on Informatics (PCI 2009), Corfu, Greece, September 10-12, pp. 197–201 (2009)
15. Sanchez, A., Barreiro, J.M., Maojo, V.: Design of virtual reality systems for education: a cognitive approach. Education and Information Technologies 5, 345–362 (2000)
16. Slater, M.: Place illusion and plausibility can lead to realistic behaviour in immersive virtual environments. Philos. Trans. R. Soc. Lond. B. Biol. Sci. 364(1535), 3549–3557 (2006)
17. Christopoulos, D., Gaitatzes, A., Papaioannou, G., Zyba, G.: Designing a Real-time Playback System for a Dome Theater. In: Proc. Eurographics 7th International Symposium on Virtual Reality, Archaeology and Intelligent Cultural Heritage VAST, Cyprus (2006)
18. Gaitatzes, A., Christopoulos, D., Papaioannou, G.: Virtual Reality Systems and Applications: The Ancient Olympic Games. In: Bozanis, P., Houstis, E.N. (eds.) PCI 2005. LNCS, vol. 3746, pp. 155–165. Springer, Heidelberg (2005)
19. Gaitatzes, A., Papaioannou, G., Christopoulos, D., Zyba, G.: Media Productions for a Dome Display System. In: Proc. ACM Symposium on Virtual Reality Software and Technology (VRST 2006), pp. 261–264 (2006)
20. Taylor II, R.M., Hudson, T.C., Seeger, A., Weber, H., Juliano, J., Helser, A.T.: VRPN: A Device-Independent, Network-Transparent VR Peripheral System. In: Proceedings of the ACM Symposium on Virtual Reality Software & Technology 2001, VRST 2001, Banff Centre, Canada, November 15-17 (2001)
21. `http://www.lua.org` (last visited September 17, 2010)
22. `http://www.openscenegraph.org` (last visited September 17, 2010)
23. Rost, R.J.: OpenGL(R) Shading Language. Addison Wesley Longman Publishing Co., Inc., Redwood City (2004)
24. `http://www.mplayerhq.hu` (last visited September 17, 2010)
25. Scanlon, E., Tosunoglu, C., Jones, A., Butcher, P., Ross, S., Greenberg, J., Taylor, J., Murphy, P.: Learning with computers: experiences of evaluation. Computers & Education 30(1/2), 9–14 (1998)
26. Klopfer, E., Osterweil, S., Salen, K.: Moving learning games forward, obstacles opportunities & openness. An Education Arcade paper (2009)
27. Tzortzaki, D.: Museums and virtual reality: using the CAVE to simulate the past. Digital Creaticity 12(4), 247–251 (2001)

28. Alexaki, M.: The Role of the Mediator in Virtual Archeaology Experiences. Master of Arts thesis, University of Athens Department of Museum Studies, Athens
29. Pujol, T., Economou, M.: Worth a Thousand Words? The Usefulness of Immersive Virtual Reality for Learning in Cultural Heritage Settings. Journal of Architectural Computing 7(1), 157–176 (2009)
30. Economou, M., Pujol, T.: Educational tool or expensive toy? Evaluating VR evaluation and its relevance for virtual heritage. In: Kalay, Y., Kvan, T., Affleck, J. (eds.) New Heritage: New Media and Cultural Heritage, London, Routledge, pp. 242–260 (2008)
31. Roussou, M.: The Components of Engagement in Virtual Heritage Environments. In: Proceedings of New Heritage: Beyond Verisimilitude - Conference on Cultural Heritage and New Media, Hong Kong, pp. 265–283 (2006)
32. Bear, J.: Finding Personality in Games. Why Mega man's jumping facial expression is More important than normal mapping. Game Developer Magazine 17(8), 13–18 (2010)
33. Fencott, C.: Content and creativity in virtual environment design. In: Proceedings of VSMM 1999, Dundee, pp. 308–317 (1999)
34. Andreadis, A., Hemery, A., Antonakakis, A., Gourdoglou, G., Mauridis, P., Christopoulos, D., Karigiannis, J.N.: Real-Time Motion Capture Technology on a Live Theatrical Performance with Computer Generated Scenery. In: 14th Panhellenic Conference on Informatics (PCI 2010), Tripoli, Greece, September 10-12, pp. 148–152 (2010)
35. Gosseling, K.: The future of gaming. A Portrait of the New Gamers. Latitide Research, `http://www.latd.com/2011/08/23/the-future-of-gaming-a-portrait-of-the-new-gamers/` (last visited September 2, 2011)
36. Hanrahan, P., Lawson, J.: A language for shading and lighting calculations. ACM SIGGRAPH Computer Graphics 24(4), 289–298 (1990)
37. Beeson, C., Bjorke, K.: Skin in the 'Dawn' demo. ACM SIGGRAPH Computer Graphics 38(2), 14–19 (2004)
38. d'Eon, E., Luebke, D., Enderton, E.: Efficient Rendering of Human Skin. Eurographics Symposium on Rendering 2007 (2007)
39. Zhang, F., Sun, H., Xu, L., Lee, K.-L.: Parallel-Split Shadow Maps for Large-Scale Virtual Environments. In: Proceedings of ACM International Conference on Virtual Reality Continuum and Its Applications 2006, pp. 311–318 (2006)
40. Zhang, F., Sun, H., Xu, L., Lee, K.-L.: Hardware-Accelerated Parallel-Split Shadow Maps. In: International Journal of Image and Graphics (2007) (in press)

Games on Prescription! Evaluation of the Elinor Console for Home-Based Stroke Rehabilitation

Per Backlund[1], Anna-Sofia Alklind Taylor[1], Henrik Engström[1], Mikael Johannesson[1], Mikael Lebram[1], Angelique Slijper[2], Karin Svensson[2], Jesper Poucette[3], and Katharina Stibrant Sunnerhagen[4]

[1] University of Skövde, School of Humanities and Informatics, Högskolevägen, SE-541 28, Skövde, Sweden

[2] Skaraborg Hospital, Department of Occupational Therapy and Physiotherapy, Lövängsvägen, SE-541 85, Skövde, Sweden

[3] Skaraborg Primary Care, Ågårdsskogens Primary Care Centre, Tallskogsvägen, SE-531 51, Lidköping, Sweden

[4] The Sahlgrenska Academy, Institute of Neuroscience and Physiology, Section for Clinical Neuroscience and Rehabilitation, Per Dubbsgatan 14, SE-413 45, Göteborg, Sweden

{per.backlund,anna-sofia.alklind.taylor,henrik.engstrom, mikael.johannesson,mikael.lebram}@his.se, {angelique.slijper,karin.eli.svensson, jesper.poucette}@vgregion.se, {katharina.sunnerhagen}@neuro.gu.se

Abstract. This paper reports the feasibility of Elinor, a game-based system for stroke rehabilitation in the home. The Elinor prototype has been positively evaluated with respect to its usability, user acceptance and motivational factors as well as its rehabilitation effect. This paper reports the findings from the whole project. To summarize the results, we find that game factors can be used to enhance motivation for rehabilitation. We had positive results with respect to many of the rehabilitation measurements employed. For example, the assessment of motor and process skills was positive as were also the self-reported improvements in daily activities. Furthermore, it seems that an increased self-efficacy with respect to the belief that the treatment can have an effect is positive and expected to increase motivation to undergo necessary rehabilitation. The usability and perceived usefulness of the system were also positively evaluated and the subjects expressed a positive attitude towards the system as well as a belief in its usefulness.

Keywords: Stroke rehabilitation, serious games, home-based rehabilitation, games for health.

1 Introduction and Background

Stroke is a clinical syndrome, typically identified by rapidly developing signs of focal or global disturbance of cerebral functions lasting more than 24 hours [1]. Approximately 30 million people worldwide have survived a stroke and the global burden and prevalence of the disease continues to grow. Many stroke survivors are

Z. Pan et al. (Eds.): Transactions on Edutainment IX, LNCS 7544, pp. 49–64, 2013.
© Springer-Verlag Berlin Heidelberg 2013

facing long-term disability. Furthermore, the incidence of stroke is predicted to rise due to a rapidly ageing population. A significant proportion of strokes afflict relatively young people. Approximately 20% of the participants in a whole population-based study in England were <65 years of age [2]. The proportion is similar in Sweden [3]. Upper extremity impairment is common. 80% of the patients suffer from it in the acute phase and 40 % suffer from it in the chronic phase. Upper extremity impairment limits the ability for intentional, coordinated, and effective movements as well as the general level of activity and participation of a person. Stroke rehabilitation is tedious and time is an important factor. In particular, it is important for patients to keep their motivation for continuous training once they are discharged from hospital. This is where the need for motivational and self-administrated systems for home-based rehabilitation arises. One potential solution is to design a game-based system for home rehabilitation. The design vision for such a system is to be accessible, highly motivating to use and self-administrated.

Virtual reality (VR) applications have emerged as a promising tool for therapy and rehabilitation over the last decade [4], [5], [6]. However, in spite of all positive reports and diversity of ideas there is still some criticism towards the concept in terms of complex technologies and a lack of studies pertaining to the motivational aspects of such systems. In relation to games, Flynn et al. [6] claim that computer games and game technology share similar behaviors as VR systems even though lacking in specificity. One proposition is hence to use games, with their often inherent ability to entertain, for certain aspects of rehabilitation. However, we have identified a lack of studies actually testing virtual reality and/or game-based rehabilitation at home. There is also a need for studies analyzing the combination of usefulness (the utilitarian aspect) and enjoyment (the entertainment aspect), i.e. so called serious games. Furthermore, in this context it is important to take usability and robustness as well as the actual integration of the home-based system in the rehabilitation process into account.

Our review of currently available technologies for VR and game-based stroke rehabilitation reveals three classes of systems. *High-end systems*, as exemplified by [7], [8], [9], intended for clinics and rehabilitation centers. This class of systems typically utilizes expensive and advanced technologies which need professional supervision and guidance. *Mid-range systems*, as exemplified by [10], [11], [12] which can be used in home rehabilitation. However, this poses requirements on robustness and usability. There is typically a lack of testing of these systems in actual home environments, which is a drawback as home rehabilitation is a feasible solution. Finally, there are examples of using commercial games, as exemplified by [6] to fully exploit the motivational power. This class of system is based on contemporary video game technology with low costs and high accessibility. One major drawback is that they are most often too advanced for impaired persons. Furthermore, the possibility to adapt games and to obtain specifically designed rehabilitation games is limited, especially for console games (e.g. PlayStation, Nintendo).

Our review of current applications for VR rehabilitation further reveals that they typically focus on the technological solution and only to some extent on motivational aspects and rehabilitation effect. In our view, the technology component is one important factor to take into account even though it does not utilize the full potential of the concept. In this paper we propose that the combination of easy-to-use technology and entertaining and challenging game content can be used for rehabilitation of stroke patients at their homes. We stress the importance of the game component as a motivational factor while keeping in mind the intended rehabilitation effect. The concept of using games for purposes other than entertainment is referred to as *serious games* [13], [14]. Although the term itself is well established in both academia and industry, there is no current single definition of the concept, see [13], [14] for some examples. However, a common component of these definitions is that it is the addition of an additional purpose of training or instruction that makes games serious. For the purpose of our work we define serious games as follows: *Serious games are games that engage users in their pursuit, and contribute to the achievement of a defined purpose other than pure entertainment (whether or not the user is consciously aware of it).* Hence, we also include purposes other than instruction and training which broadens the area of application of serious games. An important, and rapidly growing, sector within serious games is games for health [15] which aims to use cutting-edge games and game technologies to improve health and health care. This includes games for physical fitness (so called exergames), games for education and self-directed care, games for distraction therapy and games for recovery and rehabilitation. We focus on the combination of technologies and game content and we stress the importance of the game component to provide serious games for game-based stroke rehabilitation. There are some examples of evaluation of game-based systems for stroke rehabilitation [8], [9], [10], [11] but none of them has been evaluated in its intended usage situation, i.e. self-administrated home-based rehabilitation. The present study fills an important knowledge gap in game-based stroke rehabilitation at home as we are unaware of any evaluations actually made in the homes of the patients; combining entertainment factors with user acceptance and usability as well as rehabilitation results.

The purpose of this paper is to evaluate the usefulness of a game console (Elinor) specifically designed for patients suffering from stroke. The framework for evaluation is an adapted version of the widely used technology acceptance model (TAM), developed by Davis and colleagues [16], [17], [18]. One drawback of the technology acceptance model is that it only takes the utilitarian aspects into account whereas Elinor is a serious game, i.e. also incorporating an aspect of entertainment. An important factor that has been added to account for the fact that people also use these systems for recreational purposes is perceived enjoyment as introduced by [19] and [20]. In sum, people are more inclined to use a system if they perceive it as useful, easy to use and enjoyable, which is essential in the case of a serious game. In this study the original TAM has been complemented with the GameFlow model [21]. The GameFlow model comprises a set of factors pertaining to how games seem to keep the player in a state of focus and concentration allowing them to stick to the activity unaware of time and context.

In addition to the extended TAM we also evaluate the usefulness of Elinor in terms of its rehabilitation effect by employing the following tests. The National Institutes of Health Stroke Scales (NIHSS) [22] was used to measure body function. NIHSS is a standardized assessment tool after stroke. *Measures of active movement* were made using a goniometer. Movement in the shoulder was measured in flexion, abduction and outward rotation. Movement in the elbow joint was measured in flexion, extension and supination. *The finger-nose test* is a standardized test to assess coordination. *The Modified Ashworth test* [23] was used for assessing increased muscle tonus. *The Action Research Arm Test* (ARAT) [24] was used to assess upper-extremity function in different activities such as grasp, grip, pinch, and gross movements of flexion and extension at the elbow and shoulder. *Motor Activity Logs* (MAL) and *Assessment of Motor and Process Skills* (AMPS) [25] were further employed to evaluate participation in daily activities. AMPS is a standardized assessment of performance that is used to observe and evaluate a person‘s ability to perform personal and instrumental activities of daily living. The AMPS test is designed to be used to assess the quality of a person’s activities of daily living (ADL) motor and process skills and is an indicator of whether the person has the skills necessary to efficiently, safely, and independently perform ADL tasks needed to live an everyday life.

Motor Activity Log (MAL) [26] is a structured interview in which the person is to assess how well (QOM-quality of movement), and how often (AOU –amount of use) the hemi-paretic arm is used in 30 different ADL activities.

Semi structured qualitative interviews with an interview guide with different themes (experiences of the training module, reason for participating, thoughts of this training compared to other training, awareness of the impaired upper extremity) were used to capture the subjects’ perceptions and experiences of the system

This paper presents a study of how the console can be used for rehabilitation of upper extremity impairment after stroke. In the following we present usefulness; ease of use and enjoyment of the Elinor console as perceived by five patients who used the console for five weeks in their homes. In addition to the perception of the system and its usefulness we also report actual measurements of the improvement of the patients. This paper reports the entire project. The detailed findings of the different aspects have been reported in [27], [28], [29] and [30]. The paper is organized as follows: firstly we describe the Elinor console and its functionality. Secondly, we describe the overall set up of the evaluation study followed by an analysis of each of the three evaluated aspects. Finally, we close the paper by some concluding remarks.

2 The Elinor Console

Fig. 1. The Elinor console in action

The Elinor console (Fig. 1) is a system for game-based stroke rehabilitation which has been developed in a cooperative project involving game designers and rehabilitation experts. The development phase consisted of six prototype rounds; each one with a design and development phase followed by user testing. The different test groups consisted of an occupational therapist, a physiotherapist, stroke patients and nursing students.

The game console has been constructed to be a standalone and easy to use device. It is based on a lap top computer and an off the shelf game control (GameTrak).The console has a very plain interface composed of an on/off button, a volume control, a game reset button and a USB plug for personal activity logs. Fig. 2 demonstrates the main components of Elinor.

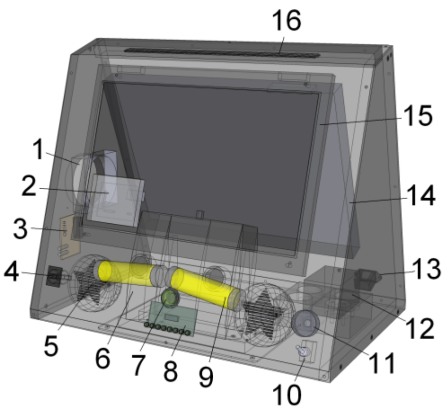


Fig. 2. Game control and overview of console components: 1. ventilation fan, 2. controller card, 3. hard reset electronics, 4. USB port, 5. speaker, 6. GameTrak, 7. pushbutton, 8. LED bar, 9. handle, 10. power switch, 11. volume control, 12. power supply, 13. AC inlet, 14. computer, 15. computer display, 16. hot air exhaustion.

Each user is equipped with a personal plug (USB memory stick) which acts as a key to the system. To use Elinor, the patient simply connects a power cord in the AC inlet, inserts a plug in the USB port and switches the power on. There are three types of users:

- User – this is a personal plug which stores game settings and logs of the patient. Irrespective of which Elinor console the patient is playing on (typically the machine at home or at the clinic) the system will look the same (same number of games, achievements, scores etc.).
- Guest –this plug allows relatives and friends to try Elinor without affecting the user logs or settings.
- Clinic –this plug puts the console in clinic mode, which makes it possible to analyze player activities and game results.

In the current version of Elinor all communication of patient data is done via the USB stick. The personal USB stick is used to communicate activities and results to care givers at the clinic.

The design is deliberately toned down to reduce complexity and risk for failure in use. The games are controlled using the GameTrak handles. The handles are attached to the console by strings and read movements in x, y and z axes, i.e. horizontal, vertical and depth. The grips are designed to be used with a full hand grip. The interaction mode allows for crude motor interaction using arm movements. All games are designed to be controlled by one hand alone or two hands in combination.

Elinor has a game library of 15 games some of which are inspired by classic video games such as Breakout, Snake and Puzzle Bubble and some are newly developed ones. There is a variation in the gameplay between puzzle, action, and sport games. Care has been taken not to introduce too complex interaction models, or to give the player a high cognitive load in terms of graphical and auditive feedback. Consequently, we have used toned down color schemes and sounds in order to adapt audiovisual stimuli to the intended users. In line with this, the system introduces a mandatory 5 minutes pause after 15 minutes of play. This helps patients avoid playing until they are completely exhausted, which is a risk for some stroke patients. Each game is also adapted to the user group in terms of speed and challenge.

Games typically start at an easy level and become harder as the player develops her/his skill. As the games have the purpose of stimulating the player to use her/his impaired arm they are designed to promote different types of arm movements. In particular, the games are adapted to challenge the impaired arm slightly more, even though seven of them are controlled by both arms. When a game is introduced, a short spoken introduction is given together with an animation showing the physical interaction required. When a game level has been successfully completed, the player is rewarded a star and the game advances to the next level. If a player fails a level five consecutive times, the game returns to an easier level. In this way it is possible for a player to be awarded stars once in a while even when they have reached the upper limit of their capacity. Players collect stars for each individual game and these are presented in the menu system. The games are ordered in the menu according to the number of stars achieved. This means that the most played game will be placed last in the menu. This helps the player to vary the gameplay (and the training). In addition, the selection of games in the menu gives the player additional training. By putting the most popular game last, the training is maximized.

System stability and reliability were considered important design rationales. The project put great effort into designing a system that would be robust enough for unsupervised home use. The success of this factor is demonstrated by the relatively few reported technical failures from 110 hours of registered play time distributed over 286 sessions. Two types of unwanted behavior were observed: internal errors unnoticed by the user and system failures directly affecting the user. The first category of errors typically consisted of delayed communication between system components and delayed system shutdowns which did not affect users or logs. The second category involved one console that suffered from a faulty graphics card. We registered seven failures due to erroneous use of personal USB sticks, most probably that the users forgot to unlock the system with their personal plug. 11 operating system errors were logged, typically leading to a delayed starting sequence for the user. Two of the consoles suffered from one internal communication error each leading to a complete reboot of the system with no further consequences for the users. Finally, one failure with unknown cause was logged where the console stopped and rebooted in the middle of a game.

3 Evaluating the Elinor Console

The Elinor console has been evaluated with respect to entertainment, usability, acceptance and rehabilitation effect by a combination of methods combining the extended technology acceptance model [18], for analysing user perceptions; measurements from rehabilitation studies, for analysing training effect, and; the game flow model for analysing player behavior. In total, five subjects (Table 1) had an Elinor console in their home for five weeks. Before the experiment started, the patients were screened by a physiotherapist and an occupational therapist to assess arm function and use in daily activities. In conjunction with the screening process, the participants had a first introduction to the system. The introductory session also included a short, semi-structured interview with questions about the subjects' experience of and attitude towards computers and computer games, which was carried out by the development team.

Table 1. Subjects' characteristics

Subject	Age	Gender	Impairment	Experience from computers	Interested in new technologies
A	64	F	Right sided	Minor	Minor
B	71	F	Left sided	None	None
C	64	F	Left sided	Minor	Minor
D	52	F	Left sided	Minor	Moderate
E	70	M	Left sided	None	None

The participants had an Elinor unit delivered to their homes and received a second introduction on how to handle the system. They also received a short manual of two pages. After the introduction of the system the participants were instructed to play as much as they liked. No recommendations with respect to how much the subjects should play were made. The study went on for five weeks and all patients met their therapists four times, i.e. once a week throughout the study. A representative from the development team was also present during these meetings, asking questions about how the subject experienced the different games. A member of the development team carried out a second interview, similar to the one made during the introduction, to investigate changes in attitudes towards the games and the system. The patients were also interviewed by a psychologist one week after the intervention. This interview was semi-structured and intended to explore motivational aspects of the subjects' experiences.

To summarize, data were collected throughout the entire study as follows:

- Personal logs: On an average 1325 minutes per player.

- Clinical tests: Pre and post experiment.
- Interviews: Two in conjunction to the clinical tests (carried out by the development team) and one after the experiment (carried out by a psychologist). The development team also asked questions during the follow-up meetings.
- Observation sessions: Four during the follow-up meetings.

The following evaluation aspects were considered:

- Rehabilitation effect: measurements and methods from physiotherapy and occupational therapy
- Game aspects: the GameFlow model
- Perceived usefulness and usability: the extended technology acceptance model

All subjects' activities on Elinor were logged on a personal USB stick. Guest activities were logged separately on a guest USB stick. Logging was enforced by the fact that the unit cannot be started without the USB stick. Data were also collected by means of interviews, observations and clinical tests. The study was approved by the Ethics Committee and all subjects gave their written informed consent. In the following we present the results of the different aspects evaluated.

3.1 Rehabilitation Effect

The rehabilitation effect was evaluated by means of a set of relevant measurements. A majority of the participants increased their capacity for active movement in the shoulder. None of the subjects showed any changes in their capacity for elbow flexion whereas the results for elbow extension and supination were mixed. Most subjects did not show any change for tonus. Finally, most subjects showed an increased capacity for coordination in the finger-nose test.

Most subjects showed an improvement in Action Research Arm Test (ARAT), Assessment of Motor and Process Skills (AMPS) and Motor Activity Log (MAL).

No persons suffered any serious adverse effects from the training. ARAT shows improvement for 3 of the subjects, 1 person showed no change and 1 person deteriorate in ARAT. The deterioration was within the range of error, while 2 of 3 persons showed a clinical definable difference. The AMPS test indicates improvements for all subjects, even though not clinically definable. The self-reported changes using MAL indicates improvements. The results from the tests employed are summarized in Table 2.

One person stated that the motivation to use Elinor was low. He became stressed when things went wrong and tired easily. He was also frustrated over the fact that the staff could see which games were used during what hours of the day. The other four described training as fun, and relaxing, which was perceived as positive for motivation. Subject C even pointed out that the imposed breaks were too long and kept a book or a cross word puzzle handy while waiting for the game to start again. The participants were well aware of the fact that they were actually carrying out training in what was perceived as a fun and motivating manner, which is not always the case.

In summary, the results indicate that the system is useful in terms of rehabilitation effect with respect to all measurements even though there are some incongruities between the subjects and the fact that some results are not clinically definable.

Table 2. Summary of the assessments. Arrow upwards indicates improvement, arrow downwards indicates deterioration. Arrow and a question mark is an indication of improvement that is not necessarily of importance.

	Subject A	**Subject B**	**Subject C**	**Subject D**	**Subject E**
NIHSS	0	3	6	1	2
Game time (total)	703 min	2187 min	1866 min	1175 min	691 min
Shoulder flexion	↑	↑?	↑?	↑?	–
Shoulder abduction	↑	↑	–	↓?	–
Shoulder rotation outward.	↑	↑	–	↑?	↑
Elbow flexion	–	–	–	–	–
Elbow extension.	↑	↓	–	↑?	↓
Supination	–	↓	–	↑	↓
Finger-nose test	↑	↑	–	–	–
Modified Ashoworth	–	↑	↓	–	–
ARAT	↓?	↑	–	↑	↑?
AMPS motor skills	↑?	↑?	↑?	↑?	↓?
AMPS process skills	↑?	↑?	↑?	↑?	↑?
MAL QOM	↑	–	–	↑?	↑?
MAL AOU	↑	–	↑?	↑?	↑

3.2 Gaming Behavior and Motivation

Rehabilitation is a time consuming and most often strenuous activity. Hence, it is believed that the utilization of gaming factors to motivate the patient has potential in stroke rehabilitation e.g. [4], [6], [7]. We have analysed the transcribed material from the clinic visits and interviews with respect to the GameFlow elements [21] and [31]. The aim is not to analyse whether each game has been enjoyable, but rather whether subjects have expectations and experiences that resembles the elements of GameFlow.

The GameFlow model as presented by Sweetser & Wyeth [21] is an ambitious attempt to unify various aspects of enjoyment of games into one single model, based on the theory of flow. A similar approach based on flow theory is presented in [31]. The GameFlow model comprises eight elements: concentration, challenge, skills, control, clear goals, feedback, immersion, and social interaction. The concentration element refers to a game´s ability to capture and retain the player's attention. Challenge means that a game should have tasks that match the player's skill level. The player skill element capture aspects related to how players should learn, develop and master the game skills. The skill element also relates to the control element which covers aspects such that players should feel a sense of control over the game shell, the characters and the actions they take. It is crucial that players feel that their actions have an impact on the game. The clear goal element means that the game, as well as sections of the game i.e. levels, should have goals that are made clear to the player. The feedback element refers to how players receive response to their actions and information about their progress towards a goal. The immersion element relates to the player's involvement and engagement in the game, demonstrated by for example, loss of concern for self and an altered sense of time. Finally, the social interaction element means that game should support social interactions through the game, such as competitions, collaboration and social communities inside and outside the game.

The subjects were not given any recommendations on when to play, what games to play or how much to play; hence it is possible to analyse subjects' gaming behaviour from the usage pattern. In some sense the goal is to make the subjects play as much as possible, within the limits of exhaustion and without causing any detriments. The medical experts expressed that a training amount of 15 minutes a day, five days a week – in total 375 minutes, would be considered suitable. The subjects spent an average of 1324 minutes playing, i.e., more than three times than expected. There was however a great variation between subjects, as can be seen in Table 3. Table 3 summarizes the total time per subject, how many distinct sessions they played, and how many days they played at least one session. From our data we conclude that there was a fair distribution of activity throughout the test period even though there was a slight decline in playing time towards the end of the project.

Table 3. Total active play time, number of sessions and days of play

Subject	Playtime (min)	Sessions	Days
A	703	40	29/35
B	2187	61	31/35
C	1866	88	33/35*
D	1175	59	26/35
E	691	38	18/35

* Subject C lost one day of play due to a hardware failure.

We also conclude that four of the subjects actually played more than five days per week on an average. When contrasting this data with interview data all subjects reported to have one or several favorite games. One subject refers to all games as being fun. The subjects' engagement in the games is frequently demonstrated in both observations and interviews. Comments such as: "You have to be very determined in order to win." (Subject D) and "You have to stay in focus." (Subject B) further support this observation. Some of our data also reveals competitiveness as a factor to take into account and two of the subjects, A and B, show clear signs of competiveness as a driving force.

There is a great variation in popularity of games. The most popular games, in terms of play-time, were Breakout and Puzzle Bubble with a total playtime of 1300 minutes each. Bike Ride is third with 650 minutes, followed by Snake with 500 minutes. The least popular "game" is Paint with 20 minutes. It is remarkable that the most popular games are all based on well-established game concepts, while the least popular is the activity that has no explicit game element. This is an observation which supports our hypothesis that games can be used as a motivational factor for this type of rehabilitation training. This observation along with the statements from the subjects and the actual rehabilitation effect forms a strong case for the approach presented in this paper.

The games have to be sufficiently challenging in order to engage the player. The games are adapted to suit the user group by means of speed, colors, sounds and interaction mode. Three subjects (A, B and D) describe that the challenge is important. For example, Subject D says that she likes Breakout: "… this is so much more than just doing your homework." and compares it to earlier experiences of playing Tetris. However, challenge does not work the same for all subjects. Subject E stops playing some games after a number of failures and seems somewhat de-motivated by them. This negative effect is balanced by the fact that the games adapt to the results and becomes easier after a number of failures. Subject E perceives one of the games more positively and he is indeed doing better in this game. The relative success seems to boost confidence and increase motivation. However, Subject E acknowledges that receiving new games over time increases motivation. Subjects A and C also find the stepwise introduction of new games motivating. In general, challenge seems to be an important motivator and Subject D actually asks for games with a harder challenge. Subject A says that she likes the challenge and that the hard games are fun. Another example is subject D who found some of the games hard but her strong sense of competition kept her motivated to try over and over again.

Subjects give several comments that reflect that they have been immersed while playing. Some examples are: "you sit down to play and then you think of nothing but the game … there is no time to think about your affliction, so to say" (Subject A); "… it's fun, time flies" (Subject B); "I think it has been very fun" (Subject C); "you get into the game in some way and the game becomes real, and then you make movements and get training without actually reflecting over it"; (Subject D). It is apparent that subjects report experiences that fit the immersion definition very well. They have experienced and appreciated that they forget about their disease and that they feel that they play rather than performing rehabilitation exercises. They experience an altered sense of time, and they are emotionally involved in games.

Real time performance feedback is often held forward as a factor increasing motivation. However, performance may be a double-edged feature. If the challenge is on a suitable level and the feedback gives the user relevant information on how to succeed it is a useful tool. Our data reveals that the participants generally understood what to do and how the games worked. However, the scoring system's adaptation to performance was not apparent to all participants (Subject B and C). There were occasions when the researchers had to inform the participants of the fact that the games became harder (or easier) based on the results over the last rounds. This feature was designed to keep the challenge at a reasonable level but at some point it also tended to make the games too hard too quickly. This type of performance feedback appealed to the competitive behavior of the participants (Subject A) at the expense of carrying out correct movements which in turn may have a negative effect on the rehabilitation effect.

Social interaction is an important feature of games. Playing games is often a means of doing things together. Our log data reveals that guests have played on all units; the average time is 180 minutes. All five of the participants had their grandchildren (or neighbor's children) playing. Interview data reveals that they did not actually play together even though Subject C sat beside them while the children were playing. The case was typically that the children played by themselves. Nevertheless, the subjects received some attention to their rehabilitation. One of the subjects (Subject D) reports that her spouse has played as well. Subject E reports that his brother has played. The brother wanted to compete against the subject but the subject declined, saying that he does not want to compete against a healthy opponent. As can be seen from this data, Elinor has, to some extent, served as a means to draw some attention to the rehabilitation process which, in turn, may be utilized as a way to engage relatives and friends to enhance motivation for training.

The comments from subjects on their experience from using Elinor, reflects all elements of the GameFlow model. Some elements, such as immersion, are reflected in an apparent, positive way – the experiences reported by subjects are spot-on to the definition of immersion. Other elements, in particular skill, can more indirectly be mapped to the definition. Still, we find it interesting that a group of five subjects, with limited experience of computer games, express opinions and report experiences that are highly relevant from a computer game design perspective.

3.3 Perceived Usefulness, Usability and User Acceptance

The technology acceptance model uses two factors as predictors for user acceptance: perceived usefulness and perceived ease of use [16]. Perceived usefulness pertains to the extent to which a person thinks that using the system will enhance his or her work performance, and perceived ease of use is pertains to the extent to which a person thinks that using the system will be free of effort. Over the years, the technology acceptance model has been adapted to various contexts and systems, including non-work related ones, e.g. the World Wide Web [19], an on-line game [32], Lego Mindstorms [20] and virtual worlds [33]. Among the factors that have been added to account for the fact that people also use these systems for recreational purposes, one factor seems to be crucial: perceived enjoyment [19], [20]. In sum, people are more inclined to use a system if they perceive it as useful, easy to use and enjoyable. According to the technology acceptance model the type of system as well as the context of use are more or less vital as predictors for acceptance. For instance, [20] found that perceived ease of use only has an indirect impact on the intention to use a system that is both useful and fun. In our review of literature, we have found no study that explores the acceptance of serious games used in (stroke) rehabilitation. This is both a very specific system and context of use, which makes it especially interesting as an object of study. Since the system is designed to elicit increased motivation for rehabilitation training, it is valuable to study what drives the patients to use the system in the first place; as recreation or as a rehabilitation tool.

In order to fit the study, the definition of perceived usefulness was rephrased to: *the extent to which a person believes that using the system will improve his or her arm function*. Four of the five participants were from the beginning very positive towards the effectiveness of the system and they kept this positive view throughout the study. Subject D stated: “one cannot take on something and think that it’s rubbish from the beginning, because then you might as well quit. But you must feel that it has some kind of effect. The most important thing must be that you have a positive attitude towards what you are supposed to do. Then half of the battle is won, I think.” Subject A not only saw benefits from a personal perspective, but also from a cost-effective perspective: “…that to have it at home. Then you relieve the pressure of the care system. I mean, a thing like that is a one-off cost. A person – that's salary and the whole lot. So I think it can be sensible."

The system's usability and ease of use should be an important predictor for acceptance, since the primary user group is people with disabilities that cannot or find it difficult to use conventional computer equipment. None of the participants claimed that they found the Elinor hard to use once it was installed at their home. Subject B thinks that the spoken instructions are good in the beginning: "It tells you exactly what one should do when there are new games (…) So it hasn't been any problems. You have to listen to what he says, but when you've played so many times, you can all that anyway."

Subject E, who also experiences a slight cognitive impairment, found the quick reference guide useful, when he could not remember how to use the system: "And then you had a reference guide and there it says in detail how one should do. He [one of the researchers] showed a little, but… it was… that's the difficult part, with me having two ears with holes in them and then it goes in here and out through the other. I forget really easily."

Ease of use is related to the accessibility of the system, which subject C describes quite well: "It's been very smooth. It's just been to go and have a seat and play a little when you felt like it. 'Now it's time to do some exercises'. That has been good." The weekly follow-ups support this conclusion, since the participants only made minor errors (such as scrolling one game too far in the menu screen) during observations. Most of the difficulties experienced were due to the specific subject's disabilities, i.e. that they found a game hard to play because they were not able to do the required movement. In those cases, the participant chose to play those games that he or she was able to play.

To summarize, the degree of perceived usefulness and ease of use seem quite high, which is positive given the impairments of the subjects. Hence we conclude that the aim to design a system that could be used with only limited instructions and a limited amount of support seem to have been reached.

4 Concluding Remarks

This study aims to explore the usefulness and effectiveness of a home-based system for rehabilitation of the upper extremity for persons suffering from stroke. The study is designed to evaluate three complementing aspects which we argue are relevant to the type of system at hand: rehabilitation effect; game aspects; as well as the perceived usefulness and usability of the system. We have designed a study to evaluate all three aspects in their actual usage context and we received positive results for all of them. In addition, the system seemed stable enough for unsupervised home use.

To summarize the results, we find that game factors can be used to enhance motivation for rehabilitation. Concerning the rehabilitation effect, we had positive results for AMPS motor and process skills. Even though these improvements were not significant they are still positive enough to motivate future work. The self-reported improvements in the motor activity log also point in a positive direction. Furthermore, it seems that an increased self-efficacy with respect to the belief that the treatment can have an effect is positive and expected to increase motivation to undergo necessary rehabilitation. The usability and perceived usefulness of the system were also positively evaluated and the subjects expressed a positive attitude towards the system as well as a belief in its usefulness. The quantitative data from the user logs, indicating a playtime exceeding the estimated by more than three times, corroborates the users' perceived usefulness of the system. Interestingly enough, we also observed player behavior in terms of immersion and engagement among the subjects even though they are not to be classified as typical gamers.

In order to be able to carry out an evaluation of all relevant aspects of a serious game we have combined methods from different disciplines. We find this approach necessary for a class of systems that combine utilitarian and recreational aspects such as serious games for health. It is therefore worth noticing that there is not only a need for interdisciplinary teams in the development process but also in the evaluation of such systems. Hence, the evaluation model developed here may form a methodological contribution to the area of usability studies of serious games even though we don´t prescribe any explicit additions to the technology acceptance model.

There are a number of initiatives for game-based stroke rehabilitation which we have classified into high-end, mid-range and commercial games/systems. Even though many of them highlight the usefulness for home-based stroke rehabilitation there is a severe lack of studies to show the feasibility of such a concept in its actuial usage context. Our study has indicated positive results for rehabilitation measurements thus demonstrating the feasibility of the concept. Furthermore, the study also demonstrates a process for carrying out game-based stroke rehabilitation in the home including how to communicate between patient administrated systems and the healthcare professionals at the clinic. Even though not evaluated in this study, the seamless integration of rehabilitation activities in the home with treatment and follow up in clinics is an essential issue to address for this type of systems.

References

1. World Health Organization (WHO) Stroke, Cerebrovascular accident, `http://www.who.int/topics/cerebrovascular_accident/en/` (retrieved November 1, 2010)
2. Rothwell, P.M., Coull, A.J., Giles, M.F., Howard, S.C., Silver, L.E., Bull, L.M., et al.: Change in stroke incidence, mortality, case-fatality, severity, and risk factors in Oxfordshire, UK from 1981 to 2004. Oxford Vascular Study. Lancet 12(363), 1925–1933 (2004)
3. National Board of Health and Welfare. Guidelines for Stroke Medical Services. In: Swedish: Socialstyrelsen. Nationella riktlinjer för strokesjukvård 2005: Socialstyrelsen (2006)
4. Rizzo, A., Kim, G.J.: A SWOT Analysis of the Field of Virtual Reality Rehabilitation and Therapy. Presence: Teleoperators and Virtual Environments 14(2), 119–146 (2005)
5. Laver, K.E., George, S., Thomas, S., Deutsch, J.E., Crotty, M.: Virtual reality for stroke rehabilitation. Editorial Group: Cochrane Stroke Group (published online September 7, 2011), `http://www.thecochranelibrary.com/view/0/index.html`, doi:10.1002/14651858.CD008349.pub2 (retrieved October 4, 2011)

6. Flynn, S., Palma, P., Bender, A.: Feasibility of Using the Sony PlayStation 2 Gaming Platform for an Individual Poststroke: A Case Report. Journal of Neurologic Physical Therapy 31, 180–189 (2007)
7. Broeren, J.: Virtual Rehabilitation – Implications for Persons with Stroke Doctoral dissertation. PhD Thesis. University of Gothenburg, Gotheburg (2007)
8. Zheng, H., Davies, R., Zhou, H., Hammerton, J., Mawson, S.J., Ware, P.M., Black, N.D., Eccleston, C., Hu, H., Stone, T., Mountain, G.A., Harris, N.D.: SMART project: application of emerging information and communication technology to home-based rehabilitation for stroke patients. In: Proceedings of the 6th International Conference Disability, Virtual Reality & Assoc. Tech., pp. 215–220 (2006)
9. Adamovich, S.V., Merians, A.S., Boian, R., Lewis, J.A., Tremaine, M., Burdea, G.S., Recce, M., Poizner, H.: A Virtual Reality-Based Exercise System for Hand Rehabilitation Post-Stroke. Presence: Teleoperators and Virtual Environments 14(2), 161–174 (2005)
10. Johnson, M.J., Trickey, M., Brauer, E., Feng, X.: TheraDrive: A New Stroke Therapy Concept for Home-based, Computer-Assisted Motivating Rehabilitation. In: Proceedings from the 26th Annual International Conference of the IEEE EMBS, pp. 4844–4847 (2004)
11. Kizony, R., Weiss, P.L., Shahar, M., Rand, D.: TheraGame – a home based virtual reality rehabilitation system. In: Proceedings of the 6th International Conference Disability, Virtual Reality & Assoc. Tech., pp. 209–214 (2006)
12. Betker, A.L., Szturm, T., Moussavi, Z.K., Nett, C.: Video Game-Based Exercises for Balance Rehabilitation: A Single-Subject Design. Archives of Physical Medicine Rehabilitation 87, 1141–1149 (2006)
13. Zyda, M.: From visual simulation to virtual reality to games. Computer 38(9), 25–32 (2005)
14. LoPiccolo, P.: Serious games. Computer Graphics World 27(2), 4–6 (2004)
15. Games for Health. The Games for health Project – Sixth Annual Games for Helath Conference, `http://www.gamesforhealth.org` (retrieved November 1, 2010)
16. Davis, F.D.: Perceived usefulness, perceived ease of use, and user acceptance of information technology. MIS Quarterly 13(3), 319–340 (1989)
17. Davis, F.D., et al.: User acceptance of computer technology: a comparison of two theoretical models. Management Science 35(8), 982–1003 (1989)
18. Venkatesh, V., Davis, F.D.: A theoretical extension of the technology acceptance model: four longitudinal field studies. Management Science 46(2), 186–204 (2000)
19. Van der Heijden, H.: User acceptance of hedonic information systems. MIS Quarterly 28(4), 695–704 (2004)
20. Chesney, T.: An acceptance model for useful and fun information systems. Human Technology 2(2), 225–235 (2006)
21. Sweetser, P., Wyeth, P.: GameFlow: a model for evaluating player enjoyment in games. Computers in Entertainment 3(3) (2005)
22. Brott, T., Jr. Adams, H.P., Olinger, C.P., Marler, J.R., Barsan, W.G., Biller, J., et al: Measurements of acute cerebral infarction: a clinical examination scale. Stroke 20(7), 864–870 (1989)
23. Gregson, J.M., Leathley, M., Moore, A.P., Sharma, A.K., Smith, T.L., Watkins, C.L.: Reliability of the Tone Assessment Scale and the modified Ashworth scale as clinical tools for assessing poststroke spasticity. Arch. Phys. Med. Rehabil. 80(9), 1013–1016 (1999)
24. Lyle, R.A.: Performance test for assessment of upper limb function in physical rehabilitation treatment and research. International Journal of Rehabilitation Research 4(4), 483–492 (1981)

25. Fischer, A.: Assessment of motor and process skills, 2nd edn. Three Star Press, Fort Collins (1997)
26. Blanton, S., Wolf, S.L.: An application of upper-extremity constraint-induced movement therapy in a patient with subacute stroke. Phys. Ther. 79(9), 847–853 (1999)
27. Alklind Taylor, A.-S., Backlund, P., Engström, H., Johannesson, M., Lebram, M.: Gamers against all odds. In: Chang, M., Kuo, R., Kinshuk, Chen, G.-D., Hirose, M. (eds.) Learning by Playing. LNCS, vol. 5670, pp. 1–12. Springer, Heidelberg (2009)
28. Alklind Taylor, A.-S., Backlund, P., Engström, H., Johannesson, M., Lebram, M.: The Birth of Elinor - A Collaborative Development of a Game Based System for Stroke Rehabilitation. In: The Proceedings of the International Conference Visualisation (Viz 2009, CGa), Barcelona, Spain, July 14-17 (2009)
29. Alklind Taylor, A.-S., Backlund, P., Engström, H., Johannesson, M., Krasniqi, H., Lebram, M.: Acceptability of Entertainment Systems in Stroke Rehabilitation. In: The Proceedings of IADIS Game and Entertainment Technologies 2009 (GET 2009), Algarve, Portugal, June 17-19 (2009)
30. Backlund, P., Alklind Taylor, A.-S., Engström, H., Johannesson, M., Lebram, M., Poucette, J., Slijper, A., Svensson, K., Stibrant Sunnerhagen, K.: Evaluation of usefulness of the Elinor console for home-based stroke rehabilitation. In: The Proceedings of the 3rd International Conference in Games and Virtual Worlds for Serious Applications, Athens, Greece, May 4-6 (2011)
31. Salen, K., Zimmerman, E.: Rules of Play: Game Design Fundamentals. MIT Press, Cambridge (2003)
32. Hsu, C.-L., Lu, H.-P.: Why do people play on-line games? An extended TAM with social influences and flow experience. Information & Management 41(7), 853–868 (2004)
33. Fetscherin, M., Lattemann, C.: User acceptance of virtual worlds. Journal of Electronic Commerce Research 9(3), 231–242 (2008)

An Analysis of the Potential to Utilize Virtual Worlds to Enhance Edutainment and Improve the Wellbeing of the Ageing Population

Ann Smith and Matthew Chilcott

University of Wales, Newport
{Ann.Smith,Matthew.Chilcott}@Newport.ac.uk

Abstract. The contribution of this paper identifies the importance of developing and designing tailored virtual places which older members of society can utilize easily and safely within this context; with a focused application of addressing social isolation, enabling wellbeing and improving edutainment. The paper also explores the complexities associated with engaging older people with virtual world's experiences alongside the potential for new commercial approaches to design and delivery edutainment and wellbeing support initiatives in 3D immersive contexts through the implementation of next generational digital services.

Keywords: Ageing, virtual worlds, well-being, design, edutainment, social isolation.

1 Introduction

The internet is now central to our economic, cultural and political lives, used as a mechanism for the delivery of public services, personal communication, and as a vast source of information and entertainment. The United Nations has estimated that by 2050, one-third of the population of developed regions and one-fifth of the population in less-developed regions will be aged 60 and over. This presents an enormous task for the government and local authorities to address the needs of people in later life. In the context of a changing demography and growing numbers of older people, a wide range of literature now recognizes that older consumers are an increasingly important market for a variety of goods and services [7].

Older people, those with lower socio-economic status, individuals with limiting disabilities and those with lower educational attainment are more likely to be digitally disengaged. An important distinction is made between digital disengagement through personal choice and through socio-economic or health constraints. Digital disengagement through personal choice is closely linked to age and life stage of an individual, with typically those who are not interested in using the internet tending to be older and retired people. There is emerging evidence that compared with the digitally excluded, internet users feel less lonely and their personal wellbeing is enhanced.

Second Life (SL), a virtual world, represents the latest and most successful frontier of online services for business, learning, training and entertainment. Online

Z. Pan et al. (Eds.): Transactions on Edutainment IX, LNCS 7544, pp. 65–80, 2013.
© Springer-Verlag Berlin Heidelberg 2013

participants have virtual lives in these worlds using personalized avatars. It is one of the latest innovations in information communication technology and has become a popular platform for communication, social networking and engagement performing e-business and conducting numerous social learning activities and collaborative practices. Real life people (in the form of an avatar) interact with each other through voice, text and instant messaging based on web 2.0 technologies. However, SL has not been widely used successfully by the elderly population but it is thought that it could benefit local communities and individuals by letting them make new friends, improve social interaction, learn new skills, find information, and help them address social isolation and support feelings of wellbeing.

This paper will discuss three studies that were conducted with the elderly members of the Gaer community (in Newport) which has the highest proportion of elderly people in Wales (2001 Census). In a culture that has traditionally viewed older people as a drain on resources, this demographic requires us to think differently about elderly people and how we can help them make a valuable contribution to their communities and increase their sense of self worth. This paper highlights the need for specialized training and continuous support for the elderly who need to feel secure in using digital technologies and potentially new solutions for social and well-being problems; and to develop responses to challenges with the 3D medium of SL. This paper explores a new opportunity offered by SL for ageing wellbeing and active living that may also offer new ways to promote inter-generational engagement using digital technology.

2 Related Work

At what age does one become an 'older user'? The age of 50 is often used to define an older adult and there is clinical evidence to suggest that some age related declines begin at this time [18]. However, people who are 50, would deny a technological similarity to people who are 80 years old. The problem lies in that the elderly are not clear about what computers can provide that cannot currently be provided through personal (including phone) contact, letter writing, and interaction with the community. For many, such personal contact is considered desirable and there is no interest in replacing phone conversations with the impersonality of email or Facebook. Older adults are generally not considered to be early-adopters of technology and their voices are generally under-represented in consideration of new technologies. Many older adults are not interested in using the Web for online services, edutainment or e-commerce. Even more striking is the fact that social networking has failed to ignite the interest of older adults to the same extent that it has done for younger users [19].

Low levels of digital inclusion tend to be in areas with relatively older populations and/or high levels of social and economic exclusion. These characteristics also reflect the roles that personal choice play on the elderly (and its relation to age and stage of life) and socio-economic exclusion in digital engagement. Researchers have also shown that internet use tends to be lower among people with lower incomes and lower educational attainment, disabled people, single parents, female homemakers, the elderly and people living in rural communities [1] [2].

2.1 Social Participation

The main advantage of virtual worlds, such as SL, is linked to social participation. Figures indicate that there is a significant decrease in social participation in old age. It seems that the proportion enjoying hobbies and pastimes increased briefly around retirement age, only to fall away again among older respondents. Similarly, retirement appears to have triggered additional newspaper reading, although this stays high for some time, before dropping gently beyond the age of 75. This appears to be reflected in the decrease in spending on recreation and culture in the over-75s. However, this decline in participation should not necessarily to be taken as simple decline in appetite for participation [27].

2.2 Virtual- Worlds - Second Life

3D virtual worlds promise to offer a number of unique pedagogic opportunities such as immersion, visualization, exploratory learning and training, networking and collaboration edutainment (e.g., multi-player educational games)/entertainment (e.g., virtual cinema theatres), that can all be tailored to the nth degree to match the needs of various users.

Many of the communities formed around SL tend to become tightly-knit with time, with people collaborating in groups and getting to know each other better and more personally through their digital identities (and sometimes also by their real identities). The relationships may even extend to real life or make important parts of real life activities, and that is why 3D virtual worlds are sometimes referred to in the literature as 3D real-virtual worlds, since they are and can be a very real part of our real lives [22]. 3D virtual worlds offer strong cognitive cues that enhance collaboration, including 3D spatial audio and avatar lip-synching, bringing pseudo-body language communication to meetings. Humans are spatial beings by nature, inhabiting feature-rich 3-D analogue spaces, so a 3D synthetic space should not be more cognitively demanding from a human-computer interface viewpoint.

2.3 Social Isolation

Another important area of application for SL is the treatment of psychological and social troubles derived from aging (as a direct or indirect consequence of cognitive changes emerging in this period of life). One of the most important and worrying age-related problems experienced by the elderly is social isolation. Because of their physical disabilities or other age-related problems, they are often isolated from their familiar environment, cultural and a social context, in a negative loop in which a physical impairment leads to a social problem, which has physical rebounds and so on. The social engagement of elderly is consequently very important to maintain and promote wellbeing, from all points of view, and videogames can offer a contribution.

SL, for example can stimulate intergenerational interactions, where grandfathers/mothers can visit virtual places with grandchildren, and the visit can be guided/assisted by parents through an Internet connection.

3 Community Engagement

Older users often need more assurance and worry about ‘breaking’ the computer. Such anxieties for older users are often exacerbated by computer problems that may be mere annoyances to younger, experienced users. For example, on the training course we observed more than one older adult stop using the computer after receiving an error message that they had performed an ‘illegal operation’. Older adults coming to the Web experience start-up difficulties not characteristic of younger learners. They typically require, for example, more practice and more instruction than younger users. Interestingly, however, there is reason to believe that the type of experience older users have will affect their ability to use the Web and the use of virtual worlds. Specifically, collaborative experiences such as workplace environments or learning as part of a class may lead to improved Web expertise [20].

3.1 Training – Introduction to SL

The training on SL that was undertaken in the Gaer Community with older learners (see later), demonstrated very clearly that for older learners training on the various technologies, and especially SL, must address their specific needs e.g. speed of training, eyesight issues, dexterity issues, explanation of terms etc.

Effective engagement was found to be one of the most crucial factors. It mattered to the older people how they were spoken to, that their needs were being met, that they were being cared for, that they and their Trainer were having fun and that the Trainer was interested in them as much as what they were learning. Without this, a firm foundation for ongoing learning would not have been created affecting the success of the project. It was apparent that the older people had it within their psyche that they were too old to learn new things, especially new technologies like SL. SL has many interfaces, fast moving graphics, its own technological language, 'real life' interactions which the older learner found terrifying initially and which would have resulted in disengagement if all of the above had not been planned for and dealt with under the project.

3.2 Cognitive and Motor Skills

Specifically, some of the problems for older users from the training session included difficulty using the menu’s in SL (including flying and walking buttons), longer times to complete tasks, longer times to select targets and links, revisiting virtual islands, and difficulties finding new information. Older users also do not do well with functionality hidden in simplified interfaces that hide functionality [21]. While navigating the Web/virtual worlds and dealing with new social interactions and dynamically changing content appears almost effortless for younger users, the cognitive declines typical of older adults make their learning and use of such technologies difficult.

Older users may have difficulty using a mouse and keyboard due to illness or injuries that limit dexterity. There is evidence that older adults use different movement strategies than younger adults, with lower peak velocities and many sub-movements, and that fine positioning over a target is a particularly difficult for this group. From the training we discovered using the mouse and looking at the screen at the same time was very challenging for the group and took time to master.

3.3 Older People's Attitudes

Some research has examined in the more detail the psychological reasons older people may have chosen not to use the internet. For instance, Reisenwitz [3] studied the internet use of older people in, and examined correlations between internet use and three personality traits. They found that older people who were more innovative, less risk averse and less prone to nostalgia were more active internet users. Researchers for the think-tank Demos developed a typology of older people according to their psychological approach toward using the internet [4]. They identified four distinct types of older person: the non-line outsider, who is not generally averse to using the internet, but is hampered by fear and uncertainty; the tech skeptic, who is critical of technology and resentful of pressure to become connected; the cautious toe-dipper, who has tentatively embraced the internet for basic tasks but is wary of trying new websites, software, or deviating from trusted brands; and the digital trail-blazer, who is adventurous about trying out new things on the internet and is evangelical in their enthusiasm.

The training provided us a diverse range of elderly personalities from the non-line outsider to the digital blazer; and from the training it clearly shows that digital blazer users were keen to discover more about SL and its potential.

3.4 Digital Divide

For those who do not go online, issues such as a lack of motivation, a fear or distrust of technology, low media literacy, cost, ageism and usability are all key barriers [15]. Research conducted by BT and Age Concern and Help the Aged [16] found that the main barrier to use of computers and the internet by older people is a lack of understanding of and confidence with 'how it works'. For some, being dismissive masks a lack of confidence and they are likely to be more open about trying the internet if they receive the right encouragement and support. The respondents made a number of suggestions for motivating and helping older people to access and use the internet. These included awareness raising by promoting relevant benefits to older people; training, preferably in small classes of older people and in a community venue; monetary incentives to encourage participation and learning; and provision of ongoing help, preferably personal help. Accessibility and usability of online (and other new media content) services frequently fail to meet the needs of an older population and are a major barrier to internet use among older people.

The physiological effects of ageing also determine how people use interactive communications channels like the Web, e-mail and instant messaging. Older people are likely to be slower and make more errors when using these channels. The reaction to color, size of images, menu systems and animation are all affected by ageing. Most interactive channels are built by the young to be used by the young, and pay scant attention to the needs of the older person [17].

The training at the Gaer provided an ideal environment to train the elderly with the use of SL and the Web; for example the training took place in their own community venue which they felt secure and so expected no surprises. At the venue they knew everyone who were taking part in the training, as well as the people who were organizing the event.

3.5 Social and Cognitive Ageing

Social or psychological ageing has been defined as 'The norms, values and roles that are culturally associated with a particular chronological age'. People act a certain way because society expects them to act in a certain way. The physical decline of an older person impacts on society and thus the attitudes of older people ('People my age don't do that sort of thing') [9]. Social or psychological ageing may have impact on the use of virtual worlds in a number of ways; for example:

- Older people assume that certain applications and services are not for them and therefore do not consider using these technologies.
- Older people may assume that certain technologies are for younger people only; and do not consider their use as a means of engaging with their social lives or to improve their edutainment.
- The media and corporate sector may strengthen the negative impact of social ageing through ageism.

However, there is also considerable research to suggest that many older people perceive themselves to be younger than their chronological age (cognitive age 'is the age [10] one perceives one's self to be and is considered an element of self-concept' [11]). For example, that 'boomers feel younger in themselves than their actual age by approximately 12 years [12]'. Being 'cognitively young' is associated with increased social participation and engagement in educational or cultural activities, while being 'cognitively old' is associated with less social activity, feelings of negativity, lower self-esteem and a belief that others view them as old.

It is not possible to unravel from the available evidence how these issues of social ageing, cognitive age and life events interact and different studies show evidence for the existence of each phenomenon. However, the overriding message is that society, including older people, has certain expectations of behavior based on age. How we as individuals choose to react to those expectations, whether by conforming to them or circumventing them by identifying ourselves as younger than we are, may vary, but the expectations themselves appear to remain largely unchallenged and are likely to have an impact on our behavior with digital technologies.

3.6 Outcome

Low interest with the Web and virtual world environments among older people in the Gaer is caused by a lack of awareness and a lack of the necessary skills. Interestingly, older adults do not tend to view themselves as ‘disadvantaged’. This fact is critical in understanding that the elderly population is still more likely than their younger counterparts to have difficulty in using special devices and software, including those built into browsers and operating systems. Particularly among novice users, there is a tendency to view problems as due to their own lack of understanding of the technology. This rarely leads to the realization that assistive technologies are available and can make computers easier to use. As such, older users tend to struggle with unmodified mainstream technology.

The training event that highlighted the use of SL to the Gaer community had a mixed review; some of the elderly found it too much to absorb as they needed more time and more coaching (e.g., one-to-one training). The remaining participants enjoyed the experience and were keen to explore further within the pilot study that we were planning to conduct.

4 Pilot Study

The focus of the small scaled study has been to evaluate how SL can be used by older people to address social isolation, support befriending, enable wellbeing and to gain edutainment by visiting historical buildings and artifacts. The study was with two over 75 year olds (Fig 1.) from the Gaer community in Newport (the largest community of elderly people in Wales) and reveals the types of experiences they enjoyed in SL and the benefits and applications they wished to highlight from engaging with this online virtual world [26].

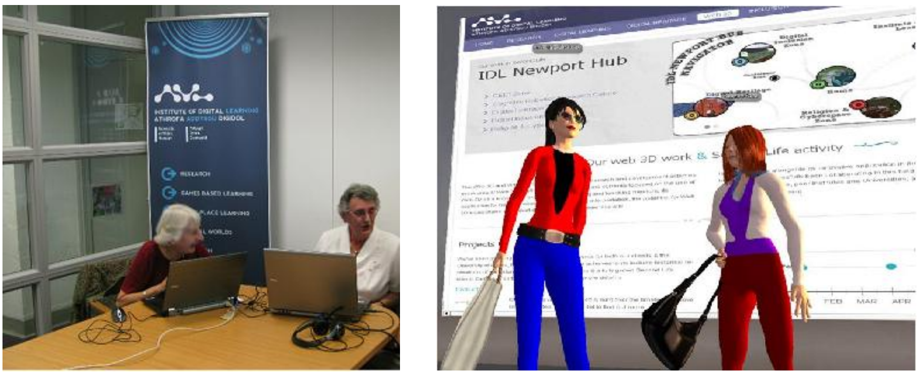


Fig. 1. The real and virtual participants using SL

4.1 Usability

In all the groups, people offered examples of products that were difficult to use, often because they did not suit the physical abilities of the user. Mobile phones, for example, were cited as difficult to use because the buttons are often shiny and too small for older people who may have limited sight and dexterity. However, it was not just modern, technological products that could be problematic, but also basic IT skills such as downloading SL and using the mouse effectively.

Time was taken to ensure participant familiarity with the navigation functions of SL. At first, participants experienced a sense of unfamiliarity with moving their avatars around but after a short while become confident and began to identify with their avatar as being part of a digital extension of themselves. As confidence grew, further time was taken in supporting the development of the 'image' of the avatar characters. The Avatars' forms of dress and apparel (in the form of handbags and shopping bags and shoes, illustrated in Fig 1.) were particularly important to participants in determining the digital identity of their character in representing themselves in the multi-user environment. In both instances, the participants developed youthful characters as preferred digital representations of themselves.

4.2 Social Interaction

The Gaer Community Centre serves as a hub for elder people to congregate and interact through shared interests such as clubs, societies and events. Ageing and mobility barriers continue to affect the extent to which real world social contact can be maintained overtime and participation in social activities is increasingly limited to day time engagements due to increasing levels of fear and anxiety caused by leaving home after dark. Previous studies identified that Skype had proved popular in supporting befriending within the community and for enriching relationships with distant relatives, however over time this Web 2 tool has only enabled limited social interaction benefit to the community [23].

The power of the community spirit is very strong in SL and offers the potential to meet so many different people [24]. The sense of presence comes when the user begins to feel that they not just watching avatars move around on the screen but they are actually one of them. The media itself becomes transparent as the person becomes, through their avatar, totally immersed in the artificial world. It is common for avatars to tour SL together and share with each other the excitement of discovery and adventure. Emotionally, they would feel the same kind of pleasures that a group of tourist might feel when exploring a tourist attraction (such as Newport Castle), by laughing and joking together in mutual enjoyment.

SL has the feel of reality and there is a sense of belonging and is ideal for cooperative group activities. It can genuinely provide a new and enriched life experience for older and disabled people, who can, through this medium, acquire a new appearance and take a vigorous and active role in the community and its many events. The ability to develop an alternative participatory representation of individuals also offers new opportunities for mental stimulation and playful social interaction. Because SL has a wide range of destinations to visit it offers elder people opportunities for collaborative exploratory virtual trips and entertainment.

4.3 Edutainment

The elder community's learning experiences in the Gaer community have taken a number of forms. A lifelong learning and participatory approach has enabled the elder community to become increasingly independent in their use of internet Web 2 tools and engagement in the digital economy through sharing experience and guidance on engaging with websites such as eBay and purchasing groceries and wider goods through wider online sources. An emphasis on digital literacy and internet safety has enabled confidence in the use of the internet. Digital Photography has also proved a popular learning programme providing a creative outlet for older members in the community. The Local History and Gardening clubs are well attended again focusing on lifelong learning in these shared interest subjects. These learning experiences are demand driven and accredited as fluid learning episodes. Their wellbeing value is regularly identified by the community workers attached to the community centre.

Virtual worlds can potentially offer solutions to many of the challenges (such as paying higher tuition fees and distance learning) faced by educational institutions. It is a common trend that learning establishments are augmenting their current curriculum with a virtual learning component or they are holding classes and entire programs exclusively in immersive learning environments in SL. The ability to utilize the interactive medium of SL also offers new opportunity for increased intergenerational lifelong learning opportunities tailored to the learning aspirations of elder community cultural interests. Opportunity is also apparent for SL to be used to address the barrier of inter-generational learning of elder citizens seeking to engage with young members of the student body in higher education in areas of shared curriculum interest.

The real world learning experience offered in the Gaer community is responsive to the communities' collective learning ambitions. It is possible using SL to extend the current provision of club based learning activities in a virtual context to dovetail with learning experienced by students at the University.

4.4 Well-Being

Within the Gaer community there is a range of existing wellbeing services offered on a face to face basis within the community. Predominately this takes the form of social support for older people which is delivered by the Gaer Community Network and the community first coordinator for the Gaer. Wellbeing is directed through social clubs and events that enable social interaction to take place during daylight hours. Popular clubs include the Gaer Computer Club, Bingo, a Luncheon Club and a generic social club. The Women's Royal Voluntary Service also offers the community a befriending service. Prime Cymru also provides social enterprise advice but there is no job seeker or employment advice for over 50s.

Fig. 2. Avatars laughing and exploring the islands in SL

SL has a unique communication environment and it differs from other Web 2 internet communication environments because it brings into play a wide range of emotions that are generally considered only possible in real life situations. People actively immersed in SL can easily become emotionally attached to another avatar. Although it is only an avatar interacting with another avatar, the emotions that can be aroused are identical to those in real life. It is very common for people to meet through their SL avatars and experience feelings of genuine friendship, even love and affection. Fig. 2 shows study participant avatars enjoying a shared joke in SL (whilst exploring the island for educational purposes), which mimics the real world experience in terms of body language and a visual reference point of shared emotion. Whilst real world provision is being accessed in the Community Centre in the Gaer these social events are not impacting on the life experiences of the vast majority of the elder community in the Gaer. This is a consequence of confidence and engagement barriers with many community residents becoming increasingly housebound in their later years.

4.5 Outcome

At first it seemed awkward for the participants when meeting other avatars. They greeted the study participants with Hello and the study group felt inexperienced in the social etiquette of the medium and experienced feelings of shyness in providing responses. Overtime confidence was forthcoming, particularly after periods of independent use of SL when meeting up in the evening and supporting each other in the communication and navigation protocols of the medium. After three weeks of support the participants were very keen to increase their SL experience in order to undertake digital trips to other parts of SL were they came into contact with other younger members of this digital community. With youthful avatars the participants experienced increased level of inter-generational communication with SL residents from around the world regularly making enquiries as to where in the world other avatars were from, what time it was there and where in SL they recommended the participants visited. Other participant led conversation focused on enquiry into what other SL residents thought about their attire with a focus on bags and shoes as self learnt conversational ice-breakers.

The study revealed that participants identified with their avatars as real rather than virtual digital beings and there was a strong sense of personification and connection with the self generated digital characters. For example, at times of immersive social interaction with other (stranger) avatars it was evident that participants on occasion experienced a sense of discomfort from unwanted interaction. This raises questions such as the need to protect future study participants from unwanted immersive engagements and supports the need for a more secure environment as a pre-requisite for further study.

Analysis identified that over time keyboard navigation to control avatars was mastered by the participants. However, to maximize the impact with wider older audiences, consideration may need to be given to the simplification of navigation interfaces through the use of larger key face assistive keyboards and or the application of joysticks or remote tools which would address potential mobility and access support needs anticipated when working with this age group. The study also indicates that some of the navigation tools in SL may be inappropriate for utilization by this age group. In particular, participants indicated feeling disorientated when attempting to fly around immersive spaces and on occasion experienced emotional responses akin to being lost with a lack of sense of place. This requires further analysis as walking was the preferred method of navigation rather than flying. This may relate to participants strong connection to avatars as real rather than virtual?

It was revealed after a period of digital trip experiences that whilst participants found many of the virtual places they visited to be stimulating the design and aesthetic experiences of these locations was often considered to be to a younger audience's taste and futuristic, thus highlighting the need for a design approach more tailored to the elder community's subcultures [25]. This is at odds with their preference for youthful looking avatars and indicates that immersive environment design is a critical factor in ensuring wellbeing and engagement in virtual spaces for this age group. This finding signposts the need for further study to investigate the impact of appearing youthful in SL in comparison to appearing as an avatar that resembles their real age to explore differences in relation to engagement and how they socialize in SL. The participants considered that SL did enable feelings of togetherness and enabled mental stimulation through exploration of SL as a virtual medium. The ease of communication at a distance and the sense of shared experience were valued. There was a strong sense of transpersonal bonding between the participants and the avatars they had created and they indicated that they somehow mapped out their self to an avatar. The experience of having a digital alter-ego was also considered to be very stimulating and fun. A more long term study would usefully evaluate the extent to which these positive outcomes could be maintained over time.

5 Aftermaths

Revisiting the participants, who were at the community training session and were involved with the pilot study, provided us with some surprising results. We were interested to hear if they were still using SL (or other digital tools) as a tool for edutainment and for the social aspect. What were the benefits, if any?

5.1 Design and Change

Since the pilot study, the interface of SL has changed dramatically and this has confused the elderly participants. With the interface change and no training; they felt uneasy again with the technology and too afraid to continue to use SL as they had before. They said that the interface was not familiar and the new design was directed towards the younger population. To support their conclusion their grandchildren had no problem adjusting to the interface change but for them it was too much and they lost interest.

This seems a common occurrence with interface and web design for the ageing population. However, there has been content designed specifically to encourage older people to use the internet, based on the use of accessible web design. For instance, the website Finerday is a social network designed to encourage older people to use it: it has a number of the functions of other networks such as Facebook, but with high contrast colors', large font and a simplified format [28]. There are also accreditation schemes for accessible web design, such as Age UKs Age OK kitemark scheme [5]. In social networking particularly, there is no evidence that older people want to network exclusively among their own age group. Indeed, there is a general lack of evidence on what types of internet content older people are more likely to be interested in.

Even real world service interfaces may have basic failings that could easily be remedied. For example, in many European cities one of the main groups eating in restaurants are those over 50, yet very few 50-year-olds are able to read a menu by candlelight without their reading glasses. That is because the menus are usually designed by young people in print shops not for senior citizens. [6]

Inclusive design offers a solution to one of the major barriers older people face; interfaces that are not designed with them in mind. Evidence suggests that 'older people neither need nor want specially designed products', but everyday products can be redesigned to suit the needs of a wider range of people with benefits for all users.

The core recommendations for design typically include for SL would be to:

- use of appropriately designed text in terms of large font sizes, and appropriate color options
- design of uncluttered web-page content with clear headings
- use of clear and consistent page navigation within a site
- use of large and clearly labeled buttons, graphics, and links

Few websites and/or virtual worlds, however, have been designed with these guidelines in mind.

5.2 Ageism

Ageism and age discrimination is the stereotyping and discrimination of someone (or a group of people) due to their age. It is not a new phenomenon, but has been increasingly studied over recent years. Research by social psychologists at the University of Kent has established that older people are stereotyped as friendly but incompetent, or 'doddery but dear' (whereas younger people are stereotyped as cold but competent [13]). These stereotypes are associated with emotions such as pity towards people in later life and are likely to feed in two types of behavior that end up causing discriminatory outcomes, such as older people being denied opportunities or services because others assume they don't want or need them [14].

5.3 Unsecure Environment

The main disadvantage of using SL from the elderly point of view (from outcomes of the pilot study and the community training) was the uncertainty of the virtual world. They did not like talking or following an avatar who they did not know, they would prefer if they knew the avatar by their real name or they could recognize them (e.g., facial features). Once again it is relating to the real world, where they would not approach a stranger and start talking to them or follow them to a café or museum.

5.4 Edutainment

One of the main components within SL is to enable avatars to talk with other avatars from anywhere in the world. Feedback from this study and the community training provided a surprising finding. Social interaction was not a priority of the group; but instead they were more interested in the educational aspects of the virtual world; visiting islands that had some historical interest value to them. They wanted to explore areas of historical interest and gain knowledge behind these historical sites. It wasn't just historical islands they liked to explore but it was visiting big cities such as Paris or New York. The reason for exploring big cities and/or the historical buildings was that in the real world visiting these places would be unpractical; such as the cost of visiting these places and problems with mobilability (e.g., wheelchair access).

Edutainment could be explored further, whereby the elderly could visit virtual health centers in big cities to find out more on for example, ageing and/or on their own disability. This area has great potential for further studies in developing edutainment for the elderly.

5.5 Further Research

Further development would give the Gaer community the facilities required to help itself, using the tools that are needed to build a more collaborative life, whilst at the same time providing each beneficiary direct access to new technology. To achieve this, would be a need to create a virtual online Gaer, which replicates key cultural and educational dimensions of the real-world Gaer environment, containing the community centre, clubs and societies it would be available for individuals to visit whenever they want. The virtual online Gaer will be completely secure, and accessible only to those individuals who are taking part in the study. All services that are offered in the real-world would be mirrored in the virtual world. Importantly, this approach would also enable social participation to continue during times of darkness when older people in the Gaer community have identified as being the time they most fear leaving their homes. The design of the online Gaer environment in SL would also take account of the design and aesthetic preferences of this age group.

6 Conclusion

More research is required for more comprehensive studies into the factors causing digital exclusion among older people, using surveys with larger sample sizes and qualitative research. The aim should be to develop a more detailed understanding of the factors influencing older people specifically. More research is needed to conduct more comprehensive studies into the influence of internet content on digital exclusion among older people. These should explore what types of content are likely to encourage older people to use the internet and the virtual world.

Website providers and virtual world's developers, including the public, private and voluntary sectors, should assess whether their content meets standards of accessibility required by many older people, and where necessary they should take steps to ensure this is the case.

However, we have considerably less information about the values, attitudes and aspirations of older people; how these are formed; whether they are affected by life transitions (such as retirement, bereavement or deteriorations in health); to what extent they are related to life stage or cohort or other influences; and how these influences might affect how they use new digital technologies and interact with virtual worlds. There is evidence that older people are less likely than their younger counterparts to engage in virtual world/gaming that has significant potential benefits, such as social interaction, well-being and edutainment. This coupled with the evidence that there are significant numbers of older people not enjoying the level of social participation they desire.

Within the context of the potential identified for next generational digital services and their implications for older and disabled people, there is a strong parallel to the call from the research community for priority to be given to formulating digital games to meet the wellbeing and entertainment needs of elderly people. It is apparent that multi-user virtual environments have also not been fully explored as a potential positive technology that can be utilized as a support tool in addressing social isolation, edutainment and support for elderly users in the information society. It is intended for further research to focus on extending existing real world social activities for older people into a tailored virtual context that will assist in determining the suitability of such approaches and identify effective methodologies to apply SL in a domestic context to meet the wellbeing and active living challenges faced by our ageing population.

Acknowledgments. Thanks are extended to all of the E-Inclusion Recycling Community Ms Julie Traynor and the Gaer Community Network as key pilot study partners; for their time and views relating to this research. Finally, we very much appreciate the time and commitment of the two lovely ladies from the Gaer who volunteered to be in the pilot study; Joan Rees and Barbara Marsh; thank you.

References

1. Räsänen, P.: Information Society For All? Structural Characteristics of Internet Use in 15 European Countries. European Societies 8(1), 59–81 (2006)
2. Wong, Y., Fung, J., Law, C., Lam, J., Lee, V.: Tackling the digital divide. British Journal of Social Work 39, 754–767 (2009)
3. Reisenwitz, T., Iyer, R., Kuhlmeier, D., Eastman, J.: The elderly's internet usage: an updated look. Journal of Consumer Marketing 24(7), 406–418 (2007)
4. Hannon, C., Bradwell, P.: Web I'm 64: Ageing, the internet and digital inclusion. Demos (2007)
5. Roberts, S.: The Fictions, Facts and Future of Older People and Technology, International Longevity Centre UK (2010)
6. Dixon, P.: Marketing to Older Consumers', interview for Marketeer (October 2008), `http://www.globalchange.com/marketing-to-older-consumers.htm`
7. Sudbury, L., Simcock, P.: Understanding Older Consumers through Cognitive Age and the List of Values: A UK-based Perspective. Psychology Marketing, 22–38 (2009)
8. Jopling, K.: Future Communities. Help the Aged, London (2008)
9. Giddens, A., Griffiths, S.: Sociology, p. 18. Polity Press, London (2006)
10. Fitzgerald, T.: The Elderly at the Millennium: Courting the Older Consumer. Ageing International 25(1), 72–86 (1999)
11. Stephens, N.: Cognitive Age: A Useful Concept for Advertising? Journal of Advertising 20(4), 37–48 (1991)
12. Leach, R.: Boomers and Beyond: Intergenerational Consumption and the Mature Imagination (2005), `http://www.consume.bbk.ac.uk/research/biggs.html`
13. Policy Unit, Age Concern England, Ageism in Britain 2006. Age Concern, London (2008)
14. Abrams, D., Eilola, T., Swift, H.: Attitudes to Age in Britain 2004–2008, London: Department for Work and Pensions. Research Report No. 599, 148 (2009)
15. Watts, T.: Major Survey Reveals Why Older People Are Increasingly on the Wrong Side of the Digital Divide (2009), `http://www.maturetimes.co.uk/digital-divide-survey-oxford-institute`
16. Age Concern and Help the Aged, Introducing Another World: Older People and Digital Inclusion. Age Concern and Help the Aged, London (2009)
17. Stroud, D., Batchelor, A.: Meet the Charmed Generation. Journal of Direct, Data and Digital Marketing Practice 10(43), 46–47 (2008)
18. Craik, F.I.M., Salthouse, T.A.: The handbook of aging and cognition (Revised Edition), 2nd edn. Erlbaum, Mahwah (2000)
19. Arjan, R., Pfeil, U., Zaphiris, P.: Age differences in online social networking. In: CHI 2008 Extended Abstracts on Human Factors in Computing Systems, pp. 2739–2744 (2008)
20. Hawthorn, D.: Training wheels for older users. In: Proceedings of OZCHI, pp. 23–25 (2005)

21. Sa-nga-ngam, P., Kurniawan, S.H.: An investigation of older persons' browser usage. In: Stephanidis, C. (ed.) HCI 2007. LNCS, vol. 4554, pp. 1000–1009. Springer, Heidelberg (2007)
22. Sivan, Y.: 3D3C Real Virtual Worlds Defined: The Immense Potential of Merging 3D, Community, Creation, and Commerce. J. Virtual Worlds Res. 1, 1–31 (2008)
23. Chilcott, M., Hartwig, J.: Newport Digital: Community Empowerment through Digital Inclusion Study. In: Proceedings of the Newport Nexus Conference, pp. 101–115 (2009)
24. Riva, G., Mantovani, F., Gaggioli, A.: Presence and rehabilitation: toward second-generation virtual reality applications in neuropsychology. Journal of NeuroEngineering and Rehabilitation 1(1), 1–9 (2004)
25. Boellstorff, T.: Coming of Age in Second Life, pp. 108–146. Princeton University Press (2008)
26. Chilcott, M., Smith, A.: Ageing Well and Learning through Online Immersive Participation Using a Multi-user Web 3D Environment. In: Third International Conference on Games and Virtual Worlds for Serious Applications, pp. 70–75 (2011)
27. Sinclair, D.: The Golden Economy: The Consumer market Place in an Ageing Society, International Longevity Centre UK (2010)
28. Berry, R.: Older People and the Internet towards a system map of digital exclusion, The International Longevity Centre-UK (2011)

Course Lectures as Problem-Based Learning Interventions in Virtual Worlds

Spyros Vosinakis[1], Panayiotis Koutsabasis[1], and Panagiotis Zaharias[2]

[1] University of the Aegean, Dept. of Product and Systems Design Engineering, Syros, Greece
[2] Open University of Cyprus, Dept. of Information Systems, Nicosia, Cyprus
{spyrosv,kgp}@aegean.gr

Abstract. Virtual Worlds (VWs) present considerable potential as future learning platforms, but further studies are required to assess their effectiveness in constructivist and collaborative learning situations. The paper investigates the suitability of VWs as a platform for hosting PBL (Problem-Based Learning) activities and explores their affordances in terms of collaboration support and learning effectiveness. We have designed an educational VW and developed a number of tools that support collaborative learning activities. Using this environment, we have conducted a PBL intervention that required from students to collaboratively design the user interface of a multimedia kiosk. We performed a thorough, formative, multi-method evaluation of the learning activity. The results reveal several encouraging findings about PBL and collaboration mediated by VWs, and lead to a series of recommendations.

Keywords: Virtual Worlds, Problem-based Learning, CSCL, User Interface Design.

1 Introduction

Virtual Worlds (VWs) are computer-generated 3D environments, in which multiple users navigate, interact and communicate having a form of embodied representation [1]. Given that this fairly new medium has significant differences and introduces novel affordances compared to traditional learning paradigms, researchers attempted to study its effects on the learning outcomes, and the circumstances and preconditions under which this new medium is to be used [2-4]. A number of prototypes and case studies have been setup in order to draw results on the use of VWs in education in the last two decades [4-7], starting from the early immersive VR systems to the current massive multi-user worlds. Although these studies vary in terms of configuration and types of educational activities tested within the 3D environment, the early results have shown that VWs have significant potential as a complementary educational medium [8,9].

Nowadays, a number of universities and high-schools are using VWs as part of their curricula.[1] However, the majority of them employ VWs simply for resource

[1] http://virtualworldwatch.net

Z. Pan et al. (Eds.): Transactions on Edutainment IX, LNCS 7544, pp. 81–96, 2013.
© Springer-Verlag Berlin Heidelberg 2013

sharing and conferencing, and the common activities that take place within the environment are text or voice-based communication, document storage and exchange, group discussions and presentations, e.g. [10-12]. These approaches do not exploit the powerful affordances of VW in presenting real-time simulations of custom environments, in which users can actively participate in an experiential and constructivist manner. We argue that VWs should be explored for setting up novel educational interventions that support and visualize evolving in-world activities with the presence and participation of people who construct and manipulate 3D objects and tools.

A learning approach that follows the principle of active and collaborative knowledge construction is Problem-based Learning (PBL). In PBL students learn by addressing authentic and open-ended problems and reflecting on their experiences, thus developing problem-solving strategies and building domain knowledge in a self-directed manner [13]. This approach has several advantages, as students are actively gaining transferable skills by investigating, explaining and resolving meaningful problems and the individual or group participation in problem-solving activities is highly motivating for them. However, there are only a few documented cases in which VWs have been used for collaborative PBL activities.

The aim of our work is to explore meaningful ways for the facilitation of collaborative PBL activities in VWs. We present the design and evaluation of a PBL intervention in a user interface design lecture. Our goal was to engage students in PBL activities through their collaborative design, experimentation and evaluation of user interface prototypes. We designed an educational environment on top of an existing VW platform, built a number of supporting tools for collaboration and prototyping, and facilitated a set of learning activities over an extensive lecture session. We performed a detailed evaluation of the activities, in which we combined various methods for data collection and analysis in order to explore dimensions of collaboration, learning and usability. The evaluation results reveal several strengths and weaknesses of VWs as potential PBL platforms.

2 Related Work

2.1 Problem-Based Learning

PBL is a learning and teaching approach that has been widely adopted in practice during the last 20 years in both traditional and online educational settings [14]. It incubates an experiential, social and active perspective to learning that contributes to the development of critical thinking skills. Typical PBL contexts require learners to work in small groups to investigate a real-life problem. The instructor acts as a facilitator of group work challenging students' learning and collaboration strategies, and may occasionally provide resources to help learners find a solution to the problem, despite that students are primarily responsible for finding their learning resources.

The collaboration and interaction of motivated students groups and the effective teacher facilitation of PBL activities are critical factors for the success of this approach in both traditional and online learning settings. With respect to online learning settings, several studies have attempted to implement PBL. Donnelly [15] integrated online PBL techniques with face to face learning. It was found that group activities were the most important parameter for the success of the process. Dennis [16] compared a face to face with online PBL settings. Results showed that the groups did not differ in learning performance but the online group spent more time for learning activities. Ozdemir [17] reported that students in a collaborative PBL environment outperformed those who were in an individual PBL environment as regards the development of critical thinking. In [18] it is suggested that only online courses with collaboration characteristics should make use of PBL.

On-line approaches to PBL may have positively assisted the processes, but there are still issues that need further improvement. Constructivist learning environments, according to Wilson [19] should contain "*a setting or a space wherein the learner acts using tools and devices, collecting and interpreting information, interacting perhaps with others, etc*". A more experiential and constructivist approach to PBL would need to include rich multimodal representations of the problem space and the ability to freely experiment with real-time interactive tools to construct possible problem solutions. VWs are a platform that can support such enhancements.

2.2 Designing PBL Interventions in Virtual Worlds

PBL has already been characterized as one of the most appropriate learning pedagogies in VWs, especially in Second Life [20]. Despite the large amount of work in combining PBL with online learning there are few studies of PBL in VWs. In [21] Second Life[2] (SL) is used to implement a collaborative PBL activity. The tutor intervened in the process by using scaffolding techniques in order to help students to achieve their goals. It has been reported elsewhere [22] that the use of scaffolding techniques is effective when collaborative learning occurs in a VW. In [23], the suitability of SL for PBL is demonstrated through the mapping of learning activities on to PBL goals as they have been stated in the framework of [13]. They used a machinima-based assessment technique where students worked in groups to create short video clips from their activities in SL. According to the findings it is supported that development of wider transferable skills can be realized effectively through VWs such as SL. In [24] findings from a case study with a PBL approach are reported, where students were tasked to create learning experiences within SL for external clients. The emphasis was mainly placed on the process of how students formed groups and created the interactive learning experiences by using the affordances of SL. It was found that SL can contribute to PBL as a pedagogical approach in several ways such as supporting the roles of tutors and students, facilitating their relationships, enhancing students' motivation and ownership of the project, as well as easing the assessment activities by the tutors. Similar results can be found in the study of [25], which refers to enquiry-based learning.

These studies are presenting the experiences of using VWs with a PBL pedagogy, however they do not yet propose an approach for the design of specific VW tools and PBL interventions in VWs or present a specific approach for the assessment of these interventions. Much of the relevant research on PBL is still conceptual and

[2] `http://secondlife.com`

information about detailed evaluation with regard to specific methods and practices is lacking.

2.3 Evaluation of PBL Activities in Virtual Worlds and Computer-Supported Collaborative Learning

The evaluation of PBL involves the intertwined dimensions of collaboration, interaction and learning, placing emphasis not only in the learning content but also in the assessment of more general skills like self-directed learning, intrinsic motivation and critical thinking. When PBL occurs in the classroom, various types of formative and summative assessment tools and methods are used encoded in complex assessment rubrics [26]. Despite that there is much work on the assessment of PBL in computer supported collaborative learning (CSCL) situations, the assessment of PBL activities in VWs is currently an open issue: there are too few studies of PBL in VWs and they have not yet developed specific tools for that purpose.

When a problem-based approach is employed in CSCL, the evaluation involves interaction analysis of the participating teams in order to clarify what types of collaborative interactions have occurred and what educational benefits have taken place [27, 28]. Interaction analysis is both a qualitative and quantitative process that can be performed with various instruments and methods that must also take into account the specific problem at hand. In [29] the evaluation of collaborative learning is described as "*placing strong emphasis on the situated nature of collaboration and the impact of certain situational factors (with a few or as little as possible a priori expectations)*". In [30] a principled framework for the study and analysis of group interaction and scaffolding is presented by combining different aspects and issues of collaboration, learning and evaluation. The evaluation happens with the qualitative inquiry of indicators about 'task performance', 'group functioning', 'social support' and 'help services'. The approach has been applied in e-learning course situations with large numbers of participants, and it can be extended to cover the particular issues of PBL interventions in single lecture situations. The interaction analysis framework of [31] is a tested approach which is appropriate for assessment of collaborative learning in the long-term on top of robust collaborative technology.

2.4 The Scope of Our Research

The aim of our work is to explore meaningful ways for the set up and facilitation of collaborative PBL activities in VWs. In our study, we describe an approach for the design and evaluation of PBL interventions in VWs that reports on the value of VWs as educational platforms for constructivist and experiential learning. The goals of this study were to: (a) design a learning intervention for a lecture of user interface design that includes a number of PBL activities; (b) facilitate the learning process, while keeping track of students' behaviour and performance; and (c) evaluate the learning process with criteria that stem out of the PBL philosophy, as well as the final outcome.

This work can contribute to the current corpus of studies that aim to discover the extent in which VWs can support constructivist activities, to evaluate the learning results of their use, and to unveil critical problems related to student collaboration and learning.

3 The VW Environment and Tools

3.1 Configuration of the Virtual World with Open Source Software

The VW implementation (Fig. 1) has been based entirely on open source software. The world server was installed in a standalone PC using the OpenSimulator platform,[3] and the FreeSwitch server[4] has been set up and connected to the environment to provide voice communication support. We have created a small island and built a number of interior and exterior places for group collaboration and whole class activities. We implemented four additional collaboration tools in the LSL Scripting language. On the client side, the Hippo OpenSim Viewer[5] was running on PCs with standard keyboard and mouse equipment plus an additional headset with microphone for voice communication. No significant decrease in efficiency or loss of quality has been detected during the whole experiment.

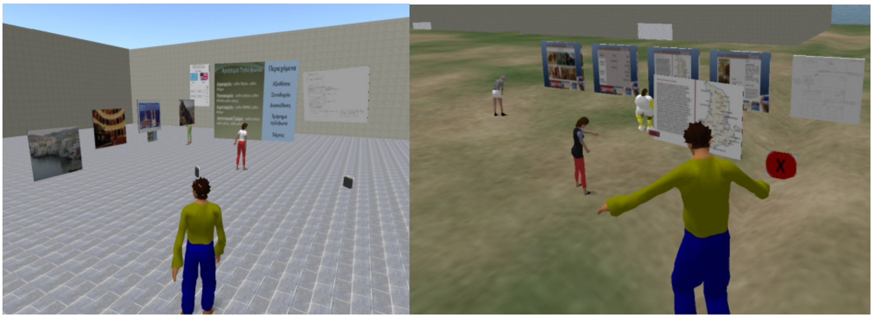


Fig. 1. Screenshots of the environment

3.2 Design and Implementation of the Supporting Tools

Group educational activities in VWs have strong requirements concerning the communication and collaboration of remote students and teachers. For our study in the area of user interface design we have identified the following tasks in group-based PBL activities:

- In the early stages, students discuss about the problem, write down facts and reveal aspects for which further knowledge may be required.
- Then, they assign roles to group members, search for and share resources, and formulate, present and explain their ideas.
- Then, they collaboratively assemble a final solution and gradually refine it.
- Finally they present it to the class for further evaluation and feedback.

To support these tasks, the educational environment should provide the appropriate means for text and voice chat, and allow students generate and share public and

[3] http://www.opensimulator.org
[4] http://www.freeswitch.org
[5] http://mjm-labs.com/viewer

private documents to exchange ideas and coordinate their activities. Also, students should be able to easily take notes during group discussions and chat sessions and share them with others. Groups should be able to collect and organize their common resources within the VW in order to use them for reference during their problem solving activities. Finally, the world should contain the building blocks to collaboratively construct a working prototype and to enhance it with further explanations about the design choices.

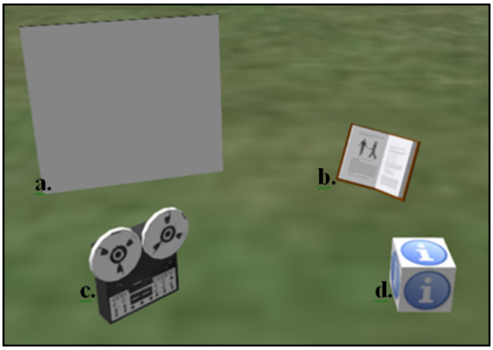


Fig. 2. The supporting tools of the environment. a. Interface Element, b. Resource, c. Comment Recorder, d. Annotation.

The platform employed for our study provided inherent support only for part of these tasks allowing for text and voice chat between users, but there is no support for offline messaging. Concerning resource management, one can only add geometric objects to the environment, whilst any other document type can only be placed in an object's contents. However, security reasons restrict other users besides the object owner to view its contents. Finally the construction of a working user interface prototype requires a lot of programming effort using the VW's scripting language, which was outside the scope of our learning session.

Thus, we have designed and implemented a number of additional tools that were available to students during the study in order to overcome these obstacles and to enhance the collaboration affordances of the environment. The implemented tools were:

- *Resource*: an object that links to external web resources
- *Comment Recorder*: a tool to record and playback user messages
- *Annotation*: an object that contains a written message
- *InterfaceElement*: an object with scripted behavior that can be used as a user interface component in the working prototype

The aforementioned objects were provided to each student on initialization and they could insert multiple copies of them inside the environment (Fig.2).

The "Resource" object is the equivalent of a hyperlink. It has the form of an open book and it opens a web resource in the default browser upon mouse click. It can be used by the teacher(s) in order to provide some initial resources to the students (guidelines, design patterns, templates, etc.) to aid them during their tasks, and by the student groups in order to share and organize the resources they found in their self-directed learning activities.

The "Annotation" object allows students to post annotations within the environment. If any user clicks on it, a text message is opened and may be saved in his/her inventory for later use. In the context of the PBL activities, annotations may be used for the asynchronous collaboration between group members (e.g. in the form of comments, notes about things to be done, role descriptions, etc.) or they may be attached to the user interface prototype as further notes or explanations of design choices.

The "Comment Recorder" object can record and playback user messages on demand by sending special commands to the text communication channel. It can be used to take notes from conversations during the early collaboration stages and also as a tool to record viewer comments during the final evaluation stage.

Finally, the "Interface Element" object is the principal design element of the user interface prototype. Its look and behavior can be configured using a simple set of commands and, depending on its configuration it can:

- show or hide in the environment as a result of an external event,
- operate as a button that will generate a batch of events if pressed, which may affect the status of other interface elements or itself
- contain a number of images (faces) that may change dynamically its appearance as a result of an external event, and
- operate as a container that forwards events to its contents

Students can combine and configure copies of the "Interface Element" object in order to design buttons, windows and image containers during the final stages of the learning activity and collaboratively construct an interactive user interface prototype.

4 Facilitation of the Learning Intervention

4.1 Course Lecture and Participants

The learning intervention was offered as an introductory, optional joint lecture for the courses of Advanced User Interfaces and Virtual Reality at the department of -omitted for anonymity-. The participants of the study were ten students (3 male, 7 female). All participants had attended a number of related lectures like HCI (Human-Computer Interaction), Interaction Design and Multimedia Design. They were all proficient computer users, but only three of them had some experiences with VWs.

The participants were allocated in three balanced groups in terms of their experience in VWs and their user interface design skills. The team mates had to perform all activities through the VW, to simulate remote collaborating conditions. Their seating positions had a certain distance from each other, and they were not allowed to communicate face-to-face during the activities. The teaching team comprised of the authors themselves, who also provided technical support and facilitated the collaboration and learning activities during the whole intervention.

4.2 PBL Activity and Goals

The lecture was organized around an authentic, ill-defined problem according to the tenets of PBL, which was given to the students in the following statement of a ‘design brief’: “*Design the user interface of a multimedia kiosk system for browsing available rooms to let in the island of Syros. The intended users are tourists (Greeks and foreigners), who can access the system from the harbour of Syros. You should take into account usability guidelines for multimedia presentations and information seeking. You should design the 5-7 most basic screens of the system, in wireframes*”. In addition, the participants were presented with an abstract work plan that included tasks that they could choose to follow with indicative times for completion.

Fig. 3. Group presentation inside the VW

The learning goals of this learning intervention were: a) to discover the usability and accessibility requirements of touch screen interfaces, b) to understand the differences in the design of such interfaces compared to other, more conventional cases, and c) to apply this knowledge in a specific practical context. Following the principles of the PBL approach, the learning session has been applied as follows:

1. The students were given an introductory session in the VW to familiarize with the interface.
2. The supporting tools were presented to the students accompanied by specific use cases.
3. Students worked in groups inside their allocated workspaces. They analyzed the problem, shared ideas and gathered resources.
4. Each group assigned roles and/or tasks to its members. They proposed and argued about concepts, designed the appearance of the user interface elements using in-world and external tools, and collaboratively constructed their prototype as a proposed solution
5. Once the group agreed on the final prototype, they attached explanatory annotations to justify their design choices and presented it to the whole class (Fig.3)
6. Students and teachers were then free to test each interface prototype themselves and leave comments and suggestions concerning the appropriateness of the solution.

5 Evaluation

5.1 Data Collection and Analysis Method

We have constructed a mixed (qualitative & quantitative) method for interaction analysis of problem-based CSCL in VWs. More specifically, we used the following methods:

- Automated monitoring of student behaviour: this was achieved by video capturing of the activity within the VW, logfile analysis with respect to the use of the tools, and observation of the state of the world during and after the exercise.
- Dialogue analysis: voice chat was recorded for most of the exercise and an analysis of utterances was performed. We followed the taxonomy of [32] who classify utterances in one of the following content categories: Procedure, task status, reference, internal state and acknowledgement.
- Students' self-reporting: we used a questionnaire that investigated several aspects of the problem-based CSCL experience, as well as follow-up discussion.
- Tutors' evaluation of the outcome: this was performed during the activity and also after the experiment taking into account all data gathered.

The method for interaction analysis explores the dimensions: 'task performance', 'group functioning', 'social support', and 'learning performance and outcome'. The first three dimensions are those proposed by [31], while the fourth dimension was added to investigate issues of particular PBL process. Thus, the evaluation method employs a number of established dimensions and indicators for interaction analysis in CSCL and extends these to support evaluation of PBL situations. A large corpus of data can be collected for each indicator involving at least two related methods, allowing for cross-examination of the results. Table 1 illustrates the dimensions and indicators of interaction analysis performed for this exploratory study and the corresponding data collection methods.

5.2 Results

The overall result of the learning interventions was that all student groups proved capable of constructing functional user interface prototypes using the in-world tools as well as of instantly testing and evaluating their solutions. The prototypes were particularly interesting, while all learning activities were conducted in an engaging, enjoyable and satisfactory manner.

The learning intervention lasted for a total time of 6.5 hours, which was about 1.5 hour more than initially estimated. The first 2 hours were devoted to the tutorial about the use of the VW. Then, a total of 3.5 hours were devoted to the activity of user interface design, presentations and follow-up, while a total of 1 hour was allocated to the breaks. Participants were asked how much time they would need to carry out the user interface task in a 'face to face' situation and deliver at the same quality: some of them answered about the same time (3.5 hours), others said about an hour less. This is a quite interesting result considering other time consuming activities in face to face situations like for example time arrangements. All teams made use of the tools provided in the environment to document on the design process (Table 2).

Table 1. Dimensions and indicators of interaction analysis and corresponding data collection methods

Interaction Analysis Indicators	Action monitoring	Dialogue analysis	Self-reporting (question-naires)	Post evaluation (observation, follow-up)
Task performance				
TP1. Problem-solving capabilities and learning outcomes	X	X		X
TP2. Contributing behaviour during tasks	X	X		X
TP3. Performance in terms of self-evaluation		X	X	
Group functioning				
GF1. Active participation behaviour	X	X		X
GF2. Social grounding	X	X	X	X
GF3. Skills that monitor and facilitate the group's well-being		X		X
GF4. Group processing		X	X	X
Social support				
SS1. Commitment toward accomplishment of the common goal	X	X	X	X
SS2. Level of peer involvement	X	X	X	X
SS3. Achievement of mutual trust		X	X	
SS4. Motivational and emotional support to peers		X	X	
SS5. Conflict resolution		X	X	X
Learning performance and outcome				
LPO1. Flexible knowledge about the problem at hand	X	X	X	X
LPO2. Effective problem-solving skills		X	X	X
LPO3. Self-directed learning skills		X	X	X
LPO4. Intrinsic motivation		X	X	X

Table 2. What tools of the VW contributed to the development of your knowledge about the problem? (Bad 1 2 3 4 5 6 7 8 9 10 Excellent)

	Average	Median	Mode	St.Dev.
Resources	5,5	7	7	2,9
Annotations	6	6	6	1,6
Comment listener	4,1	5,5	0	3,7
Interactive objects	6,7	7	7	1,4
Chat (text)	7,5	8	9	1,9
Voice Chat	9	10	10	0,5

5.2.1 Task Performance

With respect to the problem-solving capabilities demonstrated and related actions taken, we observed that students devoted a large portion of their available time to discuss about the understanding of the design problem. These were intertwined with intervals of self-directed learning, which occurred either from 'assignments' or 'requests' by other team mates (e.g. "will you find photos and content about hotels?") or from individual initiative (e.g. "I can find some text to write about Syros history"). This was also identified by the dialogue analysis (Fig. 4): most of discussion was about the procedure and task coordination (38.7%) and acknowledgements (24.0%), while less time was devoted to discuss about the task status (6.2%) and to refer to virtual objects and tools (11.0%).

Self-evaluation of individual and group performance was quite similar for all participants. The average self-rating of their individual performance regarding the use of the system was: 7 (1: Bad – 10: Excellent) (st.dev.: 1.1). That was pretty much their rating about their team's performance, i.e.: an average of: 7.1 (st.dev.: 1.3). Their responses varied more, when they were asked about their performance with respect to the task of user interface design: they rated their individual performance with an average of 6 (st.dev.: 1.8) and their team's performance with an average of 5.9 (st.dev.: 1.9). Given that we closely observed the process, we consider these as rather misbalanced self-assessments: in fact, students faced many difficulties in using the VW, and the fact that they finally achieved to make use of the tools encouraged them to rate their performance rather highly than appropriate. On the other hand, the final outcome of the process was interesting from many aspects. Students have underestimated their performance in this respect because they needed more time for improvements.

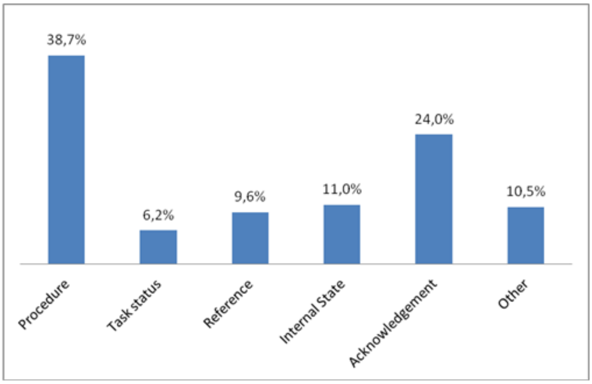


Fig. 4. Types of utterances

5.2.2 Group Functioning

With regard to active participation, we observed (mainly from dialogue analysis and self-reporting) that all participants were actively involved in the collaboration and conversations, especially in the first phases of the collaborative activity. However, there were 2 students that gradually decreased their contributions. They reported that the experiment lasted too long, and they were tired by the process, despite that they enjoyed the experience overall.

All teams exhibited active interaction skills with respect to monitoring the progress of group work; this was evident especially from dialogue analysis: a rather large number of utterances were questions about how to proceed with the activity and specific tasks (14.1%), while there were also a large number of acknowledgements (24.0%) of group work. Each team used a different style of coordination of the work. Team 1 demonstrated a totally balanced coordination scheme without someone taking up a leading role. The other two teams quickly established a leader (in both cases the person who had more experience with the use of the VW) and allocated roles during the collaboration: the main roles were those of the 'visual designer', who also sketched the layout of the screens and the 'content designer' who located and edited content (mainly images and text). All participants reported that the result of their work was a collaborative product and that the environment contributed to their collaboration.

5.2.3 Social Support

All participants and teams performed very well with respect to social support: they were all highly motivated students, who were acquainted to each other. More specifically, they rated their commitment towards the achievement of their goal at an average of 7.5 (1: Poor – 10: Excellent), and the main reason for this rating not being higher was that some of the participants got carried away out of their curiosity to explore the VW! When they had to wait for other team mates, they kept exploring the world in a playful manner.

Regarding conflict resolution, we identified that there were several disagreements during the learning task about aspects of the design. This was natural since that all participants were mature design students and had their own different opinions on the design; however, these were openly expressed, discussed and quickly resolved. This is a positive finding: the collaborating participants in the VW seem to be encouraged to individually contribute to group work as well as to constructively resolve conflicts that may arise.

5.2.4 Learning Performance

Regarding the learning performance and outcomes, the main result was that all three teams achieved the goal of the exercise (i.e. to provide the design of the user interface of an information kiosk), at a fairly satisfactory level. All teams demonstrated interesting designs that took related guidelines and content into account. However they all reported that they would need more time to elaborate their design solutions.

The participants reported that they gradually developed their knowledge about the activity at hand to a considerable extent (an average of 6; 1: Bad – 10: Excellent; st. dev.: 1.2). They also reported that they devoted about half of the time in self-directed learning: an average of 4.4 (1: None – 10: All; st.dev.:2.3). Also, when asked in which situations they best contributed to the team as with respect to whether they followed the agreed plan, they admitted that their contributions were more suitable when they stick to the plan (an average of 7.2; 1: Alone – 10: "I stick to the plan"; st. dev.:1.9).

Regarding the issue of developing problem-solving skills, students first reported on a number of problems faced: most students reported difficulties in using the VW, and a few found it difficult to document their design choices and their opinions with some of the tools provided. Then they reported on their ability to overcome these, an average of 5.2 (1: Bad – 10: Excellent; std. dev.: 2.4). The main reason for not performing better in this respect was that they had limited experience with previous use of VWs. However, we note that the final outcome of the activity, i.e. the user interface design was quite satisfactory for all teams.

6 Discussion

The experience had several positive aspects. The fact that users had a shared persistent workspace was perhaps their most important and recognized advantage of the system. All students reported that they felt engaged and motivated to work towards their common goal. They highlighted the fact that they could easily log off or postpone some of their activities in the world (especially when they performed self-directed learning activities) and they were able to see their other mates' progress when returning to the world. Discussions about the problem and tasks were easy to carry out in the VW, since they all had their own material uploaded on the shared space. Also, it was natural to compare design ideas and comment on others' work. Students arranged user interface screens in some logical order and organized discussion sessions for each screen as well as all-together. The environment was fun and kept them occupied all the time, even at times when they had to wait for other team mates to deliver their parts of the work. Finally, the increased awareness of others' work and activity and progress was also reported as a positive aspect of the environment, mainly as a motivating factor to one's own work.

On the other hand, a number of problems and drawbacks have been identified. First of all, collaboration without voice seems to be a problem. There were a couple of situations when only text chat was available (due to temporary problems of the voice server), and participants felt quite restricted in their communication. Also, a few users reported that their attention was more on the difficulties of using the environment, especially in the beginning, rather than on the user interface design task. The teams discussed and planned their activities, but they did not manage to keep track of all their coordination decisions. Despite that there were available tools in this respect, some of them found it hard to use them. Some students did not like the fact that the roles of each participant were not visible by their avatars and they also wished for more '2D functions', e.g. the possibility to embed applications from their desktop environment to the VW. Finally, perhaps the most important problem for this study was the lack of familiarity with the environment. Students faced several problems during the use of the environment and often asked how to perform certain functions. They felt that if they were more familiar the final result would be much better.

The main recommendations made by the students are the following:

1. Shared whiteboard for sketching. Sketching is an important tool for collaborative design activities and was not supported by the VW. Some participants sketched on paper and scanned their drawings, while others used

external drawing applications. The sketches were then uploaded as images in the world.

2. Tools for organization and coordination of team work. Some form of a shared agenda as well as the direct visibility of roles (e.g. as part of their appearance) would be an asset.
3. Tools for collaborative writing in the world. Shared boards for presenting and editing notes and comments would also be helpful tools during design activities.
4. More privacy. Some users reported that they needed to chat directly to their team mates, without others hearing and interfering with their discussion.

7 Conclusions

In this paper we presented the design, facilitation and evaluation of a problem-based learning activity in an open source VW platform that took place in the form of an academic lecture. The current state of the art includes too few studies of PBL in VWs with work that is still conceptual, while information about detailed evaluation with regard to specific methods and practices is lacking.

The application of PBL in VWs revealed quite encouraging results. The learning session managed to capture the attention of students, to trigger self-directed learning activities, and to foster collaboration and discourse between them. The study also highlighted a number of problems that were mostly related to activity awareness, resource sharing and coordination issues. Some of these obstacles are based on inherent deficiencies of the specific platform that we used and have been tackled in other multi-user virtual environments and in own current work (e.g. the use of a shared whiteboard); while others are still open issues. Further research is needed towards the design and evaluation of novel metaphors, tools and paradigms for student collaboration in learning activities in order to overcome these difficulties and to improve the effectiveness of VWs as learning environments.

As to the nature of the issues explored, we are applying and refining our approach in other courses and contexts that involve more VW tools, longer assessments, more student groups and different problems [33,34]. We are also working to address several of the issues identified, mainly those related to creating a more realistic context of the collaborative work situation: mainly with respect to remote collaboration work (i.e. connection to the VW from the pragmatic work environment) and real projects that will also include customers. The virtual environment presented can support real interactions and collaborative work situations and can contribute to effective constructivist learning, provided that a couple of other relevant tools are built-in and some privacy issues are addressed.

References

1. Bartle, R.A.: Designing Virtual Worlds. New Riders, USA (2003)
2. Hedberg, J., Alexander, S.: Virtual Reality in Education: Defining Researchable Issues. Educational Media International 31, 214–220 (1994)

3. Cheal, C.: Second Life: hype or hyperlearning? On the Horizon 15, 204–210 (2007)
4. Dickey, M.: Brave new (interactive) worlds: A review of the design affordances and constraints of two 3D virtual worlds as interactive learning environments. Interactive Learning Environments 13, 121–137 (2005)
5. Bricken, M.: Virtual reality learning environments: potentials and challenges. ACM SIGGRAPH Computer Graphics 25, 178–184 (1991)
6. Dede, C.: The evolution of constructivist learning environments: Immersion in distributed, virtual worlds. In: Constructivist Learning Environments: Case Studies in Instructional Design, pp. 165–175 (1996)
7. Roussos, M., Johnson, A., Moher, T., Leigh, J., Vasilakis, C., Barnes, C.: Learning and Building Together in an Immersive Virtual World. Presence: Teleoperators & Virtual Environments 8, 247–263 (1999)
8. Bell, M.W., Smith-Robbins, S., Withnail, G.: Researching Learning in Virtual Worlds. Human-Computer Interaction, 177–191 (2010)
9. Yu, T.: Learning in the virtual world: The pedagogical potentials of massively multiplayer online role playing games. International Education Studies 2, 32–38 (2009)
10. Bouras, C., Giannaka, E., Tsiatsos, T.: Virtual collaboration spaces: the EVE community. In: 2003 Symposium on Applications and the Internet, pp. 48–55 (2003)
11. De Lucia, A., Francese, R., Passero, I., Tortora, G.: Development and evaluation of a virtual campus on Second Life: The case of SecondDMI. Computers & Education 52, 220–233 (2009)
12. Monahan, T., Mcardle, G., Bertolotto, M.: Virtual reality for collaborative e-learning. Computers & Education 50, 1339–1353 (2008)
13. Hmelo-Silver, C.E.: Problem-Based Learning: What and How Do Students Learn? Educational Psychology Review 16, 235–266 (2004)
14. Pearson, J.: Investigating ICT using problem-based learning in face-to-face and online learning environments. Computers & Education 47, 56–73 (2006)
15. Donnelly, R.: Blended problem-based learning for teacher education: Lessons learnt learning. Media and Technology 31(2), 93–116 (2006)
16. Dennis, J.K.: Problem-based learning in online vs. face-to-face. Education for Health 16(2), 198–209 (2003)
17. Ozdemir, S.: The effects of individual and collaborative problem-based learning using an online asynchronized learning tool on critical thinking abilities, academic achievements, and attitudes toward internet use, Phd dissertation. Gazi University Graduate School of Educational Sciences, Ankara (2005)
18. An, Y.J.: Collaborative problem-based learning in online environments, Phd dissertation. Indiana, USA: Department of Instructional Systems Technology, Indiana University (2006)
19. Wilson, B.G.: Constructivist Learning Environments: Case Studies in Instructional design. Educational technology Publications, Englewood Cliffs (1996)
20. Bignell, S., Parson, V.: Best practice in Virtual Worlds teaching: A guide to using problem-based learning in Second Life, `http://previewpsych.org/BPD2.0.pdf` (accessed October 2010)
21. Vrellis, N., Papachristos, M., Bellou, J., Avouris, N., Mikropoulos, T.A.: Designing a Collaborative Learning Activity in Second Life An exploratory study in physics. In: Proceedings of the 10th IEEE International Conference on Advanced Learning Technologies, Sousse, Tunisia (2010)
22. Jamaludin, Y., Chee, S., Ho, C.M.L.: Fostering argumentative knowledge construction through enactive role play in Second Life. Computers & Education 53(2), 317–329 (2009)

23. Good, J., Howland, K., Thackray, L.: Problem-based learning spanning real and virtual words: a case study in Second Life. Alt-J 16, 163–172 (2008)
24. Brown, E., Gordon, M., Hobbs, M.: Second Life as a holistic learning environment for problem-based learning and transferable skills, Assessment, pp. 39–48 (2008)
25. Papamichail, K.N., Alrayes, A., Macaulay, L.A.: Exploring the Potential of Virtual Worlds for Enquiry-Based Learning. In: Lytras, M.D., Damiani, E., Carroll, J.M., Tennyson, R.D., Avison, D., Naeve, A., Dale, A., Lefrere, P., Tan, F., Sipior, J., Vossen, G. (eds.) WSKS 2009. LNCS, vol. 5736, pp. 376–385. Springer, Heidelberg (2009)
26. O'Grady, G.: Holistic assessment and problem—based learning. In: The 5th Asia Pacific Conference on Problem Based Learning: Pursuit for Excellence in Education, Singgahsana Hotel, Petaling Jaya, March 16–17 (2004)
27. Dimitrakopoulou, A., Lars, B.: State of the art of interaction analysis for Metacognitive Support & Diagnosis. Analysis 1 (2006)
28. Dillenbourg, P., Schneider, D., Synteta, P.: Virtual learning environments. Information & Communication Technologies in Education, 3–18 (2002)
29. Strijbos, J., Fischer, F.: Methodological challenges for collaborative learning research. Learning and Instruction 17, 389–393 (2007)
30. Bravo, C., Redondo, M.A., Verdejo, M.F., Ortega, M.: A framework for process–solution analysis in collaborative learning environments. International Journal of Human-Computer Studies 66, 812–832 (2008)
31. Daradoumis, T., Martinez-Mones, A., Xhafa, F.: A layered framework for evaluating on-line collaborative learning interactions. International Journal of Human-Computer Studies 64, 622–635 (2006)
32. Fussell, S.R., Kraut, R.E., Siegel, J.: Coordination of Communication: Effects of Shared Visual Context on Collaborative Work. In: Proceedings of the 2000 ACM Conference on Computer Supported Cooperative Work, pp. 21–30. ACM (2000)
33. Vosinakis, S., Koutsabasis, P., Zaharias, P., Belk, M.: Problem-based Learning in Virtual Worlds: a Case Study in User Interface Design. In: Proceedings of the 1st Global Conference: Experiential Learning in Virtual Worlds (2011)
34. Vosinakis, S., Koutsabasis, P.: A Framework for Problem-Based Learning Activities in Virtual Worlds. In: Proceedings of the 3rd Conference on Informatics in Education (CIE) (2011)

Virtual Customers in a Multiagent Training Application

Philippe Mathieu, David Panzoli, and Sébastien Picault

Laboratoire d'Informatique Fondamentale de Lille (LIFL)
Lille I University,
59655 Villeneuve d'Ascq,
France
firstname.surname@lifl.fr

Abstract. FORMAT-STORE is a serious game application designed for training salesmen and managers in the context of a retail store or a larger supermarket. In this paper, we argue that a relevant way to train a salesperson to their daily activities (e.g. customer relationship management, store management and stock control) consists in immersing them in a 3d environment populated with realistic virtual customers. The first part of this paper presents the multiagent approach we apply to the design of the intelligent customers. Specifically, we analyse the contribution of the interaction-oriented methodology IODA in facilitating the conception of a game for non computer-scientists by means of a user-friendly design tool and the automated implementation of the conceptual model. The second part describes the organisation of the game around scenarios modelled with respect to the pedagogical requirements. We discuss how the multiagent simulation is wrapped by several modules for the purpose of controlling the learning experience of the player.

1 Introduction

FORMAT-STORE is a serious game designed to train business school undergraduate students to retail trade and customer relationship management (CRM) in the context of an organic convenience store. FORMAT-STORE aims to complement a traditional learning content management system (LCMS) which contains approximately 25 thematic lectures dealing with customer welcome, information or argumentation and illustrated with practical examples of dialogues and case studies. In this context, the added value of the serious game is expressed by the following problematics: i) contextualising the knowledge from the LCMS by offering a complementary tool where the learner can apply their newly learnt skills *in situ* in a virtual store and experiment with different ways of dealing with a customer; ii) training the learner to new skills like task prioritisation or time management by means of the realistic simulation of a store, and; iii) offering a flexible evaluation of the learner, in contrast with the traditional "pen and paper" evaluation.

Z. Pan et al. (Eds.): Transactions on Edutainment IX, LNCS 7544, pp. 97–114, 2013.
© Springer-Verlag Berlin Heidelberg 2013

Assuming the role of a salesperson in a virtual replica of an actual store populated with autonomous customers, the learner is confronted to the daily tasks of managing a store and dealing with customers. Different skills are targeted by the learning game:

– Store management skills help maintaining the store functional and safe to the customers. Unpleasant to hazardous incidents can happen like a crate obstructing an aisle or a slippery stain. A salesperson must be able to notice such an event and make a decision quickly.
– Stock control is related to making sure the store is supplied with products at any time. A salesperson is expected to check the expiration date of the goods every morning and if necessary take an appropriate answer – clear the item, restock or order supplies from the wholesaler.
– Customer relationship management (CRM) is the most important aspect of a trade. It consists of looking after the customers' satisfaction, solving their problems, giving them information, etc. CRM involves a good knowledge of the products but more importantly a good practice in dealing and arguing with customers of different profiles.

The next section presents some related projects dealing with the immersion of a user in a populated virtual environment and the training of a learner to business-related activities. Section 3 introduces multiagent systems and details the interaction-oriented methodology on which FORMAT-STORE is grounded. Section 4 describes how a multiagent simulation of virtual customers is wrapped into a serious game. Some outstanding features like scenario integration and adaptive difficulty are detailed.

2 Related Work

Populating a virtual environment with artificial characters for the sake of experiential learning has been investigated in many games or projects, for example for visiting the no more existing Pennsylvania station [23] or discovering the life of Romans in ancient Pompeii [16]. In the Metropolis project [17] or Roma Nova [18], an emphasis is put on the crowd's ability to acknowledge the presence of the player, by means of gaze behaviour or basic dialogic interactions, in order for them to feel part of the population. Yet, the integration of the player is way beyond the requirements of a business training application, where expert-designed scenarios and dialogues constitute the core of the learning activity. The three following projects investigate how training may occur between a learner and a virtual customer.

Knowledge Drive is a serious game developed by Caspian Learning [1] for Volvo Car UK and aimed at replicating the experience of an actual showroom. The main objective of the game is to train salesmen to the products sold by Volvo but also raise their awareness regarding the legislation. In a 3d environment (see figure 1.a), the learner meets virtual customers and builds profiles on the basis of clues they give during dialogues. As the learner makes assumptions regarding

(a) (b) (c)

Fig. 1. (a) “Knowledge Drive” from Caspian Learning replicates the experience of an actual car showroom where the learner builds a line of argument to eventually close the deal. (b) The “Sales Game” from PIXELearning broadens the activity of the learner by enabling them to manage a professional network or a customer database. (c) The “BCV bank” project trains advisors to argue with scripted customers about financial products and bank services.

the profile and the expectations of the customers, he/she is expected to identify the right car for them and build an appropriate presentation by ruling out the irrelevant arguments from an initial argumentation. Law breaking scenarios are introduced in the argumentation; they must be identified and discarded by the learner. The Sales Game by PIXELearning [2] spans a broader range of missions related to business. In addition to sales training (depicted in figure 1.b), the learner is expected to attend virtual meetings, meet virtual colleagues, manage a professional network and build a customer database. Gaining in experience, the learner increases their knowledge and their skills and competes for the salesperson of the year election. Another game of interest has been developed by Daesign [3] for the Cantonale Vaudoise bank (BCV) to train customer advisors selling financial products and services. The game reenacts an interview with a virtual customer. Although the development of the dialogue is mostly scripted (greeting the customer, analysing the needs, arguing with the customer and closing the deal), the player must select at each significant step of the interview one option among several attitudes: analyse, elaborate, carry on (figure 1.c).

Although all the games cited in the previous paragraph focus on the CRM only, several reasons make them particularly interesting for the scope of this paper. Firstly, they point out the many advantages of teaching the relationship with the customer using an interactive simulation over relying on traditional teaching methods. Also, they demonstrate the usefulness of a game in complement of a knowledge base – the aforementioned games are bundled with a traditional learning platform – for the knowledge to be contextualised and translated into skills by the learner. Indeed, although a LCMS enables the content to be personalised to the learner, the knowledge is neither personally constructed nor applied. A game offers this opportunity as the learner uses it as a playground where new skills can be tested and old ones can be rehearsed.

The ambition of Format-Store is to combine the capacities of an immersive environment populated with intelligent customers (unpredictable expectation and needs of the customers, necessity for prioritising the tasks, ability to identify a customer needing information or help) to a tutoring system where scenarios can be designed in order for the domain experts to control the learning experience. Those requirements advocate for using a robust and modular artificial intelligence system, not only able to model the coherent behaviour of a customer but a whole crowd of them. Multiagent systems represent a suitable answer to this problematic, particularly the interaction-oriented approach detailed in the next section.

3 IODA: An Interaction-Oriented Multiagent Design Methodology

A multiagent system (MAS) [27] is an organised set of entities called agents interacting in an environment. The term was coined early in the 1990s and encompasses every simulation of a complex phenomenon where interacting particles can be identified.

Multiagent systems have been gathering an increasing interest lately as an alternative approach to mathematical modelling which more traditionally aims at modelling a phenomenon with equations. The main interest in MASs is the ability to consider a complex phenomenon as the – often emergent – result of simple agents interacting with each other. In this context, MASs offer the ability to apply a bottom-up methodology by locally defining the role and the behaviour of each agent participating in the global phenomenon. Besides, MASs help understanding the contribution of each agent whereas a mathematical model can only describe the global mechanisms of the phenomenon. Application areas of MASs include simulating natural phenomena like molecular biology [9] or ecosystems [7], animating artificial worlds [6], video games or computer-generated imagery [4], social behaviour [10], economy [5] or disaster management [19].

Individual-Centred Simulations. Whereas classical simulations aim at modelling a phenomenon with one or several mathematical equations, MASs focus on the individuals participating in the phenomenon. The question MASs address is: knowing the entities participating in a given phenomenon, what must be the behaviour of each individual for the whole phenomenon to demonstrate a desired property?

To solve Artificial Life (AL) problems, MASs have proved well suited for modelling emerging collective phenomenons composed of simple interacting individuals. Reynolds [22] has shown that a visually realistic flock of virtual birds can be obtained by applying a local behaviour composed of three simple rules to each flockmate. Another famous illustration of emergent complexity is provided by Resnick [21] with his simulation of ants and termites foraging behaviour. The common point of these simulations is the proof that self-organisation can emerge without the need for a supervising body provided the agents are locally endowed with an appropriate behaviour and means to communicate with each other.

As a consequence, simulating a phenomenon using a MAS does not require one to understand the phenomenon but merely observe the agents participating in the phenomenon. An individual-centred methodology thus offers a great advantage over a more complex approach where the global description of the phenomenon is necessary. Besides, the principle of emergence remains valid regardless of the reactive or cognitive nature of the agents.

Reactive and Cognitive Agents. All the agents in a MAS support common characteristics although some properties are inherent to different kinds of agents.

An agent is always situated in an environment according to one or several metrics attached to the environment. Possible metrics can be a distance in an Euclidean space or relationships in a social network for instance. Relying on the metric, an agent has a neighbourhood and is itself located in the neighbourhood of other agents. Each agent has a local perception of the environment, usually restrained to its neighborhood. Similarly, an agent is allowed to act or interact locally with the other agents in the environment. Interactions can be direct between two agents or indirect when a media (usually the environment itself) is necessary for a message to be conveyed from one agent to another. Finally, the behaviour of an agent is defined by an internal perception-decision-action loop. The agent is said autonomous as the decision is locally and internally taken by the agent itself.

Traditionally, a dichotomy is made between reactive and cognitive agents. Reactive agents use a trivial decision process that mostly consists in triggering an action according to a perception. Cognitive agents are proactive and plan actions in the long term to achieve internal goals or objectives. The Belief-Desire-Intention model [20] is a good illustration of how such an agent works.

The MAS design methodology involves modelling the behaviour of the agents and their interactions. Depending on where the priority is set, two methodologies exist: agent-oriented programming (AOP, [24]) and interaction-oriented programming (IOP, [25,15]).

3.1 Ioda: Interaction-Oriented Design of Agents

Ioda [14] is a multiagent simulation methodology whose originality is to focus primarily on how the agents interact instead of how they behave. This methodology is grounded on a simple observation coming from experimental experience. The design process a simulation always involves making assumptions at some point, irrespective of the phenomenon to simulate. In the context of a multiagent-based simulation – e.g. when the actual phenomenon is the obvious outcome of multiple interacting entities – the description of these interactions is the only objective assumption one can formulate. The set of processes occurring "inside" each entity and leading to the interaction can only be guessed. In order to avoid introducing a bias too early in the process, a safe approach would consider setting the foundations of every model on observable characteristics, then building on this model with the constant concern of delaying the introduction of hypothetical assumptions as long as possible. Ioda follows this principle by discarding the first hypothesis concerning the selection of which entity is an agent and which is a mere passive object, by providing a user-friendly tool (Jedi) for domain-experts to participate in the design of the interactions and by providing a tool for translating the model (Jedi-Builder) into a set of classes where the interactions can be implemented by a programmer.

3.2 Everything Is Agent

The starting point of designing a multiagent simulation consists in identifying the agents participating in the simulation. What defines an agent in a MAS is a minimum degree of behavioural autonomy and the subsequent ability to trigger autonomously an action or an interaction. Historically, "living" characters in a virtual simulation are considered as agents, "inanimate" objects like trees, furniture or items are not. Unlike other approaches, the first simplifying hypothesis of the Ioda methodology is to consider every entity involved in the simulation an agent. This choice is fully argued in [13]. Figure 2 illustrates the virtual supermarket environment used in Format-Store where every character – employee, customer – and almost every object – item, shelf, information sign, etc. – is an agent in the simulation. In the Ioda methodology, families of agents are listed in such a way that every agent in the environment strictly belongs to one family.

Considering every entity as an agent simplifies the first step of the MAS design, but also provides a convenient way to describe the interactions, as detailed in the next sections.

3.3 Interactions Made Concrete

In a similar way every entity in the simulation is represented by an agent family, any behaviour is described by an interaction in Ioda.

Unlike other MAS approaches, where an interaction is virtually expressed in the behaviour of two agents interacting together, each interaction in Ioda has a software tangibility and a central position in the design. An interaction is a rule involving two agents: it is performed by a source agent and undergone by a target agent. It is composed of two parts, a Boolean condition testing if the interaction can be triggered and an action part containing the actual interaction logic. Both these functions rely on generic primitives, left for implementation inside the agents or reused from a template library. This is formalised as follows:

`InteractionName (Src, Tgt) :=` `CONDITION:` *condition primitives*
`ACTION:` *action primitives*

The condition and action parts of the rule rely on generic perception and action primitives so that interactions are independent from the concrete implementation of the agents.

As a consequence, Ioda exhibits two unique features. Firstly, the interactions can be represented independently from the agents, as libraries of interactions for instance. Interactions are reusable from one agent family to another and from one simulation to another. They can be allocated to agents in a plug and play fashion. The other advantage of interactions reification is the ability for all the agents to be processed by a generic engine through a single iteration loop, irrespective of the nature of each agent.

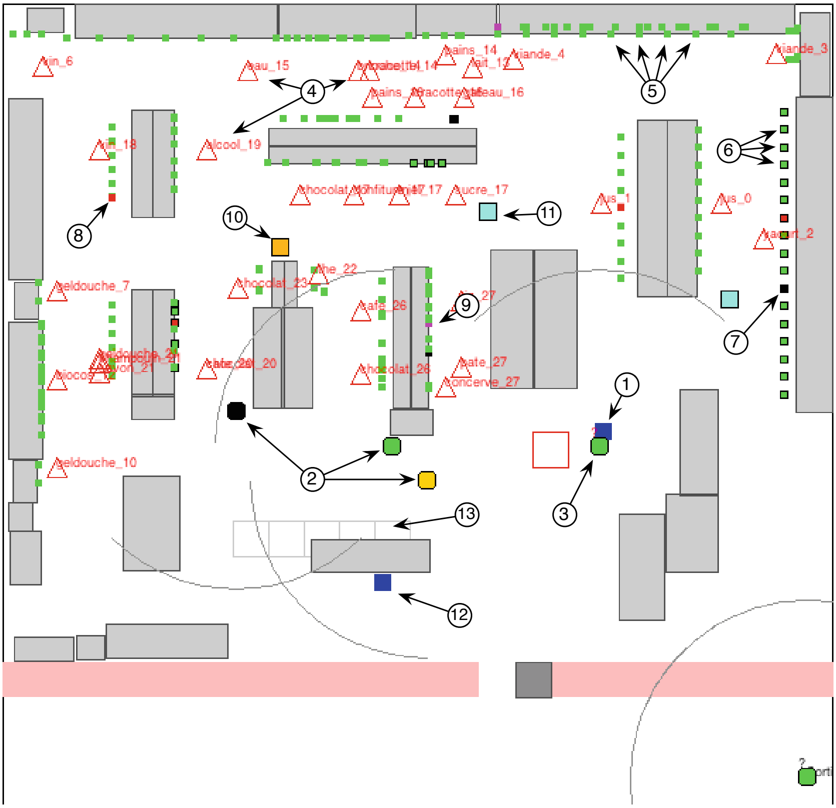


Fig. 2. Almost every object in the environment is an agent on the same account as the characters. In addition to simplifying the design process, considering every object as an agent also allows spreading the "intelligence" of the characters across the objects they interact with. The figure represents a Java AWT rendering of the multiagent simulation used during the development. The agents are: (1) employee; (2) customers with various satisfaction levels; (3) conversing customer; (4) information signs; (5) items; (6) popular items; (7) out-of-stock item; (8) damaged item; (9) expired item; (10) crate; (11) stain on the floor; (12) checkout cashier; (13) waiting queue.

3.4 Interaction Matrix

Having defined the agent families participating in the simulation, the next step of the Ioda methodology consists in describing their interactions. This is achieved by allocating the interactions in a matrix named the interaction matrix. Figure 3 presents the version of the interaction matrix used in Format-Store. All the agents participating in the simulation are listed in the matrix along with their mutual interactions. Note the avatar of the player is included in the matrix like any other agent. In the matrix, each interaction receives two additional parameters.

– the distance defines the minimum distance between the source and the target for the interaction to be allowed. For instance, seizing an item on a shelf requires the character to be standing in front of the item; a character can read an information sign from a certain distance; etc.
– the priority is used for sorting between several interactions whose preconditions are verified.

Source/Target	∅	Employee	Customer	Door	Sign	Checkout	Item	Queue	Stain	Crate
Employee	Converse(0)		StartConversation(1,0) EndConversation(1,0)				Remove(1,0) Supply(1,0) Order(1,0)		Clean(1,0)	PutAway(1,0)
Customer	Wander(0) GoTowards(1) Converse(13)	Wait(2,3)		Exit(1,12)		Pay(2,10)	Get(2,5)	StepIn(5,7) MoveOn(1,8) WalkOut(1,11)		
Door	SpawnCustomer(1)		Acknowledge(10,0)							
Sign			Acknowledge(10,0)							
Checkout			Acknowledge(10,0) CheckOut(2,0)							
Item	Expire(1) MakeStain(1) SpawnCrate(1)		Acknowledge(10,0) Upset(1,0) Ack_OutOfStock(1,0)							
Queue										
Stain			Upset(1,0)							
Crate			Upset(1,0)							

Fig. 3. The interaction matrix presents the interactions allowed for any agent family as a source towards any other agent family as a target (including self) or the environment (degenerate interaction, column ∅). How to read this matrix? For the Customer agent family for instance, the column labelled Customer lists all the interactions of which a Customer is a target. In this example, the Customer is basically informed by almost any other agent (of the location of Items by a Sign, of the price and quantity of an Item by this Item, etc). The row labelled Customer lists all the interactions of which a Customer can be the source. Priority values (n) are used when several interactions can be applied at the same time. The higher the value, the higher the priority.

Although the interaction matrix offers a simplistic representation of the simulation model, the the behaviour of each agent is exhaustively described. We argue that this representation is functionally equivalent to a more complex algorithm, put aside the difficulty for a non computer scientist to read the latter.

3.5 Computer-Aided Design for Non Experts

We have mentioned earlier in this paper the importance of involving domain-experts as far as possible in the design process. IODA follows this principle by proposing a simplified multiagent methodology made accessible to non computer scientists. The implementation stage has received the same attention with two additional elements of the IODA methodology: JEDI and JEDI-Builder. JEDI is an application programming interface (API) providing a set of Java classes for a user to ensure the rigorous implementation of their IODA conceptual model. JEDI-Builder is a Java application assuming two roles. Firstly, JEDI-Builder provides a computer-aided design tool for the user to model the interaction matrix using a user-friendly graphical interface where agents and interactions can be added in a drag-and-drop fashion. Secondly, JEDI-Builder is also able to translate such a IODA-compliant matrix into a JEDI application that can thereafter be used as a simulation or integrated in a larger project. Using these tools enables a user to implement the major part of a MAS without any particular knowledge of computer programming. At the end of the process though, a computer-scientist is required to implement the core of each interaction, if those interactions are not part of an already existing library – which is often the case as JEDI natively includes a set of predefined generic interactions.

The implementation of an interaction can take many shapes, from the most trivial to more complex tasks. For instance, in FORMAT-STORE, the player as an employee can supply an item on a shelf or converse with a customer. The interaction "Supply" simply consists in increasing the quantity value of a target item. In contrast, the interaction "Converse" involves more complex operations. The camera position is changed for a closer look on the customer. A dialogue window is opened and a script is started. This script displays a narrative text and several possible answers among which the player has to choose one. Depending on the answer, the satisfaction level of the customer decreases or not and the player is rewarded (positively or negatively).

Although complex interactions like the dialogue situation seem at first glance the most useful in terms of behaviour, the very interest of multiagent simulations lies in using multiple simple interactions, as explained in the next section.

3.6 Adaptive Behaviours

This interaction-oriented vision offers an original implementation of the concept of affordances. In his influential book [11], Gibson states that the interaction capabilities – the functions – of any object in a real environment are mainly suggested by the object itself – its shape, position, etc. One interpretation of the concept of affordances has been used many times since in character animation, where computer graphics scientists have found more intuitive to attach the animations and the algorithms necessary for an interaction to the target of this interaction. For instance, a virtual character wanting to open a door would be acknowledged by the door itself of the position it should stand at and the animation it should play. Affordances are implemented in most simulations or games featuring at least one virtual character interacting with objects in the environment. That way, every passive object describes to the character how it should be handled, under which circumstances and what is the outcome of the interaction. The character is therefore freed from that knowledge, putting a focus on managing internal goals and searching the environment for objects likely to solve these goals. Another benefit of using affordances is to avoid any glitch in the animation by carefully adjusting the positions of the interacting entities and synchronising their respective animations. In such simulations though, the character is always at the origin of every interaction. Whichever interaction is triggered is the result of a complex and often time-consuming decision process selected by a cognitive controller.

The idea that the cognitive abilities of an agent may be spread as well among the interactions offered to the agent is part of the multiagent approach, and particularly of IODA which considers that an intelligent behaviour can be expressed by a reactive agent. This argument is close to earlier research in artificial intelligence postulating that intelligence is not the result of planning and reasoning – like a cognitive agent using scripts and complex algorithms – but on the contrary that the mere appearance of such cognitive processes is the result of an agent's reactive behaviour in a dynamic environment [8].

To illustrate this principle, let us consider the behaviour of an intelligent customer shopping in FORMAT-STORE: A customer is endowed with a profile – 10 different profiles were provided initially by the client, which basically corresponds to 5 age ranges times 2 genders – and a shopping list containing items to be purchased. The number of items on the list depends on the customer's profile.

A virtual customer entering the store will consider every interaction in decreasing priority until one is realisable (e.g. the preconditions are verified). Getting out of the store is only realisable when all the items have been paid. Paying requires first to queue at the checkout (actually checking out is managed by the checkout agent). Queuing can be started when all the items in the shopping list have been retrieved. An item can be taken when the customer is actually standing at reach of hand of the item. Moving towards an item relies on standard pathfinding and navigation algorithms but requires knowing the location of the item, which is acknowledged after an interaction with one of the information signs located throughout the store. While shopping in the store, the customer seems to follow a plan. Yet, every action is independent from the following and their sequence has merely been established by selecting priorities.

The adaptive behaviour of the customers is illustrated by the way they select the items. An agent does not embed a map of the product items. The location of every item is provided by information signs located throughout the store, while the customer is wandering or shopping for already known items. That way, selecting the next item – in practice, the closest item whose location is known) depends on the customer's current knowledge, which depends in turn on the signs the customer has already interacted with. In addition, customers dynamically respond to obstacles in the store like a crate obstructing an aisle, a stain on the floor, or another customer standing in their path. They will therefore always take different paths, even when their shopping lists are similar. Figure 4 illustrate this behavioural differentiation within the virtual customers.

In addition to being adaptive, the behaviours expressed are also robust. Disturbances can be created in the game by adding/modifying/removing signs without affecting the ability for each customer to behave coherently. Ultimately, the customers can be placed in a different store and still manage to shop for goods.

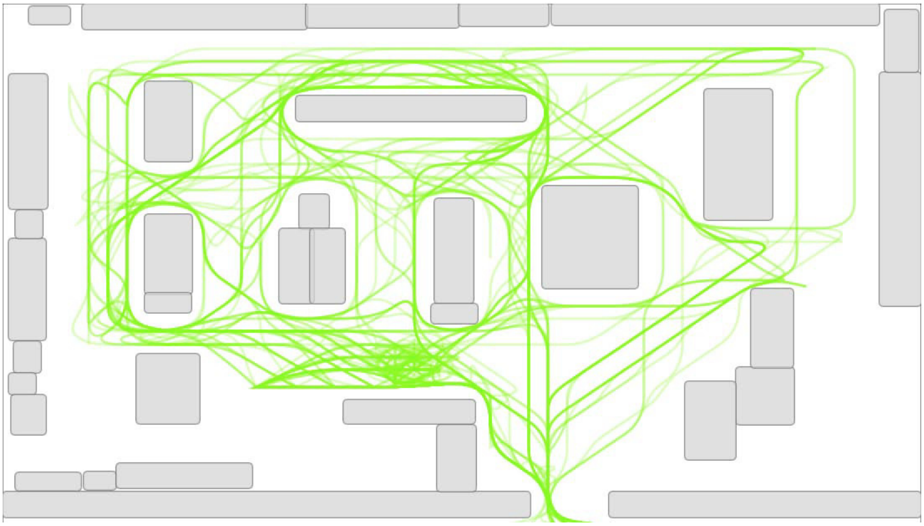


Fig. 4. Recording and drawing the position of each customer during the simulation illustrates how varied routes are obtained owing to the mere attribution of different shopping lists, however relying on IODA's adaptive planning of each customer's behaviour

4 A Serious Game for Immersive Training

The FORMAT-STORE serious game is grounded on a multiagent simulation of the customers and the virtual store. The process of transforming this simulation into a game consists in a few steps: i) integrating the player into the simulation taking the control of one agent, namely an employee, ii) wrapping the simulation using a game manager for controlling the user scenarios and iii) plugging a performance analyser responsible for the scoring and the adaptive difficulty. All these steps are detailed in the following sections.

4.1 Presentation

Towards its inclusion in the LCMS (see section 1), the FORMAT-STORE serious game is available online. The high-end graphics required by the 3d environment are provided by X3d (formerly VRML) technologies. The game requires a – freely available – plugin at the user's end but displays in return compelling 3d graphics, yet computationally efficient enough to allow for the game to run on an Internet browser.

4.2 The Human in the Loop

The result of the IODA methodology applied to the FORMAT-STORE requirements is a multiagent simulation of a store filled with goods and populated with

(a) (b) (c) (d)

Fig. 5. FORMAT-STORE features a virtual organic store populated with autonomous customers. (a) The player can control the avatar anywhere in the store using the arrows on the keyboard or graphic controls on the screen. (b) The virtual customers navigate and shop autonomously inside the store. The condition of the store and the items impacts their level of satisfaction. (c) Conversational interactions in FORMAT-STORE are rendered by a specific GUI during which the player can select appropriate answers. (d) The realistic behaviour of the customers is reinforced by their ability to avoid colliding with one another or to queue at the checkout.

customers shopping autonomously and seeking for assistance. This multiagent simulation, at the core of the serious game, is best described by the interaction matrix on figure 3 and the adaptive behaviour of the customers explained in section 3.6.

Using the arrows on the keyboard or a graphic user interface (GUI), the player is enabled to move their avatar (namely, the employee) freely in the virtual store (figure 5.a). Using the mouse, the player can also interact with many elements of the store including the product items and the autonomous characters – customers and other employees. In that latter case, a specific GUI is loaded as the player enters a conversational mode (figure 5.c) where they can select propositions.

The player is integrated in the game by means of controlling one of the agents (hence the presence of the employee in the interaction matrix) in the simulation, following a "letterbox" principle: actions from the player are captured, sent and expressed by the agent triggering the corresponding interaction. Conversely, interactions undergone by the agent are notified to the player. As a result, actions from the player are seamlessly conveyed in the simulation, preserving the autonomy of the agents and the independence of the action selection mechanism. On the other hand, the controlled agent introduces perturbations which, combined to the behavioural adaptivity of the customers, fosters a great variety on the situations presented to the player.

Having implemented the simulation part of the game, where the user is free to wander in a living replica of a store, the next part towards achieving a game is to introduce scenarios for the player to actually learn in this realistic context.

4.3 Game Management and Scenarios

Traditionally in games where the player faces virtual characters, unitary scenarios defined in accordance with the educational requirements are represented by the specific behaviour of virtual characters. For instance, in the salesmen training games presented in section 2, each character represents a pedagogical situation to explore (advise the right product in accordance with the client's profile, cope with a customer difficult to argue with, etc). In practice, each character embeds a script describing either a specific dialogue or a predefined sequence of actions/interactions.

In Format-Store, we put an emphasis on the game's scalability, namely to what extent the client will be able to mend or remove existing cases or add new pedagogical elements. We explore an original approach taking advantage of the behaviour's adaptiveness in a multiagent system. The customers wandering in the store at any time are merely going to their business – shopping for goods – trying to fulfil internal goals – purchasing items on a shopping list or querying for information – instead of following a scripted behaviour. In this context a scenario is not attached to a specific character but rather consists in attributing goals to the customers or introducing disturbances in the environment. For instance, an item a customer is looking for can be removed, a customer can be introduced in the store with a question to ask, an information sign can be misplaced, an aisle can be made impassable by an oil stain, etc. As a result of the agents' adaptive behaviour, one or several customers will be affected by the trouble introduced by the scenario and inspire a specific reaction, ranging from being upset to complaining to the employee depending on their profile.

Problem-Situations. The notion of scenario is described by the project's content manager ENACO as a problem-situation, namely a problematic situation including a context and a branching dialogue investigating the different ways for the employee to deal with it. 25 problem-situations were initially provided by ENACO, addressing various issues such as a missing item on a shelf, an aisle obstructed by a stain or a box on the floor, or sale-related questions. Integrating a problem-situation in the game raises two questions. Firstly, how should the context be represented in the customer's behaviour? Secondly, how should the dialogue be integrated as part of the agent's abilities?

Representing the context of each problem-situation is easy using the interaction-oriented methodology IODA. The way a situation arises is implemented within the preconditions of one or more interactions, which must be defined on purpose. For example, customers try to meet the employee and start a conversation when they perceive the employee and have a question to ask. The preconditions are expressed as such. The scenario consists in adding a question to an entering customer. A second example considers that customers are upset and look into complaining when they perceive an expired item. The scenario in that case consists in expiring an item in the store. Problem-situations may require more complex scenarisation, which is achieved using the exact same method. For instance, when a customer attempts to get an item, there is a chance that it is dropped and makes a stain on the floor. When it is detected, the stain as an agent has the ability to upset other customers, which is the context for a specific problem-situation. In a similar way, when an item is missing on the shelves, the employee has the ability to supply it. Doing so, a crate is left on the floor, which can also upset a customer.

Integrating the dialogue itself is equally easy since it can be simply represented as an interaction. That way, the customer is enabled to converse with an employee, the same way it is able to get an item. The dialogues provided for each problem-situations have been digitised in an *ad-hoc* XML format, representing the various branches of a dialogue and the links from one dialogue line to another. The core of the converse interaction therefore consists in using a parser to load the dialogue. Both the employee and the customer are able to converse, and each agent loads the right part of the dialogue based on its current advancement. When several branches are available, the user selects one using a dialogue GUI whereas the customer selects one randomly. When a final branch is reached, the interaction is considered terminated. Depending on the branches chosen by the player, a score is attributed at the end of the dialogue. The score values are represented in the XML structure along with the dialogue lines.

All the events (customers with questions, expired items, probabilities of dropping an item, etc.) at the origin of the various problem-situations arising in the game are controlled by a single module called the game manager and organised within a game session.

Game Session. Designing a game session is inspired by the typical day of a salesperson in a supermarket. The duration of a game session has been fixed arbitrarily to 20 minutes for practical reasons, and therefore the many activities of an average day are condensed within this short time period.

The game manager is an autonomous module operating upstream and downstream the multiagent simulation. It decides when to send a new customer in the store, whether it has a question to ask or not, which items are on its shopping list or when to trigger a new event, depending on the current level of difficulty, and based on adjustable parameters. The level of difficulty corresponds to the maximum number of customers allowed in the store at the same time: the more customers in the store, the more difficult for the player to keep the store tidy and to help every one of them.

When a customer exits the store, the game manager collects feedback information in order to evaluate the player's performance and adjust the level of difficulty (discussed next section).

Scoring and Adaptive Difficulty. In FORMAT-STORE, the player is evaluated by the virtual customers themselves. During their activity in the store, between their entrance (on the game manager decision) and their exit (when their shopping is over), each customer undergoes the influence of the other agents. Every customer entering the store is attributed a level of satisfaction which depends on their profile (grumpy, friendly, etc). Every unfortunate event like trying to buy an expired item, finding an item out of stock, stumbling across a box or a stain, etc. has a negative impact on this level of satisfaction. When the customer is removed from the store, a difference in satisfaction can be computed. Knowing that every loss of satisfaction is due to the player (expired item not replaced, out of stock item not supplied, crate not removed, stain not cleaned, etc.) the difference in satisfaction of a customer translates the performance of the player. Indeed, when the player copes with the simulation, the store is tidy and the customers helped in time; and reciprocally.

Operating as a sub-module of the game manager, the performance analyser collects in real time all these satisfaction levels, along with the dialogue scores (when applicable). Based on this series, a mathematical equation is applied in order to compute the ideal level of difficulty, e.g. the maximum level where the player stays ahead of the tasks. When the difficulty increases, more customers are allowed in the store. When the difficulty decreases, some customers are not replaced when they exit the store.

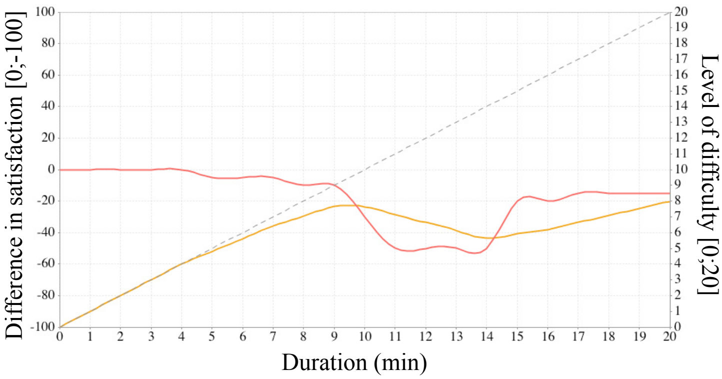


Fig. 6. The level difficulty (orange curve) is tailored to the player's performance (red curve) all along the game session. At minute 8, as the player fails coping with the simulation, the difficulty is decreased. The result is observed a few minutes later: the player gets back to grips with the game and the difficulty resumes its progression

Figure 6 shows how adaptive difficulty helps maintaining the challenge at reach of the player. When the difficulty is too high for the player to cope with, it is decreased until a normal performance is observed again. All along the game session, the player feels no frustration and is not tempted to quit the game prematurely. This is a critical point, we believe, and the next section intends to validate our approach.

Educational Validation. Despite the novelty of the technique, learning in a populated virtual environment can be supported by classical theories. In particular, two models readily apply to Format-Store.

Experiential or exploratory learning models [12] promote the free exploration and the autonomous and personalised construction of cognitive associations and understandings. The idea underlying Kolb's experiential learning is that a realistic virtual environment like Format-Store allows a contextualised learning, as opposed to the declarative and decontextualised learning provided by the LCMS. Format-Store provides means to learn from real life situations: elaborate a routine, learn to prioritise different assignments.

Socio-cultural models of learning [26] point out the fundamental role of social interactions in the development of cognition. Vygotsky's model replaces the social interaction with the teacher, and by extension with the virtual representation of a tutor, at the centre of any learning activity. In practice, this involves the constant delivery to the user of a feedback on his performance. A first requirement for the game designer is therefore to provide means to assess the player's performance and to guide them toward increasing their skills. The feedback can be delivered continuously in a game whereas it is hardly conceivable in real life. Another strong concept, which is a direct consequence of the initial idea, is to scaffold the learning by building new knowledge on top of the existing, whilst consolidating the latter. Vygotsky claims the utter importance of maintaining the learner in what he names the zone of proximal development (ZPD), situated beyond what the learner already knows – in which case the benefit is null – and below what he cannot achieve on his own – in which case no learning can happen but frustration, no matter how much help is provided. In the context of a game, adjusting the difficulty is a relevant way to control the position of the challenge in the learner's ZPD. Besides, a proportionate challenge provides a leverage on the engagement and the motivation of the player. Therefore, the second requirement commands the game designer to ensure that the level of difficulty is adaptive.

5 Conclusion and Future Work

The Format-Store project is a serious game aimed at training salesmen by immersing them in a dynamic virtual store populated with intelligent customers. In addition to its intrinsic originality, Format-Store brings two original features. Instead of scripting the agents, the educational scenarios attribute goals to the adaptive customers or define disturbances in the environment affecting their behaviour. The scalability of the game is therefore increased as changing the educational content can be handled by the content providers themselves. Another aspect concerns the uncompromising implementation of the affordances concept, allowing to simplify the visual animation of the agents, but above all to obtain intelligent behaviours in spite of the reactive nature of the agents.

Two other aspects are worth mentioning. Owing to a user-friendly design tool, domain-experts are maintained in the conception process farther than any other methodology, thus avoiding the premature introduction of biases and the translation of the formal model into a programming language is partially automated.

More generally, the methods and techniques deployed in FORMAT-STOREin terms of game management, scenarios integration and player immersion can be easily reused in other multiagent-based interaction-oriented serious games.

In the long term, we are considering the side development of a decision support tool for supermarket marketing strategy units. We intend to take advantage of our realistically-profiled crowd of intelligent customers in the context of product placement or shelves layout in a large supermarket.

Acknowledgment. The FORMAT-STORE project is supported by the French ministry of Economy, Finances and Industry under the "2009 Serious Game" scheme. In-game images are the property of Idées-3Com. The authors are grateful to Jean-Baptiste Leroy for his committed involvement in the project.

References

1. http://www.caspianlearning.co.uk/
2. http://www.pixelearning.com/services-the_sales_game.htm
3. http://www.daesign.com
4. http://www.massivesoftware.com/
5. Arthur, W.J., Holland, B., LeBaron, R., Palmer, T.P.: Asset pricing under endougenous expectations in an artificial stock market. Economic Notes 26, 297–330 (1997)
6. Bonabeau, E.: Agent-based modeling: methods and techniques for simulating human systems. In: Proc. National Academy of Sciences, vol. 99, pp. 7280–7287 (2001)
7. Bousquet, F., Le Page, C.: Multi-agent simulations and ecosystem management: a review. Ecological Modelling 176(3-4), 313–332 (2004)
8. Brooks, R.A.: Intelligence without reason. In: Proceedings of the 1991 International Joint Conference on Artificial Intelligence, pp. 569–595 (1991)
9. Desmeulles, G., Bonneaud, S., Redou, P., Rodin, V., Tisseau, J.: In-virtuo experiments based on the multi-interaction system framework: the RéISCOP meta-model. CMES, Computer Modeling in Engineering & Sciences (2009)
10. Epstein, J.M., Axtell, R.: Growing Artificial Societies: Social Science from the Bottom Up. Brookings Institution Press, Washington (1996)
11. Gibson, J.J.: The ecological approach to visual perception, Hillsdale, New Jersey, London (1979)
12. Kolb, D.A.: Experiential Learning: experience as the source of learning and development. Prentice-Hall, New Jersey (1984)

13. Kubera, Y., Mathieu, P., Picault, S.: Everything can be agent! In: der Hoek, et al. (eds.) Proceedings of the 9th International Joint Conference on Autonomous Agents and Multi-Agent Systems, Toronto, pp. 1547–1548 (2010)
14. Kubera, Y., Mathieu, P., Picault, S.: IODA: an interaction-oriented approach for multi-agent based simulations. Journal of Autonomous Agents and Multi-Agent systems (JAAMAS) 23(3), 303–343 (2011)
15. Kubera, Y., Mathieu, P., Picault, S.: Interaction-oriented agent simulations : From theory to implementation. In: Ghallab, M., Spyropoulos, C., Fakotakis, N., Avouris, N. (eds.) Proceedings of the 18th European Conference on Artificial Intelligence (ECAI 2008), pp. 383–387. IOS Press (2008)
16. Maïm, J., Haegler, S., Yersin, B., Müller, P., Thalmann, D., Van Gool, L.: Populating Ancient Pompeii with Crowds of Virtual Romans. In: Proceedings of the 8th International Symposium on Virtual Reality, Archeology and Cultural Heritage - VAST (2007)
17. O'Sullivan, C., Cassell, J., Vilhjalmsson, H., Dingliana, J., Dobbyn, S., McNamee, B., Peters, C., Giang, T.: Levels of detail for crowds and groups. Computer Graphics Forum 21(4), 733–742 (2002)
18. Panzoli, D., Peters, C., Dunwell, I., Sanchez, S., Petridis, P.: Levels of Interaction: A User-Guided Experience in Large-Scale Virtual Environments. In: Proceedings of the Second Conference on Games and Virtual Worlds for Serious Applications (VS-GAMES 2010), Braga, Portugal, March 25–26, pp. 87–90 (2010)
19. Querrec, R., Reignier, P., Chevaillier, P.: Humans and autonomous agents interactions in a virtual environment for fire figthing training. In: Virtual Reality International Conference, pp. 57–63 (2001)
20. Rao, A.S., Georgeff, M.P.: Modeling rational agents within a bdi-architecture. In: Proceedings of the 2nd International Conference on Principles of Knowledge Representation and Reasoning, pp. 473–484 (1991)
21. Resnick, M.: Turtles, Termites and Traffic Jams. Explorations in Massively Parallel Microworlds. MIT Press, Cambridge (1994)
22. Reynolds, C.W.: Flock, herds ans schools: a distributed behavioural model. In: SIGGRAPH 1987. Computer Graphics, vol. 21(4), pp. 25–34. ACM Press, Anaheim (1987)
23. Shao, W., Terzopoulos, D.: Autonomous pedestrians. Graph. Models 69(5-6), 246–274 (2007)
24. Shoham, Y.: Agent oriented programming. Journal of Artificial Intelligence 60(1), 51–92 (1993)
25. Singh, M.P.: Conceptual Modeling for Multiagent Systems: Applying Interaction-Oriented Programming. In: Goos, G., Hartmanis, J., van Leeuwen, J., Chen, P., Akoka, J., Kangassalu, H., Thalheim, B. (eds.) Conceptual Modeling. LNCS, vol. 1565, pp. 195–210. Springer, Heidelberg (1999)
26. Vygotsky, L.: Mind in Society. Harvard University Press, Cambridge (1978)
27. Wooldridge, M., Jennings, N.R.: Intelligent Agents: Theory and Practice. Knowledge Engineering Review (1994)

Game Based Early Programming Education: The More You Play, the More You Learn

Ioannis Paliokas, Chistos Arapidis, and Michail Mpimpitsos

Alexander Technological Educational Institute of Thessaloniki,
Department of Information Technology, P.O BOX 141,
57400 Sindos, Greece
{ipalioka,charap,mpimpits}@teithe.gr

Abstract. Mini-languages is a branch of Educational Software for learning programming at an introductory level. On the other hand, participation, interaction and storylines make Educational Games motivating to young learners. The study presented here examined various widely known mini-languages with an emphasis on LOGO implementations and followed a combinational route to take advantage of both Game-Based Learning (GBL) and the use of mini-languages in the design of a new LOGO-like environment. PlayLOGO 3D is a video game with defined learning outcomes aiming to support GBL activities especially designed for children aged 6-13 years in the early stages of programming education. The Expert Review Method was used for initial evaluation based on a set of heuristics for usability, game play and educational effectiveness. Although the expert team found a few violations of heuristics criteria, evaluation results are very encouraging and prove that there is enough room to make programming education more fun.

Keywords: Game Based Learning, LOGO, Edugames.

1 Introduction

LOGO is widely known as a computer programming language used for programming turtle-graphics school projects. In contrast to freshmen who learn many general-purpose programming languages as tools for writing real-world application programs, LOGO is the most common educational programming language for elementary school students. It was created by Daniel G. Bobrow, Wally Feurzeig, Seymour Papert and Cynthia Solomon at 1967 for constructivist teaching [34]. LOGO has been used in the past years in education of Mathematics, Geometry, Physics and interdisciplinary approaches in all over the world. The turtle-graphics use drawing commands followed by coordinates relative to the cursor. In most applications, the cursor is depicted by a turtle or a robot. A common set of LOGO commands includes Forward, Backward, Left and Right as well as other commands to handle lists, files, functions and even recursion. Some typical tasks assigned to students are to draw basic shapes using locomotion commands on a turtle. A lot of researchers have supported the educational use of LOGO [27]. Especially about the impact the turtle metaphor has on motivating

Z. Pan et al. (Eds.): Transactions on Edutainment IX, LNCS 7544, pp. 115–131, 2013.
© Springer-Verlag Berlin Heidelberg 2013

students, there are positive values of that flexible and universal metaphor to stimulate student's imagination, constructive, and analytical thinking [29].

On the other hand, video games constitute an alternative way to teach children of the so-called Game Generation using their own language [2]. Using a programming language to move a turtle in the screen may be fun, but this does not constitute a video game. The question is: How can we introduce programming concepts in a video game? Most educational video game designers try to emulate the commercially successful video games, but fail to gain similar success because they are resistant to change their thinking [31]. Educational software design teams usually give more importance on the visible educational and cognitive characteristics of educational applications based on their previous experience in designing educational material. This is detrimental to other characteristics that are equally important to video games like fantasy, challenge and curiosity according to the three basic game elements proposed by Malone [21]. When aimed at: A) motivating students, B) interaction with content and C) role taking, video games can increase the learning gains [11].

2 Motivation of Our Research

Why does a game-like mini-language for turtlegraphics programming is needed in schools? Mini-languages provide a sound basis for introducing programming to novices because they are small, simple, build on engaging metaphors and make user operations to be naturally visible [5]. The mini-languages approach is not entirely new. To name a few, Karel the Robot [26] was one of the first programming microworlds, Robocode [24], Gun-Tactyx [3] and Prog&Play [23] which give emphasis on Artificial Intelligence scripting. Also, Marvin's Arena [28] and MUPPETS [30] are programming games suitable for students of varying programming experience. Other implementations that could not be missing from the above list are the very well supported Alice [7] used for story-telling, animations and interactive games and C-Sheep [1] a mini-language based on a simplified ANSI C programming language.

Dealing particularly with video games, role-playing and challenge is emphasized. Interaction today is common to all contemporary educational software applications and needs no extensive analysis. Role-playing in educational activities is an established technique and when introduced in narrative interacting environments can maximize intrinsic motivation and affect positively a wide range of knowledge domains [10]. In turtle-graphics programming, role-playing can foster students to make critical choices to reach their goals. In such a scenario, students need to strategize first in order to apply knowledge to new domains as the game is going through a number of states which follow one another in a dynamic way.

Summarizing the above, most mini-languages are based on general-purpose programming languages such as Pascal, Java or C and they were designed as edutainment environments for both beginners and experienced programmers of K12 ages, while few of them were designed for university freshmen. There are rarely found competitive role-playing 3D video games based on LOGO, especially designed for elementary school students without any prerequisites or software dependencies (e.g., IDEs, compilers). The underlying motive for this implementation of LOGO was the attempt to discover new areas of interest in programming education using game-based learning scenarios.

3 LOGO Like Environments and Similar Projects

During the design phase, the development team of the PlayLOGO 3D project studied several educational software packages based on LOGO language. Studied factors include the 3D functionality, the educational orientation of the implementation, the interface design and the ease of use.

The first ‘turtle’ robot was created in 1969 at M.I.T., and it was based on a virtual turtle robot. It was designed to be an educational tool, primary for children. Its subject was the movement in two dimensions and this was done by typing words on a keyboard. The history of LOGO after that includes many different implementations, which belong to different categories. For example, Lego Logo is a special Logo implementation with an interesting human-computer interaction. It focuses on education, but uses Lego bricks, the well known children’s toy, instead of a computer simulator. Along with the classic Lego bricks there are available special bricks that contain gears, motors and sensors used to build and program a robot. Lego Mindstorms is the successor of Lego Logo and combines everything from Lego Logo and robotics. LEGOsheets made programming with Lego Mindstorms more fun by ‘continuing to reward the children with increasingly powerful abilities while requiring only small increases in the skill needed’ [14].

MicroWorlds is a pure LOGO implementation for 2D turtle graphics and it became famous in Greek Elementary and Middle schools after LCSI distributed a Greek version named MicroWorlds Pro in 2002. Its basic functionality is not limited to simple movement of the turtle or creating shapes, but extends to more complicated procedural programming. Dapontes is among numerous researchers and teachers who have become enthusiasts of MicroWorlds to support programming in Greek language [8] [13].

The variety of different implementations of LOGO is impressive. A complete list of all LOGO-like environments can be found on the LOGO-Tree Project [4]. The rest of LOGO implementations examined below extends 3D functionality.

Elica appeared in 1999 by Pavel Boytchev, professor of Sofia University of Bulgaria, as one of the first LOGO implementations with 3D functionality. It can be used to visualize mathematical subjects, animate objects and create fractals. Apart from basic LOGO programming, Elica extensions allow students to experiment with design concepts and 3D animation. The build-in 3D objects library offers a starting point for 3D design and when combined with LOGO language it offers students a complete tool to build applications like a 3D chess or Towers of Hanoi.

The StarLogo is an implementation developed byMitchel Resnick, Eric Klopfer and others at the M.I.T. [6]. The most recent version is StarLogo TNG (The Next Generation), published in June 2008. It was engineered by C and Java programming languages and uses OpenGL to result in a 3D environment. The most impressive feature of StarLogo is that language elements are represented by colored blocks that fit together like puzzle pieces. Designers describe their project as a 'programmable modeling environment for exploring the workings of decentralized systems - systems that are organized without an organizer, coor- dinated without a coordinator' [32]. This implementation can be used to model real-life phenomena like traffic jams and market economies.

AquaMOOSE 3D, by Elliott and Bruckman [12] approached mathematics education using a desktop 3D environment and let the children play with a fish avatar that follows parametric equations in 3D. As in AquaMOOSE 3D students create mathematical challenges for one another to prevent mathophobia [12], in this project they create programming challenges to prevent programmophobia.

4 Introducing PlayLOGO 3D

In this section, theory, design philosophy and methodology of PlayLOGO 3D are described. Game play characteristics, major features and level design reveal the keypoints as they were crystalized from the prototyping processes to the final version of the game.

4.1 Game Based Learning and Constructivistic Background

The narrative metaphore is an efficient tool to maximize the learner's motivation. The population of interest is consisting of very young learners who love story telling, use computers and play video games. Holzinger et al. say that '...especially small children do not make a distinction between play and learning, play and work, fantasy and reality' [17]. In a construvistic approach games are understood as mediums to develop children's mental and motor-sensory abilities, while the storytelling is working closely with emotions (as an expression of fantacy) to maximize the motivation to play the game.

Most game-based learning environments for programming implement a compiler or interpreter of LOGO or other programming language and have the same very specific purpose: to familiarize the user with the geometry, specifically the movement of an object, which is usually represented by a turtle. The proposed application maintains the purpose of familiarization with the movement in space and the use of LOGO commands, but it distances from the classical implementations in integrating three new parameters: A) the movement in three dimensional space, B) the existence of game mechanics and narrative and C) creating competition between users.

Among other immersive applications (educational or entertainment) which allow navigation in 3D space, this solution differs in the following key point: the movement is accurate. Using mouse or other hand-driven input devices, all moves are approximate in a sense that there is no arithmetic representation of the moving commands. In certain video games this characteristic is preferable because speed and ease are more important. Here, each move is given by typed commands and the movements are very accurate in units of length and degrees of rotation (given as parameters). In most cases, this accuracy in players movement will reveal the winner.

Players have to carefully estimate distances and to orientate in 3D space, and then carefully design a piece of LOGO code, given line by line, to reach their target. Players actively build an initial programming mental model concerning syntax, programming set, command order and visually seperate commands from parameters. This is the constructivistic core that underlyies the game mechanics and gives the game educational effectiveness.

The atmospheric scenes of the proposed game activate the curiosity and fantasy of the players and this is in line with Malone hypotheses about what makes games fun [21]. Regarding the third hypotheses of Malore, that is the challenge, this game makes the final outcome to be uncertain up to the last moment. Rules are very simple and clear to the players while the overall cognitive workload of students does not exceed a critical limit that otherwise could negatively influence the challenge [18].

The aim of educational video games is to maximize the total educational and entertainment benefits from dealing with it. In cases a video game is designed around specific educational scopes, such as the PlayLOGO 3D project, it can be harmoniously integrated in educational activities and can meet most of the goals and specifications set by the educational process. What changes need to be addressed by the traditional educational system in order to adopt the new philosophy of educational video games is outside of the scope of this paper.

4.2 Design and Prototyping

After carefully studying similar projects, the design team crystallized the basic educational and technological requirements at the initial phase of the development. In simple words, what wanted was: A) a LOGO-like environment to practice LOGO commands, B) a 3D immersive environment, C) a serious video game application.

The followed methodology was closer to Extreme Programming than traditional system development methods (such as SSADM or the Waterfall Model). Although a limited set of educational and technological requirements was determined at the beginning of the development, the small but flexible design team managed most programming and graphics design issues by avoiding lots of dependencies within the system to reduce the cost of changes. Early in the project's life cycle, brainstorming led to horizontal prototyping with a wide range of functions mostly concerning HCI and interface design for young children. For the sake of simplicity, a limited set of functions was finally chosen to be conceived analytically in a vertical prototyping procedure (Fig. 1). The most important and complicated element was the pseudo-interpreter, that is the internal machine to process the user-typed commands. The result was a number of versions, by which the most robust beta version was finally tested and distributed.

This project was engineered in the Lite-C programming language (GameStudio, v.A8, Conitec Datensysteme GmbH). 3D models were designed with SketchUp 7.0 (Google) and machinima videos were developed with iClone 3.2, 3DExchange and CrazyTalk 5 (Reallusion Inc). 2D graphics were processed with PhotoShop (Adobe Systems Inc). A data-driven development technique (DDD) was used to decouple content development from source code.

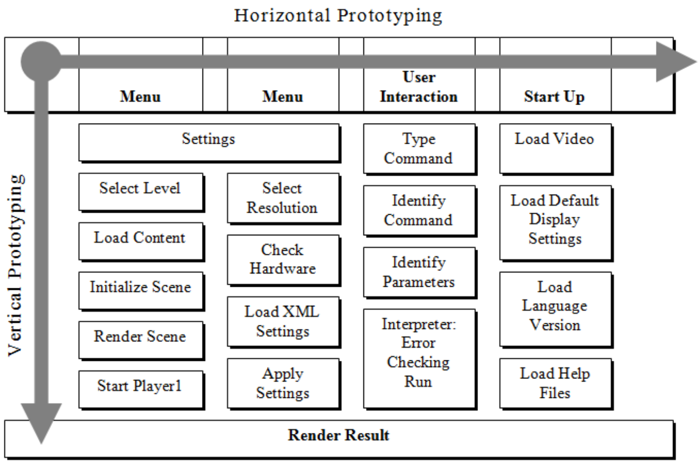


Fig. 1. Horizontal and vertical prototyping

4.3 High Concept

The scenario of the game, as described in the intro video, is a future contest for robot pilots which takes place every year in X-15 spaceship located at a constellation of Andromeda galaxy. The introductory video is used for more than one reason. Firstly, it introduces the game scenario to players (Fig. 2). This is typical to most commercial games. Secondly, the main characters (actors) explain to players the simple rules of the game in indirect way (Fig. 3). Later, players can review the help file to examine more carefully the game rules and check PlayLOGO 3D commands and syntax.

4.4 Game Overview and Features

Gameplay. Each pilot (player) drives remotely his/her robot model down in a planet's inhospitable surface (scene) while seated in an emulator at the contest platform inside the X-15 spaceship. Students play in couples and each player tries to make a collision with his/her opponent. Simple steps to reach goal are going through orientation in 3D space, lock the current position of the opponent and finally try to eliminate the distance between robots avoiding possible obstacles. Navigation is possible only by typing LOGO locomotion commands with the right syntax. During gameplay, there is no in-game vocal or textual communication between players apart from visual contact. This helps students to concentrate more on the use of LOGO locomotion commands.

The game is going through times of typing LOGO commands alternately for the two players (play in turns). After each block of commands has been typed and Enter button has been pressed, an interpretation error checking function is called. This pseudo-interpreter is also checking for data validation because some levels apply restrictions in distances and negative angles. If there are no interpretation errors, then the virtual robot executes the commands and move to a new position in 3D space. The first player who confirms a positive collision checking message from his/her robot is the winner. In this case, the other robot is destroyed and players can move to the next level. So, the collision checking of the game shows the winner depending on who sends the collision message first.

Fig. 2. Sreencapture from the intro video: The players enter the simulators

The gaming is defined as a decision making problem involving two opponent players where the outcome for each player mostly depends on the decisions taken by the other. If the current state of the game is such, one of the two players consider himself/herself as Hunter of Runaway. It is important to note that those two roles are not predefined before the game starts. Actually, it is a very sensitive and dynamically changing situation implied by the relative positions of the two players. In certain situations, one or maybe both players decide to attack because they evaluate their positions and playing order as predominant.

Avatars. Robot models are the avatars used in the game. In other words, robots are the turtles used in Microworlds and other implementations. Robots can stronger motivate the target audience and can act as a bridge between humans and machines. They are human-like in terms of body structure and at the same time they operate executing commands remotely transmitted by humans.

Fig. 3. Sreencapture from the intro video: The players enter the simulators

Camera. Since the environment is three dimensional and robots hold their orientation in space, the players cannot examine the whole virtual scene at any time. A mechanism independent of the robot's point of view was needed and this creates the sense of a target-free camera. By pressing the right mouse key a target-free camera is released to rotate the users point of view in all directions. This tool is used to scan the arena for the position of the opponent.

The Use of Keyboard. In 3D virtual environments like Second Life and also on commercial video games players use input devices like mouse and/or joystick to navigate. This kind of navigation is not precise because it simulates the physical movement of our bodies. In PlayLOGO 3D accuracy and quantification in navigation is a requirement because it simulates the result of a computer program, not a physical movement. This substantiates the choice of keyboard as the only input device to give locomotion commands and their parameters.

Table 1. The Complete List of PlayLOGO 3D Commands

No.	Command Name	Shortcut	Parameters
C1	FORWARD X	FD	X: distance
C2	BACK X	BK	X: distance
C3	LEFT F	LT	F: angle
C4	RIGHT F	RT	F: angle
C5	RISE X	RS	X: distance
C6	LOWER X	LO	X: distance
C7	SHIFTLEFT X	SL	X: distance
C8	SHIFTRIGHT X	SR	X: distance
C9	PASS	PS	{none}
C10	SET X	ST	X: distance
C11	FORWARDRISE X	FDRS	X: distance
C12	FORWARDLOWER X	FDLR	X: distance
C13	BACKRISE X	BKRS	X: distance
C14	BACKLOWER X	BKLR	X: distance
C15	PREVIOUS	PR	{none}
C16	CLEARSCREEN	SC	{none}
C17	PENUP	PU	{none}
C18	PENDOWN	PD	{none}
C19	PENCOLOR	PC	Color Name

Programming Set. This project is not full-featured for 3D design like other implementations (e.g. Elica). But there is the need to move in 3D space and thus new commands have to be included in the basic set of LOGO commands. Currently, there is no standardization for LOGO language by an international organization (like ISO or ECMA) as has been done in the past with other widely used programming languages. On the other hand, in most implementations, LOGO drawing (or moving) commands refer only to 2D space. As a solution, two more commands were imported from Elica: Rise and Lower. They need no more than a distance parameter to follow (integer data type). Note each turn moves the User Coordinate System (UCS) to the new position. The language structure, commands and parameters have intentionally been kept similar to Microworlds Pro, the most used LOGO environment in Greek schools. Currently the game is available in English and Greek. Not all LOGO commands have been used in the proposed project. The aim was not to replace any other official versions of LOGO language which are used in Greek Elementary and Middle school education, but to prepare students for later use of those environments to make school projects. A list of the available PlayLOGO 3D commands is showed in Table 1. Commands marked with an asterisk are available in a ‘plus’ version. Commands C15 to C19 are used only in design level (Raw Draw). The escape button is used to return to the Main Menu, P button for pause and right click to change camera view. Although those commands are used during the game play, it is clarified that they are used to control the game environment and they should be not considered as part of the PlayLOGO 3D programming set.

Levels. Currently, there are four levels in the game representing the corresponding arenas (Fig. 4). They are represented by futuristic scenes like surfaces of exoplanets or indoor spaceship arenas. One of them is used for training purposes before the actual contest (Raw Draw). In this extra level students can also use regular LOGO commands for drawing, plus Rise and Lower. So, the training level can be used for common LOGO drawing tasks in 3D. After a few rounds of experimentation in the training level, students get familiarize themselves with the language and syntax and can move on competition arenas.

Fig. 4. Level example: Floating Chessboard

5 Player’s Experience and Expected Educational Benefits

Primarily, an educational video game needs first to be a video game. Whether it is educational, it is by educational benefits on offer and in this example, the expected ones are:

- Familiarization with the use of a programming language. Students understand that a computer language has a predefined set of commands. No other commands -not included in the set- can be used to drive a computer when this particular language is in use.

- Learn the differences between commands used to control the environment and commands which are part of the programming language.

- Understand that each command follows some rules and those rules constitute the syntax. If the syntax of a programming language is not respected, then a compilation/interpretation error will occur.

- Understand that commands can be followed by a number of parameters. Parameters can be one or more of known data types. Parameters provide the commands with data. Although some commands (like clearscreen) do not need parameters, they still can be processed by the computer to complete a task.

- Understand that a computer cannot directly execute commands typed by the user. A compiler or interpreter needs to translate the language to machine code. If the compiler/interpreter arise an error, the user gets an error message.

- Students practice on LOGO locomotion commands. This is beneficial for later use of more formal programming tools to build school geometry, math and/or programming projects.

During game time, the optimal strategy for each player is a deterministic plan of plain locomotion commands. Those commands are typed rather than given by mouse and dictate students actions in every valid state of the game. If mouse was used to move the robots as in entertainment video games, then it would be no much educational effectiveness. In this case, the language, syntax, parameters and compiling procedure would not be visible.

6 Evaluation

6.1 Usability Heuristics

Mark Griffiths [15] argues that computer games have a very positive effect on the recreational function and a remarkable success when the games are designed to address a specific problem or reason to teach a specific skill. He also remarks some negative issues that have been taken into consideration by the design team. The first refers to the fact that video games can excite and inspire students so much that finally researchers obtain false evidence as to the motives for participation and skills of participants. Moreover, the participant’s previous experience of computer games can also affect the obtained results. This makes the evaluation of this project more challenging.

Initially, it was important to formulate a set of evaluation criteria for PlayLOGO 3D. Those should be related to usability, game play and educational effectiveness. To address the above issues, a set of 40 heuristics were developed. The Expert Review Method was used to evaluate usability of the alpha version of the game prototype. User testing and expert review methods are equally accurate in case of skilful and knowledgeable usability experts [22] [19].

Nielsen and Molichs' heuristics are of the most used usability heuristics [25] for interface design. But serious games used in education have certain differences. Moreover, Korhonen et al. imply: ‘The playability heuristic set can be extended or limited based on the needs of the evaluation' [19] and here the needs extend the pleasant gaming experience. Thus, the evaluation was mostly based on the Game Playability Heuristics (GUH) of [20] -which was implemented for Mobile Games- excluding the set of heuristics related to Mobility.

Desurvire et al. [9] proposed another powerful set of Heuristics for Evaluating Playability (HEP). Based on the hypothesis that a more extensive set of heuristics does not eliminate the chances reviewers to capture criteria violations, selected heuristics proposed by HEP were used as extensions to the current set of Korhonen & Koivisto. The selection was made having in mind the game genre of the proposed application. Although both heuristics sets are complete and powerful as standalones, finally a combination was used because some heuristics were not applicable for this kind of application.

On the other hand, Korhonen & Koivisto heuristics target only on gaming characteristics. It is widely known that educational effectiveness is hard to be proved is short periods of time and especially when important educational factors are not taken into consideration, like the curricula and teachers previous experience in GBL. Nevertheless, the educational purpose of PlayLOGO 3D prototype led the design team to add another set of heuristics in order to take feedback regarding the educational effectiveness. This does not mean that no further educational evaluation is required over time. A recently proposed methodology is Playability Heuristics for Educational Game (PHEG) which is specially designed for Educational Games [16]. From PHEG, it was used only what was missing: the subset of heuristics related to Educational-Pedagogical issues. The complete(cocktail) set of heuristics used for evaluation is shown in Tables 2, and 4. The Q40 (HEP), originally located at Game Play set of heuristics was moved to Educational-Pedagogical set with a slightly different meaning. Early in case of PlayLOGO 3D means before moving to traditional LOGO-like environments for programming tasks. Let us have in mind that the proposed video game is only the first step in a wider educational pipelined procedure related to programming and does not constitute a complete educational programming environment by itself.

6.2 Evaluation Methodology

A group of four edugame experts (and teachers by themselves) played the alpha version in couples for a few rounds to discover all of the game features. They had no more than ten minutes demonstration before actual play. This short introduction time was considered enough thanks to the simplicity and the minimalistic design of the game. Later, experts were asked to take notes with clarity and cohesion. An online survey with open-ended discussion questions directly related to selected heuristics was used to collect notes. Although the questions were translated into Greek, the original English version of the questionnaire was also available to reviewers (who have at least basic written communication skills in English) to reduce the impact of possible translation errors.

6.3 Evaluation Results

All experts mentioned that graphics and the overall interface was visually appealing. Particularly, the intro video was found very helpful in order to understand differences from the more 'traditional' LOGO environments that they had previously experienced as teachers. Although answers were given as detailed notes, in a first read they were coded as positive or negative to the related heuristic. Even in cases reviewers had given controversial answers, they were asked to take position in a positive-negative manner and they did so.

Game usability results were very encouraging (Table 2). The only not 4/4 result was related to the user manual. In Q11 ('Players do not need to use a manual to play') reviewers found that reading the user manual is necessary. One reviewer mentioned that reading the manual is not a must because the game rules are very well explained in the intro video and there is an additional in-game help screen. Regarding Q2 one expert said '...the players field of view is important for pleasure and reuse. In this game there is room for improvement'. Another reviewer advises the avatars to be visually friendlier to students, assuming that the used robot models were not.

Regarding game play (Q13-Q31) experts found some violations of the used heuristics. For example in Q18 ('The first-time experience is encouraging') half of them did not found the first experience encouraging. In Q21 ('The players can express themselves') none found that players can express themselves playing that game. This result was expected, since this project was not designed to be a full featured LOGO-like environment and application development is not possible. The same is valid for Q22 ('The game supports different playing styles'), possibly because although there are different levels, the playing style is fixed.

Table 2. Usability Evaluation Results

No.	Game Usability Heuristics	ET	E1	E2	E3	E4	Viol.
Q1	Audio-visual representation supports the game	GUH	✓	✓	✓	✓	-
Q2	Screen layout is efficient and visually pleasing	GUH	✓	✓	!	!	!
Q3	Indicators are visible	GUH	✓	✓	✓	✓	-
Q4	The player understands the terminology	GUH	✓	✓	✓	✓	-
Q5	Navigation is consistent, logical, minimalist	GUH	✓	✓	✓	✓	-
Q6	Game controls are convenient and flexible	GUH	✓	✓	✓	✓	-
Q7	The game gives feedback on the player's actions	GUH	✓	✓	✓	✓	-
Q8	The player cannot make irreversible errors	GUH	✓	✓	✓	✓	-
Q9	The player does not have to memorize things unnecessarily	GUH	✓	✓	✓	✓	-
Q10	The game contains help	GUH	✓	✓	✓	✓	-
Q11	Players do not need to use a manual to play	HEP	✓	!	!	!	!
Q12	The interface should be as non-intrusive to the player as possible	HEP	✓	✓	✓	✓	-

Q: Question number, ET: Evaluation Tool, E: Expert, Viol.: Violation found

The results of Q23 (‘The game does not stagnate’) is positive because only one reviewer found a situation where a player found obstacles resulting inability for further movements. By closing the game play evaluation, one more violation found at Q31 (‘Challenges are positive game experiences, rather than a negative experience’) where experts gave controversial results. One of them said that some times experiences are positive, while some other times are not. A second one answered positively (‘so it is true to some extend’) but with doubts.

Table 3. Gameplay Evaluation Results

No.	Game Usability Heuristics	ET	E1	E2	E3	E4	Viol.
Q13	The game provides clear goals or supports playercreated goals	GUH	✓	✓	✓	✓	-
Q14	The player sees the progress in the game and can compare the results	GUH	✓	✓	✓	✓	-
Q15	The players are rewarded and rewards are meaningful	GUH	✓	✓	✓	!	!
Q16	The player is in control	GUH	✓	✓	✓	✓	-
Q17	Challenge, strategy, and pace are in balance	GUH	✓	✓	✓	✓	-
Q18	The first-time experience is encouraging	GUH	✓	✓	!	!	!
Q19	The game story supports the gameplay and is meaningful	GUH	✓	✓	✓	✓	-
Q20	There are no repetitive or boring tasks	GUH	✓	✓	✓	✓	-
Q21	The players can express themselves	GUH	!	!	!	!	!
Q22	The game supports different playing styles	GUH	✓	!	!	!	!
Q23	The game does not stagnate	HEP	✓	!	!	!	!
Q24	The game is consistent	HEP	✓	✓	✓	✓	-
Q25	The game uses orthogonal unit differentiation	GUH	✓	✓	✓	✓	-
Q26	The player does not lose any hard-won possessions	GUH	✓	✓	✓	✓	-
Q27	There is an interesting and absorbing tutorial that mimics game play	GUH	✓	✓	✓	✓	-
Q28	The game is enjoyable to replay	GUH	✓	✓	✓	✓	-
Q29	Player should not experience being penalized repetitively for the same failure	GUH	✓	✓	✓	✓	-
Q30	Easy to learn, hard to master	GUH	✓	✓	✓	✓	-
Q31	Challenges are positive game experiences, rather than a negative experience	GUH	✓	✓	!	!	!

Q: Question number, ET: Evaluation Tool, E: Expert, Viol.: Violation found

The evaluation results related to Educational-Pedagogical heuristics (Q32-Q40) where very interesting. The first question Q32 (‘Clear goal and learning objectives’) regarding clearness of objectives gave only half positive results. Two experts found that educational objectives could be clearer. One more answered positively but mentioned that there is room for improvement. The same result comes with Q39 (‘Offers the ability to select the level of difficulty’) where two experts found that arenas truly offer varying levels of difficulty. The other two found that the level of difficulty is actually the same in all arenas or there is not enough diversity as it was expected.

Table 4. Educational-Pedagogical Evaluation Results

No.	Game Usability Heuristics	ET	E1	E2	E3	E4	Viol.
Q32	Clear goal and learning objectives	GUH	✓	✓	!	!	!
Q33	The activities are interesting and engaging	GUH	✓	✓	!	!	!
Q34	Clear and understandable structure of contents	GUH	✓	✓	✓	✓	-
Q35	Can be used as self-directed learning tools	GUH	✓	✓	!	!	!
Q36	Medium for learning by doing	GUH	✓	✓	✓	✓	-
Q37	Considers the individual differences	GUH	✓	!	!	!	!
Q38	Performance should be an outcome-based	GUH	✓	✓	✓	!	!
Q39	Offers the ability to select the level of difficulty	GUH	✓	✓	!	!	!
Q40	Player is taught skills early that you expect the players to use later, or right before the new skill is needed	GUH	✓	✓	✓	!	!

Q: Question number, ET: Evaluation Tool, E: Expert, Viol.: Violation found

The last question Q40 (‘Player is taught skills early that you expect the players to use later, or right before the new skill is needed.’) gave one criteria violation. The expert found that it is possible (this was considered as a positive answer) and another explained that he was not sure.

7 Conclusions

A new solution for applying a simplified LOGO language has been presented. With PlayLOGO 3D there are neither ready solutions, nor previously stated problems. Students try to defeat one another in an interactive narrative applying LOGO commands as ‘weapons’. Its educational effectiveness is to prepare students of Elementary Education for the actual use of LOGO language in school projects and extend the LOGO philosophy beyond two dimensions. LOGO seems to be the best choice for this project because it is widely used in Public Elementary Education as a learning programming language, most teachers can use it (especially those who have not a Computer Science background) and there is a remarkable teaching experience accumulated over the past decades. As of the final visual result, the working environment has all the characteristics of a typical video game interface and the way of use is analogous to an entertainment video game.

PlayLOGO 3D is not another typical LOGO implementation to teach advanced programming issues, but a video game about LOGO. Initially, the design team was inspired by the ‘LOGO spirit’ and ‘LOGO philosophy’ that Seymour Papert described [27]. The exuberance of a commercial computer game and the characteristics of a tight turtle graphics environment were kept in balance. Star-Logo and most of other LOGO implementations, as studied earlier, offer very sophisticated environments to build applications including video games. But those LOGO implementations are not video games in their nature. They are more like Integrated Development Environments (IDEs) as members of the LOGO family because students have to learn how to apply programming principles first. In this project, students learn the very basics of LOGO without paying conscious effort and without any prerequisites, following the principles of Game Based Learning; while having fun, they empower their spatial abilities and learn what is to drive a computer using a structured language with respect to language syntax. All of the above can be said a ‘programming pre-education’, especially designed for students who have no previous experience in any programming language.

The currently presented PlayLOGO 3D (and future versions), the users guide and instructional materials to support students and teachers can be downloaded for free at: http://www.videotutorials.gr/playlogo3d.html. The first evaluation results are encouraging and motivate the design team for future plans. Those include the distribution of a version with more levels (arenas) and a bigger set of avatars which will be constructed by users during game time, based on a library of robot components. Currently, the Artificial Intelligence of the game is under construction in order to make possible for students to play against the computer. All future versions will keep the original characteristics of the video game without downgrading its educational scope.

Acknowledgments. Game and main menu loop music was composed by the music composer Liam Bradbury especially for this project. Scenes of intro video include objects retreived by Google 3DWarehouse and the robot models are modified models retrieved by Acknex User Magazine, vol. 68.

References

1. Anderson, E.F., McLoughlin, L.: Critters in the classroom: a 3D computer-gamelike tool for teaching programming to computer animation students. In: ACM SIGGRAPH Educators Program, pp. 7–15. ACM Press, New York (2007)
2. Baer, L.: The generation gap: bridging learners and educators. J. International Digital Media & Arts Association 2, 47–52 (2005)
3. Boselli, L.: GUN-TACTYX (January 16, 2013), `http://apocalyx.sourceforge.net/guntactyx`
4. Boytchev, P.: The Logo Tree Project, versions 1.92 (2012) (January 16, 2013), `http://www.elica.net/download/papers/LogoTreeProject.pdf`
5. Brusilovsky, P., Calabrese, E., Hvorecky, J., Kouchnirenko, A., Miller, P.: Minilanguages: a way to learn programming principles. J. Education and Information Technologies 2, 65–83 (1997)
6. Colella, V., Klopfer, E., Resnick, M.: Adventures in modeling: exploring complex, dynamic systems with StarLogo. Teachers College Press, New York (2001)
7. Cooper, S., Dann, W., Pausch, R.: Teaching objects-first in introductory computer science. In: 34th SIGCSE Technical Symposium on Computer Science Education, Reno, NV, pp. 191–195 (2003)
8. Dapontes, N., Ioannou, S., Mastroyiannis, I., Tzimopoulos, N., Tsovolas, S., Alpas, A.: The Teacher as a Creator: Ideas on How to Teach MicroWorlds Pro in Kindergarten and Primary School Students. Kastaniotis Publications, Athens (2003)

9. Desurvire, H., Caplan, M., Toth, J.A.: Using Heuristics to Evaluate the Playability of Games. In: CHI Conference on Human Factors in Computing Systems, pp. 1509–1512. ACM Press, Vienna (2004)
10. Dickey, M.D.: Game Design and Learning: a Conjectural Analysis of How Massively Multiple Online Role-playing Games Foster Intrinsic Motivation. J. Education Tech. Research Dev. 55, 253–273 (2007)
11. Dondlinger, M.: Educational Video Game Design: a Review of the Literature. J. of Applied Educational Technology 4, 21–31 (2007)
12. Elliott, J., Bruckman, A.: Design of a 3D Interactive Math Learning Environment. In: International Conference on Designing Interactive Systems, pp. 64–74. ACM Press, London (2002)
13. Glezou, K., Grigoriadou, M.: Engaging Students of Senior High School in Simulation Development. J. Informatics in Education 9, 37–62 (2010)
14. Gindling, J., Ioannidou, A., Loh, J., Lokkebo, O., Repenning, A.: LEGOsheets: a Rule-based Programming, Simulation and Manipulation Environment for the LEGO Programmable Brick. In: Visual Languages, pp. 172–179. IEEE Computer Society Press, Darmstadt (1995)
15. Griffiths, M.: The Educational Benefits of Videogames. J. Education and Health 20, 47–51 (2002)
16. Hasiah, M., Bangi, S., Azizah, J.: Conceptual Framework for a Heuristics Based Methodology for Interface Evaluation of Educational Games. J. of Computer and Information Science 3, 211–219 (2010)
17. Holzinger, A., Pichler, A., Maurer, H.: Multi Media E-learning Software TRIANGLE Sase-study: Experimental Results and Lessons Learned. J. of Universal Science of Technology in Learning, 61–92 (2006)
18. Köffel, C., Haller, M.: Heuristics for the Evaluation of Tabletop Games. In: CHI Workshop: Evaluating User Experiences in Games, pp. 233–256. ACM Press, Florence (2008)
19. Korhonen, H., Paavilainen, J., Saarenenpaa, H.: Expert Review Method in Game Evaluations Comparison of Two Playability Heuristic Sets. In: MindTrek Conference, pp. 74–81. ACM Press, Tampere (2009)
20. Korhonen, H., Koivisto, E.: Playability Heuristics for Mobile Games. In: The 8th Conference on Human-Computer Interaction With Mobile Devices and Services, pp. 9–16. ACM Press, Helsinki (2006)
21. Malone, T.: Toward a Theory of Intrinsically Motivating Instruction. Cognitive & Science 4, 333–369 (1981)
22. Molich, R., Dumas, J.S.: Comparative Usability Evaluation. Behaviour & Information Technology 27, 263–281 (2008)
23. Muratet, M., Torguet, P., Viallet, F., Jessel, J.P.: Experimental Feedback on Prog&Play, a Serious Game for Programming Practice. Computer Graphics Forum, The Eurographics Association and Blackwell Publishing Ltd. 30, 61–73 (2011)
24. Nelson, M.: Robocode. IBM alphaWorks (2001)
25. Nielsen, J., Molich, R.: Heuristic Evaluation of User Interfaces. In: The CHI 1990 Conferenxe, pp. 249–256. ACM Press, Seattle (1990)
26. Pattis, R.E.: Karel the Robot: a Gentle Introduction to the Art of Programming. John Wiley & Sons (1981)
27. Papert, S.: What is Logo and Who Needs It? Logo Computer Systems Inc., Quebec (1999)
28. Pech, S.: Marvin's arena (January 16, 2013), `http://www.marvinsarena.com`

29. Petrovĭc, P.: Mathematics with Robotnăcka and Imagine Logo. In: The Eurologo, pp. 353–360. Warsaw (2005)
30. Phelps, A.M., Bierre, K.J., Parks, D.M.: MUPPETS: Multi-user Programming Pedagogy for Enhancing Traditional Study. In: Proceedings of CITC4 2003, Lafayette, Indiana, pp. 100–105 (2003)
31. Shelton, B.E., Wiley, D.: Instructional Designers Take All the Fun Out of Games: Rethinking Elements of Engagement for Designing Instructional Games. In: Annual Meeting of the American Educational Research Association, San Francisco (2006)
32. Introduction to StarLogo (January 16, 2013), `http://education.mit.edu/starlogo/`
33. Wang, D.L., Li, J., Dai, G.: Usability and Internationalization. In: Aykin, N. (ed.) HCII 2007. LNCS, vol. 4559, pp. 622–630. Springer, Heidelberg (2007)
34. Logo (January 16, 2013), `http://en.wikipedia.org/wiki/Logo_` (programming_language)

Story Guided Virtual Environments in Educational Applications

Vedad Hulusic[1,⋆] and Selma Rizvic[2]

[1] Sarajevo School of Science and Technology,
Sarajevo, Bosnia and Herzegovina
[2] Faculty of Electrical Engineering, University of Sarajevo
Sarajevo, Bosnia and Herzegovina

Abstract. Over the last few years we have witnessed a rapid development and popularisation of serious gaming. This field is becoming approved in not only education, science, medicine, religion or engineering, but also in the area of cultural heritage through serious heritage games. This can be utilised for virtual reconstructions and virtual museums and possibly used for education in the form of edutainment, comprising various techniques, such as storytelling, visual expression of information, interactivity and entertainment [19]. This paper demonstrates a new concept of using story guided virtual environments for cultural heritage virtual reconstruction, with live virtual guides in an interactive Flash format. First we compare the implementations of the same environments in x3D and Flash and then we extend the project with digital storytelling, where a user is guided through the whole application using both narrative, non-interactive, movie-like elements and interactive exploration of the virtual environment. The introduced results can be easily adopted for serious games development.

Keywords: cultural heritage, virtual reconstruction, live virtual guides, digital storytelling, interactive digital stories.

1 Introduction

For many years, games were considered only as an entertainment medium. However, serious gaming introduced several new aspects of gaming, including education, health training, scientific simulations, historical presentations, etc. Serious games, such as *Microsoft Flight Simulator* or *Global Conflict: Palestine*, can offer the same type of gameplay as commercial off-the-shelf games, or they can have different type of interaction and presentation, such as *Houthoff Buruma - The Game*.

One subcategory of serious games are serious heritage games [19]. These games involve cultural heritage objects or sites, usually in a form of virtual reconstructions or virtual museums. The former is typically used either for rebuilding

⋆ vedad.hulusic@ssst.edu.ba

Z. Pan et al. (Eds.): Transactions on Edutainment IX, LNCS 7544, pp. 132–149, 2013.
© Springer-Verlag Berlin Heidelberg 2013

non-existing objects or sites, or for digital 2D or 3D representation of an important existing object or site. The latter is used for popularisation of real museums providing world wide availability via internet or for allowing better examination of artefacts that are for example prohibited from touching. Additionally, in this type of environments, avatars or virtual guides may be employed for storytelling, which enhances the user experience by providing a higher level of immersion and presenting the important information in a narrative form.

In this paper we present a novel way of combining live virtual guides with computer generated heritage sites within Flash environment. The application is optimised for web presentations and does not require any additional software installations. Furthermore, this concept allows offline rendering. Since, even on the high-end machines it is still not possible to perform physically-based rendering at satisfactory interactive frame rates, our concept enables us to pre-render images using physically-based rendering at no cost for the user.

The paper is organised as follows: in Section 2 we give an overview of the work done so far on virtual reconstructions, virtual museums, digital storytelling and virtual guides; Section 3 briefly describes the project of the virtual reconstruction of the Church of the Holy Trinity in Mostar; in Section 4 the creation of the virtual environment in Flash is described; in Section 5 the pilot user study comparing Flash and X3D implementations is presented; Section 6 explains the proposed concept of story guided virtual environments, presents the initial user feedback and discusses potential utilisation of the concept in serious games; and finally, the conclusions and the future work are given in section 7.

2 Related Work

Virtual Reality(VR) environments, as we know it today, have been used for almost half a century, starting with the Morton Heilig's Sensorama in 1962, which is considered to be the first immersive, multi-sensory device. With the advancement of the technology and computing algorithms, the VR field progressed, producing more realistic and interactive environments and applications, such as [26,30,41]. However, these advances may not be satisfactory for conveying information to the user. This means that, although the virtual environment may be highly realistic and immersive, important information might not reach users in an understandable and desired way. An efficient way of teaching, instructing or informing users about some objects, events, facts or people is using digital storytelling [21,32]. This technique helps involve the audience as active participants and drives the development of a plot.

2.1 Virtual Reconstruction of Cultural Heritage

Virtual reconstruction of cultural heritage sites is extremely important for preservation and presentation of both non-existing and existing historical sites and stories, culture, architecture, customs and moral values. There are numerous examples of virtual reconstructions of different historical sites from all around the world [24,40,22,28,23,37,25,35]. Another type of promotion and preservation of cultural heritage is the creation of virtual museums [9,8,7,6,3,2,13,14,16,15,18,17,12,11,10]. Virtual museums allow remote visitors to experience and explore the museums using web presentations, attract potential visitors to visit the real museums and inform the visitors in the actual museums what they can see in the museum. It also provides a detailed exploration of some preserved artefacts which had to be conserved.

2.2 Virtual Guides

There are multiple forms of conveying information to a user, such as text, audio, video, animated 3D characters, real or virtual avatars, etc. The results from a user study by Sadzak et al. showed that live virtual guides are more effective and preferable for users than computer generated virtual guides. [39]. Furthermore, forms of storytelling have been proven as a good concept for information conveyance in virtual environments [38,36]. Therefore, in this paper we used digital storytelling performed by a live virtual guide, in order to tell a user the stories about the reconstructed object and its important elements.

3 Virtual Reconstruction of the Church of the Holy Trinity in Mostar

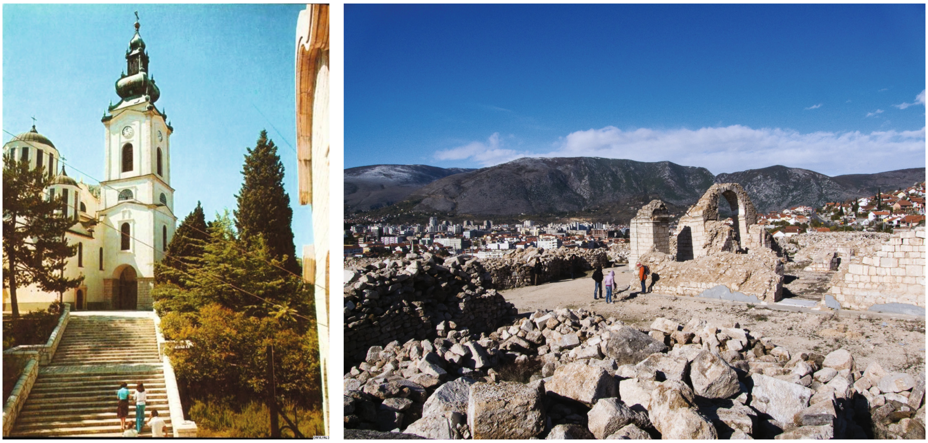


Fig. 1. The church before (left) and after the destruction (right)

The Church of the Holy Trinity in Mostar (Figure 1) was one of the most important and most significant religious and cultural objects in Bosnia and Herzegovina and the Balkans. This Orthodox church was completely destroyed in 1992, during the war in Bosnia and Herzegovina. In order to promote its beauty and significance, as well as to encourage its physical reconstruction, a virtual reconstruction has been done [31], see Figure 2. The project of virtual reconstruction included: creating 3D models of the church interior and exterior; capturing the High Dynamic Range (HDR) panoramic photograph of the church environment; developing a web and offline presentation with historical information, photo gallery, video of the church devastation, X3D interactive virtual model and a digital story with live virtual guides; and printing a 3D prototype of the model [35]. The project aroused significant media attention and the Orthodox religious community was very pleased with the results.

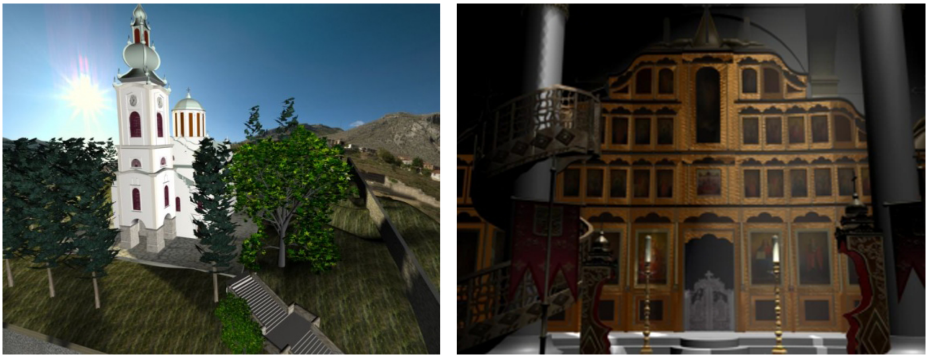


Fig. 2. Rendered images of the virtually reconstructed church: exterior (left) and interior (right)

Table 1. The comparison of the data size and download time for the X3D and our concept (Flash). The download speed used for the measurement was 3Mbps.

	Outdoor		**Indoor**		**Outdoor + Indoor**
	Total size	Initial size	Total size	Initial size	Total
X3D *(download time)*	58MB *(155s)*	32MB *(85s)*	45MB *(120s)*	28.8MB *(77s)*	103MB *(275s)*
Flash *(download time)*	5.6MB *(15s)*	5.6MB *(15s)*	5.1MB *(14s)*	5.1MB *(14s)*	10.7MB *(29s)*

However, there were a couple of drawbacks in the project. The first one was its data size. This is particularly considerable in countries with lower internet speeds, such as Bosnia and Herzegovina. Table 1 shows the model size and the download times, both for the outdoor and indoor part of the model. Although the 3D model was partially optimised, the geometry could be further optimised using polygon reduction methods. However, this process usually requires significant amount of time, as it can not be fully automated. Furthermore, the geometry itself might not be suitable for optimisation (holes in the objects, non-manifold geometry, duplicated and overlapping vetrices, edges or faces, etc.), which, therefore, requires geometry clean-up and reconstruction.

Fig. 3. The rendering view points outdoors(left) and indoors(right) marked with the red dots

The second drawback is the software requirement. Namely, a user has to install X3D player or browser plug-in for viewing the model. This is not very popular amongst internet users, as it consumes time, requires certain level of computer literacy and represents a security risk, as the user can not be sure if the program could harm his or her machine.

4 Creating the Virtual Environment in Flash

In order to improve the performance of the virtual environment, another approach was employed. Instead of exporting the 3D model to X3D, the panoramic images were rendered directly from Autodesk Maya, using the MentalRay renderer. Since Maya does not provide a panoramic lens by default, the *latlong lens* MentalRay lens shader was downloaded and used [29]. The shader enables a full 360/180 rendering of a scene in latitude/longitude format. We decided to render the indoor and outdoor scenes at three view points each, see Figure 3. During the scene set up, it was possible to further improve the appearance of the scene by adjusting the geometry, lights and materials. However, we tried to make as little intervention on the 3D model as possible.

The virtual guide was recorded against a green screen background and keyed out with the alpha channel using the Keylight plug-in in Adobe After Effects CS5, see Figure 4. The output was saved as a lossless Quicktime movie and then recoded into Flash (.swf) format using Adobe Media Encoder CS5.

For the virtual environment and user interface Tourweaver 5.00 Professional Edition was used. This software allowed us to merge the prepared images, create navigation, embed virtual guides and export it for both stand-alone (executable) and web presentation (Flash). We started with a blank new project,

Fig. 4. Keying the footage: virtual guide against the green screen (left); virtual guide keyed out with alpha channel (middle); virtual guide superimposed on the virtual model (right)

and imported the images as scenes. Then, on corresponding scenes we added Flash Media and chose the virtual guide, see Figure 5.

In the next step we added the Map Viewer and assigned it the image of the map. On each map the Radar Hotspots were set up, see Figure 3. For easier navigation, i.e. moving across the view points, Thumbnails were added at the bottom of the screen. There were three thumbnails, one for each view point, on both interior and exterior model. The final application looked as in Figure 6.

The last step in the process was to export i.e. publish the environment. Publishing options offer Flash VR and Flash VR(exe) type of exportation. The data size and the corresponding download times of the indoor and outdoor models are given in Table 1.

5 User Study: X3D vs. Flash

As shown in Section 3, using our method the data size was decreased around ten times, which speeds up the loading process dramatically. This factor might be significant, particularly for users with lower internet speeds. In order to compare how the users perceive two environments - the X3D and Flash, we made a pilot user study with 10 participants. There were 5 male and 5 female participants, aged from 24 to 43, with an average of 30. All of them had normal or corrected to normal vision and no hearing impairments. The participants were sent the instructions and the questionnaire by email. They were asked to send the questionnaire back upon completion. All participants reported to have broadband connection with varying download and upload speeds. Two participants, one male and one female, reported that they could not run the X3D environment.

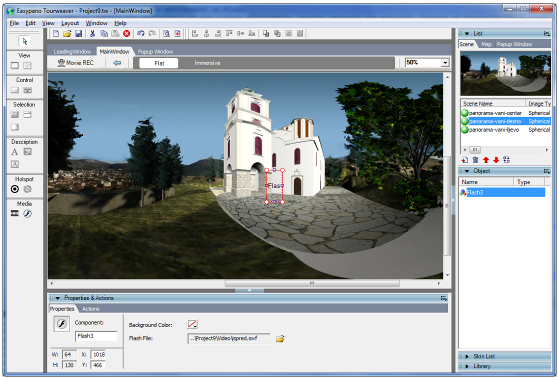


Fig. 5. The process of creating the virtual environment

This might be due to low download speed (one of the two participants reported bandwidth of 512/256 bps) or some other technical issues such as firewall. These two participants were excluded from the results evaluation.

Fig. 6. The exterior (left) and interior (right) of the virtual environment

The results gained from the pilot study are presented in Table 2. There we can see that the users were more satisfied with the Flash implementation, comparing to X3D. They reported higher loading speed, better navigation, graphics and virtual guides in the Flash application. Additionally, 50% of the participants stated that they do not mind installing new software and 50% prefer not to install any new software. Nevertheless, some users expressed concerns regarding the freedom of navigation, stating that the x3D environment offers better browsing possibilities.

Although, the pilot study was conducted with a small sample size, and therefore does not provide a statistically significant evidence, the results represent a

Table 2. The results from the pilot user study, showing the number of given answers

Question	X3D	Flash
Which application loads faster?	0	8
Which application is easier to navigate through (explore the environment)?	1	7
Which application has better graphics (looks better)?	1	7
What virtual guide (avatar) did you prefer?	1	7
Which application do you prefer in overall?	0	8
How did you feel about installing new software onto your machine:		
- I like installing new software	0	
- I do not mind installing new software	4	
- I prefer not to install new software	4	
- I never install any new software	0	

good indication of the potential of using live virtual guides in Flash environment for cultural heritage virtual reconstruction.

6 Story Guided Virtual Environment

Since the user study presented in Section 5 showed a high potential of our approach, we decided to enrich the environment, and thus user experience, introducing more elements of digital storytelling.

6.1 Concept and Implementation

By now, the presentation of the project consisted of several elements: textual information about the church (history and revival); an image gallery; a video of the church destruction; an interactive model; and a digital story [1]. Each of these elements are accessible at different locations (web pages) on the web site. Therefore, it is not guarantied that a user will visit them all and from all those various types of information be able to understand, or *feel* the whole story.

Therefore, a different concept, similar to the one proposed by the same group of authors in the Sarajevo Survival Tools project [4] (the article is in the process of submission) was utilised. This concept incorporates all the mentioned information presentation styles into one story guided virtual environment [5]. The expanded digital story guides a user through the environment, allowing him or her to interactively explore certain features and listen to the stories told by a live virtual guide.

The digital stories were created and edited with Adobe Premiere Pro CS5 software, using digital storytelling techniques introduced in [36]. The stories were then imported into Adobe Flash Professional CS5 and the controls were added, see Figure 7. During each of the three parts of the digital story, a user can skip to the end by pressing the skip button. This is introduced in case the user is exploring the project several times and does not want to listen to the

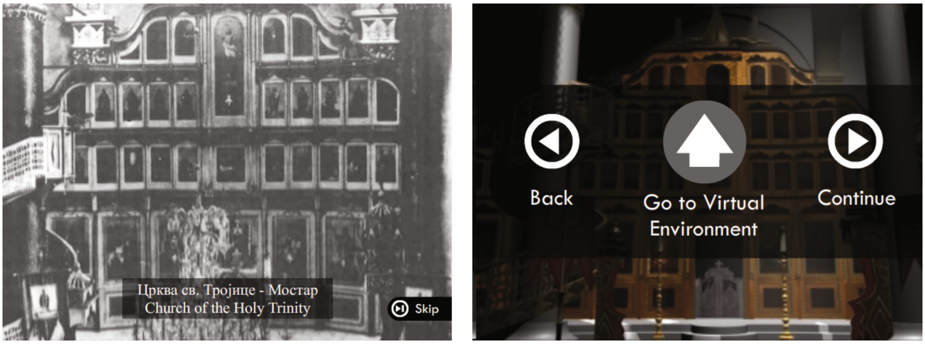


Fig. 7. Controls for navigating through the virtual environment. Skip button (left) for moving to the end of the story part being played, and the choice menu (right) presented at the end of the story parts.

Fig. 8. A screenshot of the last screen of the digital story. A user can press one of the two buttons: Restart or Home.

same story part again. Once finished or skipped, a user is given a choice: to play that part of the story again (Back button), to go to the virtual environment or the move on with the story (Continue button). At the end of the last part of the story, a user is provided with two options: to restart the whole story or to go to the project home page, see Figure 8.

Finally, an image link (Continue button) to the next part of the digital story was added to both virtual environments (indoor and outdoor), see figure 9. This way a user was enabled to go through the whole story guided virtual environment and experience it via different presentation media.

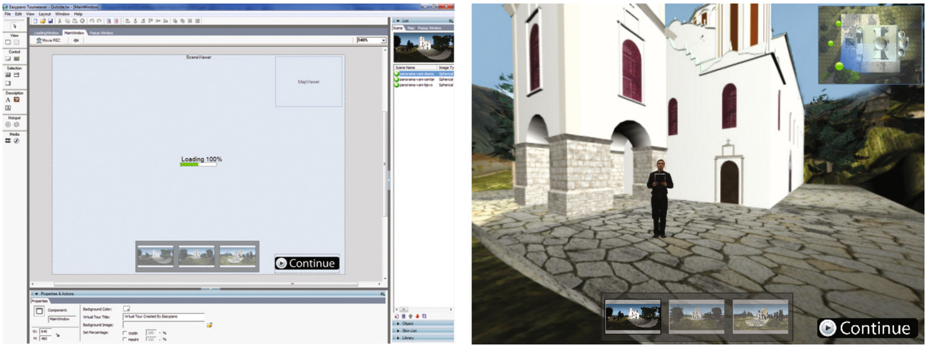


Fig. 9. Screenshots of the virtual environment exterior: development (left) and execution (right)

6.2 Initial User Feedback

In this section, the experimental design, data collection, coding and data analysis are presented. Additionally, some initial findings and ideas for potential upgrade of the system for a serious game discussed.

Experimental Design. Since practice has shown that 7 users will find approximately 80% of problems of an interface [33], a *single project study* was conducted. This is an in-depth qualitative user study, in which certain attributes are examined and possibly compared to some organizational baseline [20]. There were 9 participants, 7 male and 2 female. They aged from 25 to 68, with an average of 39. All of them reported normal or corrected to normal vision and one participant reported some hearing impairments. Therefore, the participant was excluded from the study. All the other participants reported no hearing impairments. There were two criteria for selecting participants: that they are foreigners (non-Bosnians) living outside Bosnia; and that they do not know anything about the church object.

Data Collection and Coding. For the data collection, a semi-structured questionnaire, which includes both open-ended and specific questions (see Table 3), was created and sent to the participants by email, along with the instructions document. They were asked to read the instructions, explore the story guided virtual environment, fill in the questionnaire and send back their responses.

In order to perform quantitative data analysis, data coding is necessary. This is a process of extracting qualitative data into quantitative form. In such a process the possible values of the quantitative data are created according to the given answers. Since participants often use different terms for the same phenomenon or same words for different phenomenon, it is important to perform coding as accurate as possible. In addition, one should be cautious not to lose too much information in this process.

Table 3. The questions, codes and possible values used in the study

Question	**Code**	**Possible value**
Information perception	*IP*	
Does the Church of the Holy Trinity still exist in Mostar?	**IP1**	Correct Incorrect
Who is Spasoje Vulic?	**IP2**	Correct Incorrect
What is an amvon?	**IP3**	Correct Incorrect
Who contributed to the building of the Church?	**IP4**	Correct Incorrect
Which church was used as a reference in virtual reconstruction?	**IP5**	Correct Incorrect
Immersion	*IM*	
What do you thing of the length of the interactive digital story? Is it too long, too short or long enough?	**IM1**-length	short good long
Have you experienced any problems with navigation in the virtual environments (if yes, please describe)?	**IM2**-navigation	good average bad
Presence	*P*	
Have you had a sense that you are really visiting the church, while exploring the virtual environment? Please explain why have (have not) you had that feeling?	**P1**-overall feeling of presence	Low Average High
In which environment you felt more present, exterior or interior?	**P2**-more presence int/ext	Interior Exterior both
Have you felt that the objects are real? Please explain why?	**P3**-realism	Very high High Low
Involvement	*IN*	
What kind of emotions did you feel after watching the interactive digital story?	**IN1**-emotions	
Did you enjoy yourself?	**IN2**-enjoyment	High Low
Have you felt involved in the story? Please explain why?	**IN3**-overall involvement	High Average Low
Do you think that virtual guides contributed in your involvement in the story (if yes, why)?	**IN4**-virtual guides	Yes No
How would you characterize the interactive digital story you have seen (boring, interesting, sad, other please explain)?	**IN5**-interestedness	High Low

Table 3. (*continued*)

Impressions	*IMP*	
What is your overall impression on the project?	**IMP1**-overall impression	Very bad Bad Good Very good
Do you think this project is useful and why?	**IMP2**-usefulness	High Low
Do you wish to visit the site in reality?	**IMP3**-wish to visit	No Not sure Yes
Would you like to play a serious game in this virtual environment (if yes, what could be the goal of the game)?	**IMP4**-serious game	No Maybe Yes
What you did not like in the project and why?	**IMP5**-flaws	Many A few Only one No
Have you felt dizzy while exploring the virtual environment?	**IMP6**-dizziness	No Yes

Data Analysis. The data analysis was performed in two steps: defining the hypotheses and grounding the evidence. The hypotheses were generated using the constant comparison method [27]. After coding the questions, each of them representing a particular section, we went through the data looking for patterns. From the data we built several hypotheses:

- Our concept is highly effective in conveying information about the object, providing an interesting medium for education and a good platform for edutainment;
- Our concept provides high level of enjoyment despite the nature of the presentation and evoked emotions;
- Our concept is highly important and useful for reviving the memory of the non-existing objects and attracting the visitors to the heritage site.

The aim of this study is not to prove our hypothesis, but to build up the weight of evidence supporting these propositions, that could be used as ground theories in future studies. Most of the users answered correctly on the questions related to the digital story (IP1, IP2, IP4, IP5). The reason for not having higher rate of correct answers might be the limitation of the short-term memory in humans [34]. Therefore, there were several answers “Don’t remember” (IP2, IP3, IP4, IP5). Additionally, in some countries a term “pulpit” is used for “amvon”, which could have affected the performance on this question (IP3). However, the overall rate of correct answers, presented in Table 4, represents a rather strong evidence for the first hypothesis.

Table 4. The codes and the number of answers provided

Code	Answer	Code	Answer
IP	Correct (27) Incorrect (13)	IN4	Yes (8)
IM1	Long (1) Good (6) Short (1)	IN5	High (8)
IM2	Good (4) Average (3) Bad (1)	IMP1	Very good(4) Good (4)
P1	High (2) Average (4) Low (2)	IMP2	High (8)
P2	Interior (4) Both (2) Exterior (2)	IMP3	Yes (7) Not sure (1)
P3	Very high (3) High (1) Low (2)	IMP4	Yes (3) Maybe (2) No (3)
IN1	Ashamed Sad (3) Respectful Angry (2) Impressed Curious Upset	IMP5	No (1) Only one (5) A few (3)
IN2	High (8)	IMP6	No (8)
IN3	High (4) Average (2) Low (1)		

Six out of 8 participants found the length of the story as good (IM1). One participant, that found it as long enough, also added a valuable comment: "The answer would change, of course, depending on the level of interest in the topic. It might be long for someone with a very casual interest. Someone deeply interested would probably be interested in more view with more information." This is very true and should be taken into account for any future work. The length of the presentation could also affect the mood and emotions of participants. Most of them reported feeling sad and angry/upset while watching the story (IN1). Some other emotions reported are: ashamed, respectful, impressed and curious. However, regardless of the nature of the story (savage destruction of a cultural heritage object during the war in Bosnia) and negative emotions evoked, the participants generally enjoyed the interactive digital story (IN2). A few of them reported that they enjoyed the application, but not some facts from the story. The overall impression on the project was very good (IMP1). "Well done, communicating the moral message, besides the story of destruction." and "Very good project with important content and clear message." are some of the responses (IMP1) signifying the importance of the project and the conveyed information. The usefulness of the project was rated as high (IMP2). The users found various important aspects of the project: “it uses memory to prevent us from a similar bad fate in the future, and it uses technology to repair what was damaged and-or lost. ”, “I think it helps to preserve important historical aspects.”, “Project is very important for collective memory of the local community; to explain wider public possible consequences of wars”, “I find it very useful to give the opportunity to “see” sites that are not existing and/or too far away to be visited” (IMP2).

However, semi-structured questionnaires allow for collecting not only the foreseen information but also some additional, unexpected data. Besides the mentioned, positive critics on the projects, some gaps of the proposed concept were identified. There were a few complaints and suggestions about the content: “Id like to see and learn more about e.g. the icons” (IM1), “I would like it to be made more clear who destroyed it” (IM1), “Id also like to know more about the churchs interiors and its symbolism and also how the church was used by the community by e.g. reproduce ceremonies and other celebrations” (IN3). Another element that was criticised is the navigation (IM2). A tree structured navigation, that would allow for easy wandering through the story, was proposed. Additionally, participants complained that they can not speed up the digital stories, and would like to have a timeline. One of the comments was: “provide alternative logical paths through the digital story, with the main (shorter) path and detailed branches from the main path clearly labeled” (IM1). All these are highly valuable comments and could be included in the next versions of the project in the future.

6.3 Serious Games Application

As several elements of serial gaming have been involved into this project, including virtual reconstruction of cultural heritage, entertainment, education, virtual guides, there is a potential for using such concept in serious games. The user responses were fairly positive regarding making a serious game in this virtual environment (IMP4). There were several propositions for a serious game:

– “Protect the pieces of the church during being damaged.”;
– “Something related to the icons and the stories behind them, for beginning: Memory like games.”;
– “Of course - castle defense!”.

7 Conclusions and Future Work

The technologies commonly associated with computer games are not any more solely an entertainment medium. Instead, they could be a significant factor in education, medical and scientific applications and simulations, and historical presentations. In this paper we demonstrated the effectiveness and usability of live virtual guides for storytelling purposes in a virtual cultural heritage reconstruction scenario. The live virtual guides with the alpha channel were superimposed onto panoramic images in Flash environment. The model was then compared with the same scenario implemented with the X3D virtual environment. The Flash implementation introduced many benefits comparing to the X3D application. The main advantages were data size, download speed, offline rendering capability, better image quality, and easier navigation in the environment. The main drawback of this approach is the navigational restriction to predefined view points in the environment. This, however, might be considered as a minor issue comparing to the achieved enhancements.

The results from the first user study imply that the loading speed, dependent mostly on the data size, and the possibility to run the environment is more important than the ability to freely walk around the object. Furthermore, with the Flash environment, it is possible to add as many view points as we like, and to link them to important objects or spots. That would give more freedom and spatial width to a user while browsing the environment. Additionally, for the comparison purpose, the 3D model for the Flash implementation was rendered using almost the same scene setup as for the X3D application, in order to achieve similar appearance of the outcome. Nevertheless, the models could be rendered with higher visual quality, adjusting the geometry, lighting, material and rendering settings within Maya or similar modeling and/or rendering software. Those steps could have been directly implemented in the model creation pipeline, if the model was intended for high-fidelity rendering purposes, which, in that case, would not require any additional effort.

In the second part of the project, various elements of digital storytelling and information presentation were combined into a story guided virtual environment. This type of environment guides a user through the story about the cultural heritage object, giving him or her opportunity to explore the 3D virtual model and learn more from the live virtual guides. The proposed navigation allows the user to skip any part of the digital story or move on from the virtual environment to the next story. The initial user feedback, described and presented in Section 6.2, supports the stated hypotheses. Namely, the data analysis indicates that our concept is highly effective in information conveyance in cultural heritage virtual reconstruction applications and that it could be used as a a platform for edutainment. Furthermore, its presentation style is eminently enjoyable despite the topic and emotional state of participants. Finally, the concept is highly significant and useful for reviving the non-existing objects in the memory of the visitors and for attracting the potential visitors to the heritage site.

This is, to the best of our knowledge, the first attempt to combine various digital storytelling techniques in a Flash environment for cultural heritage virtual reconstruction. Therefore, it represents a further step from the traditional non-interactive narrative content in computer games: we propose the narrator as a live person and the environment to be interactive; and common interactive non-narrative presentations as seen in traditional video games. This concept could be used for creating educational serious games, such as quizzes or puzzle games, where storytelling is important aspect of the game. Additionally, it could be exceptionally suitable for virtual reconstructions and virtual museum applications.

In the future detailed user studies with more participants will be undertaken, comparing different concepts in terms of storytelling, interactivity, information conveyance, audio-visual quality and gameplay. We would also like to apply the same concept to some other virtual environments, with similar or different scenarios and purposes. Furthermore, it would be valuable to make a serious game using our concept and evaluate the same elements through the actual gameplay.

Acknowledgments. The authors would like to thank all the volunteers who participated in the study. In addition, we would like to thank dr Amela Karahasanovic for her valuable input on designing the qualitative user study and analysing the data.

References

1. Church of the Holy Trinity, `http://crkva-sv-trojice.ba`
2. Hermitage Museum, `http://www.hermitage.ru`
3. Louvre, `http://www.louvre.fr/`
4. Sarajevo Survival Tools Project, `http://h.etf.unsa.ba/srp/project.htm`
5. Story Guided Virtual Environment, `http://crkva-sv-trojice.ba/VR-digStory/digStory1-ENG.html`
6. The Metropolitan Museum (New York), `http://www.metmuseum.org/`
7. The National Gallery (London), `http://www.nationalgallery.org.uk`
8. The Tower of Pisa, `http://www.compart-multimedia.com/`
9. The Uffizi (Florence), `http://musa.uffizi.firenze.it/`
10. The Virtual Museum of Canada, `http://www.museevirtuel-virtualmuseum.ca/`
11. Virtual Museum of the Sarajevo Assassination, `http://h.etf.unsa.ba/vmuzej-atentata`
12. Virtual National Museum of Bosnia and Herzegovina, `http://projects.etf.unsa.ba/~vmuzej/`
13. National Museum of the American Indian (2002), `http://www.nmai.si.edu/`
14. Arco (2003), `http://www.arco-web.org/`
15. Athena (2008), `http://www.athenaeurope.org/`
16. Europeana (2008), `http://group.europeana.eu/`
17. Digital Catalogue of Stecaks (2009), `http://h.etf.unsa.ba/dig-katalog-stecaka`
18. Virtual Museum of Bosnian Traditional Objects (2009), `http://www.muzejsarajeva.ba/`
19. Anderson, E.F., McLoughlin, L., Liarokapis, F., Peters, C., Petridis, P., de Freitas, S.: Serious games in cultural heritage. In: VAST 2009: 10th International Symposium on Virtual Reality, Archaeology and Cultural Heritage - VAST-STAR, Short and Project Proceedings, Faculty of ICT, University of Malta, pp. 29–48 (2009); state-of-the-Art-Report
20. Basili, V.R., Selby, R.W., Hutchens, D.H.: Experimentation in software engineering. IEEE Trans. Softw. Eng. 12, 733–743 (1986), `http://dl.acm.org/citation.cfm?id=9775.9777`

21. Brown, S., Ladeira, I., Winterbottom, C., Blake, E.: The Effects of Mediation in a Storytelling Virtual Environment. In: Balet, O., Subsol, G., Torguet, P. (eds.) ICVS 2003. LNCS, vol. 2897, pp. 102–111. Springer, Heidelberg (2003)
22. Debevec, P.: Making "the parthenon" (2004)
23. Earl, G.P.: Wandering the house of the birds: reconstruction and perception at roman italica. In: VAST 2005: Proceedings of the 2005 Conference on Virtual Reality, Archeology, and Cultural Heritage, Short Papers (2005)
24. Foni, A., Papagiannakis, G., Magnenat-Thalmann, N.: Virtual hagia sophia: Restitution, visualization and virtual life simulation. In: UNESCO World Heritage Congress Proceedings (October 2002)
25. Frischer, B., Abernathy, D., Guidi, G., Myers, J., Thibodeau, C., Salvemini, A., Müller, P., Hofstee, P., Minor, B.: Rome reborn. In: SIGGRAPH 2008: ACM SIGGRAPH 2008 New. Tech. Demos, p. 1. ACM, New York (2008)
26. Gaitatzes, A., Christopoulos, D., Roussou, M.: Reviving the past: cultural heritage meets virtual reality. In: VAST 2001: Proceedings of the 2001 Conference on Virtual Reality, Archeology, and Cultural Heritage, pp. 103–110. ACM, New York (2001)
27. Glaser, B., Strauss, A.: The discovery of grounded theory: Strategies for qualitative research. Aldine Publ. (1977), `http://books.google.com/books?id=rtiNK68Xt08C`
28. Gutierrez, D., Seron, F.J., Magallon, J.A., Sobreviela, E.J., Latorre, P.: Archaeological and cultural heritage: bringing life to an unearthed muslim suburb in an immersive environment. Journal of Cultural Heritage 5(1), 63–74 (2004), `http://www.sciencedirect.com/science/article/B6W6G-4BY308Y-7/2/5914a86f62db9bfe79a5c2844600f6e5`
29. Habel, R.: Create your own hdr environment map in maya/mental ray for image based lighting. `http://www.vi-motion.de/latlong_Tutorial/`
30. Johnson, A., Leigh, J., Carter, B., Sosnoski, J., Jones, S.: Virtual harlem. IEEE Computer Graphics and Applications 22, 61–67 (2002)
31. Karkin, Z., Rizvic, S.: Virtuelna 3d rekonstrukcija crkve svete trojice u mostaru. In: International Symposium on Digitalization of Cultural Heritage in Bosnia and Herzegovina (2008)
32. Ladeira, I., Blake, E.H.: Virtual san storytelling for children: Content vs. experience. In: VAST, pp. 223–231 (2004)
33. Lazar, J., Feng, J.H., Hochheiser, H.: Research Methods in Human-Computer Interaction. Wiley (2010), `http://www.wiley.com/WileyCDA/WileyTitle/productCd-EHEP001660.html`
34. Peterson, L.R., Peterson, M.J.: Short-term memory retention of individual items. Journal of Experimental Psychology 58, 193–198 (1959)
35. Ramic-Brkic, B., Karkin, Z., Sadzak, A., Selimovic, D., Rizvic, S.: Augmented real-time virtual environment of the church of the holy trinity in mostar. In: Proceedings of the 2009 ACM/Eurographics International Symposium on Virtual Reality, Archaeology and Cultural Heritage, pp. 141–148 (2009)
36. Rizvic, S., Sadzak, A.: Digital storytelling - representation of bosnian intangible heritage in the virtual sarajevo project. In: International Symposium on Virtual Reality, Archaeology and Cultural Heritage (VAST 2008), Braga, Portugal (2008)
37. Rizvic, S., Sadzak, A., Buza, E., Chalmers, A.: Virtual reconstruction and digitalization of cultural heritage sites in bosnia and herzegovina. In: SEEDI 2007, Review of the National Center for Digitization, pp. 82–90. Faculty of Mathematics, Belgrade (2008)

38. Sadzak, A., Rizvic, S., Chalmers, A.: The influence of storytelling quality on the human perception of computer animation. In: International Symposium on Virtual Reality, Archaeology and Cultural Heritage (VAST 2007) (2007)
39. Sadzak, A., Rizvic, S., Dalton, C., Chalmers, A.: Information perception in virtual heritage storytelling using animated and real avatars. In: Spring Conference on Computer Graphics (2007)
40. Sundstedt, V., Chalmers, A., Martinez, P.: High fidelity reconstruction of the ancient egyptian temple of kalabsha. In: AFRIGRAPH 2004: Proceedings of the 3rd International Conference on Computer Graphics, Virtual Reality, Visualisation and Interaction in Africa, pp. 107–113. ACM, New York (2004)
41. Yeh, S.C., Newman, B., Liewer, M., Pair, J., Treskunov, A., Reger, G., Rothbaum, B.O., Difede, J., Spitalnick, J., McLay, R., Parsons, T.D., Rizzo, A.A.: A virtual iraq system for the treatment of combat-related posttraumatic stress disorder. In: VR, pp. 163–170 (2009)

Rendering Technology of 3D Digital Chinese Ink-Wash Landscape Paintings Based on Maya

Xunxiang Li

College of Fine Arts and Design, Wenzhou University, Wenzhou 325035, China
Lixunxiang@163.com

Abstract. The work includes analyzing and simulating the techniques of traditional Chinese ink-wash landscape paintings, such as the ways of brush moving and ink painting, exploring the digital rendering modes with the characteristics of wrinkled-texture paintings. Besides realizing the simulation of traditional freehand ink-wash paintings by mixing the wrinkled-texture picture with the technique of modeling and rendering based on particle deposition and stacking. This study inherits and develops the traditional ink-wash painting's aesthetic theory and aesthetics, and has great significance and reference value of the change and development of the Chinese ink-wash landscape paintings.

Keywords: 3D, Chinese ink-wash landscape paintings, Modeling and rendering.

1 Introduction

At present, at home and abroad in the field of digital painting, the more successful research is mainly in the Ink-wash- effect painting (or watercolor) based on virtual two-dimensional flat space. For the Chinese ink-wash painting, especially the landscape paintings, it has great reference value about the model theory and algorithms in the effect of paper saturated with ink, but it can not be the perfect expression of the 3D dynamic ink effect [1][2][3][4][5].

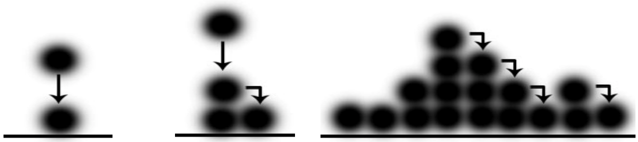
For the traditional Chinese landscape painting, the simulation is extremely difficult due to their complexity, diversity and uncertainty, the digital simulation of three-dimensional digital ink painting art is challenging. In the field of 3D ink-wash rendering, to display the ink-wash artistic effects lively and thoroughly,especially those expressed in the large freehand or small freehand ink paintings through a computer program, is a research subject of high value. Because the creation of traditional Chinese ink-wash painting is more like a manifestation of spirituality, such as the taste about flowers and birds painting, the rhyme of landscape painting, expressive portraits and so on, the proper combination and options between arts and science is crucial for expressing this "hidden or potential" artistic spirit. By using three-dimensional modeling and rendering software Maya,this paper attempts to simulate the landscape painting to solve the above problems.

Z. Pan et al. (Eds.): Transactions on Edutainment IX, LNCS 7544, pp. 150–159, 2013.
© Springer-Verlag Berlin Heidelberg 2013

2 Modeling of the Ink-Wash Landscape Painting

Of all the ink-wash landscape painting modeling,the modeling of mountain-stones is the first problem we need to solve out . Because of the self-similarity of rocks' contour, According to the most typical fractal theory, most of the 3D mountain-stones landscapes models are generated by Fractal Graphics. However, here we mainly talk about the applications of particle deposition technology in building tree and rock models or rendering in Maya[6][7].

The theory of using the particle deposition and stacking technique simulating the ink-wash wrinkle of rocks is a kind of particle deposition algorithm[8][9] . The idea of particle deposition algorithm is to make the particles fall down in order, and simulate the floating on the surface forming by the particles fell before. Enough falling particles will make it look like the viscous fluid with a flow linear structure. When a single particle falls from a high point, which is shown in Figure 1(A), putting the second particle falls above the first one and moving it until it keeps still, i.e. until all the particles next to it are not lower than it, as it shows in Figure 1(B), the other particles keep falling down and cyclically change their landing spots until a proper-sized stack effect being formed, as it shows in Figure 1(C). While particle stack is another way of deposition, it could describe the different stacking (overlapping) results made by depositions of different particle properties.



A. single falling B. both falling C. the stacking of the particles

Fig. 1. The modeling theory of the particle deposition and stacking

In the process of particle deposition and stacking (overlapping), particles could penetrate each other (make it overlap), and allow adhesive or binding (make it stack). Under the constraints of gravity, buoyancy and viscous forces, the random particles will stack up various shapes. Sometimes, the unpredictability of the shapes right meets the requirements of rich rock shapes. Figure 2 shows that under the constraints of gravity and viscous force, particles may have several potential trends in movement and conformation in the process of deposition and stacking. In addition, because of particle size, transparency, stacking (overlapping) density and boundary integration, it could make various rendering effects, Figure 3 is a group of rocks landscape views built by the author based on 3D particle stacking (overlapping) technology on Maya.

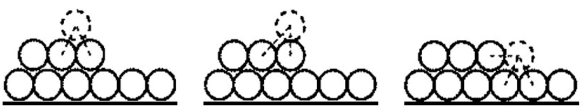


Fig. 2. Under the constraint of the gravity and viscous force, some kinds of potential movements and the structure trends of the particle deposition and stacking

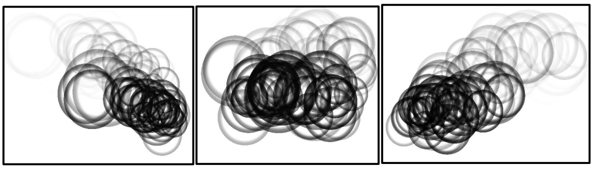


Fig. 3. A group of rocks models based on the 3D particle stacking (overlapping) technique—left, middle and right views

3 The Rendering of the Ink-Wash Landscape Painting

The 3D digital landscape painting simulation of freehand ink-wash painting need to be solved several problems: First, they must make a feature analysis of the freehand ink-wash effect based on the deep understanding about the traditional freehand landscape painting; Secondly, computer simulation process must decomposed and integrated in the process of the program processing. Traditional landscape painting can be divided into meticulous painting style (such as the green landscape painting), and freehand landscape Painting. Freehand Landscape Painting also composed by two parts, which are large freehand landscape painting and small freehand landscape painting.

Large freehand landscape painting is rare in the works of traditional landscape painting; most of these works combine these two styles to generate the atmosphere of misty rain and vigorous, lush feeling. This is a rendering technique, which combines the program texture and the program model. Therefore, 3D rendering (drawing) technology of ink-wash landscape painting contains two parts, which are the wrinkled-texture synthetic technique and the rendering technique of the rocks and trees.

3.1 The Wrinkled-Texture Synthetic Technique of the Freehand Ink-Wash Painting

The Texture synthesis is a popular technique, which can deal with the self-similar image. It uses a given input sample images to generate the output image which size is unlimited, with the naked eye you can find that the output image with the original samples are extremely similar, but not strictly the same. The author solves the problem of texture variety by using a number of the small sampling of the rocks wrinkled-texture to complete the large rocks wrinkled-texture.

The wrinkled-style is combined by linear, wrinkled, scratched, stained, and spotted, so you can get multiple combinations. Wrinkled-style usually composed by tiny stroke-lines, which combined tightly, sometime rendering layer after layer. Sometimes, the texture is just wrinkled without stained or spotted, and then it can form the texture of the local self-similarity. However, the wrinkled-texture structure is not very obvious and the shape is irregular. Because of these characteristics of wrinkled landscape painting and the complexity of the algorithm, we try Ashikhmin's natural texture synthesis method[10][11], combining with the characteristics of

Chinese landscape painting wrinkled rendering to improve and optimize this algorithm. Because the wrinkled-texture of landscape painting has a certain direction, so in the final rendering, we uses an alpha channel mask to guide the synthesis in the 3D space, designed to achieve a lively, ink-dripping, colorful artistic effect.

Ashikhmin presented the synthesis of natural texture. The so-called natural texture is the texture that combined with the small modules that very similar, but with irregular shape and size. Ashikhmin used of relevant principle, to limit the search range to the current point in the neighborhood. The method of natural texture synthesis is also using the current point of L-shaped neighborhood. The neighborhood size is Neighbor-size. It is not the direct proportion with the texture quality, the best value depends on the texture structure, too large neighborhood is not only affect the synthesis rate, but also result in a large number of repeat regions, but it needs increase the neighborhood when the texture is much smoother.

First, collecting a large number of wrinkled textures in landscape paintings saved into the system as the input sample images. For simplicity, assume that the input images and the output images have the same regular size. Using of relevant principle, the algorithm limit the search range to the neighborhood of the current point, getting the candidate pixel after inputting the offset of the corresponding position according to the L-points in the neighborhood. We define an array structure for each pixel in the output image to store the location of the pixel in the inputed image, so it is convenient to search the matching points of the neighborhood pixel. Assuming that we copy the point q in the image to the pixel p in the outputed image, and then we can build a data structure s(.), which make pixel p as the index and the equation like this:

$$S(p)=q$$

In the process of the calculation, you need to record the location of the synthetically pixel, which was shown in the input sample image. At first, the algorithm needs to initialize the array of the matching point's locations, and setting them as the random point of the inputed image, for each pixel in the outputed image, calculated in accordance with the order of scan lines. In the outputed image, considering the L-neighborhood of the current points, offsetting the corresponding location according to the location of the matching points in the array, we select this point as the candidate point to get the list of candidate points and remove the repeat points. Select the candidate point with the minimum error, which compared to the L-neighborhood of the current points in the outputed image, and copy it to the current point in the outputed image and record the location, if necessary, a second or more synthetic is acceptable, until getting a satisfactory texture. Figure 4 shows a wrinkled-texture effect of a conventional Chinese landscape painting which used the synthesis of winkled-sampling.

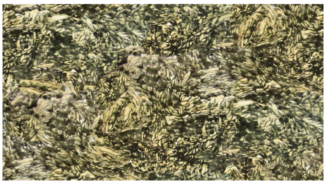


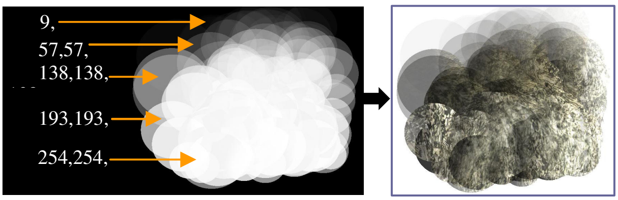
Fig. 4. The final synthesis effect of wrinkled-texture looked like folded ribbon

3.2 The Rendering Technique of the Digital Freehand Landscape Ink-Wash Effect

Using the texture synthesis algorithm just obtained the output images with a regular contour, so it is just the preparatory stage for the simulation of the digital ink-wash landscape painting. It only gets a general effect of the mountain and stone textures, and can not achieve the feeling of ink dripping, vivid spirit, if you use these textures and rocks models make mapping directly. For landscape painting of the rocks, the rock shape of the freehand landscape painting is mixed with clear and fuzzy, so it will produce a mixed effects looked like smoke or fog. Synthesis wrinkled-texture must produce a fusion effect, which combine the texture and the rocks by self-adaption. Therefore, we propose the introduction a concept of the Alpha channel mask in the process of the synthesis, thus achieving the virtual fusion effect of the irregular contours of the rocks and the wrinkled-texture.

3.2.1 The Acquisition and Definition of the Alpha Mask Channel

Alpha mask channel is an image with the similar out shape with the rocks, which is got by defining the particle density, hardness level and the degree of integration of the particle stacking (deposition) modeling. During the rendering process of the synthesis mapping about texture and the modeling, with the Participation of the alpha mask channel, for achieving a best result of the wrinkled-texture and the rocks shape, some part just shows the light ink without the wrinkle line, or (some part)just shows the wrinkle line without the light ink. In Maya, we redefined the black layer, white layer and grey layer of the Alpha mask channel and set the threshold, set the param-weight of the wrinkled-texture visibility between the grey value 0<RGB<255. the grey value of RGB changed from 255 to 0, the visibility of the wrinkled-texture is weaken gradually, and the visibility of wrinkled-texture in black area RGB（0,0,0）is 0%, the visibility of wrinkled-texture in the white area RGB(255,255,255) is 100%. Figure 5 A and B, represent the grey value of each area in the Alpha channel mask image and the corresponding display of the texture. Finally, it can generate the Alpha mask channel image by user-defining rendering.



A B

Fig. 5. A: The final synthesis effect of wrinkled-texture looked like folded ribbon. B: Texture display with the grey value.

3.2.2 The Wrinkled Rendering Synthetic of Freehand Ink-Wash Rocks

Based on the theory above, we conduct the wrinkled rendering synthetic of the 3D ink-wash rocks with the splash-ink effect. The synthetic completed in the 3D rendering engine of Maya. We enter the synthetical wrinkled-texture image, and set the rendering index of particle stacking rocks model.The system will produce the Alpha mask channel image automatically, and finally render a complete freehand landscape Painting. Figure 6(from A to B) is a mapping with the wrinkled-texture looked like folded ribbon, under the unified rocks modeling and the framework of Alpha mask channel rendering image, it formed the freehand ink-wash landscape painting effect with different wrinkled-texture.

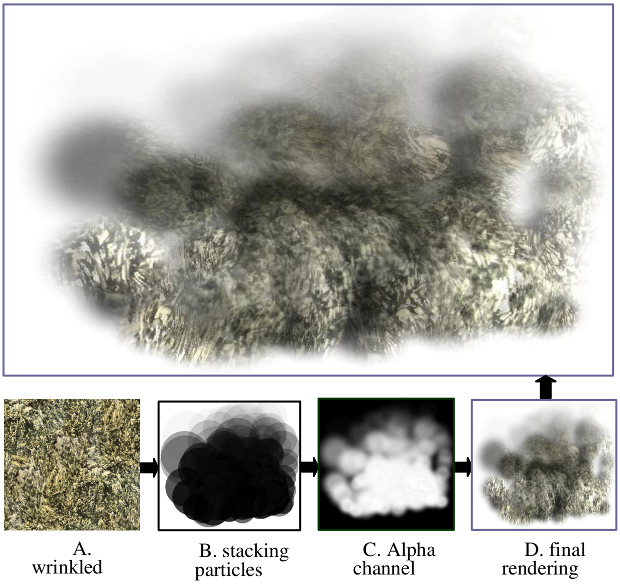


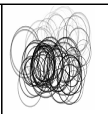
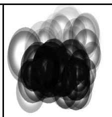
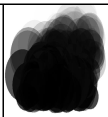
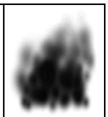
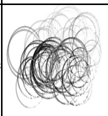
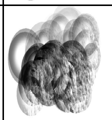
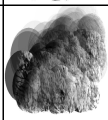
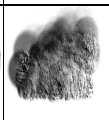
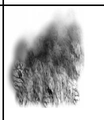
Fig. 6. The final rendering effects which produced by two different wrinkled-textures and rocks models

3.2.3 The Further Testing of the Ink-Wash Rendering Effect (Large Freehand Landscape Painting)

We conducted a further rendering tests based on the technique, which using the particle deposition and stacking for the achievement of the digital freehand landscape painting, to simulate the ink-dripping,vivid-spirit effect of the large freehand landscape painting. Before the rendering, you need to preconceive the effect you want about the digital freehand rocks based on the creation intention of the digital landscape painting, then generate the expected effect by defining the particle density, hardness level and the degree of integration of particles stacking model.

After repetitious comparing between the rocks ink-wash and the texture effects that based on the above technology, the author got the relative results that are shown in the Table 1. Each image group was divided into upper and lower Photos for the compare, the upper ones are variety of rocks with pure ink-wash rendering; the lower ones are the final rendering images with the wrinkled-texture. Figure a shows the status about particle, which softening degree is -0.9, density is 0.5, this is the process of the particle orientation and deposition. Figure b shows the status of the particle, which softening degree is -0.5 and the density is 0.5. At this status, the particles are not fusion yet, the texture and the model could not merge organically; for figure c, the softening degree of particle is 0, and the density is 1, then the particles start to merge but not very completely, there is a visible hard edge. The softening degree of particle in figure d is 0.3, density is 1, by then the particles achieve further merging, the combination of the texture and the model is more complete, but the splash-ink effect isn't clear yet. In figure e, the softening degree of the particle is 0.7, the density is 1; figure f shows the status of the particle which softening degree is 1 and the density is 1. Finally, after repetitious examination, the author thought both of figure e and figure f are successful. Their splash-ink effects have the different feeling, one is rigorous , the other is unrestrained, and you can choose any one base on the personal taste. Based on the table 1, making the further definition of the black, white and the grey layers in the alpha mask channel and threshold, setting the param-weight of the wrinkled-texture visibility between the grey value 130<RGB<255, define the texture is invisible when the grey value below 130, the rest remain the same.

Table 1. The rocks ink-wash effect rendering parameters that based on the 3D particle stacking technique (aim at the particle parameters of Softness and Stamp Density)

Without texture						
With texture						
Parameters	Figure a. -0.9; 0.5	Figure b. -0.5; 0.5	Figure c. 0; 1	Figure d. 0.3; 1	Figure e. 0.7; 1	Figure f. 1; 1

3.3 The Other Techniques Based on the Particles

The content we discussed above mainly introduced the important components of the landscape painting - the rocks modeling techniques. The next modeling techniques based on the particle deposition is only a supplement to the above modeling method and examples. It is fully consistent these ideas about modeling and the rocks modeling which produced based on the particle deposition and stacking, but some objects such as ink-wash trees, raindrop wrinkle and so on do not need the texture necessarily, or they just need a kind of colored-ink painting.

3.3.1 The Trees Rendering Based on the Particles

Assuming that the particles deposition stacks produced the particle group in accordance with the tree growth model that expanded the L system. The particle size can make self-adaptation restriction due to the tree's shape which defined by the expansion L-system, the boundaries which generated by the outline of the particle groups can produce the shape which is similar to the tree. Figure 7 shows the particles tree's shape which generated by the particle stacking. In addition, we compared normal rendering with the ink-wash rendering. It can see in the ink-wash rendering, setting the ash black as the color of the renderer, the transparency of the tube shade is completely transparent; the color is ash black, the tip fade is about 0.6, setting the softness of the brush beyond middle level.

Fig. 7. The tree's rendering illustration base on the particles

3.3.2 The 3D Wrinkled-Texture Modeling and Rendering of the Ink-Wash Landscape Painting

Ox-hair wrinkle: the ox-hair wrinkle of the traditional landscape painting is similar to the ox hair-like strokes. In Maya, we simulated the ox-hair wrinkle 3D modeling and rendering by trying to change the stacking linear form and the randomness of the complicated distribution. In addition, we set the color of the ox-hair wrinkle due to the needs. We set the hematite-color as the upper Color 1, the bottom Color 2 is black, and the two colors will integrate gradually. This effect is very similar to the traditional colored ink painting. The result is shown in Figure 8:

Fig. 8. Modeling and rendering of the ox hair wrinkle

Raindrop wrinkle: It is similar to the modeling concept of the ox-hair wrinkle, but the shape of rain wrinkle particle stacking is more symmetric and more inerratic, and it can make the groups of particles to maintain a non-uniform random state or random status. In addition, we can also make the linear controlling to the length of the raindrop wrinkle and the sharpness of the head and the end. The raindrop wrinkle just needs the shade change of the ink, and does not need the texture. The final result is shown in Figure 9.

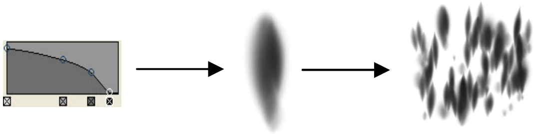


Fig. 9. The principle of raindrops wrinkle

3.3.3 The Example of the Creation About the Digital 3D Ink-Wash Landscape Painting

The For the person who has the experience about Chinese painting creation and the computer skills, creating a digital ink-wash landscape painting is quickly by using the 3D ink-wash landscape drawing modules based on above technology developed by the author. On the contrary, if you don't have any experience about such creation and the computer, you can use the templates which provide by the author to create a new work about 3D digital ink-wash landscape painting by modifying the variable values of the number, the shape, the size and the effect of the rendering. The templates include a series of theme conceptions and drawing content[12]. Figure 15 is an example about the ink-wash painting which be created by the modularized brushes.

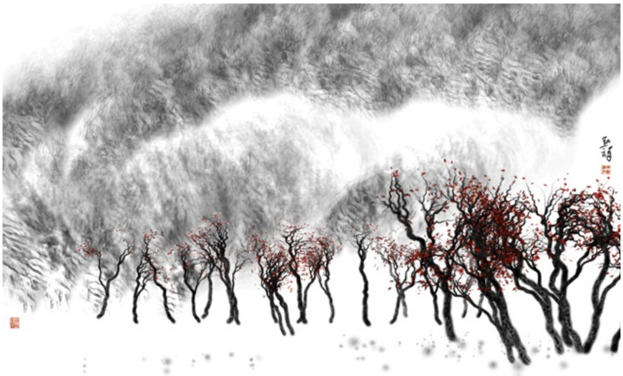


Fig. 10. A works of digital 3D ink-wash landscape painting

4 Conclusion

The work solved the problem of drawing rocks and trees of the landscape painting by using the particle stacking technique. In addition, the author made a realistic digital simulation of the traditional Chinese painting wrinkle strokes (raindrop wrinkle) by using the particle rendering technique. Compared to the pure geometric modeling and the rendering, the technique that based on the particle rendering is more powerful on the flexibility,it also has more traditional styles and characteristics. Making the rocks and trees' self customize with the Chinese painting ink-wash style by using the rendering technique to achieve the completely, thoroughly dynamic 3D ink-wash landscape painting effect. Of course, research is just the beginning, there are many other issues to be resolved. For instance, how to make the drawing of the 3D digital ink-wash landscape painting get closer to the traditional creation mode? How to simulate more traditional wrinkle texture by using the digital 3D technique? How to simulate the rocks modeling more realistically and so on, are our new problems.We will make research on these parts in the future coming digital art practice.

Acknowledgments. This paper is supported by A class key disciplines of Wenzhou University"Twelfth Five-Year Plan".

References

1. Chan, C., Akleman, E.: Two methods for creating Chinese painting. In: Proceedings of the 10th Pacific Conference on Computer Graphics and Applications (PG 2002), Beijing, pp. 403–413 (2002)
2. Way, D.-L., Shih, Z.-C.: The synthesis of rock textures in Chinese landscape painting. In: Proceedings of Eurographics 2001, Manchester, pp. 123–131 (2001)
3. Baxter, B., Scheib, V., Lin, M., et al.: DAB: interactive haptic painting with 3D virtual brushes. In: Proceedings of SIGGRAPH 2001, pp. 461–468. ACM Press, Los Angeles (2001)
4. Steve, S.: Hairy brushes. In: ACM SIGGRAPH Computer Graphics Proceedings, Dallas. Annual Conference Series, pp. 225–232 (1986)
5. Lee, J.: Simulating oriental black-ink painting. IEEE Computer Graphics & Applications 19(3), 74–81 (1999)
6. Millet, G.S.P.: The Definition and rendering of terrain maps. ACM SIGGRAPH Computer Graphics 20(4), 39–48 (1986)
7. Voss, R.D.: FRACTALS in NATURE: characterization, measurement, and Simulation. In: SIGGRAPH (1987)
8. Wang, S.-l., Zhang, L.: Game Programming All in One, pp. 451–458. Posts & Telecom Press, Beijing (2003)
9. Barabasi, A.L., Stanley, H.E.: Fractal Concepts in Surface Growth. Cambridge University Press (1995)
10. Ashikhmin, M.: Synthesizing natural textures. In: Proceedings of ACM Symposium on Interactive 3D Graphics, North Caorlina, pp. 217–226 (2001)
11. Sun, M.-J., Li, D., Sun, J.-Z.: Texture Synthesis-based System Simulation for Chinese Landscape Painting. Journal of System Simulation 16(10), 2317–2320 (2004)
12. Li, X.-X.: Creating for 3D digital Chinese ink-wash landscape paintings based on Maya. In: Proceedings of International Workshop on Digital Media and Digital Content Management (DMDCM 2011), pp. 13–17. IEEE Press (2011)

Effects of RPG on Middle School Players' Intrapersonal Intelligence

Qian Li[1], Teng Zhang[2], Bei Wang[1], and Naiyi Wang[1,*]

[1] Institute of Educational Psychology and School Counseling, Faculty of Education, Beijing Normal University, Beijing 100875, P.R. China
[2] Department of Higher Education, Graduate School of Education, Peking University, Beijing 100871, P.R. China
wangnaiyi@bnu.edu.cn

Abstract. Electronic game is a rising resource which is used to develop the players' multiple intelligences. This paper tested the intrapersonal intelligence of 192 middle school students who play role-playing game (RPG) by questionnaire, aiming at exploring the effects of RPG on intrapersonal intelligence. The results showed that RPG has a positive effect on students' intrapersonal intelligence, and the effect is subject to factors such as age, frequency of playing games and RPG type. It concluded that Role-playing game could help improving students' intrapersonal intelligence, and gave some suggestions about the development of RPG and students' intrapersonal intelligence in the future. The study highlighted a new approach of developing middle school students' multiple intelligences in the current educational environment.

Keywords: role-playing game (RPG), intrapersonal intelligence, middle school players, educational game.

1 Introduction

In recent years, with the prevalence of the theory of multiple intelligences, the question of how to apply the theory to classroom teaching has become focuses to many researchers. Moreover, outside the classrooms, researchers have begun to seek approaches to develop students' multiple intelligences. Video games, a new way of education which combines teaching with pleasure, is increasingly connected with multiple intelligences. In the book *Video games and Multiple Intelligences Cultivation* [1], the author exposited in detail the relationship between video games and the theory of multiple intelligences, and the application of video games to the development of multiple intelligences. In particular, existing experimental evidence suggests that video games have a significant effect on developing a line of individual intelligences, including interpersonal interaction [2], mathematical logic [3], body movement [4],

* Corresponding author. Institute of Educational Psychology and School Counseling, Faculty of Education, Beijing Normal University, Beijing 100875, P.R. China.

Z. Pan et al. (Eds.): Transactions on Edutainment IX, LNCS 7544, pp. 160–175, 2013.
© Springer-Verlag Berlin Heidelberg 2013

natural observation [5] and spatial intelligence [6]. The present study was based on the previous researches of video games and multiple intelligences, aiming at exploring the essential relations between role-playing game (RPG) and middle school players' intrapersonal intelligence. The study highlighted a new approach of developing middle school students' multiple intelligences in the current educational environment. Furthermore, it provided a reference for utilizing video games in quality education, and for developing games with intelligence developing functions.

2 Theory of Intrapersonal Intelligence and RPG

2.1 Intrapersonal Intelligence

As early as in 1890, William James put forward the concept of "self" for the first time. In his groundbreaking book *The Principles of Psychology*, he discussed "the consciousness of self", and suggested that the core of individuals' experiences is self [7].The founder of psychoanalytic school, Sigmund Freud, exposited that the key to mental health is self-awareness, along with the courage to face the inevitable pain and conflicts [8]. Both scholars attached much importance to the ability of being aware of oneself.

Griffin et al. [9] believed that reality can be divided into two parts. One part was assumed as an external world of objects, actions and events; and the other part would be an internal world of feelings, desires, and judgmental attitudes toward these objects and events. These latter, inner experiences can be referred to collectively as intrapersonal intelligence [9].

In *Frames of Mind: The Theory of Multiple Intelligences*, Gardner defined intelligence as an individual's capacity to fashion a product that is valued in one or more cultures; skill to work out effective solutions to real-time problems; and ability to discover new or complex problems that need to be resolved [10]. Gardner asserted that an individual has at least eight main intelligence domains, including verbal-linguistic intelligence, logical-mathematical intelligence, visual-spatial intelligence, musical-rhythmic intelligence, bodily-kinesthetic intelligence, interpersonal intelligence, intrapersonal intelligence, and naturalistic intelligence. Among these, intrapersonal intelligence was defined as the capacity to self-reflect, to have an awareness of one's strengths and weaknesses, feelings and thought processes that constitute one's knowledge of oneself [10]. Proponents of multiple intelligences suggest that an awareness of learner's strengths and weaknesses, which means the development degree of intrapersonal intelligence, can help facilitate effective instruction in education [11].

Taken together, intrapersonal intelligence is one of the multiple intelligences, through which people understand themselves and others, make plans and solve problems. It is a prerequisite to understand and well develop other intelligences and abilities. As individuals, we have to constantly re-think and re-understand our own behaviors. Thus we can adjust our behaviors and interpersonal ways in future, solve problems and understand ourselves better. A full understanding of one's inner world is the core aspect of intrapersonal intelligence [12].

2.2 Role-Playing Game

The role-playing game (RPG) is a game in which players can play one or more specific roles in a virtual world, under specific scenarios. The roles have different capabilities depending on the game plots and statistical data (such as strength, flexibility, intellectual ability, magic power, etc.). These properties vary in different plots according to the game rules. Some games improve their systems accordingly. The players play the leading roles (one or more), roaming and exploring in a given world. They could buy goods and equipments to make themselves or their teams more powerful, and fight against enemies to gain money and experiences which could upgrade themselves. RPG has the following main characteristics:

A Self-Centered Construction Activity. In the role-playing game, players are allowed to select the type or plot of a game by personal preference, to choose any partner without the restriction of age or region, and to start or exit the game whenever they want. The anonymity of RPG gives players an opportunity to attempt some behaviors which they could not carry out in a real world. Therefore, through RPG, players may gain full independence, experience a sense of satisfaction of controlling the environment, and fully express themselves [13].

Complete Freedom in Choosing a Role. Players have the autonomy to choose different roles which can tell others how the players wish to be perceived. Players can be the most powerful person or the greatest hero in the game world, even a role that is completely different from the one in real life. This, to some extent, meets the players' subconscious needs which are suppressed by the reality [14]. Further, several correlations were observed which generally support the hypothesis that individuals choose roles reflecting their own personalities. For example, individuals who scored high on extraversion tended to prefer charismatic characters. Likewise, agreeableness was positively correlated with preferences for characters with helping occupations and negatively correlated with more deviant occupations [15]. By self-identity with one's roles in RPG, the players develop a good understanding of their own personality, aspirations, status and values, therefore learn more about 'who am I', and further deepen the perception of their ideal self [16].

Various Interactions and Social Activities in a Virtual Community. Role-playing games have enormous potential to provide players with rich social experiences through various interactions along with social activities [17], such as joining a game community, team play and others [18]. The various interactions in a virtual community will facilitate the players' process of socialization through role playing. These interactions also strengthen the players' self-awareness and promote their interpersonal skills by allowing them to properly understand and evaluate their identity, strength and relationship with other players [19]. In addition, there is evidence that RPG offers a useful place where people can grow their leadership skills and potentially transfer them into the real world [17].

Multi-level System Full of Challenge and Tasks Requiring Independent Thinking. RPG supports a number of activities, such as crafting, conversation, trading, etc [20]. In the intermediate journey of the game, each individual has to become aware of their mental map, fill in aspects of their (and others') cognitive landscape by expanding current perceptions. In other words, the players have to constantly challenge and alter their pre-existing mental maps, integrate their knowledge by put it in action, look for creative solutions and address risk taking relevant to the game stakes [21]. It has been shown that RPG players score higher in divergent thinking tests than non-players [22], which indicate that RPG could be a preferable activity for the promotion of creativity through the freedom of creation and the chance to control the virtual world creatively [23].

In summary, RPG creates a virtual world in which the players could express themselves freely and independently, and gain a better understanding of their personality through self-identity with their roles in RPG. Furthermore, the rich interactions in the virtual community strengthen the players' self-awareness, as well as their relationships with others. Given the challenging tasks in RPG, the players has to constantly revisit and update their cognitive structure, and apply their knowledge into creative problem-solving processes. As a result, the educational potential of RPG has been increasingly recognized.

2.3 Intrapersonal Intelligence and RPG

In the traditional process of education, parents and schools have put much emphasis on the development of verbal-linguistic intelligence and logical-mathematical intelligence, while ignoring other types of intelligence. However, following the diversification of educational goals, the idea of developing people in an all-round way is getting more and more attention. The supporters of multiple intelligences theory advocate that people should be developed in the eight domains, and further trained to be specialists in certain fields [24]. It was suggested that instructional methods need to be varied so that students could use their intellectual strengths to better understand topics, increase their intrinsic motivation, and grasp the knowledge using their advantage on specific aspects of intelligence [24]. This calls for more diversified educational approaches and more powerful learning tools.

As a means of education, games have been widely recognized for its educational function in promoting children and adolescents' cognitive and social development. Freud proposed that games satisfy one's desire in a virtual way [25]. Alternatively, Piaget suggested that games are triggered by cognitive activities, which are strengthened by games in turn; games provide opportunities for children to consolidate their newly-acquired cognitive structure, and to develop their emotion [26]. According to Piaget's theory, various games play critical roles at each stage of children's intellectual development.

Despite their unique characteristics, video games inherit the essential attributes of games. So could video games be educational as well? Gee explored the relationship between video games and learning. He claimed that good video games combine pleasure and learning, therefore have the potential to empower people [27]. Good video games establish an innovative learning environment for students, bringing them different learning experiences from that acquired through classroom education [28].

Moreover, video games have unique advantages in developing students' multiple intelligences. For instance, puzzle games have advantages in developing logical-mathematical intelligence, and sports video games help to improve bodily-kinesthetic intelligence [29].

However, to date, studies about the effects of video games, especially RPG, on intrapersonal intelligence are scarce. As previously mentioned, the core aspect of intrapersonal intelligence is full-understanding of one's inner world. In this respect, RPG may have unique advantages since it give the players full freedom to try different sorts of roles, by which the players could reach a better understanding of themselves and their relations with the outer (though virtual) world. Hence, it seems reasonable to assume that RPG has positive effects on the players' intrapersonal intelligence.

3 Research Methods

3.1 Participants

The participants of this study were 200 middle school students from Jiangsu province of China. In order to minimize the interference of different educational levels, the participants were chosen from the schools of similar educational levels.

3.2 Research Tool

The questionnaire used in this study was based on *Middle School Student Intrapersonal Intelligence Scale* which is developed by Dr. Zhang Guoxiang of Macau University [30]. The scale consists of 15 items. Each item is used to exam the intrapersonal intelligence-related activities that the students enjoy or good at. It uses self-assessment by four-point Likert scale. If no item is excluded due to its negative discrimination value, the scale has a highest score of 60 points, a lowest score of 15 points, and a maximum score range of 45 points. The pretest showed that the actual maximum score range was about 30 points. The scale has an ideal internal consistency, with a reliability coefficient of 0.812. The validity of the scale is also high [30]. In total 200 questionnaires were given out, and 192 valid questionnaires were retrieved.

3.3 Data Analyses

Questionnaire data were analyzed using SPSS 17.0. *P* values were corrected for deviations according to Greenhouse-Geisser correction if necessary. Bonferroni correction was used for multiple comparisons.

4 Results and Discussions

4.1 A Comparative Analysis of Players and Non-players' Intrapersonal Intelligence Level

The intrapersonal intelligence levels of RPG players and non-players were shown in Figure 1. The mean scores of the 152 players and 40 non-players were contrasted by independent T test. There was a significant difference between players and non-players' intrapersonal intelligence level (t=3.90, p< .05). PRG players' intrapersonal intelligence level is higher than non- players'. This result could be explained from the following two aspects.

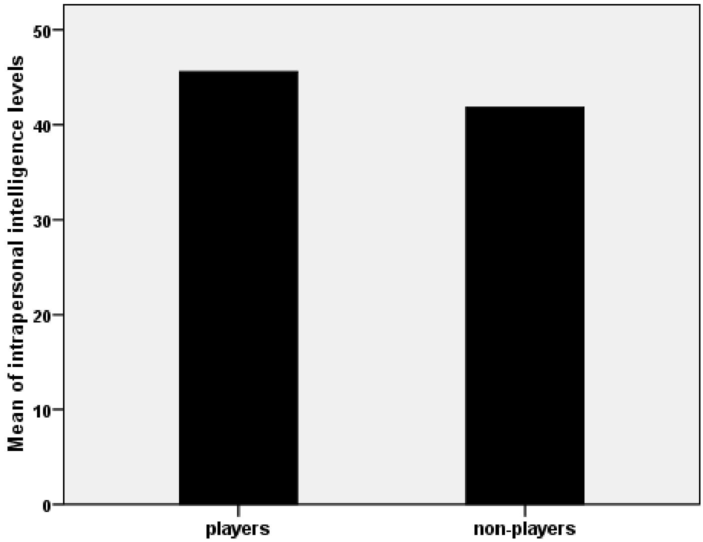


Fig. 1. The intrapersonal intelligence levels of RPG players and non-players.

Characteristics of Middle School Students' Developmental Stage. Erikson divided one's life into eight stages. Each stage is characterized by a different conflict that must be resolved by the individual. According to this theory, middle school students are at the fifth stage, where these adolescents are in search of an identity that will lead them to adulthood. Identity crisis is typified by the increasing need to find one's role in life as a productive, responsible adult, with a set of core beliefs and values. The positive side of the struggle is a sense of identity, a sense of continuity and consistency of self over time. The negative side is a sense of confusion about one's identity or role, a lack of certainty about who one is, about the part one plays in society, or more universally, in the grand scheme of life [31]. Some psychologists also suggested that adolescence is the second time period of rapid advance in self-consciousness development, in which adolescents pursuit independence and autonomy [32]. At the same time, these adolescents' cognitive ability has been substantially improved, so that they can use a more abstract and complex way to identify themselves [33]. Mowat's study demonstrated that to some extent, the majority of junior middle school students have already had developed intrapersonal intelligence [34]. It follows that middle school students are at a critical developmental stage of intrapersonal intelligence.

However, since the Chinese middle school students are under heavy academic pressure [35] and have less opportunity to contact with the outside world, it is generally hard for them to experience the feeling of success and the sense of worth. They spend most of their time following the instructions of teachers and parents, being lack of autonomy to some extent. They have relatively limited ways to achieve self-fulfillment [36]. By contrast, when playing video games, they could easily obtain a sense of autonomy and a sense of control over the external world. Lepper and Malone sorted adolescents' motivations for playing video games into two classes: personal motivation and interpersonal motivation. Personal motivation refers to motivation triggering curiosity, sense of control and fantasy, while interpersonal motivation refers to motivation triggering cooperation, competition, and sense of identity [37]. Because video games meet their needs of self-recognition and interpersonal communication, adolescents tend to spend a lot of spare time on video games.

According to "30th China Internet Development Report" released by the China Internet Network Information Center (CNNIC) in July 2012, with respect to the age distribution of Chinese netizens, 10 to 19 year-old netizens accounted for a relatively large proportion by 25.4% [38]. Further, "China Online Game Industry Report" (2012) showed that during 2006-2010, the proportion of 0 to15 year-old online players has significantly increased from 1.9% to 10.0% [39]. Therefore, as an external factor, video games have their opportunity to influence students' intrapersonal intelligence. However, to answer whether the effect is beneficial or harmful, promotive or impeditive, we need to take video games' characteristics into consideration again.

RPG's Characteristics. In his book *Teaching and Learning through Multiple Intelligence*, Campbell claimed that the most effective way to develop intrapersonal intelligence is to create an environment conducive to emotional expression. The approach to develop intrapersonal intelligence should focus on setting up of the external environment as well as guidance of individual's cognitive thinking, so that individuals could continually recognize and improve themselves [40]. In line with Campbell's idea, RPG's most significant feature is the creation of a virtual environment which could facilitate individual's self-recognition [16, 19]. Specifically, RPG involves various data such as fighting capacity, magic power, properties, etc. These data vary constantly as the game proceeds. The success of a game largely relies on the player's mastery and analyses of the data. For example, when a player's role is fighting with a monster, the player must analyze the success rate by comparing his/her own data with the monster's data, and decide to go on fighting or escape accordingly. If the player ignores the data and fights blindly, the game often ends up early and loses its attractions. In order to continue the game, the players have to learn to analyze themselves and others. In the long run, through playing RPG, middle school players' analytical ability referring to themselves and others may be improved. And the improvements might be demonstrated in their study and daily life as well. For instance, if a middle school player wants to participate in school sports meeting, he/she will analyze his/her advantages. If decides to enroll in certain sport events, he/she will think about in which aspects he/she still need to improve, and which details should be paid attention to in the course of the match, etc. In addition, he/she will analyze other competitors, making clear their strengths and weaknesses, then make a target-oriented game plan.

To sum up, the characteristics of middle school students' developmental stage determine the plasticity of their intrapersonal intelligence. And RPG provides a platform for middle school students to develop their intrapersonal intelligence. Hence, it is understandable that in the present study, the middle school RPG players had a higher intrapersonal intelligence level than non-players. However, it must be noted that while comparing players' and non-players' intrapersonal intelligence level, the impact of other factors could not be excluded. For this reason, our result does not mean that playing RPG is the necessary factor or the only reason for a higher intrapersonal intelligence level.

4.2 Other Factors Affecting Players' Intrapersonal Intelligence Levels

Gender. The mean scores of 76 female players and 76 male players were contrasted by independent T test. The result was shown in Table 1. There was not a significant difference between male and female players.

Table 1. The intrapersonal intelligence levels of female and male RPG players

	M	SD
Male	45.86	5.22
Female	45.34	7.40
t	0.45	

Age. The mean scores of 76 junior middle school players and 76 senior middle school players were contrasted by independent T test. The result was shown in Table 2. Senior middle school players' intrapersonal intelligence level was significantly higher than junior middle school players' ($p < .05$).

Physiologically speaking, although both are in their adolescence, junior middle school students are transiting from juvenile to youth. In other words, they are at the entrance of adolescence. In contrast, senior middle school students are transiting from youth to adult. In other words, they are at the exit of adolescence [41]. Therefore, senior middle school students are much better than junior ones in mental maturity. Specifically, senior middle school students show better ability in abstract thinking and logical reasoning, and ability to reflect on and monitor their thinking [41].The difference of the intrapersonal intelligence levels between junior and senior middle school students were in accord with their general mind maturity.

Table 2. The intrapersonal intelligence levels of junior and senior middle school players

	M	SD
Junior	42.19	6.02
Senior	48.68	5.14
t	6.48*	

*p<0.05.

Furthermore, despite the fact that both junior and senior middle school players play RPG for entertainment, junior players play more blindly. It is difficult for the junior middle school players to make rational analyses of the current conditions; they often operate the game impulsively and arbitrarily, with less purposiveness. In contrast, senior middle school players play game more rationally. They have more clear game goals and strategies. As a result, it is possible that RPG's promotive effects on the intrapersonal intelligence are more significant for senior middle school players than for junior ones.

Playing Frequency. In this study, the playing frequency was divided into five levels: never, occasionally, less than three days a week, more than three days a week and every day. Using the playing frequency as an independent variable, an analysis of variance (ANOVA) was carried out. The results showed that there was a significant effect of the playing frequency on the players' intrapersonal intelligence level (F=3.75, p< .05). The post-hoc multiple comparisons showed a significant difference between playing more than three days a week and playing occasionally/never play, the former resulting in a higher intrapersonal intelligence level than the latter two (ps< .05).

The development and change of middle school players' intrapersonal intelligence is not an overnight thing. It must be attributed to long-lasting stimulation of some factors. Therefore, playing RPG occasionally has little effect on the player's intrapersonal intelligence. With the playtime increases, the accumulation of the stimulation will lead to a significant raise of the intrapersonal intelligence level. From Figure 3, we can see that the players playing more than 3 days a week had the highest level of intrapersonal intelligence, while the players playing every day had lower intrapersonal intelligence level. It is probably because that playing game too frequently leads to a game addiction. In this case, the players are too immersed in the virtual world to be aware of the rules and interactions in the real world. The players could not distinguish between the virtual and the real world [42]. They may confound the game characters with the real roles, resulting in role conflicts [13]. As a consequence, the reduction of the identity of the real role will lead to a relatively low intrapersonal intelligence.

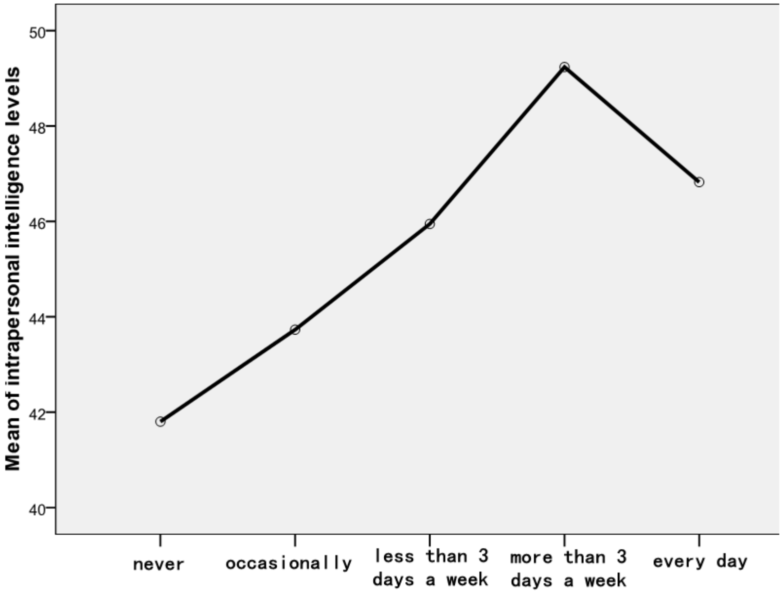


Fig. 2. The intrapersonal intelligence levels of RPG players of different playing frequencies.

The Type of RPG

Network RPG vs. Console RPG. The mean scores of network RPG players and console RPG players were contrasted by independent T-test. The result was shown in Table 3. The network RPG players' intrapersonal intelligence level is significantly higher than console RPG players' ($p < .05$).

Table 3. The intrapersonal intelligence levels of network and console RPG players

	M	SD
Network RPG	47.63	5.45
Console RPG	43.19	6.72
t	4.06*	

*$p<0.05$.

Console RPG has no real-time updating function, therefore its roles and properties are relatively fixed. The most significant feature of console RPG is that the player is playing against the computer. Due to the computer's predetermined programs, players will be familiar with the game routines after several rounds. Therefore, console RPG is not a great challenge to the players and its requirement for self-reflection is low. In contrast, network RPG supports the real-time updating, players' roles and properties are constantly changing. In addition to playing against the computer, players may also cooperate with or play against other players in a virtual community, which calls for collaboration and comprehension of self and others. In this respect, the development of the network RPG players' intrapersonal intelligence may be faster.

European, Chinese and Japanese RPG. At present, the regional backgrounds of RPG (and its roles) which Chinese students often get access to include European, Chinese and Japanese. The mean scores of the intrapersonal intelligence levels of players playing RPG of different regions were displayed in Table 4. The results of homogeneity test of variance (p>0.05) suggested that it is feasible to carry out the one way ANOVA. The results of the ANOVA were shown in Table 5.

Table 4. The intrapersonal intelligence levels of players playing RPG of different regions

	M	SD
European RPG	46.18	7.32
Chinese RPG	45.61	5.88
Japanese RPG	43.00	4.71

Table 5. The results of the one way ANOVA

	SS	DF	MS	F
Between Groups	98.06	2	49.03	1.19
Within Groups	4995.97	121	41.29	
Total	5094.02	123		

From Table 5, we can see that there was no significant effect of the game's region on the intrapersonal intelligence level. The result indicated that it is the inherent attributes of RPG, such as role playing, interpersonal interactions and challenging tasks that play an important role in enhancing middle school students' intrapersonal intelligence, rather than the region or cultural background of the RPG.

Turn-based vs. real-time RPG. The mean scores of turn-based RPG players and real-time RPG players were contrasted by independent T test. The result was shown in Table 6. The intrapersonal intelligence of real-time RPG players was significantly higher than the turn-based RPG players' (p< .05).

Table 6. The intrapersonal intelligence levels of turn-based and real-time RPG players

	M	SD
Turn-based RPG player	43.65	5.57
Real-time RPG player	47.09	6.69
t	3.05*	

*$p<0.05$.

Turn-based RPG gives players plenty of time for reflection due to the specific battle space and lacking of essential connection between two successive rounds. So its requirements of the players' analytical ability and flexibility are not high. Real-time RPG leaves players less time to think and emphasize the abilities to make instant judgment and analysis. For example, when battling with several monsters at the same time, after defeating one monster, you must analyze the remained blood and skills quickly, judging if they are enough to beat a second or third monster. Therefore, real-time RPG is more effective in training the players' ability of self-recognition and responding flexibly, and more effective in developing their intrapersonal intelligence.

5 Conclusions and Outlooks

Based on our analyses of middle school players' intrapersonal intelligence, several conclusions were drawn as follows:

(1) RPG has a significant effect on middle school students' intrapersonal intelligence. The effect is positive and promotive. Further, the players' intrapersonal intelligence is subject to factors such as age, playing frequency and RPG type. As long as a good control is conducted on these factors, RPG can be used to help improving middle school students' intrapersonal intelligence. These results are consistent with previous researches, which suggested that video games are helpful in improving the players' multiple intelligences [2-6].

(2) The mechanisms for the positive impact of RPG on the players' intrapersonal intelligence may lie in RPG's basic characteristics. That is, RPG contributes to cultivate players' composite analyzing abilities referring to themselves, others and the environment by giving them full freedom to try different sorts of roles with challenging tasks. RPG can be considered as a powerful education resource, as it can provide situational experience [17], effective social practice [17,18], a strong sense of identity [16], and shared values [19]. In addition to the entertainment and pleasure, RPG can also provide support to the development of the other types of intelligence which are neglected by classroom teaching and examination-oriented education. This conclusion might be generalized to other video games. If the developers make good use of the contents and characteristics of a certain kind of video game, it could exert its educational function on the players.

(3) Middle school period is a critical time period for the students to develop their intrapersonal intelligence. Therefore, we should explore the internal and external stimuli which are suitable for cultivating intrapersonal intelligence, and introduce them into the middle school students' study lives. Furthermore, considering the relative singleness of classroom teaching method and its limitations in developing intrapersonal intelligence, it is necessary to pursuit a combination of educational methods with different motivation and intelligence developmental patterns, and introduce more diverse approaches into education [24].

By now, the studies on the relation between RPG and middle school players' intrapersonal intelligence are still scarce. On one hand, the positive effects of RPG on the development of middle school students might not be restricted to the aspect of intrapersonal intelligence. On the other hand, there are other effective ways of improving middle school students' intrapersonal intelligence than RPG. Therefore, more researches on the relation between video games and multiple intelligences are needed in the future. In the following, the authors put forward some suggestions and outlooks for this line of studies.

(1) The present study explored the effects of RPG on middle school students' intrapersonal intelligence by contrasting players with non-players. In essence, it is a cross-sectional study, which might be interfered by many irrelevant factors. In order to increase the reliability and validity of the following studies, we suggest to combine the cross-sectional design with the longitudinal design. Specifically, by instructing the non-players of the experimental group to play certain games, and testing them after a time period, the effects of RPG could be demonstrated by a contrast of the results of the pre- and post- tests. More importantly, a contrast of the post-test results between the experimental group and the control group will rule out effects of the irrelevant factors.

(2) Existing researches have shown that RPG has positive impacts on middle school students' visual-spatial intelligence [43], musical-rhythmic intelligence [43], as well as intrapersonal intelligence. One way to generalize its positive impact to other types of intelligence is to exploit video games' advantages of individual intelligence development. In other words, we could merge characteristics of various kinds of video games into one game to make its impact more comprehensive. For instance, by adding unique effects of action games to RPG, RPG may also show a positive impact on musical-rhythmic intelligence [29,43].

(3) Not only do the entertainment and autonomy of RPG meet the needs of adolescents, the interaction in a virtual community also relieves the pressure and anxiety arising from face-to-face communications [44]. However, the addiction to RPG will lead adolescents to an isolation from the real world, being indifferent to the society and the environment, and ultimately impair their social adaptation ability [45]. Moreover, the roles, rules and situations in RPG do not constitute a one-to-one correspondence to the real life. Many adolescents are unable to transfer knowledge learned from video games to the real life. For these players, video games only provide a virtual environment in which a variety of actions could be carried out without restrictions; it is very hard to integrate the games with the realities. Hence, it is an

invaluable research area with respect to how to transfer the knowledge derived from video games to everyday life, for both educators and game developers [46].

(4) Currently, the video games in the market have far more entertaining functions than educational functions. This is partly why parents and teachers take video games as great scourges, and keep the students apart from them. On the other hand, those games with more educational functions are less popular in the market despite their good fame. Therefore, a better application of the video games to education calls for an enhancement of the educational functions of the commercial games, and an enhancement of the entertaining functions of the educational games [47]. What's more, the positive effects of RPG on the students are implicit, which could not be observed in a short term. And, only a small part of subject knowledge could be acquired from video games. Under the influence of the traditional education concept, many people have not recognized the educational potential of video games. There is still a long way to go before we can fully integrate RPG into classroom education and implement a real game-based learning [48].

Acknowledgments. This research was supported by the Fundamental Research Funds for the Central Universities (2009AAT-3).

References

1. Wang, W.: Video Games and Multiple Intelligences Cultivation, pp. 56–92. Publish House of Electronics Industry, Beijing (2009)
2. Zhang, T., Wang, B.: An Investigation of the Impact of Electronic Games on Middle School Students' Interpersonal Intelligences. Primary and Middle School Educational Technology 10, 7 (2009)
3. Lin, G.: The Impact of Educational Games on Preschool Children's Logical Mathematical Intelligence. Nanjing Normal University, Nanjing (2009)
4. Li, Y., Wang, W.: Electronic Game Classification Based on Bodily-kinesthetic Intelligence. Software Guide (Educational Technology) 1, 87–89 (2009)
5. Zhang, L., Wang, W.: The Role of Electronic Games in Promoting the Development of Naturalist Intelligence. Primary and Middle School Educational Technology 6, 28–30 (2010)
6. Yuan, L.: The Impact of Electronic Games on Middle school students' Visual Spatial Intelligence. Nanjing Normal University, Nanjing (2009)
7. William, J.: The Principles of Psychology, pp. 190–201. China City Press, Beijing (2010)
8. Esther, M.: Anna Freud's Analysis by Her Father: The Assault on the Self. Journal of Religion and Health 1, 89–95 (2001)
9. Griffin, S.: Young Children's Awareness of Their Inner World: A Neo-structural Analysis of the Development of Intrapersonal Intelligence. In: The Mind's Staircase: Exploring the Conceptual Underpinnings of Children's Thought and Knowledge, pp. 171–189. Lawrence Erlbaum Associates, Hillsdale (1991)
10. Gardner, H.: Frames of Mind: The Theory of Multiple Intelligence, pp. 76–87. China Renmin University Press, Beijing (2003)

11. Franzen, R.J.: Self-perceptions of Multiple Intelligences among Students from a Middle School in the Midwest. In: Dissertation Abstracts International Section A: Humanities and Social Science, University Microfilms, p. 82 (2000)
12. Wang, W.: Video Games and Multiple Intelligences Cultivation, pp. 15–17. Publish House of Electronics Industry, Beijing (2009)
13. Li, L.: Self-regulation Strategies of Network Game and its Impact on Teenagers' Psychological Development. Modern Primary and Secondary Education 7, 35–38 (2004)
14. Qi, W., Ge, M.: Role-playing Game and its Impact on Teenagers ' Mental Health. Chinese Journal of School Doctor 1, 109–111 (2009)
15. Park, A.E., Henley, T.B.: Personality and Fantasy Game Character Preferences. Imagination, Cognition and Personality 6, 2372–2379 (2008)
16. Jennifer, J.B.: Games People Play: Identity and Relationships in an Online Role-Playing Game. Duquesne University, Pennsylvania (2005)
17. Jang, Y.: Exploring Game Experiences and Game Leadership in Massively Multiplayer Online Role-playing Games. Brit. J. Educ. Technol. 4, 616–623 (2011)
18. Prensky, M.: Digital Game Based Learning, pp. 8–19. Paragon House Publishers, Saint Paul (2007)
19. Wang, Y.: A Research on the Influence of Role-playing Games on College Students' Role Socialization. Harbin Institute of Technology, Harbin (2011)
20. Maciuszek, D., Martens, A.: Computer Role-Playing Games as an Educational Game Genre: Activities and Reflection. In: Thomas, C., Mark, S. (eds.) European Conference on Games Based Learning, pp. 368–500. Academic Conferences Limited, Cork (2011)
21. Timplalexi, E.: A Case Study in Educational Game Designing: Junior Chemists in Action! In: Thomas, C., Mark, S. (eds.) European Conference on Games Based Learning, p. 595. Academic Conferences Limited, Cork (2011)
22. Chung, T.: Table-top Role-playing Game and Creativity. Thinking Skills and Creativity 6, 1871–1872 (2012)
23. Shaffer, D.W.: Video Games and the Future of Learning. Phi Delta Kappan 2, 105–111 (2008)
24. Akkuzu, N.: The Design of a Learning Environment Based on the Theory of Multiple Intelligence and the Study its Effectiveness on the Achievements, Attitudes and Retention of Students. In: World Conference on Information Technology, p. 3. University of Barcelona, Barcelona (2012)
25. Holowchak, M.A.: Freud on Play, Games, Sports Fanaticism. Journal of the American of Psychoanalysis and Dynamic Psychiatry 4, 695–715 (2011)
26. Yang, N.: Game Theories of Piaget. Journal of Preschool Education Studies 1, 12–14 (1994)
27. Gee, J.P.: Why video games are good for your soul: Pleasure and Learning, pp. 4–10. Common Ground Publishing, New York (2005)
28. Gee, J.P.: Good Video Games and Good Learning. Phi Kappa Phi Forum 2, 33–37 (2005)
29. Huang, Z.: How Video Games Affect the Players' Multiple Intelligences. Nanjing Normal University, Nanjing (2008)
30. Zhang, G.: Multiple Intelligences Evaluation: Two Groudbreaking School-based Experimental Studies, pp. 34–35. Educational Science Publishing House, Beijing (2007)
31. Guinee, J.P.: Erikson's Life Span Theory: A Metaphor for Conceptualization the Internship Year. Professional Psychology: Research and Practice 6, 615–620 (1998)
32. Lin, C.: Developmental Psychology, pp. 342–348. People's Education Press, Beijing (2009)

33. Harter, S.: The Development of Self-representations during Childhood and Adolescence. In: Leary, M.R., Tangney, J.P. (eds.) Handbook of Self and Identity, pp. 610–642. Guilford Press, New York (2003)
34. Mowat, J.: The Development of Intrapersonal Intelligence in Pupils Experiencing Social, Emotional and Behavioural Difficulties. Educ. Psychol. 3, 227 (2011)
35. Su, Y.D., Jia, J.F.: A Study on Middle School Students' Stresses and Coping Strategies. The Science Educational Article Collects 4, 81–81 (2004)
36. Guo, C.: Research on Adolescent Academic Self. Southwest University, Chongqing (2004)
37. Lepper, M.R., Malone, T.W.: Instinct Motivation and Instructional Effectiveness in Computer-based Education. In: Snow, R.E., Farr, M.J. (eds.) Aptitude, Learning and Instruction: Conative and Affective Process Analyses, pp. 255–286. Lawrence Erlbaum Associates, Hillsdale (1987)
38. Statistical Report on Internet Development in China, `http://www.isc.org.cn/zxzx/ywsd/listinfo-21627.html`
39. Chinese Online Game Industry Report, `http://hi.baidu.com/tangziming/item/f7e1819347691ef029164731`
40. Campbell, L.: Teaching and Learning through Multiple Intelligence, pp. 291–293. China Light Industry Press, Beijing (2001)
41. Lin, C.: Developmental Psychology, pp. 329–341. People's Education Press, Beijing (2009)
42. Matthew, G.K.: Greening the Gamescape: How Virtual Game Worlds Can Reflect Real-World Environmental Values. University of Minnesota, Minneapolis (2010)
43. Wang, W.: Video Games and Multiple Intelligences Cultivation, pp. 95–141. Publish House of Electronics Industry, Beijing (2009)
44. Jung, H.K., Chung, S.C., Jung, L.: The Effects of Escape from Self and Interpersonal Relationship on the Pathological Use of Internet Games. Community Mental Health 47, 113–121 (2011)
45. Jeffery, S.: Video Game Addiction and Depression Rates among Online Video Game Players. The Wright Institute, Berkeley (2008)
46. Zhang, L., Wang, W.: The Application Status and Problems of Electronic Game in Youth Education. Software Guide (Educational Technology) 10, 31–35 (2008)
47. Zhang, T., Wang, W.: An Investigation on the Distribution of Electronic Game Players' Multiple Intelligences. Journal of Nanjing Normal University (Natural Science Edition). Special Issue: Education and Teaching 32, 6 (2009)
48. Dickey, M.D.: Game Design and Learning: A Conjectural Analysis of How Massively Multiple Online Role-playing Games (MMORPGs) Foster Intrinsic Motivation. Educational Technology, Research and Development 3, 253–273 (2007)

Implementation of Lushan Virtual Digital Plant Museum

Jun Dai and Lifen Zhang

College of Information Science and Technology, Jiujiang University, Jiujiang 332005, China
daijun7613@qq.com

Abstract. This paper introduces the application of virtual technology in the design process of Lushan Virtual Digital Plant Museum. Based on virtual technology's characteristics of immersion, interaction and imagination,the 360-degree panorama and the three-dimension model design of virtual technology are employed to achieve the function of this museum's panoramic roaming and three-dimension plant interactive display, and demonstrate the broad prospect of virtual technology in the application of virtual digital museum.

Keywords: virtual technology, plant museum, 360-degree panorama, three-dimension model.

1 Introduction

The emergence of digital technology and the Internet followed by their swift development introduce a new dimension to the propagation of culture[1].New forms of media introduce brand new means for the propagation of cultural heritage and the promotion of museums.

Lushan plants grow in the mountains, with wreathing clouds, broad valleys and clear brooks. With more than 800 species of wild plants, Lushan landscape areas of natural vegetation accounts for 180 square kilometers, including 12 typical types of zonal distribution natural vegetation such as evergreen broad-leaved forests, deciduous broad-leaved forests, mountain coniferous forests, coniferous and broad-leaved mixed forests. The superior physiographical environment not only gives birth to the earliest plant diversity conservation base in our country, but also supplies ideal site for the research of natural succession law of forest plant. In our country's plant museum, such a superior natural condition is unique.

Therefore, it is necessary to build a web plate, which includes Lushan Plant History Museum, Special Class Museum, Happy Aquarium, Interactive Experience Area, Science Park, On-line Exchange and so on, with the help of communications, internet, and the technology of multimedia, fiction, simulation and animation. Undoubtedly, the building of Virtual Digital Plant Museum, with text, images, animation, video, fiction, simulation, interactive experience for a whole and integrated with natural landscape, history, culture and economic life, is of great significance[2]. Firstly, it is beneficial to the Lushan plant protection, inheritance and development. Secondly, it will create a convenient, rich, authoritative platform of plant science , for the public, particularly young people to form a vivid, rich

Z. Pan et al. (Eds.): Transactions on Edutainment IX, LNCS 7544, pp. 176–183, 2013.
© Springer-Verlag Berlin Heidelberg 2013

knowledge of plant science. Thirdly, it can also provide a powerful resource for the science institutions and science workers to create and develop the products, and be a basis to enhance the public scientific and cultural qualities and promote science population writing[3]. However, the development and construction of Lushan virtual plant science museum has not been founded so far.

2 Design of Lushan Virtual Digital Plant Museum

Science popularization's education under network environment is a trend. As the plant is the energy basis which the vast majority of life forms rely on and people's daily life is closely related to it, science popularization's education is extremely essential. Consequently, the plant science polarization's education based on network has become the tendency. Nowadays, plant science popularization's education taking advantage of multimedia and network technology is usually based on web forms such as words and pictures home and abroad, which can not provide the visitors with a feeling of being personally on the scene, not to mention to arouse the visitors' studying interests. In order to solve this problem, the virtual reality technology is adopted to realize the digital plant museum in the development course of this project. The home page of plant museum is shown in figure 1.

Fig. 1. Home page of Lushan Plant Museum

Lushan virtual digital plant Museum mainly by the Cultural Museum, Typical Museum, Special Museum, Science Promotion Museumand Game Pavilion of the five museums.The Cultural Museum is mainly introduced lushan plant library culture, Typical Museum is mainly introduced all kinds of plants, Special Museum mainly introduces several kinds of lushan mountain characteristics of plants, Science Promotion Museum mainly is the introduction about the plant of popular science knowledge, Game Pavilion is about plant several Flash games. Lushan virtual digital plant museum system structure is shown in Figure 2.

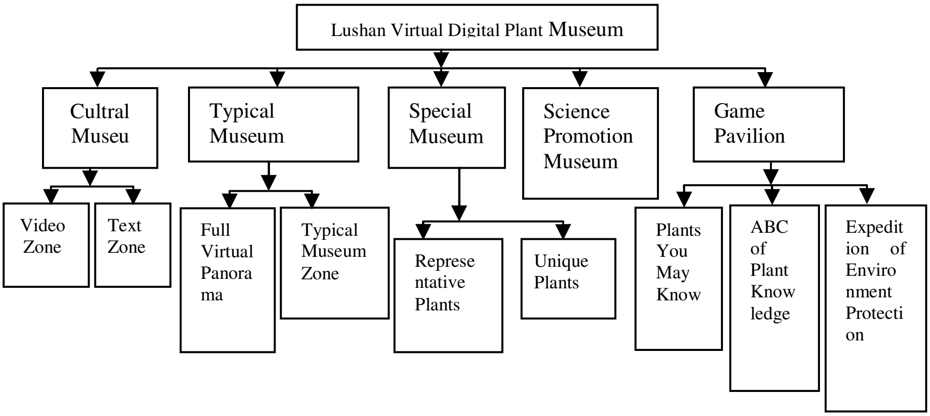


Fig. 2. System structure of lushan virtual digital plant museum

3 Implementation and Experimental Results

Virtual reality (VR for short), an immersing interactive enviroment based on calculable information, has the characteristics of immersion, interaction and imagination and and so on[4]. Virtual reality technology in plant museum is mainly composed of two aspects, one of which is the 360°Panorama technology based on images and the other one of which is the three-dimension geometry model and rendering technology based on the computer graphics.

3.1 The Applications of 360° Panorama

Based on static images, 360°Panorama is a virtual reality technology which can be achieved by computing platform. By simulating an interactive and illusive three-dimension space scene, people can carry 360°Panorama observation on the web, and by interactive operation, people can browse freely to experience personally three-dimension VR vision worlds. Compared with the previous expressive form of modeling and pictures, its advantages are as follows:

(1) It has a strong sense of reality. Based on the production and generation of true pictures, it is more believable to compare to the other modeling generating objects;

(2)It has its good interactive performance. It can express more image information than graphic pictures, and can be controlled freely;

(3) It has a strong sense of immersion. By making use of the perspective processing of image to simulate real three-dimension real scene, one would have an immersed sense;

(4) The generating is convenient, the production cycle is also short, and the production cost is low;

(5) The file is so small that the transmission is convenient, so it's suitable for Internet. Because of its diverse publish format, it's also suitable for the applications of all kinds of forms.

The steps of 360°Panorama are as follows:

(1) While taking panoramic photos, the rotation center of the ttipod must be placed in the same vertical line as the camera's rotation axis, is used to fix the camera parallelled to the ground. Meanwhile, the lens focal length being the same, shoot three cycles in the same direction, with the first cycle 60 degrees upwards, the second cycle horizontally, and the third cycle 60 degrees downwards. Shoot one picture every fixed angel. The overlapping part between adjacent pictures should be more than 15%, and the percentage of overlapping part must be the same[5].

(2) The method based on merging features whose key step is image registration, is employed to look for the feature points of the overlapping part of two images. Then, the perspective transforming relations of the feature points are used to calculate images. Finally, the images are registrated in the same coordinate. Therefore, a sphere panorama generated image is achieved by using the mosaics stitching software PTGui Pro to load the pictures, align and generate the panorama, on the basis of feature mosaic theory[6].

(3) Panoramic software Pano2VR is used to achieve the zoom function and link function between 360° rotation and hot area notes, and also used to simulate and reproduce the effect of real scene, so that the visitors can appreciate the scene and control the surrounding scene as if they are personally in the real scene. The final effect picture is shown in Figure 3.

Fig. 3. Lushan International Azalea Garden 360-degree panorama

3.2 The Creation and Display of the Three-Dimensional Virtual Plants

In order to make visitors have a real understanding of the characteristics of plants and make interactive visit possible in the introduction of special plants, special plants introduction are created and interactive display in Flash is designed by applying VRML method based on three-dimension model of virtual technology (as shown in Figure 4). Three-dimensional virtual plant model adopts two different ways: polygon model and SpeedTree plug-ins.

Fig. 4. Plants interactive display interface

(1) Polygon model
As for as a relatively simple structure of plant model is concerned, polygon model is adopted directly to create three-dimension plants[7]. Lusan stair grass is taken as an example to illustrate the creating process. First, we should create a cone, and then we transform it into editable polygon, reach the vertex level, modify the vertex, simulate roots and stems of plants. Next, we use two-dimensional graphic lines to create the outline of a leaf, add the edited polygon modifier to form the leaf's surface, reach vertex level, modify the leaf's shape, and then produce more leaves and put them into proper position. Finally, we set the paste of model using UVW paste modifier[8]. The final effect is shown in Figure 5.

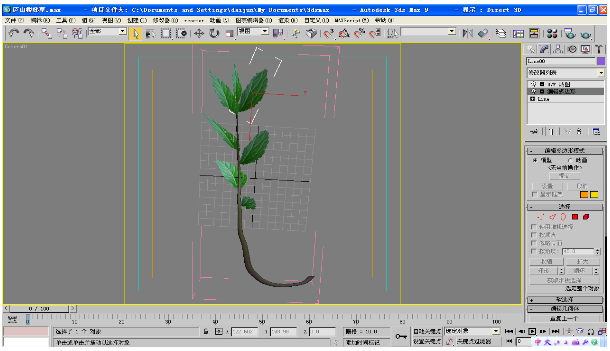


Fig. 5. Lushan stair grass model

(2) SpeedTree plug-ins model
For the plant model with a complex structure, the use of SpeedTree plug-ins can improve the efficiency of development. SpeedTree is a special three-dimension tree model software which supports big trees building and rendering quickly. The biggest characteristic of this technology is that it can simulate highly realistic situation of

trees and plants by using very few polygons. Furthermore, it contains a powerful tree gallery itself. The software is developed by American IDV company, and it is composed of three parts: SpeedTree Modeler, SpeedTree Compiler and SpeedTree SDK.Modeler. Just as its name implies, SpeedTree is mainly responsible for the modeling. Compiler is mainly used to pack the material and pinup picture for the program . SDK is used with the program to draw trees or forest modeled by SpeedTree [9]. A plant model can be divided into three parts: trunk, branch and leaves. Trunk refers to the tree itself and its biggest branch, which is basically structured by mesh. But SpeedTree doesn't store the mesh by the usual means of vertex, it uses multiple Bessel curve to store the whole mesh, which can convert to mesh at running. There are two great advantages in this way: one is that it's very convenient to make the lod of mesh, because Bessel curve is a function, which can quickly get vertex coordinates of any position, it can reduce memory usage, without the need to budget lod data like usual mesh; the other is that it's very convenient to make collision detection, it's much more efficient to check the collision by the use of curve than by the use of usual mesh [10]. Brunch refers to the smaller stems, and there are lots of branches on a real tree. If we use mesh to model this part, the burden of rendering is considerable. Therefore, for this part, the SpeedTree uses a simplified model, in which each branch is structured by two quad meshes and these two quads can form a v-shaped fissure, then render the pinup picture of branch on the quad. Usually the main branch is located in the access of two quad joints, so there will be a V-type three-dimension model when finishing the alpha blend. For the leaves, SpeedTree process is a billboard with leaves textures, which will not bear such a heavy burden when substantial rendering happens. SpeedTree can not only model plant, but also set the effect of shadows and lighting[11]. By setting the parameter, it will get the final plant model. Figure 6 shows SpeedTree interface and Figure 7 shows the final effect of Lushan brocade.

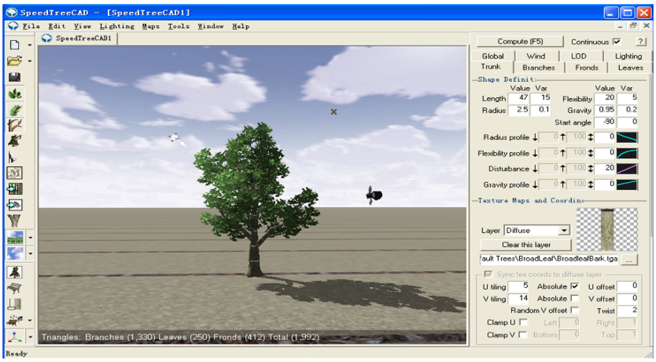


Fig. 6. SpeedTree interface

Fig. 7. Lushan brocade cuckoo rendering

4 Conclusion

By making use of the characteristics of virtual reality, virtual reality technology was applied in Lushan virtual digital plant museum in this paper, with the help of 360°Panorama technology based on images and the three-dimension geometry model technology based on computer graphics, to enable Lushan virtual herbarium's research and construction possible and meanwhile to make it the powerful resources for science popularization institutions, researchers, works and production of science popularization. Virtual reality technology also enables Lushan virtual digital Plant Museum to be the base for improving public scientific and cultural qualities and promoting science popularization creation. It is believed that along with the continuous development of virtual reality technology, nowadays, the digital virtual museum will continue to make progress and development[12], and the virtual technology, being the key technology, will also become irreplaceable in the construction of digital plant museum.

References

1. Li, W.: Museum Culture and New Media Communication. In: Preservation and Presentation of Cultural Heritage in Digital Age, pp. 46–50 (2010)
2. Yang, W., Bo, X.: Digital protection in the study of virtual reality optimization analysis. Journal of Yunnan University 32(5), 1522–1536 (2010)
3. Pei, L., Zhang, R.: Research on visual LOD technology of the Digital Museum. Electronic Test 09 (2012)
4. Luo, Y., Jun, Z., Yan, L.: Virtual reality technology and its application. Science and Technology Information (1), 63, 53 (2010)
5. Feng, H.: Virtual panorama space generation technology research and implementation. University of Electronic Technology 4 (2007)
6. Wu, Z.: Virtual panorama space and binocular vision research and implementation of 3d reconstruction. University of Electronic Technology 5 (2008)
7. Chen, S., Juan, Z.H.: Virtual forest park system modeling method research and implementation. Computer Application and Software 3, 58–61 (2010)

8. Liu, X., Hu, C.B., Xing, Z., Yan, H.: Plant growth modeling and visualization - retrospect and prospect of. Automation 27(6), 816–835 (2001)
9. Jimh, S.A.D.: Visual interface for virtual plantmodelling [J]. International Workshop on Functional Structural Plant Models 11(7), 273–276 (2004)
10. Rossignac, J., Bowel, P.: Multi-resolution 3D approximation for rendering complex scenes. In: Proceedings of the 20th Conference Modeling in Computer GraPhic, Berlin, pp. 453–465 (1993)
11. Qu, G.: Digital museum in the application of virtual reality technology. J. Museum Research (3), 1139 (2008)

Pedestrian Detection Based on Kernel Discriminative Sparse Representation

Keyang Cheng[1,2], Qirong Mao[2], and Yongzhao Zhan[2]

[1] School of Computer Science & Technology, Nanjing University of Aeronautics & Astronautics, Nanjing, Jiangsu, China, 210016
[2] School of Computer Science & Telecommunications Engineering, Jiangsu University, Zhenjiang, Jiangsu, China, 212013
kycheng@ujs.edu.cn

Abstract. This article puts forward a novel framework for pedestrian detection tasks, which proposing a model with both sparse reconstruction and class discrimination components, jointly optimized during dictionary learning. We present an efficient pedestrian detection system using mixing sparse features of HOG, FOG and CSS to combine into a Kernel classifier. Results presented on our data set show competitive accuracy and robust performance of our system outperforms current state-of-the-art work.

Keywords: Kernel Discriminative Sparse Representation, Pedestrian Detection.

1 Introduction

Pedestrian counting in public places plays a key role in many applications, such as evacuating from a dense region to a sparse one when an emergency happens, or optimizing the design of traffic infrastructures to provide better transportation services. Furthermore, social security and surveillance strongly depend on the effectiveness of pedestrian counting. A wide variety of pedestrian detection methods have been proposed [1-6].

Sparse representations have recently drawn much interest in signal, image, and video processing. Under the assumption that natural images admit a sparse decomposition in some redundant basis (or so-called dictionary), several such models have been proposed, e.g., curve lets, wedge lets, band lets and various sorts of wavelets [7]. Interestingly, while discrimination is the main goal of these papers, the optimization (dictionary design) is purely generative, based on a criterion which does not explicitly include the actual discrimination task, which is one of the key contributions of our work. In [8], a discriminative method is introduced for various classification tasks, learning one dictionary per class; the classification process itself is based on the corresponding reconstruction error, and does not exploit the actual decomposition coefficients. In [9], a generative model for documents is learned at the same time as the parameters of a deep network structure. In [10], multi-task learning is performed by learning features and tasks are selected using a sparsity criterion. The framework we present in this paper extends these approaches by learning

Z. Pan et al. (Eds.): Transactions on Edutainment IX, LNCS 7544, pp. 184–195, 2013.
© Springer-Verlag Berlin Heidelberg 2013

simultaneously a single shared dictionary as well as models for different signal classes in a mixed generative and discriminative formulation (see also [11], where a different discriminative term is added to the classical reconstructive one). Similar joint generative/discriminative frameworks have started to appear in probabilistic approaches to learning, e.g., [12, 13, 14, 15, 16], and in neural networks [17], but not, to the best of our knowledge, in the sparse dictionary learning framework.

The remainder of this paper is organized as follows. In Section 2, we describe the procedure of feature extraction, and in Section 3, we present a formulation for learning a dictionary tuned for a classification task, which we call discriminative sparse learning. Section4 gives the optimization procedure of discriminative sparse learning .Experimental results are provided and analyzed in Section 5. Finally, Section 6 concludes this work.

2 Feature Extraction

Obviously, the choice of features is the most critical decision when designing a detector, and finding good features is still largely an empirical process with few theoretical guidelines. We evaluate different combinations of features, and introduce a new feature based on the similarity of colors in different regions of the detector window, which significantly raises detection performance. The pedestrian region in our detection window is of size 48*96 pixels.

Histograms of oriented gradients (HOG) are a popular feature for object detection, first proposed in [18]. They collect gradient information in local cells into histograms using trilinear interpolation, and normalize overlapping blocks composed of neighboring cells. Interpolation, local normalization and histogram binning make the representation robust to changes in lighting conditions and small variations in pose. HOG was recently enriched by Local Binary Patterns (LBP), showing a visible improvement over standard HOG on the INRIA Person data set [24]. In our experiments we compute histograms with 9 bins on cells of 8*8 pixels. Block size is 2*2 cells overlapping by one cell size.

HOF Histograms of flow were initially also proposed by Dalal et al. [19]. We have shown that using them (e.g. in [19]'s IMHwd scheme) complementary to HOG can give substantial improvements on realistic datasets with significant ego motion. Here, we introduce a lower-dimensional variant of HOF, IMHd2, which encodes motion differences within 2*2 blocks with 4 histograms per block, while matching the performance of IMHwd (3*3 blocks with 9 histograms). Fig. 2(d) schematically illustrates the new coding scheme: the 4 squares display the encoding for one histogram each. For the first histogram, the optical flow corresponding to the pixel at the ith row and jth column of the upper left cell is subtracted from the one at the corresponding position of the lower left cell, and the resulting vector votes into a histogram as in the original HOF scheme. IMHd2 provides a dimensionality reduction of 44% (2520 instead of 4536 values per window), without changing performance significantly. We used the publicly available flow implementation of [20]. In this work we show that

HOF continues to provide a substantial improvement even for flow fields computed on JPEG images with strong block artifacts (and hence degraded flow fields).

Several authors have reported improvements by combining multiple types of low-level features [21, 22, 23]. Still, it is largely unclear which cues should best be used in addition to the now established combination of gradients and optic flow. Intuitively, additional features should be complementary to the ones already used, capturing a different part of the image statistics. Color information is such a feature enjoying popularity in image classification [24] but is nevertheless rarely used in detection. Furthermore, second order image statistics, especially co-occurrence histograms, are gaining popularity, pushing feature spaces to extremely high dimensions [25, 22].

We propose to combine these ideas and use second order statistics of colors as additional feature. Color by itself is of limited use, because colors vary across the entire spectrum both for people (respectively their clothing) and for the background, and because of the essentially unsolved color constancy problem. However, people do exhibit some structure, in that colors are locally similar—for example (see Fig. 1) the skin color of a specific person is similar on their two arms and face, and the same is true for most people's clothing. Therefore, we encode color self similarities within the descriptor window, i.e. similarities between colors in different sub-regions. To leverage the robustness of local histograms, we compute D local color histograms over 8*8 pixel blocks, using trilinear interpolation as in HOG to minimize aliasing. We experimented with different color spaces, including 3*3*3 histograms in RGB, HSV, HLS and CIE Luv space, and 4*4 histograms in normalized rg, HS and uv, discarding the intensity and only keeping the chrominance. Among these, HSV worked best, and is used in the following.

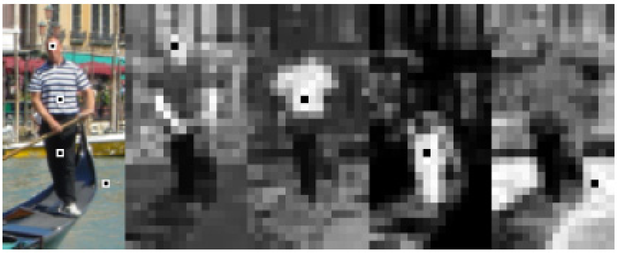


Fig. 1. Self-similarity encodes relevant parts

3 Supervised Dictionary Learning

We present in this section the core of the proposed model. In classical sparse coding tasks, one considers a signal x in R^n and a fixed dictionary D = [d_1, . . . , d_k] in $R^{n\times k}$ (allowing k>n, making the dictionary over complete). In this setting, sparse coding with an $\ell 1$ regularization amounts to computing

$$R^*(x,D)=\min_{\alpha\in R^k} \| x-D\alpha\|_2^2+\lambda_1\|\alpha\|_1 \tag{1}$$

It is well known in the statistics, optimization, and compressed sensing communities that the ℓ1 penalty yields a sparse solution, very few non-zero coefficients in α, although there is no explicit analytic link between the value of λ1 and the effective sparsity that this model yields. Other sparsity penalties using the ℓ0 regularization can be used as well. Since it uses a proper norm, the ℓ1 formulation of sparse coding is a convex problem, which makes the optimization tractable with algorithms such as those introduced in [26, 27], and has proven in practice to be more stable than its ℓ0 counterpart, in the sense that the resulting decompositions are less sensitive to small perturbations of the input signal x. Note that sparse coding with an ℓ0 penalty is an NP-hard problem and is often approximated using greedy algorithms.

In this paper, we consider a setting, where the signal may belong to any of p different classes. We first consider the case of p = 2 classes and later discuss the multiclass extension. We consider a training set of m labeled signals $(x_i)m_i=1$ in R^n, associated with binary labels $(y_i \in \{-1, +1\})m_i=1$.Our goal is to learn jointly a single dictionary D adapted to the classification task and a function f which should be positive for any signal in class +1 and negative otherwise. We consider in this paper two different models to use the sparse code α for the classification task:

(i) linear in α: $f(x,\alpha,\theta)=w^T\alpha+b$, where θ = {w∈$R^k$, b∈R} parametrizes the model.

(ii) bilinear in x and α: $f(x,\alpha,\theta)=x^T w\alpha+b$, where θ= {W∈$R^{n\times k}$, b ∈R}. In this case, the model is bilinear and f acts on both x and its sparse code α.

The number of parameters in (ii) is greater than in (i), which allows for richer models. Note that one can interpret was a linear filter encoding the input signal x into a model for the coefficients, which has a role similar to the encoder in [28] but for a discriminative task. A classical approach to obtain for (i) or (ii) is to first adapt D to the data, solving

$$\min_{D,\alpha}\sum_{i=1}^{m}\| x_i - D\alpha_i \|_2^2 + \lambda_1 \| \alpha_i \|_1 \tag{2}$$

Note also that since the reconstruction errors $\| x_i - D\alpha_i \|_2^2$ are invariant to scaling simultaneously D by a scalar and αi by its inverse, we need to constrain the ℓ2 norm of the columns of D. Such a constraint is classical in sparse coding [29]. This reconstructive approach provides sparse codes αi for each signal xi, which can be used a posteriori in a regular classifier such as logistic regression, which would require to solve

$$\min_{\theta}\sum_{i=1}^{m} c(y_i f(x_i,\alpha_i,\theta)) + \lambda_2 \| \theta \|_2^2 \tag{3}$$

where C is the logistic loss function ($C(x) = \log(1 + e^{-x})$), which enjoys properties similar to that of the hinge loss from the SVM literature, while being differentiable, and λ2 is a regularization parameter, which prevents over fitting. This is the approach chosen in [30] (with SVMs). However, our goal is to learn jointly D and the model parameters θ. To that effect, we propose the formulation

$$\min_{D,\theta,\alpha}(\sum_{i=1}^{m} c(y_i f(x_i,\alpha_i,\theta)) + \lambda_0 \| \| x_i - D\alpha_i \|_2^2 + \lambda_1 \| \alpha_i \|_1) + \lambda_2 \| \theta \|_2^2 \tag{4}$$

where λ0 controls the importance of the reconstruction term, and the loss for a pair (x_i,y_i) is

$$S^*(x_i, D, \theta, y_i) = \min_{\alpha} S(\alpha, x_i, D, \theta, y_i) \quad (5)$$

Where $S(\alpha, x_i, D, \theta, y_i) = c(y_i f(x_i, \alpha_i, \theta)) + \lambda_0 \| x_i - D\alpha_i \|_2^2 + \lambda_1 \| \alpha_i \|_1$ In this setting, the classification procedure of a new signal x with an unknown label y, given a learned dictionary D and parameters θ, involves supervised sparse coding:

$$\min_{y \in \{-1;+1\}} S^*(x, D, \theta, y) \quad (6)$$

The learning procedure of Eq. (4) minimizes the sum of the costs for the pairs (x_i,y_i)m_i=1 and corresponds to a generative model. We will refer later to this model as SDL-G (supervised dictionary learning, generative). Note the explicit incorporation of the reconstructive and discriminative component into sparse coding, in addition to the classical reconstructive term (see [31] for a different classification component).

However, since the classification procedure from Eq. (6) compares the different costs $S^*(x, D, \theta, y)$ of a given signal for each class y = −1, +1, a more discriminative approach is to not only make the costs $S^*(x_i, D, \theta, -y_i)$ small, as in (4), but also make the value of $S^*(x_i, D, \theta, -y_i)$ greater than $S^*(x_i, D, \theta, -y_i)$, which is the purpose of the logistic loss function C. This leads to:

$$\min_{D,\theta} \left(\sum_{i=1}^{m} c(S^*(x_i, D, \theta, -y_i) - S^*(x_i, D, \theta, y_i))\right) + \lambda_2 \| \theta \|_2^2 \quad (7)$$

As detailed below, this problem is more difficult to solve than (4), and therefore we adopt instead a mixed formulation between the minimization of the generative Eq. (4) and its discriminative version (7), (see also [32])—that is,

$$\left(\sum_{i=1}^{m} \mu c(S^*(x_i, D, \theta, -y_i) - S^*(x_i, D, \theta, y_i)) + (1-\mu) S^*(x_i, D, \theta, y_i)\right) + \lambda_2 \| \theta \|_2^2 \quad (8)$$

where μ controls the trade-off between the reconstruction from Eq. (4) and the discrimination from Eq. (7). This is the proposed generative/discriminative model for sparse signal representation and classification from learned dictionary D and model θ. We will refer to this mixed model as SDL-D, (supervised dictionary learning, discriminative). Note also that, again, we constrain the norm of the columns of D to be less than or equal to one.

All of these formulations admit a straightforward multiclass extension, using softmax discriminative cost functions $c_i(x_1, ..., x_p) = \log(\sum_{i=1}^{p} e^{x_j - x_i})$, which are multiclass versions of the logistic function, and learning one model θi per class. Other possible approaches such as one-vs-all or one-vs-one are of course possible, and the question of choosing the best approach among these possibilities is still open. Compared with earlier work using one dictionary per class [33], our model has the advantage of letting multiple classes share some features, and uses the coefficients of the sparse representations as part of the classification procedure, thereby following the works from[34, 35, 30], but with learned representations optimized for the classification task similar to [31, 36].

Our bilinear model with $f(x,\alpha,\theta)=x^T w\alpha+b$ does not admit a straightforward probabilistic interpretation. On the other hand, it can easily be interpreted in terms of kernels: Given two signals x_1 and x_2, with coefficients α_1 and α_2, using the kernel $K(x_1,x_2)=\alpha_1^T\alpha_2 x_1^T x_2$ in a logistic regression classifier amounts to finding a decision function of the same form as f. It is a product of two linear kernels, one on the α's and one on the input signals x. Interestingly, Raina et al. [30] learn a dictionary adapted to reconstruction on a training set, then train an SVM a posteriori on the decomposition coefficients. They derive and use a Fisher kernel, which can be written as $K'(x_1,x_2)=\alpha_1^T\alpha_2 r_1^T r_2$ in this setting, where the r's are the residuals of the decompositions. In simple experiments, which are not reported in this paper, we have observed that the kernel K, where the signals x replace the residuals r, generally yields a level of performance similar to K′ and often actually does better when the number of training samples is small or the data are noisy.

4 Optimization Procedure

Classical dictionary learning techniques (e.g., [30, 37, 38]), address the problem of learning a reconstructive dictionary D in $R^{n\times k}$ well adapted to a training set, which is presented in Eq. (3). It can be seen as an optimization problem with respect to the dictionary D and the coefficients. Although not jointly convex in D, it is convex with respect to each unknown when the other one is fixed. This is why block coordinate descent on D and performs reasonably well [30, 37, 38], although not necessarily providing the global optimum. Training when μ = 0 (generative case), i.e., from Eq. (4), enjoys similar properties and can be addressed with the same optimization procedure. Equation (4) can be rewritten as:

$$\min_{D,\theta,\alpha}\left(\sum_{i=1}^{m} S(x_j,\alpha_j,D,\theta,y_i)\right)+\lambda_2\|\theta\|_2^2, s.t. \forall j=1,\ldots,k, \|d_j\|_2\leq 1 \tag{9}$$

Block coordinate descent consists therefore of iterating between supervised sparse coding, where D and θ are fixed and one optimizes with respect to the α's and supervised dictionary update, where the coefficients α_i's are fixed, but D and θ are updated. Details on how to solve these two problems are given in sections 4.1 and 4.2. The discriminative version SDL-D from Eq.(7) is more problematic. To reach a local minimum for this difficult non-convex optimization problem, we have chosen a continuation method, starting from the generative case and ending with the discriminative one as in [33]. The algorithm is presented in Figure 2.

Input: n (signal dimensions); (x_i, y_i) m_i=1 (training signals); k (size of the dictionary); λ_0, λ_1, λ_2(parameters); $0 \le \mu_1 \le \mu_2 \le \ldots \le \mu_m \le 1$ (increasing sequence).
Output: $D \in R^{n\times k}$ (dictionary); θ (parameters).
Initialization: Set D to a random Gaussian matrix with normalized columns. Set θ to zero.
Loop: For $\mu = \mu_1, \ldots, \mu_m$,
Loop: Repeat until convergence (or a fixed number of iterations),
• Supervised sparse coding: Solve, for all i = 1, . . . ,m,

$$\begin{cases} \alpha_{i,-}^{*} = \arg\min_{\alpha} S(\alpha, x_i, D, \theta, -1) \\ \alpha_{i,+}^{*} = \arg\min_{\alpha} S(\alpha, x_i, D, \theta, +1) \end{cases}$$

• Dictionary and parameters update: Solve

$$\min_{D,\theta}(\sum_{i=1}^{m} \mu c((S(\alpha_{i,-}^{*}, x_i, D, \theta, -y_i) - S(\alpha_{i,+}^{*}, x_j, D, \theta, y_i)))$$

$$+ (1-\mu) S(\alpha_{i,y_i}^{*}, x_i, D, \theta, y_i) + \lambda_2 \| \theta \|_2^2) s.t. \forall j, \| d_j \|_2 \le 1$$

Fig. 2. SDL: Supervised dictionary learning algorithm

4.1 Supervised Sparse Coding

The supervised sparse coding problem from Eq. (6) (D and θ are fixed in this step) amounts to minimizing a convex function under a ℓ1 penalty. The fixed-point continuation method (FPC) from [27] achieves good results in terms of convergence speed for this class of problems. For our specific problem, denoting by g the convex function to minimize, this method only requires ∇g and a bound on the spectral norm of its Hessian Hg. Since the we have chosen models g which are both linear in α, there exists, for each supervised sparse coding problem, a vector a in Rk and a scalar c in R such that

$$\begin{cases} g(\alpha) = c(a^T \alpha + c) + \lambda_0 \| x - D\alpha \|_2^2 \\ \nabla g(\alpha) = \nabla c(a^T \alpha + c) a - 2\lambda_0 D^T (x - D\alpha) \end{cases}$$

and it can be shown that, if $\| U \|_2$ denotes the spectral norm of a matrix U(which is the magnitude of its largest Eigen value), then we can obtain the following bound, $\| H_g(\alpha) \|_2 \le | H_c(a^T \alpha + c) | \| a \|_2^2 + 2\lambda_0 \| D^T D \|_2$

4.2 Dictionary Update

The problem of updating D and θ in Eq. (11) is not convex in general (except when μ is close to 0), but a local minimum can be obtained using projected gradient descent (as in the general literature on dictionary learning, this local minimum has experimentally been found to be good enough in terms of classification performance). Denoting E(D,θ) the function we want to minimize in Eq. (11), we just need the partial derivatives of E with respect to D and the parameters θ. When considering the linear model for the α's, $f(x, \alpha, \theta) = w^T \alpha + b$, and θ = {w∈Rk, b∈R}, we obtain

$$\begin{cases} \dfrac{\partial E}{\partial D} = -2\lambda_0 (\sum_{i=1}^{m} \sum_{z=\{-1,+1\}} \omega_{i,z} (x_i - D\alpha_{i,z}^{*}) \alpha_{i,z}^{*T}), \\ \dfrac{\partial E}{\partial w} = \sum_{i=1}^{m} \sum_{z=\{-1,+1\}} \omega_{i,z} z \nabla c(w^T \alpha_{i,z}^{*} + b) \alpha_{i,z}^{*}, \\ \dfrac{\partial E}{\partial b} = \sum_{i=1}^{m} \sum_{z=\{-1,+1\}} \omega_{i,z} z \nabla c(w^T \alpha_{i,z}^{*} + b), \end{cases}$$

Where $\omega_{i,z} = -\mu z \nabla c(S(\alpha_{i,-}^{*}, x_i, D, \theta, -y_i) - S(\alpha_{i,+}^{*}, x_i, D, \theta, y_i) + (1-\mu) 1_{z=y_i}$

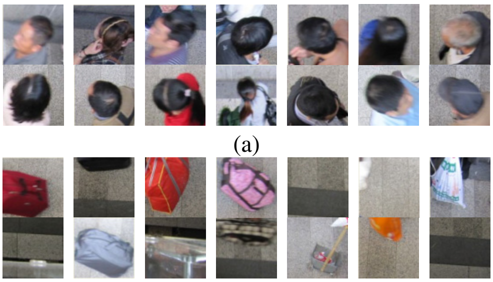
Partial derivatives when using our model with multiple classes or with the bilinear models $f(x,\alpha,\theta) = x^T w\alpha + b$ are not presented in this paper due to space limitations

5 Experiments

To evaluate the performance of the proposed algorithm, we carry out a series of experiments on a dataset extracted 500 images of size 48*96 from a video. If the image is contain a pedestrian, the label of it will be 1, otherwise -1. Fig. 3(a) shows several images with label 1. Fig. 3(b) shows several images with label -1. 100 images from the dataset are selected as the test examples. Different number images of the dataset are selected as the training examples to compare the accuracy rate.

Fig.4 shows the compare results of recognition between with HOG, HOF and Color features respectively and with the corresponding sparse features.Fig.5 shows the result of using mixing features to compare the two methods. As shown in the graph, our method performs better than the method directly using HOG, HOF and Color features to recognition. In addition, with the increasing number of training samples, our method performs better.

Fig.6 shows the result of these two methods using shading images to test. Compared with the traditional method, our method has better recognition accuracy and shows good robustness.



(a)

(b)

Fig. 3. Images with label 1(a) and images with label -1(b)

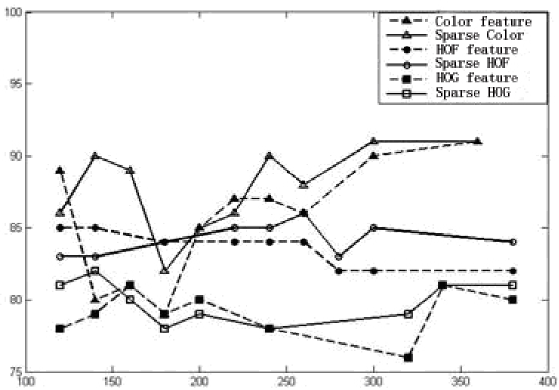


Fig. 4. The compare results of recognition between with HOG, HOF and Color features respectively and with the corresponding sparse features

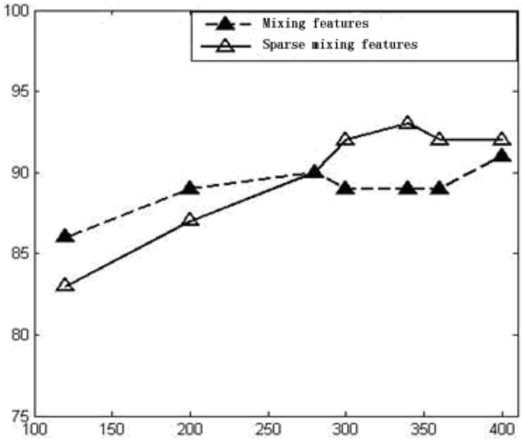


Fig. 5. The result of using mixing features to compare the two methods

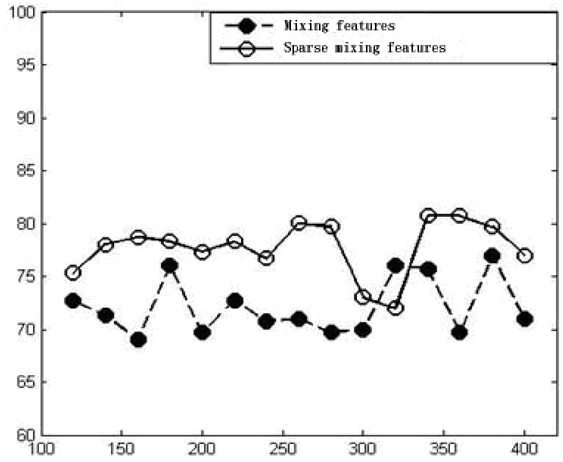


Fig. 6. The result of these two methods using shading images to test

6 Conclusion

We proposed a system for pedestrian detection with very good accuracy. To achieve good classification performance, we put forward a novel framework for pedestrian detection tasks, which proposing a model with both sparse reconstruction and class discrimination components, jointly optimized during dictionary learning. We present an efficient pedestrian detection system using mixing sparse features of HOG, FOG and CSS to combine into this a Kernel classifier. Results presented on our data set show competetive accuracy and robust performance of our system outperforms current state-of-the-art work. Although we use the system for the detection of pedestrians, the general idea can be applied to the detection of other object classes as well.

Acknowledgments. This research is supported by the national science foundation of China (NFSC) No. 61170126, 60273040, provincial universities natural science foundation of Jiangsu province (11KJD520004).

References

1. Andreas, E., Konrad, S., Bastian, L., Luc, G.: Object detection and tracking for autonomous navigation in dynamic environments. International Journal of Robotics Research 14, 1707–1725 (2010)
2. Anthony, C., Yin, J., Sergio, A.: Crowd monitoring using image processing. Electronics and Communication Engineering Journal 1, 37–47 (1995)

3. Kim, C., Human, B.: Gait analysis using Self Organizing Map. In: The 2009 Chinese Conference on Pattern Recognition and the 1st CJK Joint Workshop on Pattern Recognition, pp. 888–891. IEEE Press, Piscataway (2009)
4. Ma, G., Ioffe, A., Stefan, M., Anton, K.: A real time object detection approach applied to reliable pedestrian detection. In: IEEE Intelligent Vehicles Symposium, pp. 755–760. IEEE Press, Piscataway (2007)
5. Marana, A., Cavenaghi, M., Ulson, R., Drumond, F.: Real-time crowd density estimation using images. In: Bebis, G., Boyle, R., Koracin, D., Parvin, B. (eds.) ISVC 2005. LNCS, vol. 3804, pp. 355–362. Springer, Heidelberg (2005)
6. Kong, C., Yang, J., Nie, J.: A study on pedestrian detection models based on the analysis on real accident scenarios. Qiche Gongcheng/Automotive Engineering 11, 977–983 (2010)
7. Mallat, S.: A wavelet tour of signal processing, 2nd edn. Academic Press, New York (1999)
8. Mairal, J., Bach, F., Ponce, J., Sapiro, G., Zisserman, A.: Learning discriminative dictionaries for local image analysis. In: 26th IEEE Conference on Computer Vision and Pattern Recognition, pp. 1–8. IEEE Press, Piscataway (2008)
9. Ranzato, M., Szummer, M.: Semi-supervised learning of compact document representations with deep networks. In: 25th International Conference on Machine Learning, pp. 792–799. ACM Press, New York (2008)
10. Argyriou, A., Evgeniou, T., Pontil, M.: Multi-Task Feature Learning. In: 20th Annual Conference on Neural Information Processing Systems, pp. 41–48. Neural Information Processing System Foundation, Vancouver (2006)
11. Rodriguez, F., Sapiro, G.: Sparse representations for image classification: Learning discriminative and reconstructive non-parametric dictionaries. In: IMA Preprint, p. 16 (2008)
12. Blei, D., McAuliffe, J.: Supervised topic models. In: Advances in Neural Information Processing Systems, pp. 208–213. Curran Associates Inc., New York (2007)
13. Holub, A., Perona, P.: A discriminative framework for modeling object classes. In: Conference on Computer Vision and Pattern Recognition, pp. 664–671. IEEE Press, Piscataway (2005)
14. Lasserre, J., Bishop, C., Minka, T.: Principled hybrids of generative and discriminative models. In: 2006 IEEE Computer Society Conference on Computer Vision and Pattern Recognition, pp. 87–94. IEEE Press, Piscataway (2006)
15. Raina, R., Shen, Y., Ng, A., McCallum, A.: Classification with hybrid generative/discriminative models. In: Advances in Neural Information Processing Systems, pp. 109–113. MIT Press, British Columbia (2004)
16. Salakhutdinov, R., Hinton, G.: Learning a non-linear embedding by preserving class neighbourhood structure. In: The 11th International Conference on Artificial Intelligence and Statistics, pp. 412–419. Microtome Publishing, Brookline (2007)
17. Larochelle, H., Bengio, Y.: Classification using discriminative restricted boltzmann machines. In: The 25th International Conference on Machine Learning, pp. 536–543. ACM Press, New York (2008)
18. Dalal, N., Triggs, B.: Histograms of oriented gradients for human detection. In: Computer Society Conference on Computer Vision and Pattern Recognition, pp. 886–893. IEEE Press, Piscataway (2005)
19. Dalal, N., Triggs, B., Schmid, C.: Human detection using oriented histograms of flow and appearance. In: Leonardis, A., Bischof, H., Pinz, A. (eds.) ECCV 2006. LNCS, vol. 3952, pp. 428–441. Springer, Heidelberg (2006)

20. Werlberger, M., Trobin, W., Pock, T., Wedel, A., Cremers, D., Bischof, H.: Anisotropic Huber-L1 optical flow. In: The British Machine Vision Conference, pp. 123–128. Elsevier Ltd., Oxford (2009)
21. Doll, P., Tu, Z., Perona, P., Belongie, S.: Integral channel features. In: The British Machine Vision Conference, pp. 777–780. Elsevier Ltd., Oxford (2009)
22. Schwartz, W., Kembhavi, A., Harwood, D., Davis, L.: Human detection using partial least squares analysis. In: 12th International Conference on Computer Vision, pp. 24–31. IEEE Press, Piscataway (2009)
23. Wojek, C., Walk, S., Schiele, B.: Multi-cue onboard pedestrian detection. In: Computer Society Conference on Computer Vision and Pattern Recognition Workshops, pp. 794–801. IEEE Press, Piscataway (2009)
24. Sande, K., Gevers, T., Snoek, C.: Evaluation of color descriptors for object and scene recognition. In: 26th IEEE Conference on Computer Vision and Pattern Recognition, pp. 1582–1596. IEEE Press, Piscataway (2008)
25. Watanabe, T., Ito, S., Yokoi, K.: Co-occurrence histograms of oriented gradients for pedestrian detection. In: Wada, T., Huang, F., Lin, S. (eds.) PSIVT 2009. LNCS, vol. 5414, pp. 37–47. Springer, Heidelberg (2009)
26. Efron, B., Hastie, T., Johnstone, I., Tibshirani, R.: Least angle regression. Annals of Statistics 32, 407–451 (2004)
27. Hale, E., Yin, W., Zhang, Y.: A fixed-point continuation method for l1-regularized minimization with applications to compressed sensing. In: CAAM Tech. Report,TR07-07 (2007)

The Design and Implementation of Computer-Aided Chinese Medicated Diet System

Xiao-Juan Hu[1], Zheng He[1], Hong-Hai Zhu[2], and Jun Dong[2,*]

[1] Software Engineering Institute, East China Normal University, Shanghai 200062, China
[2] Interdisciplinary Division, Suzhou Institute of Nano-tech and Nano-bionics, Chinese Academy of Sciences, Suzhou, 215125, China
jdong2010@sinano.ac.cn

Abstract. Chinese Medicated Diet (CMD) which combines the properties of food and medicine not only prevents and cures diseases but also prolongs lives. It is becoming more and more popular in modern society and in great demand. Facing the shortage of medicated diet knowledge in common people, this paper focuses on the design and development of an online software system for CMD: Chinese Medicated Diet System (CMDS). The design and implementation of CMDS are discussed as key issues including persistence layer and database access. The proposed demo provides an efficient way to popularize and internationalize the CMD. It should be enriched step by step and apply in daily life in the future.

Keywords: Chinese Medicated Diet, Therapy, Computer-aided Analysis.

1 Introduction

Chinese Medicated Diet (CMD), which combines the properties of food and medicine as the treasure of the Traditional Chinese Medicine (TCM), is becoming more and more popular. The popularization and internationalization of CMD is put forward in "The State Council on Supporting and Promoting Development Certain Opinions" [1,2]. CMD originated in ancient is incorporated into TCM research by the national administration of TCM in 1985 [3]. Recently, as the improvement of life level, people have higher requirements in health maintenance and disease prevention and treatment. Because of not only CMD's edibility but also its usefulness for keeping the body healthy, it has gained much favor. However the development of CMD can't satisfy the growing needs about people's health because of some shortages, such as the single transmission way, passive acceptance by readers with the single form of books [4,5,6] and the limitations of Chinese description which is disadvantageous to promote the CMD to the world.

Some computer-aided systems for CMD have been put forward. Zhiyi Chinese Medicinal Diet System [7] provides diet therapy prescriptions through estimating

* Corresponding author.

Z. Pan et al. (Eds.): Transactions on Edutainment IX, LNCS 7544, pp. 196–204, 2013.
© Springer-Verlag Berlin Heidelberg 2013

user's physical by inputting symptoms. Zhang [8, 9] designs the system applied in restaurants and related to food service industry, by which users can choose seats, order and check out. Nevertheless it can't accurately provide diet therapy prescriptions just based on symptoms [7]. Chinese Medicated Diet System (CMDS) combines the results of computer-aided diagnostic terminal and the symptoms by user input as the basis of diet therapy prescriptions.

This paper is organized as following. The research idea of CMDS is introduced in section 2 including the computer-aided diagnosis and intelligent processing techniques based on the architecture of internet of things. CMDS requirement, the detailed design of structure and the core functions are presented in section 3. Demo is illustrated in section 4. Finally, the conclusions are given in section 5.

2 Original Intentions

The portable diagnostic equipment is becoming more and more popular, by which users' health status can be diagnosed from remote doctors. Sub-health state is always ignored therefore it becomes diseases in the near future. The computer-aided system for CMD not only provides auxiliary treatment for diseases and sub-health state, but also makes for the popularization of CMD knowledge [10].

Integrating computer-aided diagnosis and treatment system, the outputs of the computer-aided diagnosis system is taken as the inputting of treatment system that can increase the effectiveness of treatment. Based on the thought above [11], CMDS take the classification results from computer-aided diagnosis system and the symptoms by user's inputting as the basis of diet therapy prescriptions.

CMDS provides medicinal diet suggestion to guide the user's physiotherapy for health preserving, not to replace the prescription drugs and professional doctor. Combining the process of the classical Chinese medicine diagnosis and treatment, the relationship of the results of computer-aided diagnosis, symptoms, disease and diet therapy prescription are illustrated in Fig 1.

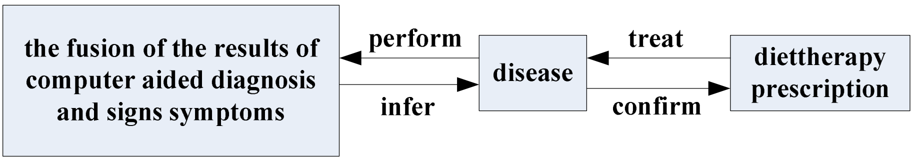


Fig. 1. the relationship of signs and symptoms, disease and diet therapy prescription

The results of computer-aided diagnosis, one of the inputting in CMDS, are critical for the effectiveness of diet therapy prescription. A real-time multi-parameter body signal acquisition and intelligent monitoring based on the architecture of internet of things (see Fig 2), which is being developed, consists of patient terminals, consultation terminals and server groups (including hospitals).

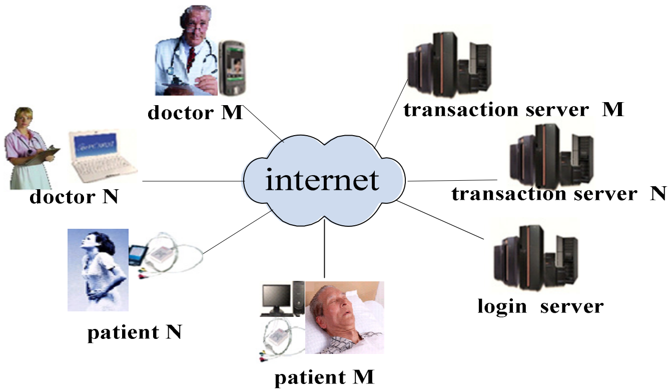


Fig. 2. An internet of things' application on real-time multi-parameter body signal acquisition and intelligent monitoring

Patient terminals, multi-parameter simultaneous acquisition for Electrocardiogram (ECG), respiration, body temperature, blood pressure, physical activity, body position and pulse signals, transmit the multi-parameters to mobile phones or computers by Bluetooth or USB. Dealing with kinds of operations related to monitoring, it is up to the server groups, such as patient information acquisition, case acquisition, diagnosis request and so on. Consultation terminals, acquiring multi-parameters information from patient terminals, take the intelligent algorithm as the core part, which make the doctor's diagnose more easily.

Algorithm research needs a lot of data and accumulation. Because of the problems existing in ECG's algorithm research, Chinese Cardiovascular Disease Database (CCDD) [12] and its management tool have been completed, in which the ECG data has been increased. Some methods for ECG classification have been attempted [13,14]. The accuracy of ECG classification can be improved by the method based on expert's experience and morphological characteristics by some experiments [15,16]. The algorithms fusing ECG, respiration, temperature, blood pressure, action and other parameters are being investigated. These are the foundation of the CMDS.

3 The Design and Implementation of the CMDS

The users get diet therapy prescription suggestion based on the classification results from computer-aided diagnosis system and the symptom input by user in CMDS. CMDS also realizes user participation functions, such as evaluation questions and so on. The function of the system is as shown in Fig 3.

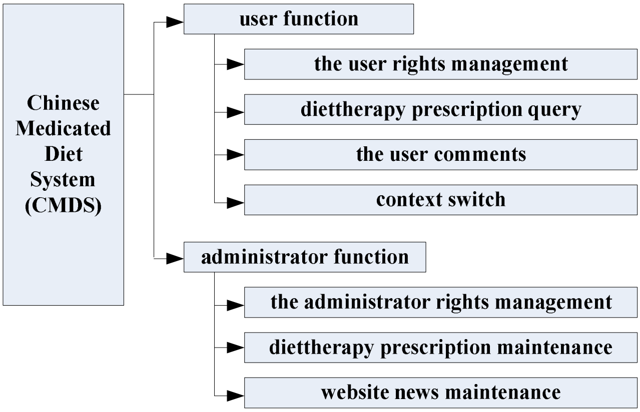


Fig. 3. The functions of the system

3.1 The Core Functional Requirements

Based on the analysis of its functions, there are two core functions in CMDS:

Case name: Diettherapy Prescription Suggestions
Executor: the user of CMDS
Pre-conditions: none
Post-conditions: none
The basic path:

1. CMDS Provides a series of disease signs and symptoms to the users.
2. The users choose one or more signs and symptoms.
3. The users confirm to start to query.
4. CMDS shows one or more disease names matching signs and symptoms and displays one or more diettherapy prescriptions.
5. Providing diettherapy prescriptionSuggestions is over, the user come back to step 1.

Alternative path:

3a. Signs and symptoms re-election.
  3a1. 1) The users add one or more signs and symptoms.
    2) Remove one or more signs and symptoms selected .
  3a2. CMDS update signs and symptoms sets.
  3a3. Come back to step 3.
4a. Inappropriate diettherapy prescriptions in CMDS.
  4a1. CMDS prompt no diettherapy prescriptions to show.
  4a2. Come back to setp 5.

Fig. 4. Cases description of diet therapy prescription suggestions

Firstly, CMDS offers suggestions of CMD according to signs, symptoms and the classification results from computer-aided diagnosis system. For the diagnosis and treatment of common disease, both traditional Chinese and western medicine theory support premises: the same disease has the similar symptoms on different patients; the same prescription is effective for different patients with the same disease. Cases description is described in Fig 4.

Secondly, CMDS provides kinds of languages to satisfy users' needs in different countries and regions. The internationalization of CMDS could make the software remain unchanged when it is transplanted to other regions with different language. The international task in CMDS contains two parts: interface internationalization and data internationalization. Interface internationalization supported by international module of Struts2 is realized only by a locale class and a resource pack. Base on accomplishing Struts2 international functions, data internationalization provides options for switching language, judges the current selected language and accesses the data of corresponding language by context.

3.2 CMDS Architecture

CMDS based on the frame of Struts2, complies with the model of Model, View and Controller (MVC). MVC, which is low coupling, high reusability and easy to rapid deployment and maintenance, compels to keep the application's inputting, processing and output apart. MVC is divided into three core components (Model, View and Controller) to deal with their respective tasks. In CMDS implementation, JSP and JSP Tags are as the Views; the Controller is Action inheriting Struts2; and the Model are composed with the domain objects realized by persistent object (PO), Service components maintained by Spring frame and the entity of persistent object (See Fig 5).

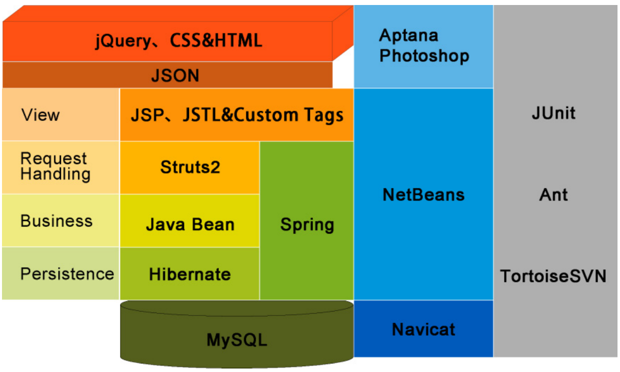


Fig. 5. CMDS architecture

The data flow diagram in Fig 6 shows execution paths of user requests passing parts of the system and creating the responses finally. User requests sent by the browser arrive at filter and interceptor, which realizes an agency model in Servlet standard and

Structs2 frame, respectively. User requests after performing preprocessing get to the response handle: Action, in which legitimacy is verified. According to the request parameters, session object and configuration parameters in System, Action calls Service component to execute business logic. During the execution process of Service component, DOA coordinates the completion of data access. The results of Service component execution received by Action are converted to the transmission format. Users need to return to the browser though choosing appropriate View Result by Action. So far the responses of user requests are completed.

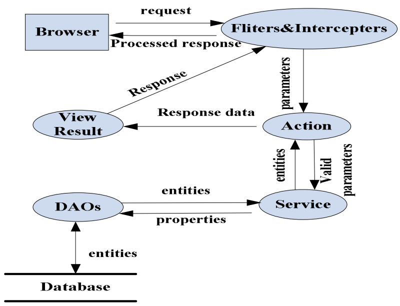


Fig. 6. Processing a typical user request

3.3 Database Access Design and Implementation

In consideration of the shortcoming of Java Data Base Connectivity (JDBC)[17], CMDS uses Object/Relation Mapping (ORM), Hibernate frame and Java persistence API (JPA). The biggest advantage of ORM is providing a simple and easy way for database interaction [18]. Hibernate could reduce the workload of the database operation and the code of data extraction. JPI has a standard interface which is easy to use and owns strong flexibility.

Fig 7 shows the database table structure. It is mentionable that there are only a surrogate keys in some tables, which is unreasonable in the common database design. For example, in the table Recipe Entity, there has only one field id. Since only one table can replace some fragmentary data table with JPA describing the inheritance relationship between entities, seemingly unreasonable tables are kept for subsequent development needs.

The relationships between classes are showed in Fig 8. In the right of Fig 8, Abstract class AbstactEntity extracts out common attributes of all objects. Abstract class MultilangEntity defines basic attributes of international objects. At last, an Entity class inherited from the AbstractEntity or MultilangEntity corresponds with the same table or field in the same table by annotation in Java to map object relational database. In the left, Interface AbstractDAO makes the basic Create, Read, Update and Delete (CRUD) method standard. DAOImpl implements the interface. DAO interface offers convenient access methods with hidden database details. Separating DAO class interface and implementation decouples Hibernate implementation and upper level system, and reflects that JPA is neither a kind of standard nor the implementation.

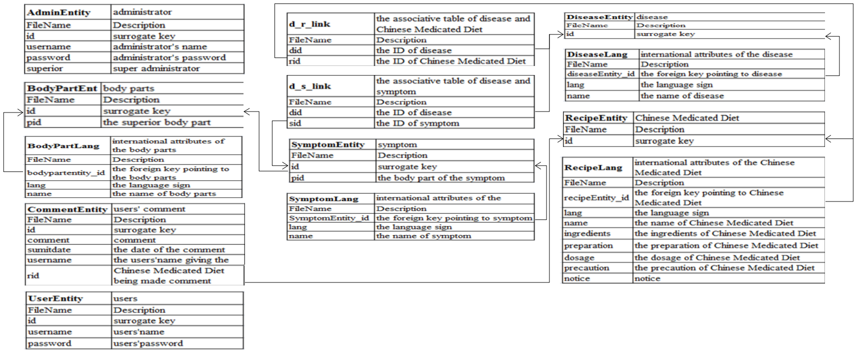


Fig. 7. Database table structure

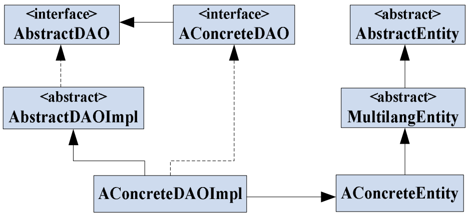


Fig. 8. Persistence layer diagram

The user interaction design based on jQuery frame [19] follows principles of using Ajax technology [20] to load the page data, reduce page jump and make interface concise.

4 CMDS Demo

With "thoracic spinal deformity" and "bone itch" symptoms of male user as an example, the functions and the user operations are demonstrated. (i) Choose sex in gender selection page, and then click the choice men. (See 1 in Fig 9). (ii) Choose specific organs from menu after click the body parts corresponding with thoracic spinal deformity symptom: chest. (See 2 in Fig 9). (iii) Checked corresponding symptoms in the popup page of all possible symptoms as 3 in Fig 9, then click on the "OK" button. (iv) Repeat step ii to iii, choose the second symptom: bone itch. The result is as 4 in Fig 9. (vi) Click on “Began inquires”. The system diagnoses user with osteoporosis very likely, and recommends the corresponding diet therapy prescription: carrots porridge (see 5 in Fig 9), then show formula content by clicking (see 6 in Fig 9). After viewing the "Carrots porridge", users can comment on the contents of the formula for future reference of the other users.

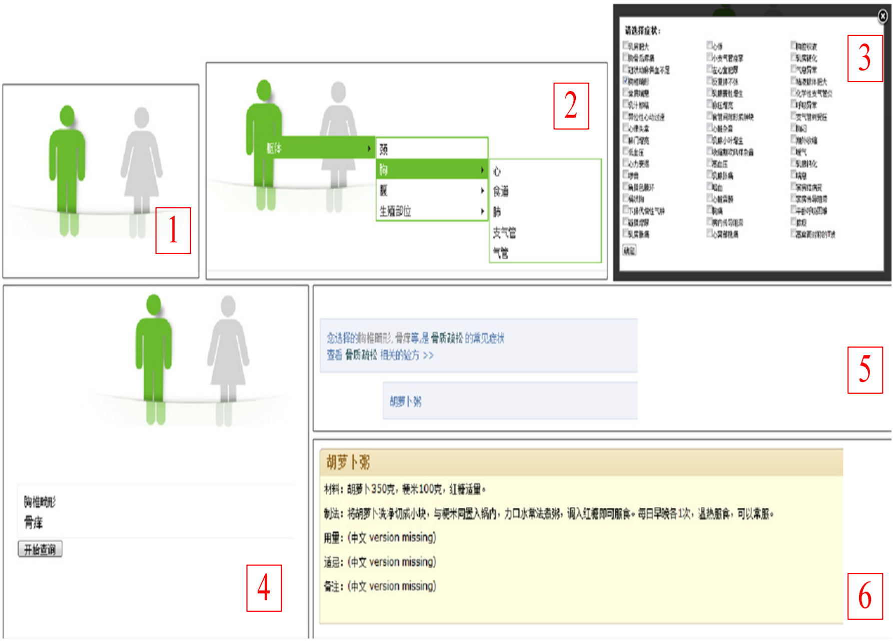


Fig. 9. CMDS Demo

5 Conclusions

This paper introduces the design and implementation technologies of CMDS in detail, such as persistence layer's design, database access's implementation and so on. CMDS has two key problems to be solved in the future. One is sorting and translating diet therapy prescription. The other is the inspection of CMDS in the realistic environment. Based on the original intentions of CMDS, a lot of works will be done for integrating computer-aided diagnosis and treatment system. In the future, we will not only try to solve above problems in CMDS, but also try to improve the accuracy of computer -aided diagnosis system.

References

1. Traditional Chinese Medicine Research Institute, Guangzhou institute of traditional Chinese medicine: Chinese Medicine Dictionary. People's medical Press, Beijing (1995)
2. The State Council on Supporting and Promoting Development Certain Opinions, `http://www.gov.cn/zwgk/2009-05/07/content_1307145.htm`
3. State Administration of Tradition Chinese Medicine of the People's Republic of China, `http://www.satcm.gov.cn/web2010/zhengwugongkai/xingyekuaixun/gedikuaibao/2010-10-07/4248.html`

4. Zhang, W.-G.: Chinese Medicated Diet. Shanghai University of Traditional Chinese Medicine Press, Shanghai (1990)
5. Xiang, P.: Entire Gather of Prescriptions about Dietetic Therapy of Traditional Chinese Medicine. People's Medical Publishing House, Beijing (2000)
6. Wang, S.-Y.: Zhonghua Medicated Diet. Tianjin Ancient Books Publishing House, Tianjin (2007)
7. Traditional Chinese Medicine Medicinal Food Nutrition System, http://www.wlc168.com/
8. The Analysis and Design of Medicine Diet and Diet Therapy Information System, http://www.paper.edu.cn
9. Zhang, L., Lv, J.-N., Zhang, S.-Q.: The Analysis and Design of Information System of Chinese Meditative Diet for Food Therapy. World Sci-tech R&D 31(1), 89–91 (2009)
10. Yu, S., Li, S.-Z., Chen, J.-X., et al.: Research on Dietetic Therapy of TCM by Computer-aided Design. In: 2009 IEEE International Symposium on IT in Medicine & Education, pp. 1085–1090. IEEE Press, Jin'nan (2009)
11. Hu, D.-H., Li, D.-H., Guan, J.-H., et al.: Character Space of the Four Diagnostics and of the Syndrome Differentiation in the Chinese Traditional Medicine. Journal of Biomedical Engineering 22(4), 286–289 (2003)
12. Zhang, J.-W., Dong, J.: An Enhanced Standard ECG Database with Its Management & Annotation Tools. International Journal on Artificial Intelligence Tools 21(5), 1–26 (2012)
13. Dong, J., Tong, J.-f., Liu, X.: The Abnormal vs. Normal ECG Classification Based on Key Features and Statistical Learning. In: Graña Romay, M., Corchado, E., Garcia Sebastian, M.T. (eds.) HAIS 2010, Part I. LNCS (LNAI), vol. 6076, pp. 136–143. Springer, Heidelberg (2010)
14. Wang, L.-P., Shen, M., Tong, J.-F., Dong, J.: An Uncertainty Reasoning Method for Abnormal ECG Detection. In: 2th IEEE International Symposium on IT in Medicine & Education, pp. 1091–1096. IEEE Press, Ji'nan (2009)
15. Dong, J., Zhang, J.-W.: Experiences-based Intelligence Simulation in ECG Recognition. In: International Conference on Computational Intelligence for Modeling Control & Automation, pp. 796–801. IEEE Press, Austria (2008)
16. Zhang, J.-W., Hu, X.-J., Dong, J., et al.: A Framework for ECG Morphology Features Recognition. In: 23rd IEEE International Symposium on Computer-based Medical System, pp. 85–91. IEEE Press, Perth (2010)
17. Weng, Y.C., Hsieh, S.H., Lai, F.P., et al.: Wireless/Wired Collaborative Remote Consultation Emergency Healthcare Information Systems Framework. In: IEEE Region 10 Conference, pp. 1–4. IEEE Press, Hong Kong (2006)
18. Mayr, C., Zdun, U., Dustdar, S.: View-based Model-driven Architecture for Enhancing Maintainability of Data Access Services. Data & Knowledge Engineering 70(9), 794–819 (2011)
19. jQuery, http://jquery.com/
20. Jesse James Garrett: Ajax: A New Approach to Web Applications, http://www.adaptivepath.com/ideas/essays/archives/000385.php

Music-Driven Emotion Model Applied in Digitalized Dance Performance of Sacrificial Ceremony for Confucius

Desheng Lv[1], Yuting Wang[1], Chen Guo[1], Zhigeng Pan[2], and Haibin Shi[1]

[1] Department of New Media & Arts, Harbin Institute of Technology, Harbin, China
[2] Digitial Media and HCI Research Center, Hangzhou Normal University, China
{Deshengl,Wangyt,Guochen,Shihaibin}@hit.edu.cn,
zgpan@cad.zju.edu.cn

Abstract. The sacrificial ceremony for Confucius is one of the most important Chinese cultural heritages which could be protected by means of digital technology. A digitalized system, which was developed to reconstruct the musical dance performance of the ceremony, includes music analysis based on emotion model and action movement library based on a virtual model for the choreography with the motion capture technique, and the action correlation with music-dance match. The simulation results of the digitalized dance performance of the ceremony shows that the system allows people to make a better understanding of the cultural evolution.

Keywords: Emotion model, digitalized cultural heritage, sacrificial ceremony for Confucius.

1 Introduction

Musical dance performances are the heart of the sacrificial ceremony for Confucius, which was an ancient ritual to commemorate Confucius's birthday. This ceremony, which was called as *national ceremony*, had been paid more and more attention by Ancient Chinese Emperors since it firstly held in 478 B.C. and perfectly demonstrate the traditional Chinese culture.

Kallmann [1] presents a feature modeling approach to define behavioral information, paying special attention to the capabilities of interaction. Sheng [2] designed a Motion Compilation System in Bianzhong Choreography, and a methodology is presented to achieve the motion connection by analyzing the characteristic action unit and related restriction. As early as in 1996, Goto and Muraoka [3] implemented a virtual dancer "*Cindy*" which can dance in rhythm to musicians' playing. Kovar and Gleicher [4] constructed a directed graph called a *motion graph* for creating realistic, controllable motion, and the general framework they presented can be applied to the specific problem of generating different styles of locomotion along arbitrary paths. Mori and Ohta [5] realized the virtual hip-hop performance which allows users to determine the action sequences sets and builds

Z. Pan et al. (Eds.): Transactions on Edutainment IX, LNCS 7544, pp. 205–212, 2013.
© Springer-Verlag Berlin Heidelberg 2013

relationships between music and movement characteristics. Reidsma and Nijholt [6] designed a virtual rap actor that can dance with music rhythm.

To reconstruct the musical dance performance of the sacrificial ceremony for Confucius, we firstly collected dance movement data so as to build a music-driven digitalized dance animation system based on emotion model. Then, we created three-dimensional human body models with the synthesis of a certain amount of features on the dance unit and judged actions related and music action match degrees. In the end, through simulation platform within the emotional characteristics of dance and music, the intangible cultural heritage found its new space in the contemporary evolution under the conditions of full development in the new era.

2 System Framework

This music-driven dance system relies on four core modules: action movement library module, music emotion analysis module, action choreography module, and music and dance output module, which is shown in Figure 1. The action movement library includes the model setting of figures and the basic dance movements of dance performance. The music emotion analysis module includes the correlation analysis of music and music emotion matched to music. The action choreography module combines music emotion and dance movements by analyzing the dance grammar to establish a network action unit. The last module is based on the former work and makes a comprehensive analysis to the emotional characteristics of dance and music. It also determines to what extent the emotional attributes of music and dance can match each other.

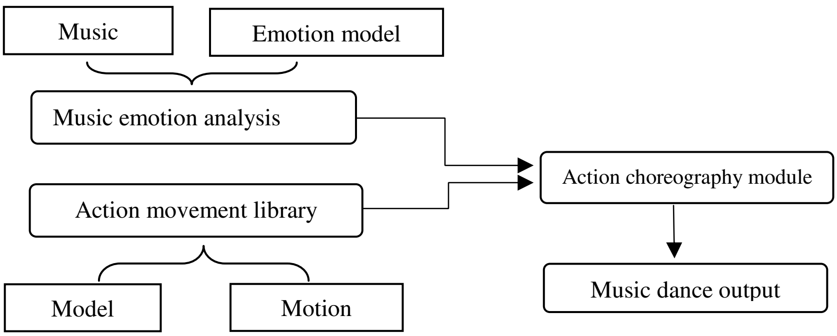


Fig. 1. Music-driven emotion model dance system framework

3 Music Emotion Analysis

To analyze the rhythm of the music, we selected the representative music fragment in dance performance of sacrificial ceremony for Confucius for rhythm resolve. The staff notation and the parsed rhythm are shown in Figure 2 and Figure 3.

Fig. 2. The staff notation of the music fragment

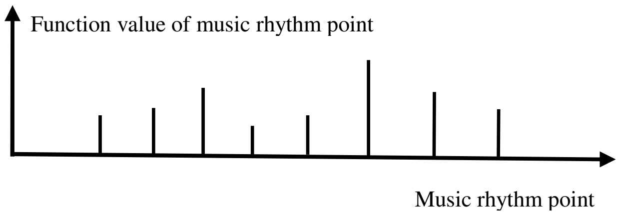


Fig. 3. The parsed rhythm of the music fragment

We choose the Hevner affective ring [6] as emotion model. Hevner affective ring, which is shown in Figure 4, was utilized to fully express the emotional content of music. As a psychological model of the emotional content of music, Hevner affective ring is widely used complying with the connotation of music emotion inherent laws. But Hevner's research aims at the West group. After all, the emotional sub-classes in the vocabulary are not necessarily consistent with China's real situation. Therefore, in this paper, the emotional expression model is based on the ring, but not the same. As picking out eight emotion adjectives groups which are appropriate for the ceremony music, our Emotional Expression Model is shown in Table 1. Also, each type of the Model is a representative of one-eighth of the region in Hevner affective ring.

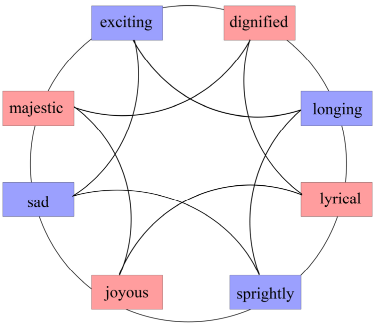


Fig. 4. Hevner affective ring[6]

Table 1. Emotional Expression Model

Category	Example
Dignified	Awe-inspiring Sober
Longing	Longing Tender
Lyrical	Lyrical Leisurely
Sprightly	Graceful Light
Joyous	Bright Joyous
Sad	Heavy Mournful
Majestic	Exalting Majestic
Exciting	Passionate Agitated

4 Action Movement Library

By means of bone skin animation technology, a single dancer role consists of a single skin mesh and bones. Every vertex of the skin is under the influence of one or more bones. We can get the vertices in the correct location in the coordinate system of the world by weighting each bone for its impact on the vertices and then follow the skin mesh and the corresponding weight information.

Through previous collections such as writing records, shoot pictures and videos, we get numerous materials for choreography. Dance movements using motion capture equipment can be recorded. After getting the dance motion data, a dancer character model was established, as shown in Figure 5.

In order to get available processing data, performers are required to put on a one-color, tight-fitting clothes. At key parts of the body, such as joints, hip, elbow, wrist, we will post some of special Marker on them, as shown in Figure 6.

Fig. 5. A dancer character model

Fig. 6. Distribution of markers in motion capture

5 Action Choreography

After building music and dance movement database, we classify the movement styles based on emotion model, identify dance styles, summarize the characteristics of each move unit, and conduct further research about the movement association.

In order to make a division of joint action about the signature style of action units in the physical layer characteristics, a motion equation can be expressed as:

$$Fi = (\Sigma Ai \ \& \ \Sigma Bi \ \& \ \Sigma Ci \ \& \ \Sigma Di) \tag{1}$$

Where Fi means the whole body, ΣAi, ΣBi, ΣCi, ΣDi are four major joint actions, ΣAi refers to head movements, ΣBi refers to trunk movements, ΣCi refers to arm movements, ΣDi refers to the lower limbs. In each category of the style actions, according to the frequency of joints of the body unit, we define the priority factor and determine the optimal solution of the choreography, close to the mark action.

5.1 Action Correlation Analysis

Because of different dance directors, the sacrificial dances will have their own characteristics and styles. However, there are strict standards of sacrificial music and dance. Dancers must observe certain rules and procedures. Dancers' steps, postures, direction and cycles are clearly defined.

If the joints group occurs without meeting the dance aesthetics and logic principles of relative restriction, it must be discarded. Figure 7 shows the movements which are head to left & body left sideways & stretch right hand & reach out right foot, then the following movements are head to right & body right sideways & stretch left hand & reach out left foot, as shown in Figure 8. Considering about aesthetics and logic aspects, movements like bow & body sideways & cross hand & right foot should be avoided, as shown in figure 9.

Fig. 7. Proper movements

Fig. 8. Following movements

Fig. 9. Improper movements

5.2 Music and Action Matching

The key point of music-driven dance is how to calculate the matching degree between the bottom of the extraction music physical characteristics and the rhythm of the dance movement characteristics. Here we use dynamic time warping (DTW) to measure the matching degree. Based on dynamic programming, DTW combines the time warping and distance computational measure to find a calculation method which has the shortest distance between two vectors.

By means of DTW algorithm, and the action of a given music unit series, we can combine the best match actions with a music feature point path. For any random action music clips combination (R, T), R is the action characteristic vector sequence and T is the music characteristic vector sequence. R and T are defined as:

$$R = \{R(1), R(2),\dots, R(M)\}, \qquad T = \{\{T(1), T(2), \dots, T(N)\} \tag{2}$$

Where M is the total frame of action model and N is the total frame of music model.

The computation formula for the accumulating distance is as follows:

$$\lambda(j, k) = d(j, k) + \min\{\lambda(j-1, k-1), \lambda(j-1, k), \lambda(j,k-1)\} \ (1\leq j\leq M,\ 1\leq K\leq N) \tag{3}$$

Where d (j, k) is Euclidean distance between T(j) and R(k). The smaller the Euclidean distance is, the smaller the distortion is, and hence the greater the similarity will be. Through the recursive type we can calculate the matching degree [7].

6 Questionnaire

There are many ways of user experience experimental testing. We have taken a laboratory test method which combine the questionnaire and sound thinking. Sound thinking method is also called the sound thinking test. People to be tested need to incessantly express their own opinion during operating system and express their inner thought. This method has the advantage of stable flexible and low cost. In the interactive user experiment of Confucius memorial ceremony, we guide the user timely and encourage them to express their inner feelings and problem while doing the experience process. We listen to the outward users and induce the endoscopic user to communicate and capture the user's actions and the emotions through words in time.

In order to evaluate and test our approach, 10 undergraduate students are invited to do the user experience research. They all passed the music sensitivity test which proves that they have good music sensitivity. After watching the combination of action-music clips, each participant was required to fill in the questions as listed in the table 2, the category of questions in the questionnaire scores on the satisfaction of the two experiences. 5 points stands for very satisfied, 4 for satisfied, 3 for generally, 2 for not satisfied, 1 for very dissatisfied with.

Table 2. Users' satisfaction results of Confucius memorial ceremony animation

animation	Overall experience	Reality	Easy	Natural	Fun	Flexible
average	3.7	3.5	3.9	3.5	3.2	3.1
N	10	10	10	10	10	10
standard deviation	0.82327	0.52705	1.19722	0.70711	0.91894	1.28668
animation	Attractive	Culture understand	Cultural entertainment	Cultural learning	Initiative participate	
average	3.2	4.1	3	3.7	2.8	
N	10	10	10	10	10	
standard deviation	0.91894	0.56765	0.94281	1.1595	1.0328	

In addition to the questionnaire quantitative study, we also get some valuable opinions from participants. Some have proposed that we should add more mark actions, in order to compile various kinds of dance. Some participants suggested that we can elaborate the dance from the aesthetic view. Moreover, some users proposed that we should create a rich interaction so that people could dance with virtual dancers which mean the synchronization between them.

7 Conclusions and Future Work

In this paper, we developed a music-driven system to digitalize the musical dance of sacrificial ceremony for Confucius. In this system we analyze the music based on Hevner affective ring emotion model, use motion capture technology to build the action movement library and accomplish the digital production through the action choreography. This system to achieve the fit of music and dances in the ceremony for Confucius, and has reference value in how to complete preservation of traditional culture and heritage. Meanwhile, it has further exploration in music-driven dance choreography and interactive aspect between watcher and system on the basis of digital recovery and performance.

However, our work considered relatively limited. It is a really complex problem to measure the matching degree between action and music. One possible solution is to use an autonomic machine learning model, which can get the music connection rules from large scale sample database within the global harmony consideration.

Acknowledgments. This research is co-funded by MOE (Ministry of Education in China) Project of Humanities and Social Sciences (Grant NO: 09YJCZH027) and Heilongjiang Provincial Social Science Funds of China (Grant NO: 08E056).

References

1. Kallmann, M., Thalmann, D.: Modeling Behaviors of Interactive Objects for Real-Time Virtual Environments. Visual Languages and Computing 13, 177–195 (2002)
2. Sheng, J.H., Shao, W., Sun, S.Q.: Exploring of the Motion Compilation System in Bianzhong Choreography. Journal of System Simulation 17(3), 631–634 (2005)
3. Goto, M., Muraoka, Y.: A Virtual Dancer "Cindy" Interactive Performance of A Music Controlled CG dancer. In: Proceedings of the Lifelike Computer Characters, p. 65 (1996)
4. Kovar, L., Gleicher, M., Pighin, F.: Motion graphs. ACM Transactions on Graphics 21(3), 473–482 (2002)
5. Mori, H., Ohta, S.: Automatic Dance Generation from Music Annotation. In: Proceedings of the CHI 2004 International Conference on Advances in Computer Entertainment Technology, pp. 352–353 (2004)

6. Reidsma, D., Nijholt, A., Poppe, R., et al.: Virtual Rap Dancer: Invitation to Dance. In: Proceedings of CHI 2006 Extended Abstracts on Human Factors in Computing Systems, pp. 263–266 (2006)
7. Henver, K.: Experimental studies of the elements of expression in music. American Journal of Psychology 48(2), 246–268 (1936)
8. Lee, H.C., Lee, I.K.: Automatic Synchronization of Background Music and Motion in Computer Animation. Computer Graphics Forum 24(3), 353–362 (2005)

Real-Time Rendering Framework in the Virtual Home Design System

Pengyu Zhu[1], Mingmin Zhang[1], and Zhigeng Pan[2,*]

[1] State Key Lab of CAD&CG, Zhejiang University, Hangzhou, 310027
[2] Digital Media and HCI Research Center, Hangzhou Normal University, Hangzhou, China
{zhupengyu24,zhigengpan}@gmail.com

Abstract. This paper introduces a home design system with its great functions and framework design, including the scene management based on the Cell&Portal system, improved variance shadow mapping and the recently popular real-time rendering framework called deferred lighting. In the implementation details, we put in some useful improvements, such as compressing the Geometry Buffer and Lighting Buffer to decrease the video memory and bandwidth occupation with which the multi-render-target limitation has been dislodged, using the light volume stencil culling which is similar to the shadow volume algorithm to identify the lit pixels and modifying the physically correct shading model based on Fresnel term to adapt to the deferred lighting framework.

Keywords: Scene Management, Soft Shadow, Deferred Lighting, Light Volume.

1 Introduction

With the real estate industry blooming as well as the housing prices continuing higher in recent years, in order to live a life with comfort and economy, people have to consider a variety of factors before purchase and renovation of housing, such as the size and practicality, the cost estimate of decorating and the beauty and comfort of the room layout. Meeting the above various needs, we have developed a user friendly, flexible and powerful virtual home design software. Distinct from the traditional design systems such as AutoCAD, 3ds Max which not only require the users to have a strong design skills but also need a few hours' design cycle, there already exist some home design systems with foolproof operations and powerful functions such as 72xuan, Sweet Home 3D. However these systems still have some weakness, including lack of efficient scene management and poor rendering effects. Premise of powerful features, our system aims at providing users with comfortable operations and efficient rendering effects.

When there are a lot of geometries in the scene, the real-time rendering system needs to cull the geometries which are outside the view frustum efficiently to minimize the geometric information dumped into the pipeline. There exist lots of visibility

* Corresponding author.

Z. Pan et al. (Eds.): Transactions on Edutainment IX, LNCS 7544, pp. 213–224, 2013.
© Springer-Verlag Berlin Heidelberg 2013

algorithms, such as bounding volume test, space partition [1]. Haumont [2] has come up with a visibility test algorithm called Cell&Protal Graph which is especially suitable for interior scenes.

Real-time shadow can greatly improve the image authenticity, but a majority of home design software generally uses the model-bounded texture to simulate pseudo shadow. The most popular real-time shadow generation algorithms include Shadow Volume [3] and Shadow Mapping [4]. The shadow volume algorithm needs to calculate the models' silhouette edges, whose time complexity is associated with the geometric information among the scene, which may become the bottleneck of loading models. The shadow mapping is more widely used among real-time rendering areas due to its simplicity, generality and high speed, but it suffers from aliasing artifacts because of perspective projection and under-sampling errors. Simple solution is to increase the shadow map resolution or use Percentage Closer Filter (PCF) which may affect performance. Donnelly [6] came up with the Variance Shadow Map (VSM) algorithm in 2006, which can produce soft shadow efficiently by blurring the shadow map immediately. VSM is exciting but still has some problems, such as light-bleeding artifact due to variance jittering when the scene is complicated. Wojciech [6] has reduced the lighting-bleeding artifact greatly by combining VSM with Exponential Shadow Map (VSM).

The time complexity of pixel-lighting in traditional real-time rendering framework (forward rendering) is $O(m * n * s)$ where m represents geometry number, n represents light number and s represents pixel number. Therefore it's unable to meet the real-time need when there're multi lights in the scene. Deferred Shading was a real-time rendering framework published by Deering [7] in 1988 which has been widely used in games such as Startcraft2, Battlefield and Dota2 in recent years [8-10]. Deferred Shading stores pixel-related position, normal and material information in multiple textures called Geometry Buffer (GBuffer). The time complexity of pixel-lighting decreases to $O(m + n * s)$ because of decoupling lights and geometries. However, Deferred Shading still has some weakness:

- Fat GBuffer costs so much video memory and bandwidth.
- Rendering to GBuffer needs Video Cards with Multi-Render-Target feature.
- Hard to handle semi-transparency geometries.
- Unable to utilize hardware accelerated MSAA.
- The whole rendering pipeline can only use single rendering equation.

The rendering equation among Deferred Shading can be expressed as follows:

$$L_o(v) = \sum_{i=1}^{n} f_{shade}(B_{L_i}, l_i, v, n, C_{diffue}, C_{specular}, S) \tag{1}$$

B_{L_i}: light's intensity and color information, l_i: incident light vector, v: vertex position, n: vertex normal vector, C_{diffue}: surface diffuse material, $C_{specular}$: surface specular material, S: shininess.

According to Equation (**2**), the GBuffer needs to store normal, position and material. This information may conquer 3 to 4 viewport-sized floating textures, which consume so much video memory and bandwidth. How to slim down the GBuffer becomes the key of improving the performance of Deferred Shading. Our system implements a meliorated framework called Deferred Lighting which was proposed by Naty [11] in 2009. Deferred Lighting has been widely used in video games and game engines such as Crysis2 and Unreal3 [12-13]. Compared to Deferred Shading, Deferred Lighting takes an additional forward geometry pass but reduces the GBuffer size. It also stores lighting result into textures called Lighting Buffer, with which we can use multiple rendering equations among shading. We take some optimizations when implementing this framework to decrease GBuffer and Lighting Buffer to single texture each and eliminate the Multi-Render-Target limitation.

2 Cell&Portal Scene Management

Our system uses Cell&Portal to do scene management. In the implementation details, we maintain a room list and a model list. At First, we compute the room areas as Cells in terms of the walls' information, then identify the doors and windows belong to the associated rooms as Portals, refer to figure 1. Using Equation (**2**) we can efficiently identify the model's belonging room:

$$\vee_{k=1}^{m}\{\wedge_{i=0}^{n-1}(\boldsymbol{v}_{k,i}-\boldsymbol{p})\times(\boldsymbol{v}_{k,(i+1)\%n}-\boldsymbol{p})>0)\} \tag{2}$$

$\times$: represents cross product, m: convex polygon in the room, n: vertex number, $\boldsymbol{p}$: viewpoint position, $\boldsymbol{v}$: represents vertex position.

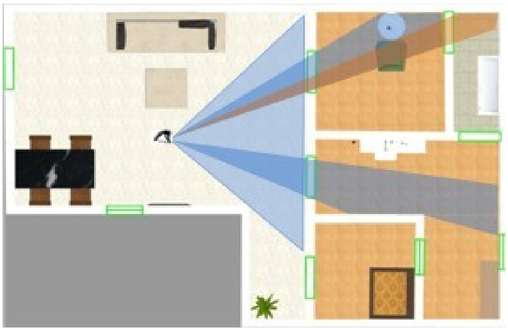


Fig. 1. Cell&Portal in the System

At last, we use the Gribb [14] to extract the world space's frustum information from View-Projection Matrix. With above information, we can compute the models that need to be bumped into the pipeline recursively as the following pseudo code:

```
program Cell_Portal (cell, frustum): model_list
    var model, portal, new_frustum;
  begin
    for(model in cell.model_list)
      if(Intersect(model.OBB, frustum))
        model_list.add(model);
    for(portal in cell.portal_list)
      if(Intersect(portal, frustum))
        new_frustum = Cull(frustum, portal);
        Cell_Portal(portal.other_cell, new_frustum);
  end
```

3 Real-Time Soft Shadow

Variance Shadow Mapping can produce plausibly soft shadow. Due to our small range indoor scene, one median sized resolution shadow map is sufficient. VSM also allows filtering the shadow map immediately, which is much more efficient than the other shadow mapping algorithms that need multi sampling.

We supply three different resolution shadow map due to the current scene range, including 256*256, 512*512 and 1024*1024. VSM uses Chebyshev Inequality (Equation (**3**)) to simulate the shadow factor, so we can filter the shadow map immediately (Mipmap, anisotropic filtering and Gaussian blur) .

$$\begin{aligned} &P(\mathrm{O} \geq \mathrm{R}) \leq p(\mathrm{R}) = \frac{\sigma^2}{\sigma^2+(R-u)^2}\,, u < R \\ &u = E(\mathrm{O}) \\ &\sigma^2 = E(O^2) - E(\mathrm{O})^2 \end{aligned} \tag{3}$$

R: pixel depth, O: associated texel depth, σ^2: texel depth variance, u: texel average depth.

Then we can present the shadow function as follows:

$$\mathrm{S(R)} = \begin{cases} p(\mathrm{R}) & u < R \\ 1 & u \geq R \end{cases} \tag{4}$$

When $u \geq R$ the pixel is lit.

3.1 Light Bleeding

When the occlusions are complex in the scene, light can emerge in some wrong shadow areas due to depth variance jittering, refer to figure 2.

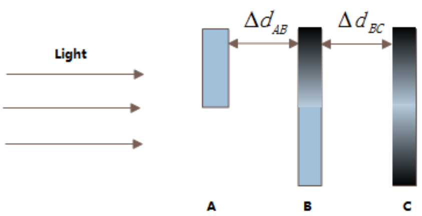


Fig. 2. Light Bleeding: the penumbra of object B bleeds onto C because of variance jittering when $\Delta d_{AB}/\Delta d_{BC}$ becomes big

We add a negative tail value after $p(\mathrm{R})$, this can reduce the bleeding light's intensity, but may cause band artifact, refer to Equation (**5**).

$$p'(\mathrm{R}) = \max\left(\frac{\sigma^2}{\sigma^2+(R-u)^2} - \gamma\,, 0\right), u < R \tag{5}$$

Wojciech solves this problem by combining another shadow mapping algorithm based on statistics called Exponential Shadow Map (ESM) with VSM. This combination can handle most situations except when VSM and ESM both fail, but that hardly happens. Then we get the shadow function as follows:

$$\begin{aligned} &S'(R) = \begin{cases} p''(e^{cR}) & u < R \\ 1 & u \geq R \end{cases} \\ &p''(e^{cR}) = \max\left(\frac{\sigma'^2}{\sigma'^2 + (e^{cR} - u')^2} - \gamma, 0\right) \\ &\sigma'^2 = E((e^{cO})^2) - E(e^{cO})^2 \\ &u' = E(e^{cO}) \end{aligned} \tag{6}$$

c is a control parameter, its value is as big as possible in theory, but too large c may cause floating point precision, we set c 30 in our system. Taking the exponential of depth makes $\Delta d_{AB}/\Delta d_{BC}$ gets smaller.

4 Deferred Lighting

Modified from Equation (**1**), we get a set of equations as follows:

$$\begin{aligned} &\mathbf{L_{diffuse}} = \sum_{i=1}^{n} f_{diffuse}(B_{L_i}, \mathbf{l_i}, \mathbf{n}) \\ &\mathbf{L_{specular}} = \sum_{i=1}^{n} f_{specular}(B_{L_i}, \mathbf{l_i}, \mathbf{n}, \mathbf{v}, S) \\ &\mathbf{L_o} = \mathbf{C_{diffuse}} \circ \mathbf{L_{diffuse}} + \mathbf{C_{Specular}} \circ \mathbf{L_{specular}} \end{aligned} \tag{7}$$

°: tensor product, $\mathbf{L_{diffuse}}$: pixel's diffuse light intensity, $\mathbf{L_{specular}}$: pixel's specular light intensity.

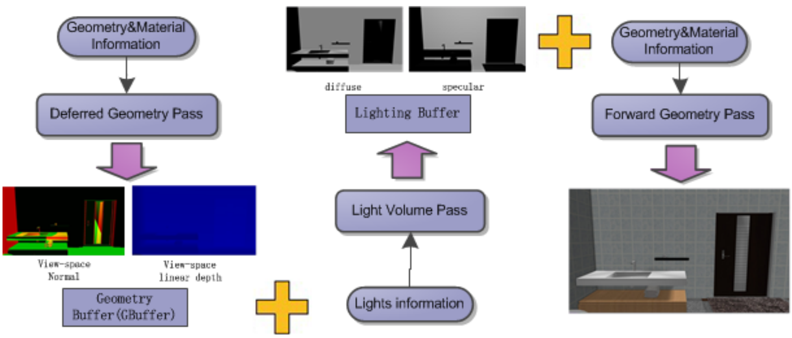


Fig. 3. The Framework of Deferred Lighting

According to Equation (**7**), our GBuffer only needs to store n, v, S, and we can get the two material terms $C_{diffuse}$ and $C_{Specular}$ which occupy two textures in the naïve deferred shading framework from the second geometry pass. Because of the decoupling of lighting pass from shading pass, Deferred Lighting is also called Partial Deferred Shading, refer to figure 3.

In terms of figure 3, there is an additional rendering pass in order to get the Lighting Buffer which stores $\mathbf{L_{diffuse}}$ and $\mathbf{L_{specular}}$. The naïve memory layout still needs two textures which consumes so much video memory and bandwidth, we must compress it into single texture with some optimizations.

4.1 Buffer Compression Optimization

In the implementations, GBuffer uses GL_RGBA16F pixel format and Lighting Buffer uses GL_RGBA8. GBuffer contains $\mathrm{n, v, S}$, among which n and v occupy *xyz* three terms each. We adopt "Spheremap Transform" [16] from CryEngine3 to compress n into two terms and extract the original when needed refer to Equation (**8**).

$$\begin{aligned} \boldsymbol{G_n} &= normalize(\boldsymbol{n}_{xy}) * \sqrt{\frac{\boldsymbol{n_z}*2+1}{2}} \\ \boldsymbol{n}_z &= length\left(\boldsymbol{G}_{\boldsymbol{n}_{xy}}\right) * 2 - 1 \\ \boldsymbol{n}_{xy} &= normalize\left(\boldsymbol{G}_{\boldsymbol{n}_{xy}}\right) * \sqrt{1 - \boldsymbol{n_z}^2} \end{aligned} \tag{8}$$

$\mathrm{G_n}$: the compressed normal information, n : the original normal vector. This compression scheme consumes 18 pixel shader instructions, but compared to some other schemes, it has a better balance between performance and effect.

In fact we can easily reconstruct the view position from pixel's *xy* information in NDC space. So we just store the view space linear depth in the GBuffer, and construct View-Ray to rebuild the view position by similar triangle theorem in very small errors (Equation (**9**)). By means of the above compression schemes, the GBuffer is compressed into a single texture.

$$\boldsymbol{v} = \frac{\boldsymbol{r}*\boldsymbol{f_z}*\boldsymbol{v_z}}{\boldsymbol{r_z}} \tag{9}$$

$\mathrm{f_z}$: the far clip plane's depth in view space, r: the View-Ray which is produced by the rasterization from the far clip plane corners when rendering full screen quad or the pixel view position when rendering light volume.

Lighting Buffer stores pixels' lighting intensity information, including $\mathbf{L_{diffuse}}$ and $\mathbf{L_{specular}}$ which contains three terms each. Assuming that $\mathbf{L_{diffuse}}$ and $\mathbf{L_{specular}}$ have the same color information, we use an approximate solution which just stores the luminance of $\mathbf{L_{specular}}$, and reconstruct its color information from $\mathbf{L_{diffuse}}$ (however specular color information is lost when $\mathbf{L_{diffuse}}$ and $\mathbf{L_{specular}}$ have different color information).

$$\mathbf{L_{specular}}' = \mathbf{L_{diffuse}} \circ \left(\frac{\mathrm{lum}(\mathbf{L_{specular}})}{\mathrm{lum}(\mathbf{L_{diffuse}})+\varepsilon}\right) \tag{10}$$

ε represents a small constant (0.0001 in our implementation) in case of dividing zero.

4.2 Light Volume

In the Deferred Lighting framework, the lighting computation of global lights such as the parallel light from sun is fulfilled by a full screen quad rendering pass. However, there also exist some local lights which may just affect limit number of pixels, it'll be a waste of time to render a full screen quad pass as well. General solution is to use light volume (sphere as point light and comicalness as spot light) to mark out the pixels affected by local lights. Fabio [16] made use of the hardware accelerated sissor test to mark out pixels by projecting the light volume onto screen viewport, but this method may mark out some pixels incorrectly. Our system uses a method similar to the shadow volume algorithm which is able to identify the pixels affected by lights accurately no matter the eye point is inside or outside the light volume. This algorithm can be divided into two passes whose pipeline states are taken out in Table 1.

Table 1. The pipeline states of Light Volume

State	Pass 1	Pass 2
Backface Culling	Back	Front
Color Mask	None	RGBA
Depth Writable	False	True
Depth Test	LessEqual	GreateEqual
Stencil Test	True	False
Stencil Op	Z-Fail Incr	Equal Ref=0
Clear Stencil	False	True

According to the rendering states in Table 1, in Pass 1 we cull the back face of light volumes, the blue pixels pass the depth test and keep the stencil buffer, red pixels fail the depth test and increase the stencil buffer value. In Pass 2 front face is culled, the green pixels pass the depth stencil test and marked as lights affected pixels refer to figure 4-a. When eye point is within the light volume, this method still marks out the pixels correctly seen from Figure 4-b.

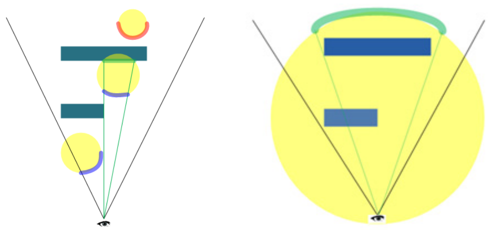


Fig. 4. Light Volume algorithm: (a) eye point outside the light volume (b) eye point inside the light volume

4.3 Rendering Equation

We use the rendering equation with physical correctness published by Sloan [1] to do shading:

$$f_{shade}(B_L, \mathbf{l}, \mathbf{n}, \mathbf{v}, S) = B_L \circ \left(\mathbf{C_{diff}} + R_F(\mathbf{C_{spec}}, \mathbf{l}, \mathbf{h}) \frac{8+S}{8} (\mathbf{n} \cdot \mathbf{h})^S\right) (\mathbf{n} \cdot \mathbf{l}) \tag{11}$$

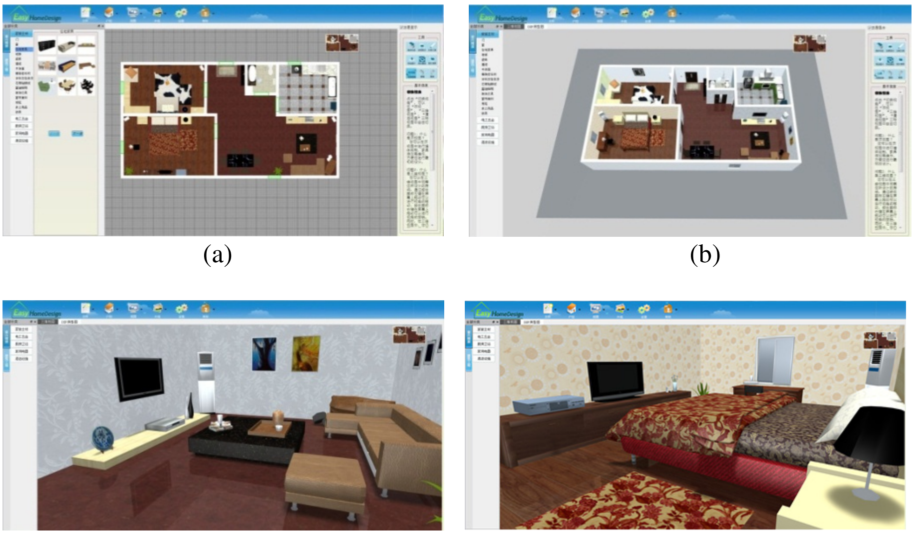
B_L: the light diffuse and specular reflection intensity, $\mathbf{l}$: the incident light, $\mathbf{v}$: the eye vector, $\mathbf{h}$: the half vector in Blinn Phong model, R_F: Fresnel reflection function.

Due to the fact Fresnel function R_F depends on $\mathbf{l}, \mathbf{h}$, which also depend on the light information, this leads the failure to decouple lighting from shading, so we use an approximate solution by modifying R_F as follows:

$$R_F{}'(\mathbf{C_{spec}}, \mathbf{n}, \mathbf{v}) \approx \mathbf{C_{spec}} + (\mathbf{1} - \mathbf{C_{spec}})(1 - \mathbf{n} \cdot \mathbf{v})^{\mathbf{5}} \tag{12}$$

Although this alterative solution is not physically correct, when $\mathbf{v}$ is coinciding with $\mathbf{l}$, the value of $\mathbf{n} \cdot \mathbf{v}$ is the same as $\mathbf{h} \cdot \mathbf{l}$. Then $R_F{}'(\mathbf{C_{spec}}, \mathbf{n}, \mathbf{v})$ no longer depends on the light, it can be applied to our shading framework with unnoticeable artifact.

5 Results



(a) (b)

(c) (d)

Fig. 5. (a) top orthogonal design view (b) 3D design view (c)(d) 3D navigation view

Our software has three different views (figure 5), including the top orthogonal view for wall construction, operating models, laying floors, etc, 3D design view for operating models, pasting wallpaper, etc, and the 3D navigation view for walking through the indoors. The results below are running on a PC with Intel Core2 Quad 2.5G Hz CPU, 4GB RAM, Nvidia GTX460 video card and Windows7 OS.

5.1 Scene Management

We take three variations as experiments when Cell&Portal system is on and off. The results show that the frame rate has a big raise when Cell&Portal is on though the performance still depends on the scene distribution, refer to table 2.

Table 2. Frame rate test under Cell&Portal Management

House Type	Triangle Number	On	Off
West Lake Type A	46987	102 fps	42 fps
Wanan Court B	34959	105 fps	56 fps
Wanan Court C	29871	151 fps	65 fps

5.2 Soft Shadow

In the implementation details, the system automatically choose the shadow map resolution from three candidates (256*256, 512*512 and 1024*1024) according to the scene range. And by setting the value of γ and c reasonably, the light bleeding is alleviated effectively. Seen from figure 6-left, when the scene's occlusion relationship is complex, the original VSM introduces noticeable light bleeding artifact due to variance jittering. On the other hand, too big γ may cause banding as well (figure 6-middle).

Fig. 6. Light Bleeding: left: VSM causes light bleeding middle: VSM with $\gamma = 0.2$ causes banding right: VSM+ESM with$\gamma = 0.1, c = 30$

5.3 Deferred Lighting

The video memory consumed by Deferred Shading and Deferred Lighting with our compress optimization respectively is displayed below (the resolution of viewport is 1072*768). According to table 3, Deferred Lighting uses just the half memory comparing to Deferred Shading and needs only one texture each rendering pass which saves bandwidth.

During the process of rendering, GBuffer contains a GL_RGBA16F float point texture which uses Spheremap Transform to compress normal into RG channels and stores the linear depth in channel G and shininess in channel A, refer to figure 7.

Table 3. Memory Consumption of Deferred Shading and Deferred Lighting

Buffering ǀ Channels	**Deferred Shading**	**Deferred Lighting**
Normal ǀ 2	√	√
Linear Depth ǀ 1	√	√
Shininess ǀ 1	√	√
Diffuse Material ǀ 3	√	X
Sepcular Material ǀ 3	√	X
Emissive Material ǀ 3	√	X
Diffuse Lighting ǀ 3	X	√
Specular Lighting ǀ 3	X	√
Textures	GL_RGBA16F*3	GL_RGBA16F*1 GL_RGBA8*1
Memory(MB)	18.84	9.42

(a) (b) (c)

Fig. 7. Geometry Buffer (GBuffer) (a) Spheremap Transformed normal (b) Linear Depth (c) Shininess

Lighting Buffer contains a GL_RGBA8 texture which stores diffuse lighting information in RGB channel and specular lighting luminance in channel A (figure 8).

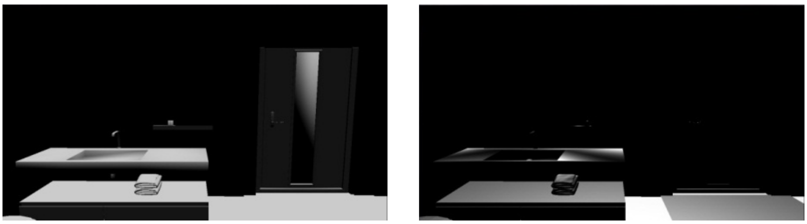


Fig. 8. Lighting Buffer (a) Diffuse Intensity (b) Specular Luminance

When entering the shading pass, geometries, materials and textures are dumped into pipeline, and Lighting Buffer is viewed as a shader resource. After the pixel lighting and shadow generation, the result of the bathroom is showed below:

Fig. 9. Rendering Result

6 Conclusion

This paper introduces a virtual home design system, and describes the next-gen indoor rendering framework with its optimizations in details. In addition to pixel lighting, we can add some features effectively such as screen space ambient occlusion (SSAO), high dynamic rendering (HDR) and depth of field (DOF) in the future.

Acknowledgements. This research work is co-supported by the following NSFC projects: grant no: 61003197, 60970076, 61170318.

References

1. Tomas, M., Haines, E.: Real-Time Rendering, 3rd edn. A.K. Peters Ltd. (2008)
2. Haumont, D., Debeir, O., Sillion, F.: Volumetric cell-and-portal generation. J. Computer Graphics Forum 22(3), 303–312 (2003)
3. Crow, F.C.: Shadow algorithms for computer graphics. In: Proceedings of SIGGRAPH 1977, pp. 242–248. ACM Press, Barzel (1977)
4. Williams, L.: Casting curved shadows on curved surfaces. In: Proceedings of SIGGRAPH 1978, pp. 270–274. ACM Press, Atlanta (1978)
5. Donnelly, W., Lauritzen, A.: Variance shadow maps. In: Proceedings of the 2006 ACM SIGGRAPH Symposium on Interactive 3D Graphics and Games, pp. 161–165. ACM Press, New York (2006)
6. Wojciech, S.: Variance Shadow Maps Light-Bleeding Reduction Tricks. GPU Pro2 Advanced Rendering Techniques. A.K. Peters Ltd. (2011)
7. Deering, M.S., Winner, B., Schediwy, C., Duffy, Hunt, N.: The Triangle Processor and Normal Vector Shader: A VLSI system for High Performance Graphics. J. Computer Graphics 22(4), 21–30 (1988)
8. Shishkovtso, O.: GPU Gems 2: Deferred Rendering in S.T.A.L.K.E.R. Addison-Wesley Professional (2005)
9. Valient, M.: Deferred Rendering in Killzone 2. Presentation, Develop Conference, Brighton (2007)

10. Filion, D., Mcnaughton, R.: Effects & techniques. In: SIGGRAPH 2008: ACM SIGGRAPH 2008 Classes, pp. 133–164. ACM Press, New York (2008)
11. Naty: Deferred Lighting Approaches, http://www.realtimerendering.com/blog/deferred-lighting-approaches/
12. Mittring, M.: Finding Next-Gen: CryEngine 2. In: SIGGRAPH 2007: ACM SIGGRAPH 2007 Courses, pp. 97–121. ACM Press, New York (2007)
13. Samaritan: Unreal Engine 3 Showcase. Technical report, Epic Games (2011)
14. Gribb, G., Hartmann, K.: Fast Extraction of Viewing Frustum Planes from the World View Projection Matrix. Online document (2001)
15. Martin. M.: A bit more Deferred – CryEngine 3. Technical report, Triangle Game Conference (2009)
16. Fabio, P., Franciso, F.: Deferred Shading Tutorial, http://fabio.policarpo.nom.br/docs/Deferred_Shading_Tutorial_SBGAMES2005.pdf

Multiple-Cue-Based Visual Object Contour Tracking with Incremental Learning*

Aiping Wang[1], Zhi-Quan Cheng[1,2,**], Ralph R. Martin[3], and Sikun Li[1]

[1] School of Computer, National University of Defense Technology, P.R. China
cheng.zhiquan@gmail.com
[2] Guangzhou OkGis Institute, Guangdong Province, P.R. China
[3] School of Computer Science and Informatics, Cardiff University, Wales, UK

Abstract. This paper proposes a visual object contour tracking algorithm using a multi-cue fusion particle filter. A novel contour evolution energy is proposed which integrates an incrementally learnt model of object appearance with a parametric snake model. This energy function is combined with a mixed cascade particle filter tracking algorithm which fuses multiple observation models for object contour tracking. Bending energy due to contour evolution is modelled using a thin plate spline (TPS). Multiple order graph matching is performed between contours in consecutive frames. Both of the above are taken as observation models for contour deformation; these models are fused efficiently using a mixed cascade sampling process. The dynamic model used in our tracking method is further improved by the use of optical flow. Experiments on real videos show that our approach provides high performance object contour tracking.

Keywords: Tracking, Snake model, Particle filter, Mixed cascade.

1 Introduction

Visual object contour tracking is a fundamental problem in computer vision, and has been widely investigated because of its usefulness in many fields, such as video surveillance, object recognition, 3D reconstruction, and medical diagnosis.

Active contour, or snake, models [1] are a popular approach to contour tracking. The basic idea is to drive evolution by minimizing energy functional for the object contour; this energy depends on internal spline forces and external image forces. Numerous active contour algorithms have been proposed; both parametric active contours [2,3,4,5,6,7] and geometric active contours [8,9,10], are used, with different representations for the contour curve. In the former, the contour is approximated by an explicit parametric model, typically using a set of control points [1,2,5]; B-splines are often used [3,4]. In the second case, the contour is typically represented by an implicit function, as in the level set method [9,10]. In general, parametric contour methods are more efficient, and are thus more suitable for contour tracking in real-time.

* This work was supported by the National High Technology Research and Development Program (863, No. 2012BAH09B02) and National Science Foundation of China(No.60970094,61103084).
** Corresponding Author.

Z. Pan et al. (Eds.): Transactions on Edutainment IX, LNCS 7544, pp. 225–243, 2013.
© Springer-Verlag Berlin Heidelberg 2013

Traditional snake models suffer from a serious limitation when used for tracking in image sequences: the convergence of results is very sensitive to the initial contour location. To deal with this problem, various estimation tools, such as the Kalman filter and particle filters, can be used to update parameter values over the sequence. For example, [2,7,11,12] used a Kalman filter to track a fixed number of marker points, or parametric values, such as a B-spline's control points. but, the Kalman filter assumes linear system and measurement models, which is unsuitable for many applications.

In order to track contours with non-Gaussian and nonlinear state densities in cluttered video sequence, Isard and Blake [4] introduced the condensation algorithm. They used a B-spline representation for object contours, and particle filters to track the curve parameters given noisy observations. Since their approach only allows affine deformations of the contour, it is unsuitable for deforming objects, undergoing local deformations. Rathi et al. [13] combined a particle filtering algorithm with the geometric active contour framework to give an approach that can be used for tracking moving and deforming objects. However, they directly track affine deformations, while using an approximately linear observer to estimate any non-affine deformation of the object contour. Thus, their method cannot deal with complex contour deformations. Vaswani et al. [14] proposed a further algorithm, *Deform PF-MT*. They suggest that in most real problems, much of the contour deformation depends on a few parameters, while the deformation in the rest of the state space is small. Hence they use the deformations at a small sub-sampled set of locations along the contour as an effective basis space for particle filtering. However, they still explicitly track the contour deformations. In the presence of complexity and uncertainty of object deformations, their method is error-prone. Furthermore, the above approaches only use simple observation models, which which do not provide stable tracking target in the presence of large shape changes or significant occlusion.

Recently, discriminative learning methods, especially incremental ones, have received much attention; they aim to find a decision boundary that can best separate the object from the background, instead of just building a model which only describes the target. Incremental learning methods have been shown to be suitable for tracking objects with appearance variations or within cluttered environments. Wei et al. [15] integrated discriminative methods into a level set framework. However, this approach uses a set of off-line-trained weak classifiers. These cannot cope well with variations in object appearance, or noisy backgrounds. Furthermore, the tracking framework used simply builds on the detection results output by the classifiers, so is unstable due to the accumulation of classification errors.

In this paper, we propose a multi-cue based discriminative object contour tracking algorithm with two major contributions as follows.

1. We give a mixed cascade particle filter tracking algorithm using multiple observation models to improve the accuracy and stability of object contour tracking. The *incremental extremely random forest classifier* (IERF) [16] is adopted as an incremental learning approach to modelling target appearance. We describe contour deformation using separate inter-frame and intra-frame contour deformation models. The former is defined by using bending energy based on the thin plate spline (TPS) model [17], and multiple order graph matching [18] between contours in consecutive frames. The intra-frame deformation model is based on an energy which

depends on the contour evolution process within the current frame. To fuse these observations efficiently, a *mixed cascade* importance sampling process is used. Our method combines a series of observers in multiple stages of importance sampling. However, we do not use a simple cascade. We allow multiple observer outputs to be blended during a single stage of the importance sampling process, and then combine multiple stages in cascade mode. Therefore, we refer to our method as *mixed cascade importance sampling*. We use optical flow information to further improve the accuracy of the particle filter model.

2. We give a novel contour evolution energy which integrates an incremental learning discriminative model with a parametric snake, giving improved performance in contour evolution.

Results on challenging video sequences demonstrate the effectiveness and robustness of our method.

The rest of the paper is organized as follows: Section 2 describes the IERF classifier for object tracking, our discriminative parametric snake model is presented in Section 3. The multi-cue mixed cascade particle filter tracking algorithm is discussed in Section 4. Experimental results are given in Section 5 and conclusions in Section 6.

2 Incremental Extremely Random Forests

First we briefly discuss the incremental learning discriminative model used in the paper: the *incremental extremely random forest* (IERF), originally introduced in [16]. The IERF classifier is suitable for online learning and classification of streaming data, and is especially useful in object tracking problems.

Extremely randomized trees provide a tree-based ensemble method for supervised classification. However it has to be trained in off-line mode: the entire training data must be given in advance. Using an incremental extremely random forest (IERF) leads to a tree-based ensemble method which can deal with online learning with streaming data.

An IERF classifier $H(x)$ builds an ensemble of decision trees, $H(x) = \{h_i(x)\}_{i=1}^{N}$, where x is a sample with unknown label, and N is the total number of decision trees. Each decision tree is grown according to the classical top-down procedure without sample replacement. For each leaf node, an example list is maintained in memory and the a count is kept of the number of items in each labeled class derived from the incoming examples. When a new labeled example arrives, it is routed down the current tree to a leaf node, based on its attribute values. If certain splitting conditions are satisfied, then the leaf node is split into two, and the tree extended.

Let $c(x) \in C = \{1, 2, \cdots, M\}$ be the true label for sample x, where there are M possible labels. Each tree $h_i(x)$ outputs a value $\hat{c}_i(x)$ from label set C as its proposed classification for x. The final output of $H(x)$ is the mode of $\{\hat{c}_i(x)\}_{i=1}^{N}$.

Provision of training data for online learning is an important issue for IERF classifiers. A typical solution is co-training [4], a commonly used approach in semi-supervised learning. The basic idea is that two classifiers classify the unlabeled data, and use this newly labeled data to update each other.

(a) (b)

Fig. 1. (a): target within the white rectangle, (b): confidence map provided by IERF classifiers

An IERF classifier can be easily applied to object tracking by integrating it into the co-training framework [16]. In this paper we use two IERF classifiers working separately in the color histogram space and the space of histograms of oriented gradients (HoG) [19]. In the first frame, the target is identified and a number of overlapping square subwindows are extracted at random positions from the target region and the background area. Then the two classifiers are initialized by using these few pieces of training data. In the following frames, both classifiers classify the incoming data, and and the co-training framework is used: the color classifier chooses some most confident samples to update the HoG classifier, and vice-versa.

The outputs from each IERF classifier are expressed as two independent confidence maps for the pixels, where the classification margin is used as the confidence measure. Large values in the confidence map are likely to belong to the target. We add these two confidence maps together to create a final confidence map: see Fig. 1.

As Fig. 1 shows, this tracking method using IERF classifiers can coarsely locate the target, narrowing it down to a small region of the original video frame. We take advantage of this characteristic to design a novel parametric snake model, which converges more accurately than an ordinary parametric snake model, as we explain in the next section.

Note that the IERF classifiers used in our tracking method can update themselves using the constraint of the object contour determined in each frame, which makes the tracking robust.

3 Discriminative Parametric Snake Model

A *parametric* snake model is used for efficiency. Our model represents a curve defined by a set of discrete contour points (*snaxels*) $\{v_i\}_{i=1}^{N_s}$, where N_s is the number of contour points and $v_i = [x_i, y_i]^T$, the coordinates of v_i in the image. The snaxels are connected by line segments to create the curve. Then, following the usual approach, the contour evolution energy E_{snake} is represented as

$$E_{\text{snake}} = \sum_{i=1}^{N_s} \left(E_{\text{int}}(v_i) + E_{\text{ext}}(v_i)\right) \tag{1}$$

where E_{int} is the *internal energy* produces forces to make the snake contour smooth, while E_{ext} is the *external energy* produces forces attracting the snake to desired image features such as lines and edges.

The internal energy E_{int} at each snaxel is typically defined as $E_{\text{int}}(v_i) = (\alpha|v_i^{'}|^2 + \beta|v_i^{''}|^2)/2$, where $v_i^{'}$ and $v_i^{''}$ are the first and second (discrete) derivatives of v_i, given positive weights α and β respectively. The external energy E_{ext} at each snaxel is typically defined as $E_{\text{ext}}(v_i) = -|\nabla(G_\sigma * I(x_i, y_i))|^2$,where G_σ is a Gaussian kernel with standard deviation σ, and ∇ stands for the gradient operator. $G_\sigma * I(x_i, y_i)$ is the convolves the image with a Gaussian filter to smooth it. These energies produce forces on the snake, causing its contour to move and change shape to minimize the energy E_{snake}.

In practice, it is not easy to use such a definition for the external energy in natural cluttered video sequences. Convergence of the contour to the correct solution can easily be affected wherever parts of the target and background have similar appearance, or wherever image contrast is weak. To help correctly locate the contour, we add a further term, the gradient of the confidence map, into the external energy. This is now defined as:

$$E_{\text{ext}}(v_i) = -\gamma|\nabla(G_\sigma * I(x_i, y_i))|^2 - \lambda|\nabla(G_{\sigma'} * I_{\text{conf}}(x_i, y_i))|^2 \tag{2}$$

where I_{conf} is the confidence map produced by the IERF classifiers (see Section 2), and γ, λ are weighting factors that control the effect of each energy term. Because the confidence map is less noisy than the original video frame, the Gaussian filters have different deviations, and $\sigma^{'} < \sigma$.

In summary, weights α and β control the relative importance of the elasticity and rigidity of the curve, while the γ and λ control attraction of the snake towards the target boundary under as determined by gradients of the video image, and by the gradient of the confidence map. Each weight is in $[0, 1]$.

Since the confidence map provides a much more accurate location for the target than the video pixels, the new external energy more successfully attracts the snake to the right place, as we show in our experiments later.

4 Multi-cue Fusion Mixed Cascade Particle Filter

To track the object contour, we adopt a Bayesian inference framework to estimate each target state sequentially. Below we first review the standard particle filter, and then describe our new multi-cue mixed cascade particle filter tracking framework.

Particle filtering is a sequential importance sampling algorithm for estimating properties of hidden variables in a hidden Markov model, given observations.

Given some set of observations of feature values $O_t = (o_1, \cdots, o_t)$ for a target up to time t, the aim of a particle filter system is to estimate the posterior $p(x_t|O_t)$, where x_t is the state of the target at time t, based on the *observation* model (the *likelihood*) $p(o_t|x_t)$ and the *dynamic* model $p(x_t|x_{t-1})$:

$$p(x_t|O_t) \propto p(o_t|x_t) \int p(x_t|x_{t-1}) p(x_{t-1}|O_{t-1}) dx_{t-1} \tag{3}$$

The tracking result is obtained as the *maximum a-posteriori* (MAP) estimate, which is: $x_t^* = \arg\max p(x_t|O_t)$.

The particle filter approach approximates the integral in Equ. 3 by using a set of weighted samples (particles) $\{x_t^i, w_t^i\}_{i=1}^N$, where each x_t^i is an estimate of state and w_t^i is the corresponding weight. These particles are generated during the initialization stage, and evolve continually.

In our tracking system, the object state is represented as

$$x_t =< c_t, c_{t-1}, s_t, s_{t-1}, S_t, S_{t-1} >,$$

where c_t, s_t, and S_t, are the coordinates of the centroid of the target region (defined as a rectangle here), the size of the target region, and the set of snaxels at time t. The target region rectangle is initialised by drawing it on the first frame.

4.1 Multiple Observation Models

Combining multiple cues can significantly improve tracking performance, as different cues can complement each other and thus overcome failures of individual cues. Therefore, we use multiple observation models which represent both the object's appearance and and contour.

We assume the likelihood probability to be

$$p(o_t|x_t) = \prod_{i=1}^{3} p(o_{t,i}|x_t) = p_{\text{classifier}}(o_{t,1}|x_t) p_{\text{intra\_deform}}(o_{t,2}|x_t) p_{\text{inter\_deform}}(o_{t,3}|x_t) \quad (4)$$

where $o_t = \{o_{t,1}, o_{t,2}, o_{t,3}\}$ represents the state at time t, including the appearance of the object, the intra-frame deformation and the inter-frame deformation separately, and $p_{\text{classifier}}$, $p_{\text{intra\_deform}}$, $p_{\text{inter\_deform}}$ are the corresponding likelihoods.

Given the output H_F of the IERF classifiers described in Section 2, the observation likelihood $p_{\text{classifier}}$ is defined as:

$$p_{\text{classifier}}(o_{t,1}|x_t) = \frac{1}{1+\exp(-\tau H_F(x_t))}, \quad (5)$$

where τ is a control parameter.

Deformation of the object contour can be considered in two ways, within the current frame, and between consecutive frames, for which we use separate observation models. The likelihood of intra-frame deformation is defined according to the snake energy E_{snake} given by Equ. 1:

$$p_{\text{intra\_deform}} \propto \exp(-E_{\text{snake}}/\sigma_s^2), \quad (6)$$

where σ_s is a user defined weighting factor.

The likelihood of inter-frame deformation is calculated using the bending energy of a thin plate spline (TPS) model.

$$p_{\text{inter\_deform}} \propto \exp(-E_{\text{bending}}/\sigma_b^2), \quad (7)$$

where σ_b is a user specified weight.

To find E_{bending}, first, we find correspondences ψ between the sets of snaxels for contours in consecutive frames, using a multiple order graph matching method [18]. A multiple order approach leads to more consistent relationships with more accurate and robust results than using a single order. On the other hand, multiple order graph matching requires fewer iterations than bipartite graph matching.

We solve a linear equation to find the TPS coefficients. Let the two sets of snaxels be $S_{t-1} = \{s_{t-1}^i\}_{i=1}^{N_{t-1}}$ and $S_t = \{s_t^i\}_{i=1}^{N_t}$ in frames at times $t-1$ and t, with N_{t-1} and N_t snaxels respectively. We use two independent interpolation functions f_{m_x} and f_{m_y} to model the coordinate transformations between them. These functions satisfy $x_t^{\psi(i)} = f_{m\_x}(x_{t-1}^i, y_{t-1}^i)$ and $y_t^{\psi(i)} = f_{m\_y}(x_{t-1}^i, y_{t-1}^i)$, where $(x_{t-1}^i, y_{t-1}^i) \in S_{t-1}$ is the image coordinate of the snaxel s_{t-1}^i at time $t-1$, while $(x_t^{\psi(i)}, y_t^{\psi(i)})$ is the image coordinate of snaxel $s_t^{\psi(i)}$ which is the corresponding point to s_{t-1}^i. Because $E_{\text{bending}} \approx b^T K b$ [17], and letting $K_{ij} = U(\|s_{t-1}^i - s_{t-1}^j\|)$, where $U(r) = r^2 \log r^2$, the TPS coefficients b can be obtained by solving the following linear equation:

$$\begin{bmatrix} K & D \\ D & 0_{3\times3} \end{bmatrix} \begin{bmatrix} b \\ a \end{bmatrix} = \begin{bmatrix} v \\ 0_{3\times2} \end{bmatrix} \tag{8}$$

where the i^{th} row of D is $(1, x_{t-1}^i, y_{t-1}^i)$, and the i^{th} row of v is $(x_t^{\psi(i)}, y_t^{\psi(i)})$, from the interpolation conditions $(f_{m\_x}(x_{t-1}^i, y_{t-1}^i), f_{m\_y}(x_{t-1}^i, y_{t-1}^i)) = v_i$. The first and second columns of b satisfy

$$\sum_{i=1}^{N_{t-1}} b(i,1) = \sum_{i=1}^{N_{t-1}} b(i,2) = 0 \tag{9}$$

$$\sum_{i=1}^{N_{t-1}} b(i,1) x_{t-1}^i = \sum_{i=1}^{N_{t-1}} b(i,1) y_{t-1}^i = 0 \tag{10}$$

$$\sum_{i=1}^{N_{t-1}} b(i,2) x_{t-1}^i = \sum_{i=1}^{N_{t-1}} b(i,2) y_{t-1}^i = 0 \tag{11}$$

4.2 Two-Stage Mixed Cascade Importance Sampling

Having introduced multiple individual cues above, a standard particle filter can be directly adopted by updating every particle weight using the observation model: $p(o_t|x_t) = \prod_{i=1}^3 p(o_{t,i}|x_t)$. Instead of updating particle weights directly according to the above observation model, we use two stages of importance sampling to fuse all the observations efficiently in a similar way to the cascade method used in [20], but we do not do so simply in sequence.

As shown in Fig. 2, in the first stage, the particle weights are calculated by using the appearance model and the intra-frame deformation model. We combine the particle weights obtained from the appearance model, and those from the intra-frame deformation model, $p_{\text{classifier}} \cdot p_{\text{intra\_deform}}$, allowing us to preserve the particles which have contours close to the real location. Then in the second stage, we update the particle weights again using the intra-frame deformation model and inter-frame deformation models, because the inter-frame deformation model $p_{\text{inter\_deform}}$ will be meaningful when two contours in two particles are not quite different.

Our method avoids much unnecssary computation, compared to the obvious method of computing all three likelihoods for every particle at each step of updating the particle weight. Because our method does not fuse all observation models fully in sequence like the method in [20], we call it *mixed cascade* importance sampling. See Algorithm 1.

The output $\{x_{t,stage2}^i, w_{t,stage2}^i\}_{i=1}^{N_2}$ of Algorithm 1 is used for state estimation.

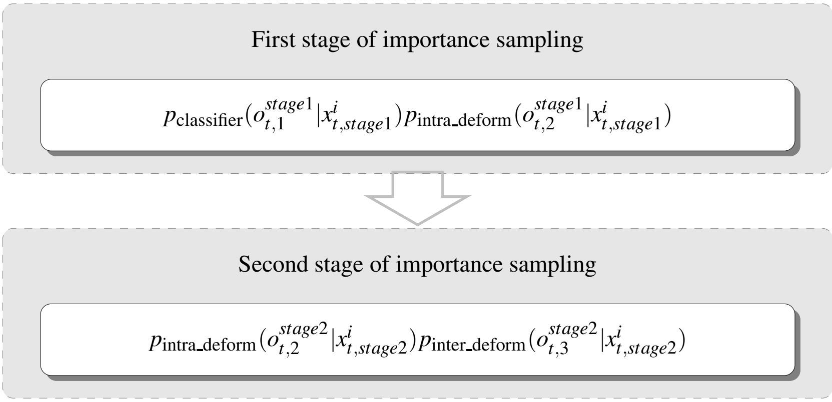


Fig. 2. Mixed cascade importance sampling, showing the two different stages of the importance sampling process

---

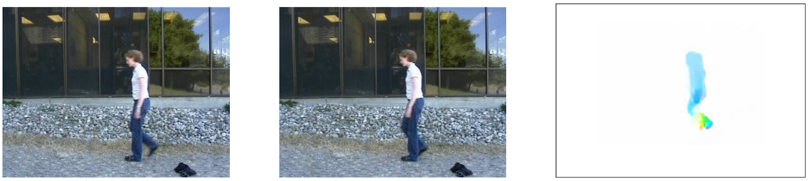
Algorithm 1. Mixed cascade importance sampling (IS)

Calculate the weight $w_{t,stage1}^{i}$ of each particle in the first stage of importance sampling:
    for $i = 1 \cdots N_1$
        $w_{t,stage1}^{i} = p_{\text{classifier}}(o_{t,1}^{stage1}|x_{t,stage1}^{i})p_{\text{intra\_deform}}(o_{t,2}^{stage1}|x_{t,stage1}^{i})$
    endfor
Resample according to the weights $w_{t,stage1} = \{w_{t,stage1}^{i}\}_{i=1}^{N_1}$:
    Generate l_j from $\{w_{t,stage1}^{i}\}_{i=1}^{N_1}$, and replace $\{x_{t,stage1}^{i}, w_{t,stage1}^{i}\}_{i=1}^{N_1}$ by $\{x_{t,stage2}^{l_i}, 1/N_2\}_{j=1}^{N_2}$

Disturb the states of all particles in the first stage:
    for $i = 1 \cdots N_2$
        Get $x_{t,stage2}^{i}$ from $g(x_{t,stage2}|x_{t,stage1}^{i})$
        Let $\lambda^i = g(x_{t,stage2}^{i}|x_{t,stage1}^{i})$, where g is a 0-mean Gaussian function
    endfor
Calculate the weight $w_{t,stage2}^{i}$ of each particle in the second stage of the importance sampling process:
    for $i = 1 \cdots N_2$
        $w_{t,stage2}^{i} = p_{\text{intra\_deform}}(o_{t,2}^{stage2}|x_{t,stage2}^{i})p_{\text{inter\_deform}}(o_{t,3}^{stage2}|x_{t,stage2}^{i})$
    endfor
Normalize $\sum_{i=1}^{N_2} w_{t,stage2}^{i} = 1$.
Note: $o_t = \{o_{t,1}^{stage1}, o_{t,2}^{stage1}, o_{t,2}^{stage2}, o_{t,3}^{stage2}\}$ and $x_t = \{x_{t,stage1}, x_{t,stage2}\}$ are the observations and states at time t. $o_{t,1}$ is the appearance model. $o_{t,2}$ and $o_{t,3}$ are the intra-frame deformation model and the inter-frame deformation model.

4.3 Improved Dynamic Model

An auto-regressive (AR) process [21], like the first order AR model in Equ. 12, is usually used as the dynamic model in particle filter algorithms. A is the state transition matrix, and B defines the transfer radius of the particle. r_{t-1} represents multivariate



(a) (b) (c)

Fig. 3. Optical flow field for two consecutive frames. Image (a) and (b) are two consecutive frames. The color in (c) displays the angle and brightness while white denotes no motion.

Gaussian random noise.

$$x_t = Ax_{t-1} + Br_{t-1}. \tag{12}$$

The main drawback of such a model is that it does not take into account the most recent observation, and a constant velocity model is usually not suitable for real applications. Here we use information obtained from optical flow to revise Equ. 12.

As shown in Fig. 3, optical flow represents pixel motion in a video sequence. In recent years, optical flow computation has become more and more accurate, and GPU computation has provided a number of real-time optical flow algorithms [22,23]. The optical flow in a frame at time t gives a pixel motion vector: $v_{\mathrm{img}} = \{v_x, v_y\}$, where v_x, v_y are optical flows along the x and y axes. An improved first order AR model is given by:

$$x_t = Ax_{t-1} + \delta x_t + Br_{t-1}, \tag{13}$$

where δx_t is the update obtained from v_{img}. The centroid of the target region c_t satisfies

$$c_t = Ac_{t-1} + \bar{v}_{\mathrm{img}}(c) + Br_{t-1} \tag{14}$$

where $\bar{v}_{\mathrm{img}}(c_t) = \{\bar{v}_x(c_t), \bar{v}_y(c_t)\}$ is the average optical flow around point c_t. Each snaxel $d \in S_t$ also satisfies

$$d_t = Ad_{t-1} + \bar{v}_{\mathrm{img}}(d_t) + Br_{t-1} \tag{15}$$

where $\bar{v}_{\mathrm{img}}(d_t) = \{\bar{v}_x(d_t), \bar{v}_y(d_t)\}$ is the average optical flow around point d_t.

Our overall multi-cue based discriminative contour tracking method is shown in Algorithm 2.

We use the first frame to initialise all particles. A bounding rectangle is drawn around the target by the user. The region including the target plus the contour is regarded as a particle. All particles in the first frame share the same region parameters and same contour. The region outside the target rectangles is marked as background.

5 Experimental Results

In this section, we first demonstrate the performance of our discriminative parametric snake model and then the improved dynamic model. Finally, we verify the whole multi-cue based discriminative contour tracking method.

Algorithm 2. Multi-cue based discriminative contour tracking

At $t = 0$, label the initial target region R_0, extract the initial contour point set S_0, initialize the particles $\{x_0^i, w_0^i\}_{i=1}^{N_1}$.
for $t > 0$
  Detect the target by IERF classifiers, and get the confidence region R_t.
  Resample according to the weights of the particles.
  if $t = 1$
    Generate l_j from $\{w_{t-1}^i\}_{i=1}^{N_1}$;
    Then replace $\{x_{t-1}^i, w_{t-1}^i\}_{i=1}^{N_1}$ by $\{x_{t-1}^{l_j}, 1/N_1\}_{j=1}^{N_1}$
  else
    Generate l_j from $\{w_{t-1}^i\}_{i=1}^{N_2}$;
    Then replace $\{x_{t-1}^i, w_{t-1}^i\}_{i=1}^{N_2}$ by $\{x_{t-1}^{l_j}, 1/N_1\}_{j=1}^{N_1}$
  end if
  Calculate the optical flow $v_{\text{img}} = \{v_x, v_y\}$ between frames at time t and $t-1$.
  Predict the new states of the particles using the dynamic model in Eqns.14–15:
  Generate x_t^i from $p(x_t | x_{t-1} = x_{t-1}^{l_j})$ and get $\{x_t^i, 1/N_1\}_{i=1}^{N_1}$
  Calculate the weights w_t^i of the particles using Algorithm 1
  Estimation the state of the target as $\hat{x}_t = \sum_{i=1}^{N_2} w_t^i x_t^i$
  Update the IERF classifiers.
end for

Note: N_1 and N_2 are the total numbers of particles in two importance sampling stages.

For all experiments, we used the same configuration of IERF classifiers. For classification, each video frame is divided into 9×9 patches. An HSL color histogram with 8 bins per patch is used to represent color features and a HoG histogram with 9 orientation bins per patch is used to represent the HoG features. 30 trees are used for each feature space. Our least-recently-used strategy is to remove samples more than T frames old from the forest. Here we choose T from 10 to 50 depending on object appearance and background changes. The GPU implementation of the Huber-L1 method in [22] was used for optical flow calculation.

For all experiments, The model parameters were set to $\tau = 1.0$, $\sigma_s = 0.85$ and $\sigma_b = 0.08$. The deviation σ of the Gaussian filter used for the original video frame was set to 0.5, while the deviation σ' for the confidence map was set to 0.3.

5.1 Test of Discriminative Parametric Snake Model

To test the performance of our discriminative parametric snake model, we implemented two different contour tracking methods. Method 1 shares the same tracking framework proposed in this paper but using the usual parametric snake model (using image gradients as the external energy). Method 2 is our proposed multi-cue fusion mixed cascade particle filter tracking method with the proposed snake model.

The weights used for the snake model in method 1 were $\alpha = 0.4$, $\beta = 0.2$, $\gamma = 0.6$; the weights in method 2 were $\alpha = 0.4$, $\beta = 0.2$, $\gamma = 0.6$, $\lambda = 0.8$. Both methods shared the same particle filter tracking framework, with 20 particles used in each importance sampling stage.

Fig. 4 shows tracking results on a pedestrian sequence. The background has a strong texture which seriously affects tracking methods. Method 1, using the usual snake model, failed on frame 12, with a small region of cluttered background inside the contour on the person's leg. In subsequent frames 14 16, the contours did not converge correctly, and the part in error got larger. However, our method avoided the distraction provided by the cluttered background. The discriminative model helps the extracted contours closely follow the real edge of the pedestrian.

Fig. 5 shows tracking results on another pedestrian sequence. Around frame 254 the target person occludes another person moving across. In method 1, the contour is disturbed by the strong vertical edge of the occluded person. Again, in our method, the contours are quite accurate and the curves enclose only the target person: the improved external energy term forces the contour to converge towards the real edge of the target.

These and other experiments have verified the stabilizing effect of our discriminative parametric snake model.

5.2 Test of Improved Dynamic Model

We next tested the performance of the improved dynamic model used in the particle filter framework. We implemented two different contour tracking methods. Method 1 uses the same tracking framework but using the ordinary AR model. Method 2 was our proposed tracking method with the improved dynamic model.

Both methods used the same particle filter tracking framework, with 20 particles used in each importance sampling stage. The weights of the snake model were $\alpha = 0.4$, $\beta = 0.2$, $\gamma = 0.6$, $\lambda = 0.8$ for each model.

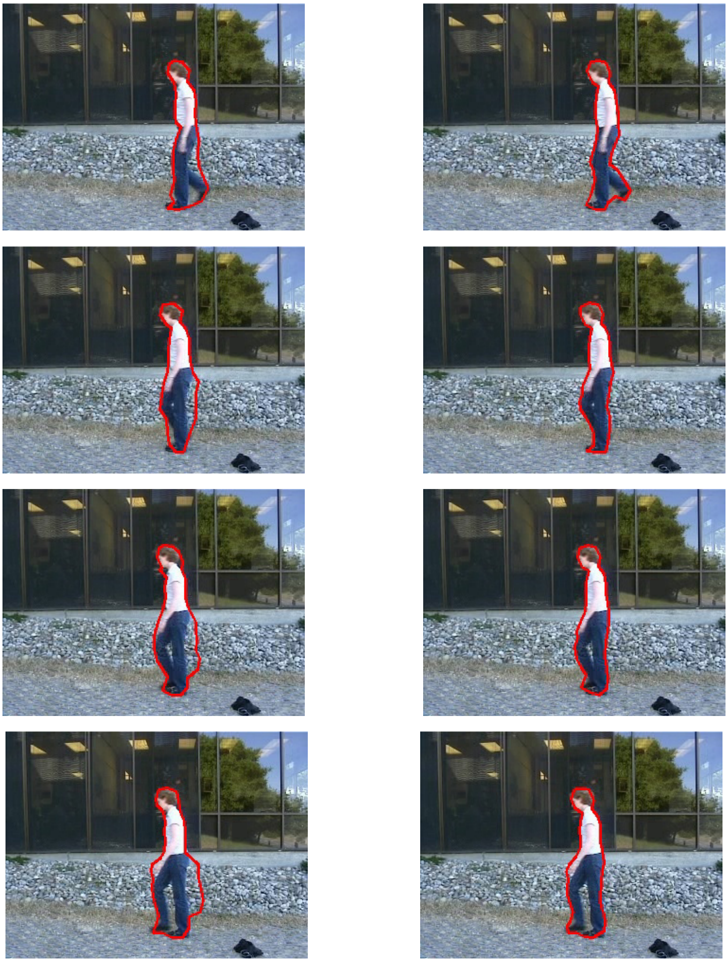
The targets in the video sequences used in this section all moved quickly. Fig. 6 contains a girl's head with a fast motion and Fig. 7 contains a running person. It is clear that method 1 failed to track these targets with fast or irregular motion. From frame 18 in Fig. 6, the face contour includes part of the background region and totally loses the target in the successive frames. In Fig. 7 the person changes location very fast and method 1 cannot find the right location just a few frames after the start. The improved dynamic model used in the tracking framework in method 2 can predict correct locations in successive frames, giving accurate tracking results as seen in Figs. 6 and 7.

5.3 Test on the Proposed Contour Tracking Method

This section demonstrates the performance of our proposed algorithm on several different challenging video sequences. For comparison, we also implemented the *Deform PF-MT* tracking method [14]. In both methods, the initial contour was always located manually in the first frame.

For our method, the weighting factors of the snake energy were again $\alpha = 0.4$, $\beta = 0.2$, $\gamma = 0.6$, $\lambda = 0.8$, and the total numbers of particles in two importance sampling stages were both set to 20.

Fig. 8 shows the tracking result on a face sequence. The camera is moving, so the constant velocity assumption in [14] finds it difficult to predict the object state. Furthermore, the illumination changes significantly, so the observation model in [14] finds it hard to describe the appearance of the target. Thus the Deform PF-MT method soon



(a) Result of method 1 (b) Result of method 2

Fig. 4. Pedestrian sequence with cluttered background. The red closed curve indicates the contour. From top to bottom are frames 12, 14, 15 and 16. Column (a) is the result of method 1, (b) is the result of method 2 (our method).

failed, starting from frame 8. Because of the IERF classifiers, our method is adaptive to appearance changes in the target. Furthermore, the improved dynamic model predicts the state well, making our tracking method more robust to camera shake.

Fig. 9 shows the tracking result on another face sequence. The target in the video sequence is occluded by a magazine in many frames, which makes it hard for the the Deform PF-MT method to find the correct region of the target. However, our method is robust to the occlusion as it uses an adaptive discriminative model.

Fig. 10 shows results on a vehicle sequence. The camera is again unfixed, and the car in the scene moves quickly, making it is difficult for the the Deform PF-MT method to predict the object state. Furthermore, illumination changes significantly (due to the

(a) Result of method 1 (b) Result of method 2

Fig. 5. Pedestrian sequence with occlusion. The red closed curve indicates the contour. From top to bottom are the frames 254, 260, 265 and 272. Column (a) is the result of method 1, (b) is the result of method 2 (our method).

bridge), making it hard for the the Deform PF-MT method to describe the appearance of the target. The Deform PF-MT method failed as the illumination changed, and the contour drifted to another region in the frame—see frame 234. Because of the IERF model, our method is adaptive to the illumination change, and the improved dynamic model also predicts the state well, allowing our tracking method to track the fast moving car.

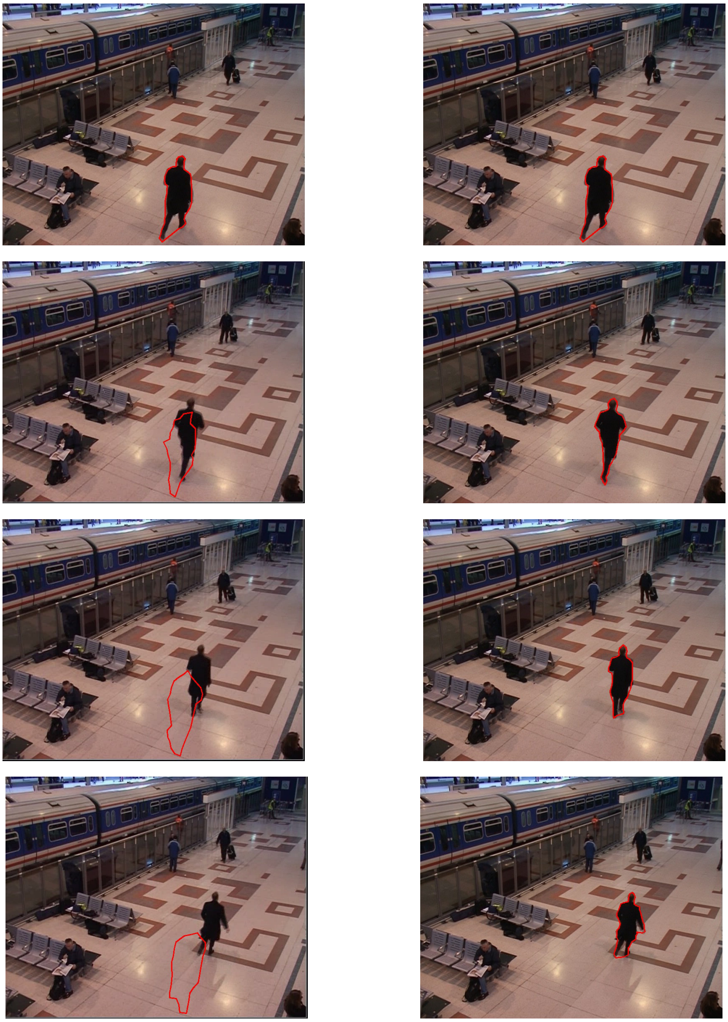
The number of frames successfully tracked for each these challenging sequences is given in Table 1, clearly demonstrating the advantage of our new approach.

(a) Result of method 1 (b) Result of method 2

Fig. 6. Head sequence with fast motion. The red closed curve indicates the contour. From top to bottom are the frames with No.0, 18, 22 and 25. Column (a) is the result of method 1, (b) is the result of method 2 (our method).

Table 1. Number of successfully tracked frames for various sequences, and total number of frames

	Our method	Method in [14]
Face sequence 1	50/50	5/50
Face sequence 2	62/62	6/62
Vehicle sequence	55/55	8/55



(a) Result of method 1 (b) Result of method 2

Fig. 7. Running pedestrian sequence. The red closed curve indicates the contour. From top to bottom are the frames with No.1200, 1206, 1213 and 1220. Column (a) is the result of method 1, (b) is the result of method 2 (our method).

Our method was implemented using un-optimized C++ code and all experiments were run on a 2.3GHz, 1GB PC. The average speed was 3 4 fps for colour sequences of 320×240 resolution. However the IERF classifier can be parallelized, and our method could thus run much faster.

(a) Result of method [14] (b) Result of our method

Fig. 8. Face sequence with camera shake and varying illumination. The red closed curve indicates the contour. From top to bottom are the frames 1, 8, 18 and 38. Column (a) is the result of the the Deform PF-MT method, (b) is the result of our tracking method.

6 Conclusion

This paper has presented a multi-cue discriminative object contour tracking algorithm with two main contributions.

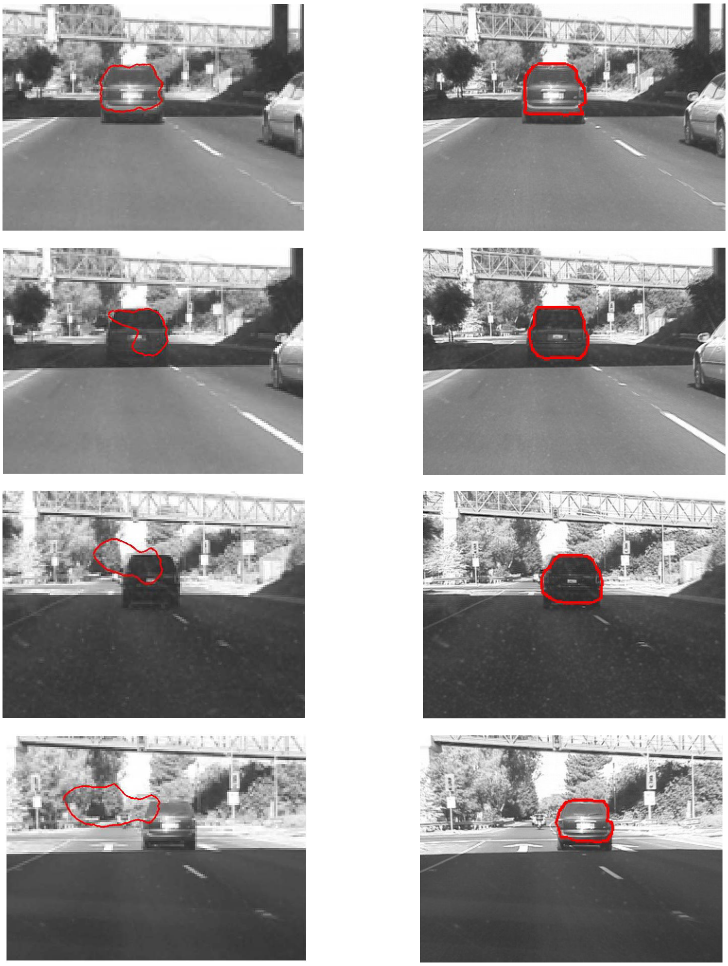
The first is a mixed cascade particle filter tracking algorithm using multiple observation models to improve the accuracy and stability of object contour tracking. An incremental extremely random forest classifier (IERF) is used to provide an incremental model of target appearance. To describe the deformation of the object contour, we use both inter-frame and intra-frame deformation models. The former is based on thin plate spline bending energy, using multiple order graph matching to determine

(a) Result of method [14] (b) Result of our method

Fig. 9. Face sequence with occlusion. The red closed curve indicates the contour. From top to bottom are the frames 264, 272, 276 and 302. Column (a) is the result of method [14], (b) is the result of our tracking method.

correspondence between contours in consecutive frames. The latter is given by the contour evolution energy in the current frame. To fuse these observations efficiently, a mixed cascade importance sampling process is used. Optical flow is used when updating the dynamic model in the particle filter algorithm to achieve further improvements.

The second contribution is a novel contour evolution energy. We integrate an incremental learning discriminative model into the parametric snake model to improve the performance of the contour evolution process.



(a) Result of method [14] (b) Result of our method

Fig. 10. Vehicle sequence with camera shake and seriously varying illumination. The red closed curve indicates the contour. From top to bottom are the frames with No.180, 188, 210 and 234. The column (a) is the result of method [14], while the column (b) is the result of our tracking method.

Results on challenging video sequences demonstrate the effectiveness and robustness of our method.

References

1. Kass, M., Witkin, A., Terzopoulos, D.: Snakes: Active contour models. International Journal of Computer Vision 1(4), 321–331 (1988)
2. Terzopoulos, D., Szeliski, R.: Tracking with kalman snakes. In: Active Vision, pp. 3–20. MIT Press (1992)

3. Blake, A., Isard, M.: Active Contours. Springer (1998)
4. Isard, M., Blake, A.: Condensation – conditional density propagation for visual tracking. Journal of the Society for Industrial and Applied Mathematics (1998)
5. Xu, C., Prince, J.: Snakes, shapes and gradient vector flow. IEEE Transactions on Image Processing 7(3), 359–369 (1998)
6. Frank, Y.S., Kai, Z.: Locating object contours in complex background using improved snakes. Comput. Vis. Image Underst. 105(2), 93–98 (2007)
7. Peterfreund, N.: Robust tracking of position and velocity with kalman snakes. IEEE Trans. Pattern Anal. Mach. Intell. 21(6), 564–569 (1999)
8. Caselles, V., Kimmel, R., Sapiro, G.: Geodesic active contours. International Journal of Computer Vision 22(1), 61–79 (1997)
9. Sethian, J.A.: Level Set Methods and Fast Marching Methods. Cambridge University Press (1999)
10. Osher, S., Fedkiw, R.: Level Set Methods and Dynamic Implicit Surfaces. Springer (2003)
11. Dambreville, S., Rathi, Y., Tannenbaum, A.: Tracking deformable objects with unscented kalman filtering and geometric active contours. In: American Control Conference, pp. 2856–2861 (2006)
12. Chen, Q., Sun, Q.-S., Heng, P.A., Xia, D.-S.: Two-stage object tracking method based on kernel and active contour. IEEE Transactions on Circuits and Systems for Video Technology 20, 605–609 (2010)
13. Rathi, Y., Vaswani, N., Tannenbaum, A., Yezzi, A.: Tracking deforming objects using particle filtering for geometric active contours. IEEE Trans. Pattern Anal. Mach. Intell. 29, 1470–1475 (2007)
14. Namrata, V., Yogesh, R., Anthony, Y., Allen, T.: Deform pf-mt: particle filter with mode tracker for tracking nonaffine contour deformations. IEEE Trans. Image Processing 19(4), 841–857 (2010)
15. Wei, L., Xiaoqin, Z., Jun, G., Weiming, H., Haibin, L., Xue, Z.: Discriminative level set for contour tracking. In: Proceedings of the International Conference on Pattern Recognition (ICPR) (2010)
16. Wang, A., Wan, G., Cheng, Z., Li, S.: An incremental extremely random forest classifier for online learning and tracking. In: IEEE International Conference on Image Processing (ICIP), pp. 1449–1452 (2009)
17. Donato, G., Belongie, S.: Approximate Thin Plate Spline Mappings. In: Heyden, A., Sparr, G., Nielsen, M., Johansen, P. (eds.) ECCV 2002, Part III. LNCS, vol. 2352, pp. 21–31. Springer, Heidelberg (2002)
18. Wang, A., Li, S., Zeng, L.: Multiple Order Graph Matching. In: Kimmel, R., Klette, R., Sugimoto, A. (eds.) ACCV 2010, Part III. LNCS, vol. 6494, pp. 471–482. Springer, Heidelberg (2011)
19. Dalal, N., Triggs, B.: Histograms of oriented gradients for human detection. In: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, pp. 886–893 (2005)
20. Li, Y., Ai, H., Yamashita, T., Lao, S., Kawade, M.: Tracking in low frame rate video: A cascade particle filter with discriminative observers of different life spans. IEEE Trans. Pattern Anal. Mach. Intell. 30, 1728–1740 (2008)
21. Brockwell, P., Davis, R.: Time Series: Theory and Methods (Springer Series in Statistics). Springer (2009)
22. Manuel, W., Werner, T., Thomas, P., Wedel, A., Cremers, D., Horst, B.: Anisotropic huber-l1 optical flow. In: Proceedings of British Machine Vision Conference (BMVC), pp. 1–11 (2009)
23. Marzat, J., Dumortier, Y., Ducrot, A.: Real-time dense and accurate parallel optical flow using cuda. In: Proceedings of The 17th International Conference in Central Europe on Computer Graphics, Visualization and Computer Vision, WSCG 2009, pp. 105–111 (2009)

Building Virtual Entertainment Environment with Tiled Display Wall and Motion Tracking

Liangjie Zhang[1], Xiaohong Jiang[1], Kaibin Lei[2], and Hua Xiong[3]

[1] College of Computer Science and Technology, Zhejiang University
[2] College of Computer Science and Technology, Southwest University for Nationalities
[3] Augmented Reality Group, Hangzhou Realtime Digital Media Tek. CO., Ltd.
zlj100012000@163.com, jiangxh@zju.edu.cn, leikaibin@swun.cn,
xionghua.cadcg@gmail.com

Abstract. Presented in this paper is an immersive and interactive entertainment environment which integrates tiled display wall and motion tracking techniques. A fast calibration method is proposed to achieve geometry alignment and color consistency of the tiled display wall. A robust motion tracking algorithm is designed to obtain the player's moving direction and speed from video streams captured by a web camera. The tracking results are finally used to control the avatar in a virtual scene. The proposed system provides a low-cost and easy to install human-computer interaction environment.

Keywords: virtual environment, tiled display wall, motion tracking.

1 Introduction

With the development of computer hardware, 3D graphics and motion sensing technology, virtual entertainment (VE) has been widely applied for its immersive 3D display and interactive motion sensing. Immersive visual experience is one the most important part for VE. Multi-projector display technology is widely used for its high resolution and brightness. Human-computer interaction is another important feature of VE. Therefore, sensing devices are needed in VE environment to track and analyze a player’s motion and reflect it in the virtual world. Common sensing devices include database gloves, digital pen, touch panel, remote sensing controller, motion tracking camera and so on.

There are many solutions for the implementation of VE. Virtual sports games, such as tennis [1] and table tennis [2], can avoid the demands on sports fields and facilities. Special environment simulations, such as space station [3] and musical instrument [4], provide users with wonderful visual and interactive experience. Some solutions focus on creating systems allowing scalability across different applications [5, 6]. In this paper, a VE environment built with tiled display wall and motion tracking techniques is presented. The main features of the environment include using low-cost devices, easy and scalable installation, fast calibration, and real-time tracking. The PC cluster, projectors and calibration camera of the tiled display wall are all for home use, and

Z. Pan et al. (Eds.): Transactions on Edutainment IX, LNCS 7544, pp. 244–251, 2013.
© Springer-Verlag Berlin Heidelberg 2013

only a low-cost web camera is needed for motion tracking. Projectors placement is not strictly restricted and they can directly project on a planar wall. Projectors can be easily added or removed from the environment according to the need of applications. Geometry and color calibration of the tiled display wall is achieved by a web camera, and can be finished in less than 5 minutes. Motion tracking is achieved in real time with a web camera. Tracking errors can be automatically detected and corrected.

This paper is organized as follows. Section 2 introduces the construction of the multi-projector tiled display wall and Section 3 describes the motion tracking algorithm. Experimental results are given in Section 4 and we conclude in Section 5.

2 Constructing the Tiled Display Wall

2.1 Hardware Architecture

As shown in the Fig. 1, the hardware components of the tiled display wall are:

- A PC cluster consists of 5 PCs, each with a Intel 2.93G duo-core CPU and a GeForce GTS 250 display card for image processing.
- 4 DLP projectors, connected to the display ports of the PC cluster for immersive display.
- 2 web cameras, one with high resolution (1600x1200), used for display wall calibration, and the other with low resolution (640x480), for motion tracking.
- A fast Ethernet switch connecting the PC cluster.

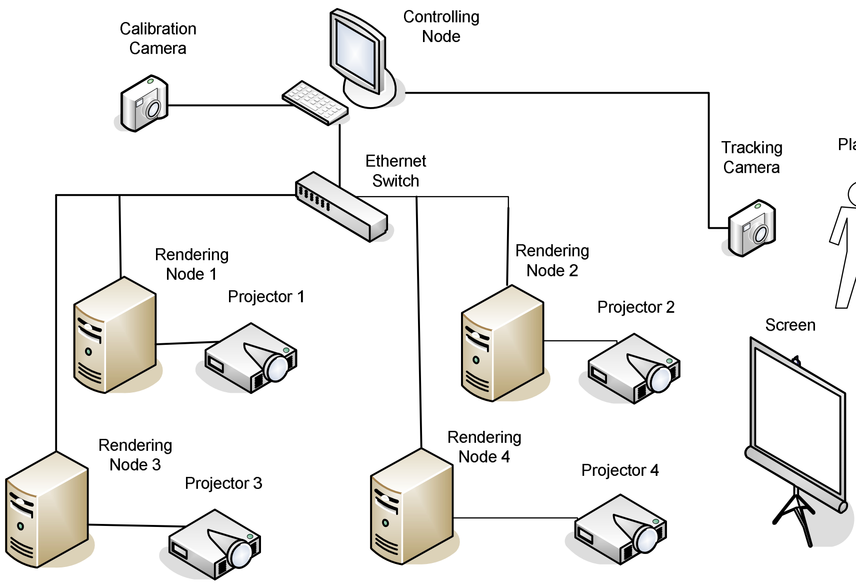


Fig. 1. Hardware architecture of the tiled display wall

2.2 Software Architecture

The software of the tiled display wall adopts the client-server mode. The client end software runs on the controlling node, while the server end software runs on the rendering nodes. The software architecture, as shown in the Fig. 2, consists of the calibration, rendering, display modules, and the motion tracking module. These modules are designed as the middleware between the lower hardware and the upper applications.

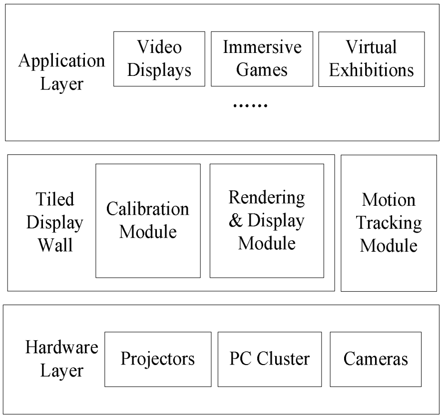


Fig. 2. Software architecture of the tiled display wall

2.3 Display Wall Calibration

Geometry Calibration. Geometry calibration can remove mismatch between the edge boundaries of adjacent projection regions caused by casual placement. As the projection image and display wall are both planar, it is in fact a perspective transformation when an image is projected on the wall. The goal is to find the transformation which warps the image and projects it in the rectangular inner bounding box of the projection region on the wall. According to [7], the transformation can be computed with a planar homography. It is a 3x3 matrix with eight degrees of freedom and can be computed from at least four point correspondences by projecting a pattern image on the display wall and capturing it with a camera. Then the image is pre-warped with the homography and is projected on the wall to get a seamless display.

Color Calibration. Color calibration solves the color inconsistency of the tiled display wall. First, adjacent projectors may have different color characteristics. Secondly, the overlapped projection regions of two or more projectors may cause highlighted bands. As mentioned in [8], color difference between projectors can be described as chrominance and luminance differences. And the difference is mainly caused by luminance when the projectors are of the same brand and type. Therefore, the projection image is first converted from RGB to HSV space, in which Hue (H) and Saturation (S) channel represents chrominance respectively and Value (V) channel represents luminance. Then supervised by a web camera, all the projection images iteratively adjust their Value channel to get a uniform color display. Color blending techniques described in [8] are used to solve the highlighted bands problem. When displaying, the pixels projected in the overlapped region are set as the product of actual value and a weight between 0-1, to assure the total intensity of each pixel in the overlapped regions is equal to the non-overlapped regions.

2.4 Digital Content Rendering and Display

For real-time rendering, a copy of the data, including the 3D models of the scene and the avatar, is saved on the local disk of each rendering node to eliminate the data transmission time. For rendering and display synchronization of different rendering nodes, the rendering process is divided into two passes: the rendering pass and the display pass. In the first pass, the 3D scene is rendered with local data and is saved in the back frame buffer of each rendering node. When all rendering nodes finish the rendering pass, the master node asks all rendering nodes to switch content from the back frame buffer to the front frame buffer.

3 Motion Tracking and Analysis

3.1 Basic Motion Tracking Algorithm

Motion tracking is achieved by using a low-resolution web camera. Focused on the efficiency and accuracy of tracking object in continuous video frames, we adapted the CAMSHIFT algorithm in [9]. The core idea of the CAMSHIFT algorithm is as follows. First the target object is manually selected in a frame and its distribution histogram is computed. The Hue (H) channel of the HSV color space is used for computation, as it's a stable feature when lighting environment changes. Then the window with maximum density distribution for the target object is located by iteratively searching in the current video frame. This process is repeated frame by frame to achieve a continuous target tracking. The CAMSHIFT algorithm can also self-adaptively adjust the center and the size of the search window.

3.2 Improvements of the Motion Tracking Algorithm

The basic CAMSHIFT algorithm is not robust as target missing sometimes occurs, for example, when the player moves outside of the range of the camera. The tracking algorithm is improved to elliminate the cumulative errors in tracking process and to restore the position of the target object.

Combining with Edge Orientation Histogram. Besides color, shape is also a prominant feature of an object. We use the edge orientation histogram (EOH) [10], i.e. the distribution of the orientation of an object's edge points, to describe the object's shape. It is utilized together with the color histogram to get more robust tracking results.

False Target Detection and Relocation. Target object's initial EOH is computed when it is manual selected by the player. During tracking, whenever the object moves to a new search window, a new EOH is computed and is compared with the initial version. If the difference between the two histograms exeeds a threshold, the target object is considered as a false tracking object. And a relocation process for the target object is performed immediately. The algorithm in [11] is used to extracted the contour from the video frame and to compare the content in the contour with the initial object in size, color and EOH. Finally, the best-match contour is selected as the new location of the target object.

3.3 Motion Analysis

During the tracking process, the motion parameters including the direction, distance and speed is analyzed. Moving speed can be computed by moving distance and frame time. Fig. 3 shows the speed curves of two target objects clung to the player's legs in up and down directions, in which positive value means moving up and negative value means moving down. From the speed curves, step frequency of the player can be evaluated and motion state of the player such as working or jumping can also be determined. These motion analysis results are finally used to control the avatar in the virtual scene.

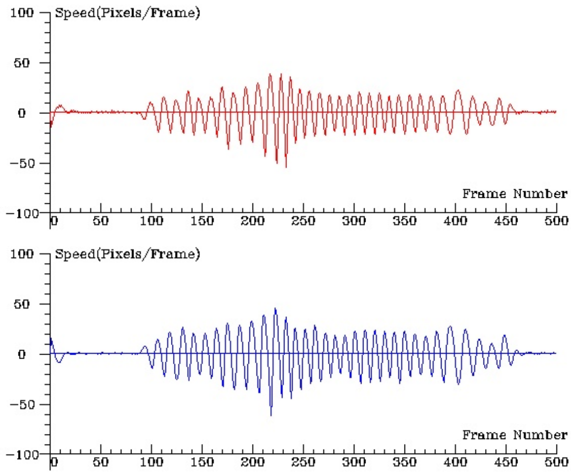


Fig. 3. Speed curves of two target objects

4 Experimental Results

4.1 Hardware Setup

As illustrated in Fig. 4, our VE environment can be easily set up with a PC cluster and projectors. Without projection screen, planar wall is enough for display. The setup process can be finished in 30 minutes.

Fig. 4. Before (left) and after (right) hardware setup of a VE environment

4.2 Display Wall Calibration

As shown in Fig. 5, the geometry and color calibration of a 1x4 projector array can be automatically achieved using a high resolution web camera. Less than 5 minutes is needed for the whole process.

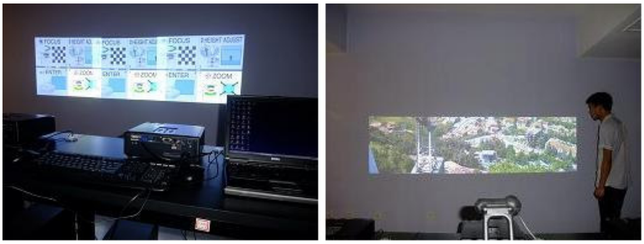


Fig. 5. Before (left) and after (right) calibration of the VE environment

4.3 Performance

According to our experiments, the frame rate of the VE environment achieves 30 frames per second, which meets the requirement of real-time applications. The average processing time for the motion tracking in a frame are listed in Table 1.

Table 1. Average processing time for motion tracking

	Motion Tracking	False Target Detection	Target Relocation
Average Time (ms/frame)	20.3	15.6	30.1

4.4 Application Design

We have developed a 3D virtual touring application named “Water Village”. In this VE application, the avatar of the player can wander in a virtual water village. The walking speed of the avatar follows that of the player. There are 3 modes of wandering: forward, backward and rotation, which are switched by jumping of the player. Fig. 6 illustrates the visual effect of the VE application.

Fig. 6. Virtual touring of the “Water Village”

5 Conclusion

This paper presents an immersive and interactive virtual entertainment environment built on multi-projector tiled display wall and motion tracking techniques. The tiled display wall has the advantage of easy setup, fast calibration and real-time display. The motion tracking process is designed based on low cost device, i.e. physical markers and a web camera. The motion tracking algorithm can perform real-time tracking and self-correction. The proposed system is an ideal low cost platform for experience and development of various VE applications.

Acknowledgments. This work is supported by Grand Science and Technology Projects of Zhejiang Province under Grant No. 2008C11105, Research and Development Program for Application Technology of Sichuan Province under Grant No. 2009GZ0156 and Innovation Fund for Technology based Firms of China under Grant No.09C26213303965.

References

1. Xu, S., Song, P., Chin, C.L., Chua, G.G., Huang, Z.Y., Rahardja, S.: Tennis Space: An Interactive and Immersive Environment for Tennis Simulation. In: Fifth International Conference on Image and Graphics, pp. 652–657. IEEE Press, New York (2009)
2. Li, Y.Z., Shark, L.K., Hobbs, S.J., Ingham, J.: Real-Time Immersive Table Tennis Game for Two Players with Motion Tracking. In: 14th International Conference Information Visualisation, pp. 500–505 (2010)
3. Courter, D., Springer, J.P., Neumann, C., Cruz-Neira, C., Reiners, D.: Beyond Desktop Point and Click: Immersive Walkthrough of Aerospace Structures. In: IEEE Aerospace Conference, pp. 1–8. IEEE Press, New York (2010)
4. Berthaut, F., Hachet, M., Catherine, M.D.: Piivert: Percussion-based Interaction for Immersive Virtual Environments. In: IEEE Symposium on 3D User Interfaces, pp. 15–18. IEEE Press, New York (2010)

5. Yapo, T.C., Sheng, Y., Nasman, J., Dolce, A., Li, E., Culter, B.: Dynamic Projection Environments for Immersive Visualization. In: Computer Vision and Pattern Recognition Workshops, pp. 1–8 (2010)
6. Roman, P., Lazarov, M., Majumder, A.: A Scalable Distributed Paradigm for Multi-User Interaction with Tiled Rear Projection Display Walls. IEEE Transactions on Visualization and Computer Graphics, 1623–1632 (2010)
7. Sukthankar, R., Stockton, R., Mullin, M.: Smarter Presentations: Exploiting Homography in Camera-projector Systems. In: International Conference on Computer Vision, vol. 1, pp. 247–253 (2001)
8. Brown, M., Majumder, A., Yang, R.G.: Camera-based Calibration Techniques for Seamless Multi-projector Displays. IEEE Transactions on Visualization and Computer Graphics 11, 193–206 (2005)
9. Bradski, G.: Computer Vision Face Tracking as a Component of a Perceptual User Interface. In: Proceedings of IEEE Workshop Applications of Computer Vision, IEEE Press, New York (1998)
10. Liu, W.J., Zhang, Y.J.: Edge-Color-Histogram and Kalman Filter-Based Real-Time Object Tracking. Journal of Tsinghua University (Sci. & Tech.) 48, 1104–1107 (2008)
11. Suzuki, S., Abe, K.: Topological Structural Analysis of Digital Binary Images by Border Following. Computer Vision, Graphics and Image Processing 30, 32–46 (1985)

A Calibration Method for Removing the Effect of Infrared Camera Self-radiance on the Accuracy of Temperature Measurement

Shidu Dong[1], Song Huang[2], He Yan[1], and Qun Jiang[1]

[1] Computer College, Chongqing University of Technology, Chongqing, 400050, China
[2] Computer Department, Chongqing Education College, Chongqing, 400065, China
dongshidu@yahoo.com.cn

Abstract. The radiance coming from the interior of an uncooled infrared camera has significant effect on the accuracy of temperature measurement. To remove the effect, this paper proposes a novel three-phase method, which is outlined below. First, from a set of samples, how the pixel value of each blackbody varies with the camera temperature is determined. Second, at the given camera temperature, the calibration function, describing how the pixel value varies with the temperature of blackbody is formed. Finally, with the aid of the calibration function, the temperature of the probe object can be determined by its pixel value. Experimental results show that with the method, the camera can measure the temperature of the object with an error of no more than 1°C.

Keywords: infrared image, camera radiance, two-point-correction, curve fitting.

1 Introduction

Apart from the radiance coming from the object being focused, the pixels on the focal plane array (FPA) in a camera also receive the unfocused flux from the interior of the camera (*camera self radiance* for short), has a significant effect on the measured value of the temperature of the object [1, 2, 3]. In addition, the camera radiance varies with the interior temperature of the camera (*camera temperature*), resulting from the change of the ambient temperature. Even the temperature of the object keep constant, its' the pixel value vary enormously with ambient temperature [4]. In general, there is an error between the measured value of the temperature of an object and the temperature itself, and as ambient temperature around the camera ranging from -5°C to 45°C, this error may be as high as approximately 100°C without regard for the effect. In order to remove the effect, a compensation method [2] had been proposed by us. But, due to the slight change of the responsivity of the pixel with the radiance, when the temperature of the object is far from the temperature of the reference blackbody and the ambient temperature around the camera is far from the ambient temperature where camera calibration data are obtained, the accuracy of temperature measurement is lowering [2].

Z. Pan et al. (Eds.): Transactions on Edutainment IX, LNCS 7544, pp. 252–261, 2013.
© Springer-Verlag Berlin Heidelberg 2013

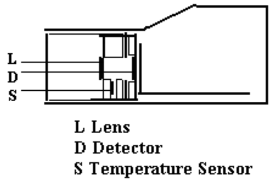


Fig. 1. Illustration of an infrared FPA camera

For improving the accuracy of temperature measurement, this paper presents a three-phase scheme. First, at different camera temperatures, the images of the blackbodies with different temperature are collected. And from these sample data, how the pixel value of each blackbody varies with the camera temperature is determined. Second, in the measurement process, the pixel values of the blackbodies are calculated by camera temperature, and the calibration function at the camera temperature, describing how the pixel value varies with the temperature of blackbody, is formed. Finally, with the aid of the calibration function, the temperature of the probe object can be determined by its pixel value. Experimental results show that the proposed method gains higher accuracy than the previous compensation scheme proposed by [2].

2 Background Knowledge

Fig.1 plots an infrared focal plane arrays (FPA) camera. The detector used in the camera is a LW IRCMOS uncooled integrated microbolometer detector referenced as UL 01 01 1 made by ULIS [5] (see Fig.2), which has 320×240 pixels and a spectral response ranging from 8μm~14μm.

The radiance reaching the camera can be modeled as

$$E_B = \tau\varepsilon L_B(T_S) + \tau(1-\varepsilon)L_A(T_A) + (1-\tau)L_{ATM}(T_{ATM}), \quad (1)$$

where L_B is the blackbody emittance, τ the atmosphere transmission coefficient, ε the emissivity, T_S the temperature of the object being focused, T_A the ambient temperature of the object, and T_{ATM} the atmospheric temperature [2, 6].

In what follows, we adopt the following two hypotheses.

(H_1) The camera is close to the object, and τ is unity. This implies

$$E_B = \varepsilon L_B(T_S) + (1-\varepsilon)L_A(T_A) \; . \quad (2)$$

Fig. 2. UL 01 01 1 detector [5]

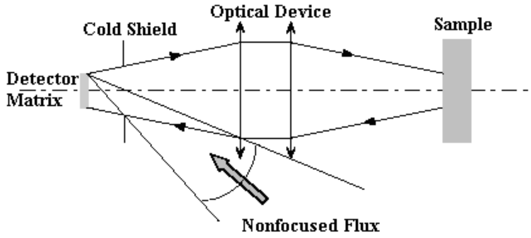


Fig. 3. Radiance received by a pixel [1]

(H_2) The object is a blackbody with emissivity ε=1. Then

$$E_B(T_B) = L_B(T_B) = L_B(T_S)\ , \tag{3}$$

where T_B is the temperature of the blackbody.

Despite a metal packaging is placed in front of the FPA, it can not remove the flux which does not come from the object via lens [1, 2] (see Fig.2 and Fig.3). Therefore pixels on FPA receive not only the radiance from the object being focused (*object radiance*), but also the camera self radiance [1, 2, 6].

$$E_F(T_B, T_C) = \alpha_1 E_B(T_B) + \alpha_2 E_C(T_C)\ , \tag{4}$$

where T_C denotes the interior temperature of the camera (*camera temperature*), i.e. the readout of the temperature sensor S (see Fig.1), E_C camera self radiance, α_1 and α_2 constants. It should be indicated that experimental results show that the camera temperature is always higher than the ambient temperature around the camera.

For simplicity, only the central pixel is considered. Suppose the responsivity of the center pixel is R. Then, the value is determined by the following equation:

$$\begin{gathered} D(T_B, T_C) = R(\alpha_1 E_B(T_B) + \alpha_2 E_C(T_C)) + L_{off}\ , \\ D(T_B, T_C) = R_1 E_B(T_B) + R_2 E_C(T_C) + L_{off}\ , \end{gathered} \tag{5}$$

where $R_1 = R\alpha_1$ and $R_2 = R\alpha_2$. As L_{off} is constant, the value of the central pixel is completely determined by the object radiance, i.e. $E_B(T_B)$, and the camera self radiance, i.e. $E_C(T_C)$.

For some reasons, the values of different pixels on the IRFPA are different in spite of the identical radiance received[8,9], which are commonly corrected by calling the two-point or one-point correction algorithm[10,11]. For our purpose, we describe a modified two-point correction algorithm, which takes the value of the central pixel as standard reference; when other pixels receive the same radiance as the central pixel, their values are corrected to that of the central pixel, thus the estimated temperature values are equal to that of the central pixel.

The modified two-point correction algorithm is related as follows:

$$D'_{i,j} = \frac{D_C^H - D_C^L}{D_{i,j}^H - D_{i,j}^L} D_{i,j} + D_C^L - \frac{D_C^H - D_C^L}{D_{i,j}^H - D_{i,j}^L} D_{i,j}^L \,, \tag{6}$$

where $D_{i,j}$ is the intensity of the (i, j) pixel, the $D'_{i,j}$ its corrected value, $D_{i,j}^H$ the intensity in the high temperature blackbody image, $D_{i,j}^L$ the intensity in the low temperature blackbody image, and D_C^H and D_C^L the intensities of the central pixel.

3 Proposed Method

3.1 Fitting Polynomial of the Pixel Value of a Blackbody with Camera Temperature

At different ambient temperatures, the images of the blackbodies with different temperatures are collected.

For each blackbody, as T_B is fixed, $R^1 E_B(T_B) + L_{off}$ in Eq. (5) keeps constant. As a result, $D(T_B,T_C)$ can be regarded as the function of T_C, which is often approximated by the following polynomial of degree n[2, 12, 13].

$$D(T_B, T_C) \approx F^n(T_C) = \sum_{k=0}^{n} \alpha_k T_C^k \,. \tag{6}$$

The values of $\alpha_n,\ldots,\alpha_0$ can be determined by means of least squares fitting by using the pixel values in collected infrared images of the blackbody at more than n+1 different camera temperatures (*training sample data*). In view of the working temperature range of the cameras from 268.15K to 313.15K (-5°C to 40°C), which departs from origin, the normal equations of the fitting polynomial may be ill-conditional [13, 14]. To overcome the drawback, the following Equation, a translation of Eq.(13), can be adopted.

$$F^n(T'_C) = \sum_{k=1}^{n} \alpha_k \left(T'_C\right)^k + D(T_B, T_M)\,, \tag{7}$$

where, $T'_C = T_C - T_M$, T_M is the median of the camera temperatures of training sample data and $D\left(T_B, T_M\right)$ is the pixel value at T_M.

Let the pixel value of the blackbody at the camera temperature T_C be collected to measure the precision of the fitting function (*testing sample data*). The mean squared error (*MSE*) $\delta^n(T_C)$ is determined by

$$\delta^n = \underset{i,j}{Avg}\left[F^n(T_C^i) - D\left(T_B^j, T_C^i\right)\right], \tag{8}$$

where, i and j denote i^{th} camera temperature and j^{th} blackbody respectively.

3.2 Choice of the Degree of the Polynomial

Experimental results show that the precision of the fitting function varies with the degree of the polynomial. If the training sample data at E different ambient temperatures are collected, then the degree of the polynomial n ranges from 0 to E-1. So, the optimal degree d, which causes the minimal average fitting error, can be determined by

$$\overline{\delta}^{d} = \min_{n\in[0,E-1]}(\overline{\delta}^{n}) . \quad (9)$$

For the purpose, the pixel values of 10 blackbodies with different temperature at 10 different ambient temperatures are collected, seven as the training sample data and the rest as the testing sample data, i.e. E=7. From Table1, it can be seen that, the average MSE of third-degree polynomial reaches minimal value. In addition, the values of a_5 and α_6 of the polynomials beyond degree 5 are less than 10^{-6}. Thus the following third-degree polynomial is adopted.

$$D(T_B,T_C) \approx F^3(T_C) = a_3T_C^3 + a_2T_C^2 + a_1T_C + a_0 . \quad (10)$$

Eq.(6) can be written as

$$D(T_B,T_C) \approx \sum_{k=1}^{3} a_k (T_C - T_M)^k + D(T_B,T_M) . \quad (12)$$

For each blackbody, the values of a_3, a_2 and a_1 can be solved by means of least square fitting with the pixel of the blackbody at least 4 camera temperatures.

Fig. 4 depicts the collected pixel values (see "*") and the fitting curves corresponding to Eq.(12), which approximately represents the dependency of $D(T_B,T_C)$ on T_C.

Table 1. Average MSE and the coefficients of the 1-6th degree polynomials for the testing sample data

degree	*avgMSE*	a_0	a_1	a_2	a_3	a_4	a_5	a_6
1	316787.6	-135.37	12.752					
2	415.44	9.354	51.234	0.154				
3	327.56	1.901	52.434	0.232	0.000289			
4	351.58	1.987	52.502	0.213	-0.000169	-0.000002		
5	383.49	2.451	52.609	0.206	-0.000457	-0.000001	0	
6	458.63	11.015	50.351	0.174	0.00623	-0.000008	0	0

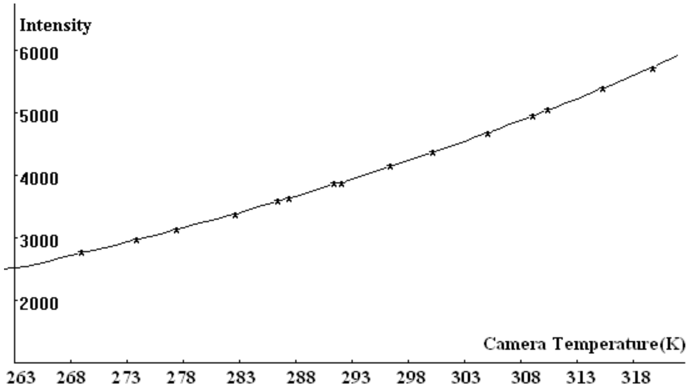


Fig. 4. Pixel value-camera temperature curve of the 40°C blackbody

At the given camera temperature, the pixel values of all blackbody can be calculated by Eq. (12).

3.3 Calibration Function

At the given camera temperature, as T_C is fixed, $R^1 E_C(T_C) + Off$ in Eq. (5) keeps constant. As a result, $D(T_B,T_C)$ can be regarded as the function of T_B (*Calibration function*), which is often approximated by the following polynomial of degree n[2,10].

$$D(T_B, T_C) \approx F^n(T_B) = \sum_{k=0}^{n} \beta_k T_B^k \ . \tag{13}$$

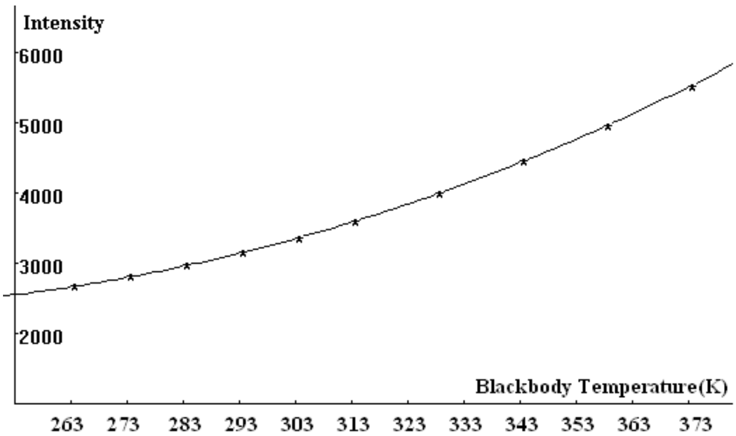


Fig. 5. Pixel valueies are gained from Eq.(12t-blackbody temperature curve at the camera temperature 13.06°C (ambient temperature 10°C)

Similar to the previous section, the optimal degree of the polynomial can be determined, and experiment result shows that the second-degree polynomial is optimal. Thus,

$$D(T_B, T_C) = \sum_{k=0}^{2} \beta_k T_B^k \ . \tag{14}$$

At given camera temperature T_C, the parameter β_k can be solved by means of least square fitting with the pixel values of different temperature blackbodies calculated by Eq. (12).

Fig.5 depicts the pixel values of blackbodies (see "*") and the fitting curves corresponding to Eq. (14), which approximately represents the dependency of $D(T_B,T_C)$ on T_B.

3.4 Process of Temperature Calculation

Based on the previous discussions, we present a complete description of the temperature calculation process as follows.

1) At different camera temperatures, the infrared images of the blackbodies with different temperatures are collected.
2) The parameters α_1, α_2 and α_3 for each blackbody are solved with Eq.(12)
3) Let the image of the object and its' camera temperature T_C be collected.
4) The pixel value of each blackbody at T_C is figured out.
5) At T_C, the parameters β_0, β_1 and β_2 are solved with Eq.(14)
6) The T_B of the object is calculated with Eq.(14)
7) If the object being measured is not a blackbody, its temperature is governed by the rule[17]

$$T_S = \left(\frac{T_B^4 - (1-\varepsilon) T_A^4}{\varepsilon} \right)^{\frac{1}{4}} , \tag{15}$$

where ε and T_A stand for the emissivity and the ambient temperature of the object, respectively, provided τ = 1.

4 Experimental Results

In our experiments, blackbodies referenced as M345X [15] are chosen, and a temperature chamber that holds the cameras is used to change the camera temperature, while the blackbodies to be measured or to calibrate the camera are placed outside of the chamber. The cameras are focused on the blackbodies through a hole that is opened as required.

Due to that the permissible measured values of the temperatures of the cameras range from -10°C to 100°C, 10 blackbodies, whose temperature are –10, 0, 10, 20, 30, 40, 55, 70, 85 and 100°C respectively, are used in the calibration process. In view of the working temperature range of the cameras from -10°C to 45°C and taking the precision requirement into consideration, the pixels of each blackbody at the 10 ambient temperatures (i.e. –5, 5, 10, 15, 20, 25, 30, 35, 40 and 45°C) are collected. In addition, the temperatures of the high temperature blackbody and the low temperature blackbody, which will be used in the two-point correction, are set to be 80°C and 40°C, respectively. In test process, blackbodies are chosen as objects to be probed, whose temperatures are -6, 6, 16, 22, 33, 61, 71, 91 and 101°C respectively, and the ambient temperatures are -4, 1, 5, 10, 14, 27 and 39°C respectively.

Table 2 shows the measured values and errors of the temperatures of the blackbody at various ambient temperatures, where the first row gives various blackbody temperatures, the first column is about various ambient temperatures, and the second column is the corresponding camera temperature. it can be observed from Table 2 that, with the proposed scheme and at an ambient temperature ranging from –4°C to 39°C, the camera can measure the temperature of the object ranging from -6°C to 101°C with a prescribed precision of 1°C.

For comparison, in the experiment of the scheme proposed by [2], the temperature of the reference blackbody is 40°C and the reference ambient temperature is 25°C.

Fig. 6 illustrates the comparison of temperature measurement error between the proposed method (I) and the method presented by [2] (II). Avg error denotes the average temperature measurement error of an object at different ambient temperature, and max error denotes the corresponding maximal measurement error. From Fig. 6, one can see that, for the temperature of the object ranging from -6°C to 101°C, the proposed method gain higher accuracy than the scheme proposed by [2].

Fig. 7 Plots the comparison of temperatures measurement error varying with camera temperature between the proposed method (I) and the method presented by [2] (II). Avg error denotes the average temperature measurement error of the different temperature objects at a certain camera temperature, and max error denotes the corresponding maximal measurement error. From Fig. 7, it can be shown that, for the camera temperature ranging from 0°C to 43°C (ambient temperature from -5 to 40°C), the proposed method gain higher accuracy than the scheme proposed by [2].

Table 2. Measured values and errors of the temperature of the blackbody (°C)

Ambient Temperature	Camera Temperature	-6	6	16	22	33	61	72	91	101
39	40.44	-5.46	6.61	16.25	22.51	32.85	61.38	71.12	91.65	100.57
27	28.63	-6.10	5.87	15.60	22.25	32.79	61.64	71.99	91.33	101.25
14	15.88	-5.86	6.86	15.65	22.17	32.76	61.31	72.57	91.22	101.06
10	13.13	-5.70	6.31	16.21	22.27	32.77	61.09	71.60	90.69	101.70
5	10.44	-5.61	6.18	15.69	22.27	32.81	61.32	72.56	91.31	101.81
1	6.19	-5.11	6.78	16.38	22.3	32.77	61.47	72.6	91.85	101.92
-4	1.13	-5.14	6.30	16.27	22.47	32.78	61.32	72.35	91.78	101.83
Average Error		0.46	0.45	0.31	0.32	0.21	0.36	0.48	0.49	0.57
Maximal Error		0.89	0.86	0.40	0.51	0.24	0.64	0.88	0.85	0.92

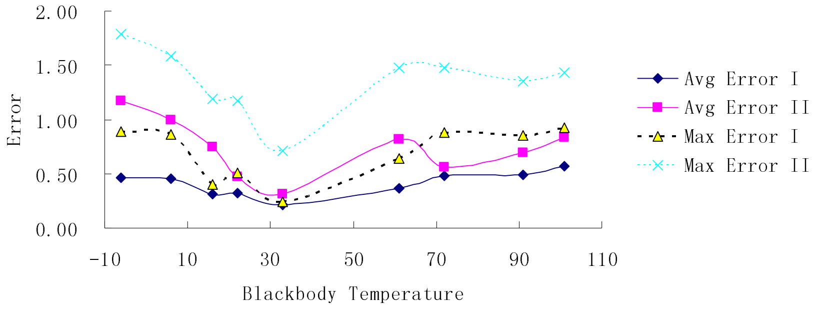


Fig. 6. Comparison of temperature measurement error varying with blackbody temperature between the proposed method (I) and the method presented by [2] (II)

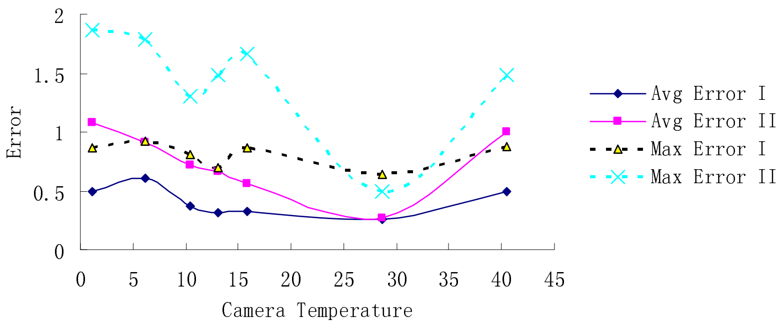


Fig. 7. Comparison of temperature measurement error varying with camera temperature between the proposed method (I) and the method presented by [2] (II)

5 Conclusion

This paper has proposed a scheme for removing the effect of the radiance from the interior of an infrared temperature-measuring camera. The method utilizes curve fitting to approximate the function of the pixel of a blackbody with the camera temperature and the calibration function. Through the two functions, the effect of camera-self radiance can be removed, and the temperature of the object can be figured out. Experiment results show that the proposed scheme gain higher accuracy of temperature measurement than the method presented by [2].

Acknowledgments. The authors are grateful to the anonymous reviewers for their valuable comments and suggestions. This work is supported by Science and Technology Research Project of Chongqing's Education Committee (KJ110806).

References

1. Horny, N.: FPA camera standardization. Infrared Physics & Technology 44, 109–119 (2003)
2. Dong, S., Yang, X., He, B., Liu, G.: A transplantable compensation scheme for the effect of the radiance from the interior of a camera on the accuracy of temperature measurement. International Journal of Infrared and Millimeter Waves 27, 1517–1528 (2006)
3. Fieque, B., Crastes, A., Tissot, J.L., Tinnes, S.: 320×240 uncooled microbolometer 2D array for radiometric and process control applications. Detectors and Associated Signal Processing 5251, 114–120 (2004)
4. Jurgen, H.: High-temperature measurement techniques for the application in photometry, radiometry and thermometry. Physics Reports 469, 205–269 (2009)
5. SOFRADIR, 320×240 LWIR Uncooled microbolometer detector technical specification, ID ML073-V3/28.03.02/NTC Issue 9, 1–15 (2002)
6. Kang, C.: Computer simulation of compensation method for infrared focal plane array. Journal of Beijing Institute of Technology 9, 330–333 (2000)
7. Tissot, J.L.: IR detection with uncooled sensors. Infrared Physics & Technology 46, 147–153 (2004)
8. Scribner, D.A., Kruer, M.R., Killiany, J.M.: Infrared focal-plane array technology. Proceedings of the IEEE 79, 66–85 (1991)
9. Milton, A.F., Barone, F.R., Kruer, M.R.: Influence of nonuniformity on infrared focal plane array performance. Optical Engineering 24, 855–862 (1985)
10. Zhou, H., Lai, R., Liu, S., Jiang, G.: New improved nonuniformity correction for infrared focal plane arrays. Optics Communications 245, 49–53 (2005)
11. Dong, S., Yang, X., Yang, W., Yan, H., Wang, Y.: A method for removing the effect of the camera radiance on the infrared image. International Journal of Infrared and Millimeter Waves 29, 499–507 (2008)
12. Lu, M., Zhao, H.: Modeling and Optimizing of Joint Inventory in Supply Chain Management. In: Pan, Z., Cheok, A.D., Müller, W., Yang, X. (eds.) Transactions on Edutainment V. LNCS, vol. 6530, pp. 71–79. Springer, Heidelberg (2011)
13. Richard, L.B., Douglas, F.J.: Numerical analysis, 7th edn., pp. 104–107. Brooks Cole, Pacific Grove (2000)
14. Xiong, Y., Li, G., Han, G.: Mean Laplace–Beltrami Operator for Quadrilateral Meshes. In: Pan, Z., Cheok, A.D., Müller, W., Yang, X. (eds.) Transactions on Edutainment V. LNCS, vol. 6530, pp. 189–201. Springer, Heidelberg (2011)
15. Mikron, M345X Series: Large area precision blackbody calibration sources for test and measurement applications, `http://www.lumasenseinc.com/EN/products/calibration-sources/large-area-calibration/`
16. Shao, B., Zhang, M., Mi, Q., Xiang, N.: Prediction and Visualization for Urban Heat Island Simulation. In: Pan, Z., Cheok, A.D., Müller, W. (eds.) Transactions on Edutainment VI. LNCS, vol. 6758, pp. 1–11. Springer, Heidelberg (2011)
17. Hetsroni, G., Mosyak, A., Pogrebnyak, E., Rozenblit, R.: Infrared temperature measurements in micro-channels and micro-fluid systems. International Journal of Thermal Sciences 50, 853–868 (2011)

Enjoying of Traditional Chinese Shadow Play – A Cross-Culture Study

Yan Shi[1], Fangtian Yin[2], and Jinhui Yu[1]

[1] State Key Lab of CAD&CG, Zhejiang University, 310027 Hangzhou, China
[2] Department of Design Arts, Zhejiang University, 310027 Hangzhou, China
{hzshiyan,yingft}@gmail.com, jhyu@cad.zju.edu.cn

Abstract. Piying, the world's intangible culture heritage, is an old Chinese art form and one of the origins of modern movie, which encountering the risk of extinction. The spirit of traditional Piying is to express rich emotion and stories through action change by artists. A cross-culture study was conducted to investigate the effect of Piying-induced emotion on heart rate and heart rate variability during Piying perception and performance. The result confirmed that Piying performance was far more effective in emotion induction than Piying perception. The results suggested Chinese are more fond of traditional Piying elements and American prefer experiencing interaction of Piying while the result of Japanese is between those of Chinese and American, they hope that there is a rule to follow in Piying show. Our approach is the first explorative emotion study on Piying art, which could be used as design mechanism to re-create and inherit Piying culture.

Keywords: Emotion, Culture heritage, Performance, Perception, Heart rate variability.

1 Introduction

Piying, a traditional Chinese Shadow Play, is a characteristic Chinese folk art form, which is regarded as one of the origins of modern movie. [1],[2] In Piying, the shadow of fur made 2D characters with delicate carving could be seen by audience in front of the curtain. The artists behind the curtain control the actions of shadow using sticks fastened to the characters. [5] But now, Piying gradually fades away in people's life, encountering the risk of extinction. [3],[4] It is time that it be supposed to consider the way to reform and recreate Piying art form and to reserve a foothold of Piying among the fierce art performance market.[1]

The spirit of traditional Piying performance is to express rich stories and emotion through action change controlled by the artists. [5] The successful approach to protect Piying culture is determined by the extent to which its attractiveness enables people to have a positive experience and increase the immersion of Piying culture content, which means the various emotion induced by Piying culture.

Z. Pan et al. (Eds.): Transactions on Edutainment IX, LNCS 7544, pp. 262–277, 2013.
© Springer-Verlag Berlin Heidelberg 2013

The culture legacy of Piying art including Piying perception and performance. Piying art is also universally enjoyed by both performers (instrumentalists and/or player) behind the scene and those who watch Piying performances (active or passive audience) in the front stage. Piying performance is distinctly different from Piying perception because the performer can integrate voluntary sensorimotor actions that convey his or her unique artistic expression and emotions during the performance of a Piying story.

Piying art is Intangible cultural heritage (ICH) which is tangible and can be interacted with. ICH emphasizes people-oriented skill, experience, spirit, with features of active state, flowing and change. If already the world may mean so many different culture styles, how can we expect one country's culture treasures be enjoyed and interacted with abroad, how can traditional Piying art can be enjoyed by modern people? People's cultural background is also often mentioned to influence people's feeling of a kind of culture impression. Moreover, after culture shock, foreigners may be enlightened by Piying art, and may be able to take his or her experience home. Great ideas to Piying art may also be enriched by foreign influence. Few Piying studies investigated the relationship between people's cultural background and Piying art induced emotion. Thus, with this study, we aim to investigate the perception effect and performance effects of Piying-induced emotion in a cross-culture study then explore the basic elements for Piying art related essence which is worth inheriting and re-creating.

2 Related Work

2.1 Culture Chinese Shadow Play--Piying

With a history of two thousand years, Chinese Shadow Play-Piying is a splendid art, combined with delicate hand craftsmanship and folk drama, which means Piying is the sole graphic art performed at the stage in Chinese folk arts. [1] As the world's intangible culture heritage, Piying is commonly viewed as a traditional edutainment approach to express a Chinese historical socio-cultural message of faith and customs, such as a specific lifestyle, a traditional festival, through the show's interactive performance. [4],[6] It was spread to Turkey, Germany, France, Italy, Russia, Thailand and other countries around world in 13th century. [4],[1],[8]

2.2 The Situation of Piying

However, with the development of film, television, PC games, Piying gradually fades way in people's life, encountering the risk of extinction. [7] People have sense of strangeness and distance for the traditional and ancient art, Piying. [3],[4] Chinese government has established cultural protect associations to protect Piying, such as Piying museum and Piying based research groups. [5],[3] Such way has protected and inherited Piying but cut off connection between Piying and primitive background. Some researches present a set of techniques developed for turning Piying show into electronic form or made Piying animation system and a set of techniques developed for turning Chinese Shadow Play into electronic form. [10],[11],[12] Hsu use motion planning method to design an animation system that can generate the motions of a character in a shadow play automatically [13],[14] And new computer software has been designed to create Piying animation characters. [9],[15],[16] A few programs focus on interactive performance under mixed real environment, generate new digital media and animation creation plan for Piying animation. [17],[5],[18] In foreign countries, there are associations and curriculums focusing on Piying on campuses [19] and a special gallery about Piying in the museum in Bursa of Turkey showing many kinds of Piying props and production supplies in different periods. People there have applied Piying to TV drama and published many books about Piying. [1]

However, the method of preserving is limited to several common computing technologies, like the screen animation. Although via these methods, people could receive partial feeling and information about Piying, the vivid performance experience of Piying show and the improvised emotional expression presented during the performance cannot be delivered to people.

2.3 The Relationship between Emotion and Culture

Culture is shared, is adaptive or has been adaptive at some points in the past, and is transmitted across time and generations. [20] Culture plays a central role in shaping emotional and cognitive experiences. Cultural norms have profound implications for the ways in which emotions are constituted, experienced, expressed, and managed. [21][22][23] Thus, emotions vary across cultures.

From the aspect of social-culture, in the Chinese way of life emphasis is put on the individual's appropriate place and behavior among his fellowmen. [24] Chinese and Westerns are two opposite cultures: interdependent cultures versus independent. From the art-culture aspect, Chinese are implicit and extroverted in aesthetics so that Chinese design enjoys the unique art style of formal beauty. However, western culture determines the thought of western design. The design idea of the west is extroverted aiming at expressing form and shape. Impacted by different cultures, people are different in thinking and dealing with things. Thus there is difference in selecting and perceiving emotion.

Culture is closely related to human life. It cannot be called culture without emotion. Even though China and the west have different cultures, connotative culture can surpass such difference to reach uniformity of emotion and culture.

2.4 Heart Rate Variability (HRV)

Changes in emotions generated by exposure to Piying art undoubtedly involve interactions with central neural network responsible for autonomic nerve functions, which can in turn affect peripheral cardiovascular functions. Heart rate variability (HRV), as an index of cardiac vagal tone, is associated with autonomic flexibility and emotional responding. [25],[26],[27] In that way, HRV has been presented as a good and accessible research tool to study the understanding of emotional effects of Piying art. There are no literature-revealed investigations of the emotion-associated autonomic cardiovascular responses of Piyng culture yet.

2.5 Present Study

In the present study, we first want to get Heart rate (HR) and Heart rate variability (HRV) information about the emotion-related autonomic and physiological responses to Piying perception and performance. Take Chinese and USA participants in our study as optimized cultural diversion and choose Japanese as a country with similar cultural background. Hofstede [28] provides an empirical framework of culture by defining several dimensions of culture. China and the USA differ substantially on all dimensions and Japan almost ranks between this two countries at all dimensions. (see Table 1)

All subjects have not got in touch with Piying show before. We conducted a 3x2 one-between-one-within-subjects mixed experiment to investigate the effects of emotional input on HR and HRV when participants from three countries watching one Piying show in perception and performance conditions.

Table 1. HOUFSTEDE's culture dimension scores for Chinese,USA and Japanese

	Chinese	USA	Japanese
Power distance	80H	40L	54M
Individualism	20L	91H	46M
Masculinity	50M	62H	95H
Uncertainty avoidance	60M	46L	92H
Long-term orientation	118H	29L	80H

H, top third; M, medium third; L, bottom third (among 53 countries and regions for the first four dimensions; among 23 countries for the fifth).

Make questionnaire for each subject after HRV experiment. The questionnaire design pays attention to the feeling of subjects for Piying perception and performance, and designs questions as per essences of traditional Piying perception and performance.

In this framework of emotion-induced exploration, we defined the following two research questions: (a) The levels of emotion attainable with performance may be necessarily higher than those experienced during Piying perception; (b) What influence has the subject's cultural background on the subject's experience on Piying culture.

The apparent dearth of information about the emotion-related responses to Piying show is surprising. It is the first time to research the inheritance and innovation of Piying culture from emotion acpect; and focus on the value of Piying performance and its influence on Piying culture development. In our view, the result can be a powerful design driver that helps connecting culture, emotion and art aspects. Such information may be potentially more valuable than information only concerning Piying perception. The study of cultural background differences in emotional response to Piying art can provide considerable insight into the culture connotation and intercultural acceptance. Such information is also of potential value for cultural educator who tries to combine key elements of Piying art then turn it to edutainment, also benefit people who have economic interests in these matters. Such thinking mode of cultural protection and inheritance can be applied in protection process of other intangible cultural heritage.

3 Methods

3.1 Subject

There are twenty native Chinese and forty participants coming from America and Japan respectively, half men and half women. All subjects' age ranged from 20 to 31 years (mean±SD= 23.8±4.9 years). All students studied at the Zhejiang University, and have been paid to participate in this study. None of them have watched Piying show before. Specifically, take Chinese young adults who are familiar with Chinese culture and corresponding American and Japanese who are unfamiliar with Chinese culture as subjects. All subjects have gone through standard medical health checkups which include the measurement of resting ECG, BP, chest X-ray, haemogram, and liver function, and all were classified as being "healthy".

All of the Chinese subjects have at least 8 years of Chinese Arts training. Sixteen of the subjects focus on Chinese painting and calligraphy, and the remaining 4 have an experience of Chinese musical instrument, such as Zither and pipa. All subjects from America and Japan come to China with the hope of feeling Chinese culture. It is the first time for them to be in China, with an average length of stay of 18 days.

It was recognized that the validity of the present investigation was dependent upon each subject's ability to monitor his or her own emotions during performance. This was facilitated in part by the selection of a Piying show that was not a challenge of understanding for Chinese and westerner participants (see below). Subsequently, the procedure for subject selection was rigorous. A letter describing the general nature of our proposed culture psychological study was sent to four colleges of Zhejiang University, especially college of International Exchange Education and college of Art in order to solicit interest in participating in the investigation. The Ethics Committee of Zhejiang University Faculty of Information Technology approved all procedures.

3.2 Experimental Piying Show

The Piying show selected for the present study was the classic Piying show named 'Story of turtle and crane '. The story is a fable which is about turtle and crane that tried various ways of using their advantages to eat each other. This Piying show has won awards in international Puppet festival in 1965 and 2000.

This Piying show was chosen for several reasons. First, since a media content that is acceptable in one culture can be perceived inappropriate, rude or offensive in another, [29] so the content was chosen to be as culturally neutral as possible. Second, it needs representative story of traditional Piying. Third, reduce the understanding barrier in Piying and the emotional disturbance for figure. Fourth, only background music to break through language constraint. Fifth, suitable for eliciting stronger emotion-related cardiovascular response in the participants.

Five Piying shows eventually selected by us. In our earlier survey of understand resulting from watching these five videos of different shows, Chinese art-major students (N=25) as well Chinese non-art-major students (N=20), European students(N=49) nearly equally rated the video named the 'Story of turtle and crane ' as highest degree of understanding(average±SD=7.8±1.6 points) among these five videos on a 9-point scale.

3.3 Experimental Tasks

We have built simple Piying stage in experiment room, and invite two Piying artists to perform for the experiment. Three experimental tasks were examined in the present study. These were, (1) resting in a stationary sitting position (the baseline situation). (2) the perception task : the Piying show watching at the front stage. (3) the performance task : watching behind scene performance process of this Piying show. The order of task 2 and task 3 were randomly assigned for each of the subjects. It has been recommended that segments of HR should at least last for 5min for purpose of HRV analysis. [30] Therefore, the three tasks are lasted for 5min.

During the experiment, subjects were requested to avoid non-essential movement. Comparable restrictions on movement of the trunk, head, and upper limbs during the experiment were requested because movements of the large segments of the body (mainly the trunk and head) can significantly modulate the HR response. Thus, the subjects were requested to minimize motions of any portion of the body and make sure to keep their eyes.

3.4 Apparatus

The apparatus used for data collection was assembled in a laboratory that provided a comparable environment for an acoustically controlled and temperature-controlled recording studio. The experimental set-up consisted of a headset electrocardiogram (ECG, Nanjing heart - wing VISHEE, ltd, China) a Piying stage and Chinese folk musical instruments. There is light projected on the top of Piying stage. Instrument such as Erhu, bamboo flute and hand gong are carried by performers.

3.5 Procedures

All the subjects have not appreciated Piying on site or in movie before. Therefore, we interpret Piying culture for them before formal experiment. One day before scheduled data collection, each subject came to the experiment room in the laboratory. We interpret the history of Piying and its status in Chinese historical culture and they have 15 minutes to contact the Piying stage we built, Piying character and feel the operation method of Piying. They are required not to understand and appreciate Piying through other channels before formal experiment. The subjects were also given additional instructions, including avoidance of strenuous exercise and maintaining customary diet for 24h preceding the day of data collection, and to refrain from ingesting food, alcohol and caffeine for at least 2h preceding the data collection session.

After arriving at the experiment room on the day of data collection, the head ring testing HRV is on their heads, and electrode clip testing HR is carried on their eardrop. The subject was then requested to maintain a seated position in front of the screen for approximately 10 min to achieve a stable and calm cardiovascular status then a baseline reading was taken for 5 min. The subject then performed the rest main experiment tasks. Adaptation periods of approximately 5-10 min between the tasks allowed HR to return to the baseline.

3.6 Subjective Evaluation Measures

At the conclusion of data collection for each of the tasks the subject was asked to provide a subjective rating for his or her responses to the experimental Piying show using 10-point rating scales. The scales were used to designate the levels of valence (1=very unpleasant, and 10=very pleasant) and arousal (1=lowest arousal, and 10=highest arousal). They were also requested to identify the portion or measures where the highest pleasant emotions were perceived. In addition, subjects were asked to report if they were able to elicit emotions in response to Piying perception task and Piying performance tasks.

In addition, after all tasks subjects were asked to finish a semi-open structured questionnaire, in order to gain further insights into their experience. The questions were classified into five groups based on the features of traditional Piying art: (a) a general impression of Piying art; (c) image impression; (b) interactive form of Piying characters; (d) music impression; (e) the real control of backstage Piying art.

3.7 Heart Rate Data Treatment

From the onset of each experiment task, ECG were continuously monitored and recorded using a personal computer via an A/D converter at a sampling rate of 1000HZ. ECG data were translated into bpm data for each R-R interval. They were then re-sampled using cubic-spline interpolation to obtain an equally sampled time series.

3.8 HRV Analysis

HRV measures are derived by estimating the variation among a set of temporally ordered interbeat intervals. [25] From the onset of each experimental task, consecutive R–R intervals were extracted for 5 min HR data. The frequency-based technique of power spectral analysis is a sophisticated approach we choose to quantifying HRV. The frequency bands for LF, HF were 0.04–0.15 Hz, 0.15–0.40 Hz respectively. The LF component is affected by sympathetic influence, while HF component primarily reflects cardiac parasympathetic influence, many researchers have reported the ratio of LF to HF power as an index of "sympathovagal balance". [30],[31]

3.9 Statistical Analysis

Depending on the purpose of the comparison, we used a one-way ANOVA with repeated measures. Independent variables were nationality (Chinese versus American versus Japanese), and task (perception versus performance). If the sphericity assumption was violated, then Greenhouse–Geisser degrees of freedom corrections were applied. Post-hoc analyses were conducted using Least-Significant Difference procedure. $P<0.05$ level was taken as the evidence of a significant difference.

4 Result

4.1 Heart Rate

HR and HRV data were evaluated using 5-min data from entire experimental period. Figure.1 shows changes in the 10-s mean values of HR for three different nationality subjects during 5 min for three experimental tasks. According these three culture different subjects, the performance task produced constantly higher HR values than perception tasks during the entire testing period. The phase fluctuation was also greater during the performance task than that during perception tasks, and thus it had the largest range of HR. The HR of the three groups in ending performance is basically the same with that in setup of show. The findings suggested that perception is less attractive for long term. Americans' has slightly declined in Performance while the other two countries' retain stable. At the beginning, Americans have the maximum HR. The three countries all have highest emotional point e. The two points are the moment of Piying characters' close interaction in Piying show. (Figure. 1)

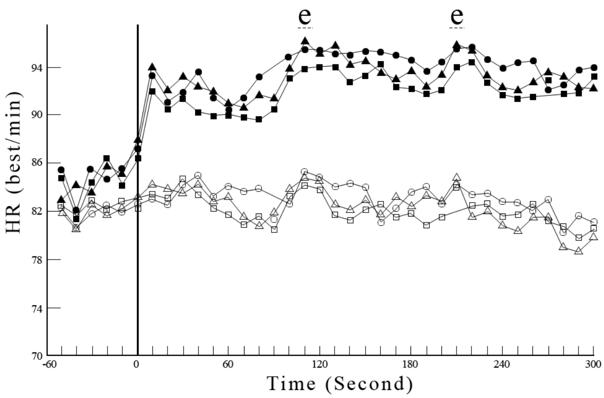


Fig. 1. Time course of the 10-s mean value of HR during the pre-and experimental period for all experimental conditions. ●, Performance task of Chinese; ▲, Performance task of American; ■, Performance task of Japanese; ○, Perception task of Chinese; △, Perception task of American; □, Perception task of Japanese.

4.2 Heart Rate Variability Difference Caused by Different Culture Background

In quiet state, the associated probability of LF, HF, LF/HF of Chinese, Americans and Japanese are all larger than significance level (p>0.05). That means there is no significant difference in average value of three indexes for these three groups. There was a significant effect for HF in perception task. The difference between Chinese and Americans (p<0.001) is more significant that of Americans and Japanese (p=0.04) while there is no significant difference between Chinese and Japanese. (Figure. 2) There is no significant difference in the three indexes during performance. The table below gives the results of the HRV analyses.

Table 2. Five-minute data for HRV among three different culture subjects

	Perception	Performance
LF	7.8	7.3
HF	14.9***	13.7
LF/HF	25.4	9.2

The values are the F values for all subjects

* p<0.05 ** p<0.01 *** p<0.001.

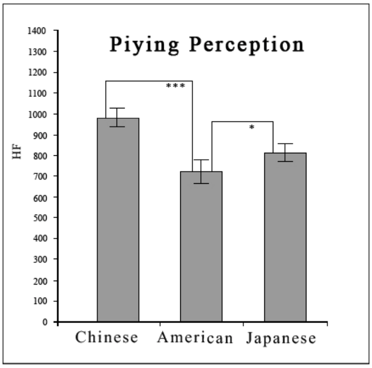


Fig. 2. Mean LF of Chinese, American and Japanese for Perception task

4.3 Heart Rate Variability Difference Caused by Perception and Performance

Make statistical analysis for LF, HF and LF/HF under baseline, perception and performance of three nationality subjects. Except for HF of Chinese, LF of Americans and HF of Japanese, the associated probability of other indexes of the subjects in the three countries are all smaller than significant level(p<0.01). Therefore, the three states have significant difference in at least one group or the other two, or in all the 3 states. (Table 3) The result of LSD multiple comparisons: in group with significant difference, the significant effect of baseline and performance are higher than the indexes of baseline and perception. This indicates a higher emotion-induced arousal level during watching performance than that during watching Piying perception. (Figure. 3)

Table 3. Five-minute data for HRV between perception and performance

	Chinese	American	Japanese
LF	16.6***	5.8	11.2***
HF	3.3	7.9***	3.1
LF/HF	13.8***	4.9*	4.9***

The values are the F values for all subjects

* p<0.05 ** p<0.01 *** p<0.001.

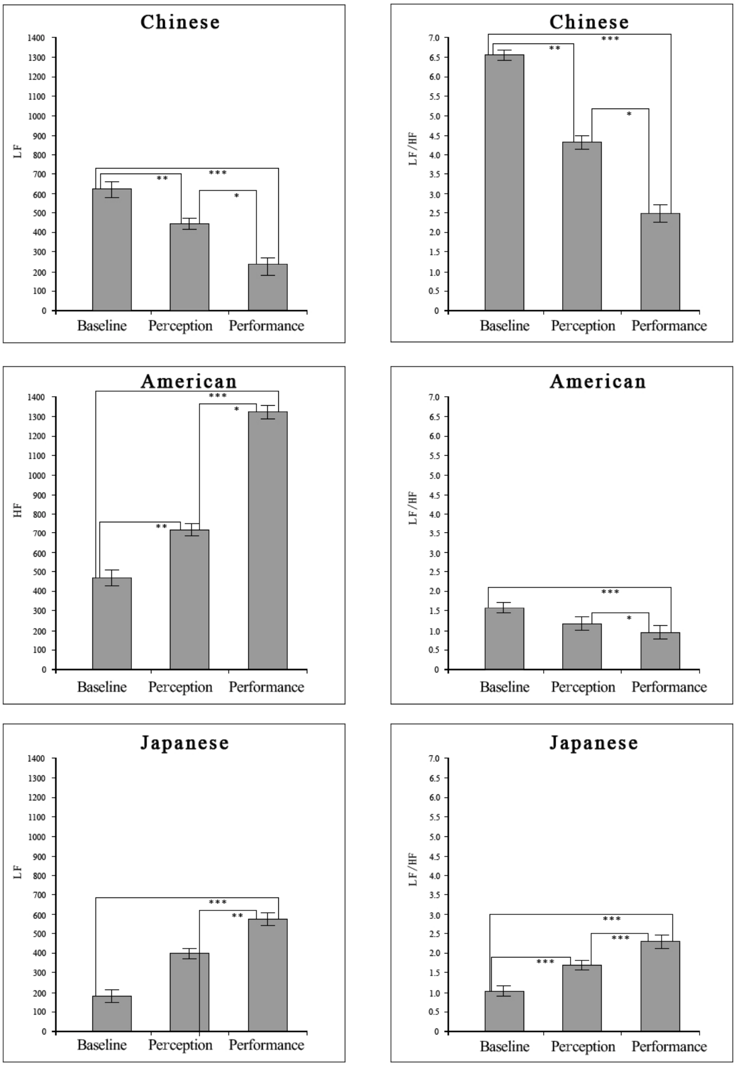


Fig. 3. Mean LF, HF or LF/HF of baseline, perception task and performance task for Chinese, American or Japanese

4.4 Subjective Measures

The mean values of the valence and arousal levels of all conditions for all subjects are reported in Table 4. ANOVA revealed a significant task effect for the mean value of valence and arousal level, indicating a higher emotion-induced arousal level during performance than that during perception. Significant main effects of culture difference were also found for arousal levels. The highest level of pleasant feeling of Chinese, Japanese and Americans was commonly reported to have been experienced during the

120-121th and 194-195th measures around the close interaction between Piying characters for twice. (Figure.1). Some reported the highest level of pleasant feeling at the 116th and 122nd bars, and some at the 183rd bar.

Table 4. Subjective evaluation measures of emotion-induced by Piying art

Culture	Chinese		American		Japanese		ANOVA F-values	
Task	Perception	Performance	Perception	Performance	Perception	Performance	Culture	Task
Valence	5.9±0.8	7.7±1.2	5.2±0.8	8.0±1.1	4.5±1.3	7.2±1.4	6.5	12.9*
Arousal	5.1±0.9	7.3±1.0	4.4±1.3	8.9±0.8	4.9±1.2	6.7±1.4	9.9*	7.8***

The values are the F values for all subjects
＊ p<0.05 ＊＊ p<0.01 ＊＊＊ p<0.001.

In order to develop more robust indicators from the experiment, we asked subjects to give a rate to each feature of Piying art (1=do not like, 10=like very much). To better understand the influence, subjects were asked to give a ranking to these eight elements that clustered to five main features.(Table 5) The mean values of impress of interaction between artiest and Piying characters are very high among three different culture subjects. It was believed that the performance is popular among people and can well arouse their emotion. It has been proved in ranking given by subjects. There is great difference in 'shadow'. The ranking of selection by the east is among the top few while Americans ranks the end. Image impress of 2D Piying characters is on the top while is ranked end in terms of image Piying color of three cultural backgrounds.

Table 5. Subjective evaluation measures of features of traditional Piying art

		Chinese		American		Japanese	
		Mean	rank	mean	rank	mean	rank
a	General impress of "shadow"	8.3	1	2.3	8	6.8	2
	General impress of Piying created atmosphere	6.8	3	4.5	6	4.2	5
b	Image impress of 2D Piying characters	7.2	4	6.7	2	7.3	1
	Image impress of color of Piying characters	2.1	8	3.2	7	2.4	8
c	Interactive form of Piying characters	4.6	6	5.2	4	5.9	4
d	Music impression	6.8	5	5.1	5	3.7	7
	Music impress of instruments	6.4	7	6.8	3	5.8	6
e	Impress of the real control of backstage Piying art	7.9	2	8.8	1	7.1	3

5 Discussion

The results of the present investigation yielded two novel findings: (1) the result confirmed that Piying performance was far more effective in modulating emotion-related activity than Piying perception. (2) Chinese are more fond of traditional Piying elements and American prefer experiencing interaction of Piying while the result of Japanese is between those of Chinese and American, they hope that there is a rule to follow in Piying show.

5.1 The Effects of Emotion Induction by Different Culture Background Subjects

In perception task there is significant difference between Chinese and Americans, American and Japanese in HF of HRV and the HF average value of Chinese is higher than that of Americans and Japanese. (Figure.2) The difference in HF shows the activation degree of vague. Based on HR figure, it can be seen that the initial position of HR values for Americans and Japanese are all high. Americans is higher than Chinese but it decreases in later period, which indicates that Piying is fresh to foreign subjects but lacks persistent appeal.

The questionnaire after HRV experiment also supports HRV data analysis. All American subjects reported that at the beginning Piying perception gave them a higher level of arousal as well as surprising feeling, indicating that the subjects must have received a higher level of positive emotions.

American and Japanese give the rank of “Image impress of 2D Piying characters” second grade and first grade, respectively. The image of Piying character is fresh to Westerner. The Piying character make Chinese subjects associate with the expression method of traditional Chinese painting.

As for interactive form of Piying character, subjects all feed back that the interaction form is monotonous. Foreign subjects say that since Piying is a performance without expression and they cannot well read the story just by means of interaction among Piying characters and the performance is monotonous sometimes. Combining with the value figure of HR, it can be inferred that the highest emotional point e appears in interaction among characters.

As for background music, most subjects know it is folk music of China with features. But 8 Americans subjects and 12 Japanese subjects mention that the background music is too noisy and wish it be powerful but not noisy. Hence, the ranking of music impression is at the end few. In music instruction, domestic and foreign subjects are all astonished with traditional Chinese instrument and the ranks the top compared with the former. 6 Americans subjects mention that instrument familiar by the world may be added such as drum. Also they wish Piying show mainly adopts traditional Chinese instrument.

In performance task, the significance of HF, LF and LF/HF is not obvious. Based on HR value, the HR of subjects of the three countries are all high during performance. It may attribute to two reasons: (1) One explanation for this may be the presentation of performance behind the curtain is fresh to both Chinese and western people, with fluctuated mood. Therefore, the difference is insignificant. (2) The presentation of performance behind the curtain reflects the interaction process between Piying artists and characters and among artists so that subjects are not strange. Thus, the difference is insignificant.

The analysis above complies with the result of questionnaire after experiment. All subjects reported that performance show gave them a higher level of arousal as well as valence (pleasant feeling), indicating that the subjects must have received a higher level of rewarding and positive emotions under this condition. 14 subjects wish to try Piying performance. Especially that 6 people have mentioned the word of “game”. In their mind, it is a game and can interact with Piying characters or their partner to immerse them into the interaction. Though Chinese subjects are not as excited as western ones when describing the mood when they watch the performance, they express that it is the first time to know the actual performance process behind the curtain, which is fresh and interesting.

Piying character is presented as shadow in front of the curtain and entity behind the curtain. "Shadow" is the showing way of Piying character with deficiency and excess state. Factors frequently exist in Chinese art, e.g. uninterrupted mountains in Chinese painting represented by pieces of ink marks, the artist conception with interrupted writing but uninterrupted conception in Chinese calligraphy. There is great difference in evaluation for the typical Piying element among the three countries. China and Japan rank it as the first and second while the last in American ranking. The national culture difference can thus been seen. More than half of American subjects become aware of the concept of "shadow" only in questionnaire. Moreover, according to Hofstede's culture dimensions [28], in our case, we might speculate that the long-term orientation in Chinese culture would result in persist core spirit of traditional culture, while Americans find values rather than fulfilled enjoy of the moment. (Table 1)

As for Piying color, the evaluations from the three countries are all at last. So is the feedback: In perception task, it is hard to distinguish and note the color of Piying character which only presents dark hue. In performance task, it can be seen that Piying character is pained by diverse colors but is not impressive and without rich colors. It is hoped to employ modern and bright multi-colors to vitalize the color image of Piying character.

5.2 The Effects of Piying Perception and Piying Performance in Inducing Autonomic Nerve Response to Emotion

There is significance in LF, HF and LF/HF values under baseline, perception task and performance task for people from the three cultural backgrounds, indicating that Piying performance can attract and arouse emotion, which has supported the foregoing deduction: in performance task, such form is fresh and attractive to the three types of subjects, with distinct mood fluctuation. Therefore, there is no significance among races.

The reason may be that the performance behind the curtain of Piying can arouse the resonance of subjects. There is significance between baseline and perception task. The subjects are significant in arousing and interest in Piying but the interaction performance presented in performance task give stronger sense of participation to subjects and arouse their interest. We may reasonably assume that Piying perception present fresh and curious feelings to subjects while interaction of Piying performance propel them to have a try. In this way, the subjects and Piying culture positively interact with each other not just one-way culture display and influence of publicity. Instead, it is two-way communication and interaction with Piying culture from emotion. Thus, Piying performance facilitates emotional enthusiasm of the subjects. It is like taking intangible culture heritage into tangible interaction which means taking advantage of natural physical affordance to achieve a heightened legibility and seamlessness of interaction between people and information [32] in order to build an immersive impression.

The result concluded from questionnaire after experiment supports the analysis. Four American subjects say they admire that ancient Chinese can design such nice hand-made characters. Five American subjects mention that Piying makes them associate with puppet show which is also to display story by hand-made control. Four foreigner subjects said 'the overall feeling of Piying show (perception) is not as fun as they thought it was'. Thus, the 2D performance mode of Piying is the presentation way of traditional Piying but is limited in the effect of arousing audience's emotion. In performance task, the performance way of Piying artist makes western subjects have the idea to try.

Chinese subjects also show interest in Piying. Some said that "they feel it vivid, interesting and of rhythm", "a marvelous visual art", "the character features are unique and interesting", "the plot is traditional", "the joints can move, which is quite interesting". In terms of Piying performance, some subjects say that: such performance is of strong rhythm just like kung fu to attract audience. In effect, nearly 90% domestic and foreign subjects mention that it is the most exciting moment when they see Piying artist talks and sings Piying character singly. It indicates that the interactive performance process is approved by subjects and can highly arouse their emotion.

The observation that Piying performance was far more effective in modulating emotion-related response than Piying perception should have implications for education and entertainment. From time to time, museums will perform and promote traditional Piying show for people in order to deepen people's impression on traditional culture. Our data indicated that greater effects of emotion modulation would be expected as a result of interact with Piying character or playing an interactive Piying show, rather than watching Piying art. Immersion experience of Piying art can also practice self-expression ability and cooperative ability of children.

In the ranking of 'Impress of real control of Piying art in background', American subjects rank the first, followed by Chinese and Japanese. On the one side, it indicates that Piying performance has high arousing degree and subjects wish to immerse and participate in Piying performance. Moreover, the fourth dimension of Hofstede's culture dimensions appears relevant to this element's rate. [28] The fourth dimension is labeled as Uncertainty Avoidance, and it can be defined as the degree to which people in a country prefer structured over unstructured situations. Japanese show high score on uncertainty avoidance, people show more nervous energy, while Americans with low score mean easy going. Seven Japanese subject mentioned they hope that there is a rule to follow in Piying show. We suggest interaction of Piying may mean curious to Americans but mean a certain degree of out of rules to Japanese.

According to the atmosphere built by Piying show, Chinese gave a top rank, but the other two countries did not. Chinese people prefer to act as members of groups rather than individuals, which means collectivism. While American learns very early to think of itself as "I" instead of as part of "we". [28] Therefore, Chinese like the get-together atmosphere created by Piying show.

References

1. Wu, Q., Lin, X.: The development and heritage of Chinese Piying. Heilongjiang Science and Technology Information, Heilongjiang, pp. 229–229 (2010)
2. Georges Sadoul, M.: Movie History. Basic Books, France (1946)
3. Hu Huang, J.: The gospel of Piying— Road of protecting Daoqing Piying show, vol. 4, pp. 70–71. People's Music, Beijing (2009)
4. Zhengchang Xue, J.: Inheritor of Piying "intangible cultural heritage": Zhang Jinxu and his Piying family, vol. 6, pp. 127–131. Social Science in Ningxia, Ningxia (2010)
5. Yan, S., Yao, L., Ji, X., Ying, F.: Integrating old Chinese shadow play-Piying into tangible interaction. In: Proc. TEI 2004, pp. 375–375. ACM Press, New York (2010)
6. Gao, Y.: Master thesis, Study on modeling, color and cultural connotation of Piying in Luoshan, Hunan University (2009)
7. Wang, X.: Piying: pioneer of modern movie, pp. 28–28. Invention and Innovation, Hunan (2001)
8. Ozturk, S., Karagoz, M.: Co-Opted: Turkish Shadow Theatre of the Early Republic. Asian Theatre Journal 23, 292–313 (2006)
9. Gao., L., Cai, J.: Analysis and realization of figure motion in Piying animation. Computer Engineering & Design 31(10), 2335–2338 (2010)
10. Zhu, Y.B., Li, C.J., Shen, I.F., Ma, K.L., Stompel, A.: A new form of traditional art: visual simulation of Chinese shadow play. In: Proceedings of SIGGRAPH 2003 Conference on Sketches & Applications, pp. 10. ACM Press (2003)
11. Paul, B., Forlines, C., Ramesh, R.: Handheld Projectors for Mixing Physical and Digital Textures. In: Proceedings of Computer Vision and Pattern Recognition (CVPR), pp. 110–112. IEEE Press, San Diego (2005)
12. Gudukbay, U., Erol, F., Erdogan, N.: Traditon Offers Artistic Possibilities for New Media Technologies: An Animation System for Shadow Theatre. In: International Symposium on Electronic Art, pp. 86–96. ACM Press (2000)
13. Hsu, S.W., Li, T.Y.: Planning Character Motionsfor Shadow Play Animations. In: Proceedings of Computer Animation and Social Agents, pp. 39–45. IEEE Press (2005)
14. Hsu, S.W., Li, T.Y.: Generating Secondary Motions in Shadow Play Animations with Motion Planning Techniques. In: Proceedings of SIGGRAPH 2005 Conference on Sketches & Applications, pp. 35–35. ACM Press (2005)
15. Yali Liu, J.: Charm of traditional Piying and innovation of animation, vol. 5, pp. 67–68. Art Criticism, Beijing (2008)
16. Güdükbay, U., Erol, F., Erdogan, N.: Beyond Tradition and Modernity: Digital Shadow Theater, Leonardo, pp. 264–265 (2000)
17. Wei, J.: Master thesis, Shadow Animation System Based on Wii Remote Controller, Xi'an Electronic and Science University (2012)
18. Sun, Y.: The Traditional shadow play, and interactive media design. Journal of Nanping of Teacher College, 132–134 (2006)
19. Intangible Heritage-Piying, N.: `http://baby.sina.com.cn/edu/08/3101/1840104616.shtml`
20. Triandis, H.C.: Culture and Social Behavior. McGraw-Hill, New York (1994)
21. Keltner, D.: Expression and the course of life: studies of emotion, personality, and psychopathology from a social-functional perspective. In: Annals of the New York Academy of Sciences, pp. 222–243 (2006)

22. Batja, M.: Culture and emotion: different approaches to the question. In: Mayne, T.J., Bonanno, G.A. (eds.) Emotions: Currrent Issues and Future Directions, New York, pp. 214–250 (2001)
23. Mayne, T.J., Bonanno, G.A.: Emotions: Current Issues and Future Directions. New York, Guilford Press (2001)
24. Tseng, W.-S., Hsu, J.: Chinese Culture, Personality Formation and Mental Illness. International Journal of Social Psychiatry 16, 5–14 (1970)
25. Appelhans, B.M., Luecken, L.J.: Heart Rate Variability as an Index of Regulated Emotional Responding. General Psychology 10(3), 229–240 (2006)
26. Thayer, J.F., Lane, R.D.: A model of neurovisceral integration in emotion regulation and dysregulation. Journal of Affective Disorders 61, 201–216 (2000)
27. Thayer, J.F., Siegle, G.J.: Neurovisceral integration in cardiac and emotional regulation. IEEE Engineering in Medicine and Biology 21, 24–29 (2002)
28. Hofstede, G.: Cultural constraints in management theories. Academy of Management Executive 7, 81–94 (1993)
29. Lu, M., Walker, D.F., Huang, J.: Do they look at educational multimedia differently than we do? A study of software evaluation in Taiwan and the United States. International Journal of Instructional Media 26, 31–42 (1999)
30. Task Force of the European Society of Cardiology and the North American Society of Pacing and Electrophysiology, Heart rate variability: standards of measurement, physiological interpretation and clinical use. Circulation 93, 1043–1065 (1996)
31. Eckberg, D.: Sympathovagal balance: A critical appraisal. Circulation 96, 3224–3232 (1997)
32. Noman, D.A.: The Design of Everyday Things. Basic Books, USA (1988)

A Patch-Based Data Reorganization Method for Coupling Large-Scale Simulations and Parallel Visualization

Li Xiao, Zhiwei Ai, and Xiaolin Cao

Institute of Applied Physics and Computational Mathematics Beijing, China
{xiaoli,azw,xiaolincao}@iapcm.ac.cn

Abstract. The scale of some datasets generated by simulations on tens of thousands of cores are gigabyte or larger per output step. It is imperative that efficient coupling of these simulations and parallel visualization. A patch-based data reorganization method was presented for this coupling through a parallel file system. Based on the method, simulation data sets in application codes are reorganized by patch and written into many files in parallel. These datasets in these files can be read directly by visualization software with low I/O overheads. For two real simulations on above 30000 cores, large-scale datasets have been generated and visualized efficiently.

Keywords: parallel visualization, large-scale datasets, patch-based data reorganization method.

1 Introduction

Supercomputers have been indispensable for scientific research. In the past ten years, the growth of supercomputer peak performance has increased by one thousand times [1]. These supercomputers consist of ten of thousands of processor cores. For example, TianHe-1A [2] has 7168 compute nodes and 86016 processor cores.

With the development of supercomputers, the scalability of application codes has also increased quickly. Now, some application codes have been run efficiently on thousands of processors cores, even more [3,4]. These datasets generated by these simulations are measured in gigabytes or even terabytes. For example, a typical simulation of laser plasma interactions on 36000 cores consists of 0.768 billions of cells and 20 billion particles. The numerical simulations output 32GB datasets per output step and 1.6 TB time-varying datasets of 50 steps in total.

For these large scale datasets, parallel visualization must be used so that we can look at the full extend of the data as high as possible resolution [5]. Visit [6] is used to visualize and analyze simulation results on many of the world's top supercomputers. It is a free interactive parallel visualization and graphical analysis tool for viewing scientific data. It can handle very large dataset whose size is the tera-scale range.

Coupling simulations and visualization through a parallel file system is the most common situation [7, 8]. Specifically, the simulation write data to multiple files in

Z. Pan et al. (Eds.): Transactions on Edutainment IX, LNCS 7544, pp. 278–289, 2013.
© Springer-Verlag Berlin Heidelberg 2013

parallel, then the files generated by these simulations are read directly by visualization tools. For this coupling way, the main performance bottleneck is reading data from many files to the memory of rendering processors [9, 10]. With the growth of data size and the number of variables, the time for data transmission from storage to memory becomes significant [11]. Peterka et al. in paper [12, 13] shows that I/O can consume up to 90% of the total visualization time.

Yu and Ma present two parallel I/O methods for parallel visualization of large-scale data in a high-performance computing environment [14]. Their test results demonstrate that their I/O methods effectively remove the I/O bottlenecks commonly present in time-varying data visualization. Storage researchers attempt to improve the I/O bandwidth and latency of the parallel file systems through techniques such as advanced I/O forwarding [15]. We will discuss another effort about parallel I/O.

Because I/O is such a critical part of the large-scale parallel visualization, our team devotes considerable effort to its study. With knowledge of the underlying I/O system, access patterns, and data structure, simulation data sets in application codes can be stored to make reading "easier" for parallel visualization software. Therefore a patch-based data reorganization method has been designed to write data in the way that parallel visualization software is favorite of. It can effectively improve the I/O performance of VisIt software.

In the following three sections we discuss our efforts in visualization for extreme scale data. In section 2 we will depict coupling method of large-scale simulations and parallel visualization. Section 3, we describe our work in data reorganization for obtaining efficient I/O during visualization. In section 4 we show two images of parallel visualization about large-scale simulation results running on ten of thousands of cores. The data set generated by simulations is stored in an effective way designed by us.

2 Coupling Method

2.1 Patch-Based Data Structure

For most parallel program based on domain decompose method, the whole computational domain is divided some sub-domains. This sub-domain is called patch in this paper. For structure mesh application, the patch is a logically rectangular region on which the simulated system and the related physical variables can be well defined [16]. The patch is the basic unit of data storage, load balancing and numerical computation. This paper only discusses structure mesh application. It can be generalized to the others applications.

Fig. 1 depicts patch-based data structure. It shows a two-dimensional structured mesh consisting of 20×20 cells. It is decomposed into seven patches and each patch is defined by a logical index box. These patches are ordered and are distributed between two processors. The left four red patches belong to the first processor and the right three green patches belong to the second processor. In each patch, physical variables exist in the form of an array of patch data. Patch data is defined on the region covering the ghost box. Patches and its patch data are distributed among processors. The number of patches usually is larger than the number of cores. Each core has usually several or tens of patches.

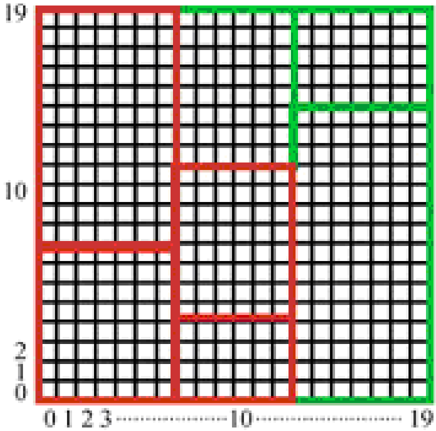


Fig. 1. A two-dimensional structured mesh consisting 20×20 cells. The mesh is portioned to 7 patches.

The data field composed of a lot of sub-domains can be visualized by VisIt software. The data in one sub-domain is called data-block.When dealing with large-scale data fields, VisIt's engine can be run in data parallel. Each engine process reads, deals with and renders one data-block by one data-block. Finally, a complete synthesis of the piece results created by each data-block is presented to the research analysts. When reading each data-block, VisIt reads the variables needed by the visualization one by one.

In order to visualize the results of large-scale simulations, the patch in parallel program need be mapped to the sub-domain in VisIt. Patch can be defined the basic unit of data output in parallel program and data input in VisIt.

2.2 Coupling by Files

Note that the ideal data organization and structures used by the application code might not be ideal for parallel visualization. Generally, the number of cores used to run the simulation is usually much larger than the number of cores used to parallel visualization. Now, some parallel program can be run on thousands of cores, even more. But visualization software can run on tens or hundreds of cores. Therefore It is imperative that efficient coupling of these simulations and parallel visualization.

Figure 2 depicts typical coupling simulations and visualization through a parallel file system. It includes two sections. First, simulation data sets in application codes are reorganized by patch and written to many files in parallel. These data in these files can be read directly in an efficient manner by VisIt. Second, in VisIt software, each process read evenly some data-blocks in a few files in order to improve I/O performance.

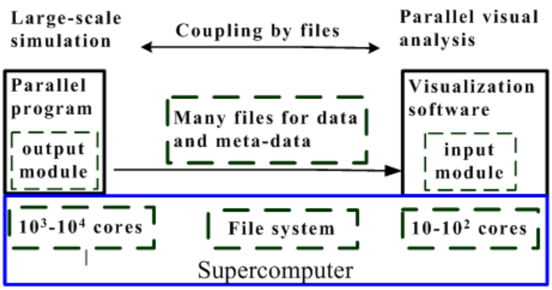


Fig. 2. Coupling of simulations and visualization

For efficient coupling of simulations and visualization, a output module in parallel program and a input module must be designed elaborately. For this coupling way, the main performance bottleneck is reading data from many files to the memory of rendering processors. So the key is how to organize and write data into files that visualization software can read with low I/O overhead.

2.3 Three Kinds of Files

When it comes to visualize simulation results, application codes usually write data of single time step in many files at regular intervals of time. In order to make the files understandable for Visit, these simulation results have to be organized and written into three kinds of files.

Figure 3 presents these files layout and their organization in the file directory structure. There is a file named "time_varying.record" in top directory. This file includes the name and storage location of each metadata file for each output step. When Visit software opens the file, it can get metadata files's location in the file system of all outputs step and reconstructed the time varying datasets.

In directory of each output step, there is one metadata file named "metadata". Metadata file stores summary information of all patches. This file contains information related to variable, such as plot variables name, the type of the variable (scalar, vector, tensor…), the position of variables defined, etc. This file also includes the information related to patch, such as the range of logical index, the range of spatial position for each patch and the range of value of all plot variables defined on the patch. The data size of metadata file is directly proportional to the number of patches.

In directory of each output step, there are many raw data files named "raw_data_xxx.dat", where xxx is serial number of process. Traditionally, application programs often write out one file per process. Each process writes all values of variables of all local patches into corresponding raw data file. Where, patch is a basic unit of dataset output. For each patch, values of each plot variable are written in turn. When node has more and more cores, several processes share a raw data file in order to reduce parallel I/O conflict and improve I/O efficiency.

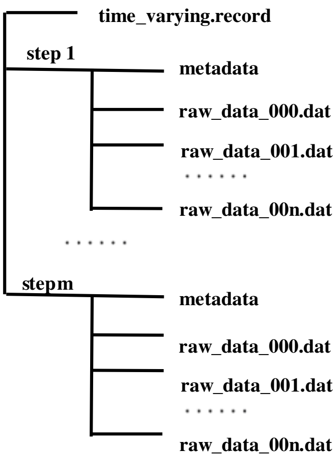


Fig. 3. Three kinds of files and their organization structures in file system

3 Patch-Based Data Reorganization Method

3.1 Data Reorganization

The domain decomposition used by the simulation code does not always match what the parallel visualization code designed to use. As a result, simulation results named plot data,in application codes must be reorganized in memory before writing to files. It consists of two objectives. One is increasing the size of patch and thus reducing the number of patches. The other is reducing the number of raw data files. Algorithm 1 has been designed for reorganizing plot data in order to adapt data layout used by visit software.

In step 1), an independent level named plot level is created. The plot level has the same resolution with the computing level for numerical calculations. But the size and process mapping of patches are different between two levels. It creates plot level according to some parameters given by user. Table 1 lists these control parameters.

In step 2), the communication schedule from computing level to plot level is created. The logical parameter need_schedule is true only when the output is initial step or computing level has changed for load balancing.

In step 3), memory space storing plot data on patches on plot level is allocated. In step 6), the memory space is released after writing data sets. It is very important for reducing memory overhead in numerical calculations phase.

In step 4), it transfers plot data from computing level to plot level based on communication schedule created above. The method is very similar with post process method for organizing plot data from files. But it is more efficient because these plot data is still in memory.

In step 5), it writes plot data and its auxiliary information to raw data files and meta-data files.

Algorithm 1: reorganize_plot_data

```
      // create plot level based on parameters
  1)  if(initial_step)create_plot_level();
      //   schedule from computing level to plot level.
  2)  if(need_schdule)create_schedule().
      // memory for storing plot data   on plot level
  3)  allocate_memory_on_plot_level();
          // transfer data from computing level to plot
level
  4)  transfer_data();
      // write plot data on plot level
  5)  write_plot_data();
  6)  deallocate_memory_on_plot_level();
```

These datasets in raw data files and meta-data files can be read directly in parallel way by VisIt software. Each process reads, deals with and renders data one patch by one patch. Each process may read tens of patches, which may lie irregularly in some files. The more the number of opening file is, the more expensive the I/O overhead is. Therefore, a load balancing algorithm in input module was designed so that each process only need open several files for reading patches. First, this algorithm sorts patch number by its global number and file number. Then, it computes data size of each patch. Finally, it assigns evenly patch among visualization processes. This algorithm can reduce efficiently the number of opening files.

3.2 Control Parameter for Data Reorganization

Algorithm 1 has been implemented in output module. The module supplies some control parameters for users to optimize the layout of data. These parameters are designed to control the layout and amount of the output data. Table 1 lists these parameters. The size of patch is controlled by two parameters, largest_patch_size and smallest_patch_size. The parallel output module will try to ensure that every patch size is the largest_patch_size. At same time, it will be as far as possible to ensure that the data size is not less than the smallest_patch_size. When the number of cells in each dimension can be divided exactly by the largest patch width in the corresponding dimension defined by largest_patch_size, the size of patch on plot level is equal to the largest_patch_size. The number of processes shared the same raw data file is controlled by the parameter num_procs_per_file.

Table 1. Control Parameters

Parameters	Default	Description
largest_patch_ size	40×40×40	the size of largest patch
smallest_patch_size	20×20×20	the size of smallest patch
num_procs_ per_file	1	the number of processes shared a common raw data file
sample_ratio	(1,1,1)	sample ratio of grid point

The latter two parameters are used to achieve other special purpose. According to the parameter sample_ratio, the parallel I/O module will create a low-resolution representation of the dataset to improve performance of visualization. Even further, according to this parameter, the parallel output module can build a multi-resolution representation of the dataset in a group of output files. So the zoom in operation can be achieved at the level of file system.

Actually, the combinations of these parameters can be used. For example, the output module will simultaneously reduce the amount of data and optimize the patch size. Tuning these parameters can ensure that the output model output the data in the layer that is suitable to the file systems, the storage systems, the mode of input data of visualization software and many other features. Good data reorganization will be helpful to enhance the data read and write performance. It is very important because visualization software often spends a lot of CPU time waiting for data read completely.

3.3 Performance Results

I/O is the main performance bottleneck of visualization tool. The reading performance of data file using our layout was tested in VisIt. The test data is from the LPI (laser plasma intersection) simulations [17]. The LPI code runs on 4000 cores of 400 nodes and output data in different patch size or the number of the raw data files for same simulation resolution.. The simulation consists of 800×1600×800 cells and 5.36×109 particles. Five scalar variables and two vector variables defined on cell center are output for plotting. Visit is running on 50 cores of 10 nodes for parallel visualization by volume rendering. The I/O costs were tested for two cases at a single output time step.

Table 2. Variety of file size as the Patch size changes

Files	Patch size	Metadata size (M)
1000	10×10×10	442.05
1000	20×20×20	56.11
1000	40×40×40	7.09
1000	80×80×80	0.91

Table 2 shows the variety of file size as the patch size changes. We vary the patch size from 10×10×10 to 80×80×80 by doubling the size along each dimension with 4 processes sharing a raw data file. As the patch size increases, the size of metadata files decreases proportionally. The metadata file stores information of each patch, thus the amount of data stored in metadata file is basically proportional to the number of patch. As the patch size increases, the number of patches decreases quickly with 8 times. It proves that the size of meta-data file is directly proportional to the number of patches. In raw data files, some description information about data array and hierarchical structure is provided, which is also proportional to the number of patches. Total size of output data is the sum of the size of meta-data file and the size of all raw data files. So the total size of the output date decreases as the patch size increases.

Table 3. Variety of I/O costs as the patch size changes

Patch size	Mean time(s)	Min time(s)	Max time(s)
10×10×10	216.18	60.18	309.28
20×20×20	56.23	12.69	82.55
40×40×40	25.47	10.73	33.62
80×80×80	13.68	8.91	16.09

As shown in Table 3, I/O costs decrease quickly as the patch size increase. The patch whose size is 80×80×80 has the best I/O performance with mean times 13.58. There are three reasons for this phenomenon.

Firstly, as noted earlier, patch is the basic unit that VisIt read data. The smaller the patch size will lead to the smaller the I/O granularity. When I/O granularity is small, IO bandwidth will not be fully utilized.

Secondly, the smaller the patch size will lead to the more the number of the patches. When the number of the patches is very large, the number of accessing to the files will be increased significantly, so the time used for opening and closing the files is significantly increased too.

Thirdly, the smaller the patch size will lead to the more the number of the patches in one file. In the case of parallel I/O, when the number of patches in one file is large, the file access conflict will be increased.

Table 4. Variety of I/O costs as the number of files changes

Files	Mean_time(s)	Min_time(s)	Max_time(s)
4000	59.46	52.05	67.40
2000	35.79	29.50	41.16
1000	25.47	10.73	33.62
500	19.03	12.66	26.42

Table 4 shows the variety of I/O costs with the variety of the number of files. When one process wrote a raw data file, 4000 files were generated. When 4 processes share a raw data file, 1000 files were generated. In this case, the 4 processes wrote data into the same file in turn. The patch size is 40×40×40. As shown in Table 4, I/O cost is decreased as the number of files is decreased. When the file number is 500, the best I/O performance can be got. At the same time, the I/O time is imbalance. We will develop a finer load redistribution method and study how to reduce data read overhead as much as possible.

4 Parallel Visualization for Real Applications

Two large-scale data sets of real application have been generated by two application codes. These programs have been run on tens of thousands of cores with several or tens of hours. In order to improve I/O performance, four or six cores share a raw data file. The patch size is mostly 40×40×40. These results have been visualized by using VisIt on 72 cores of 12 nodes for helping physicists get new insight.

1) LARED-P is a three-dimensional code modeling long pulse laser prostration in large scale plasmas. It can be used for modeling LPI (laser plasma intersections) process in a cone-guided geometry of fast ignition. It uses the method of Particle-In-Cell (PIC). Electrons and ions are distributed in the cell of a uniform rectangular mesh. The Maxwell electromagnetic equations coupled with particle movement equations are solved. Particles interact with the electromagnetic fields.

A typical simulation of laser plasma interactions on 36000 cores consists of 0.768 billions of cells and 20 billion particles. The numerical simulations output 32GB datasets per output step and 1.6 TB time-varying datasets of 50 steps in total. When 6 processes share a raw data file, 6000 files were generated per output step.

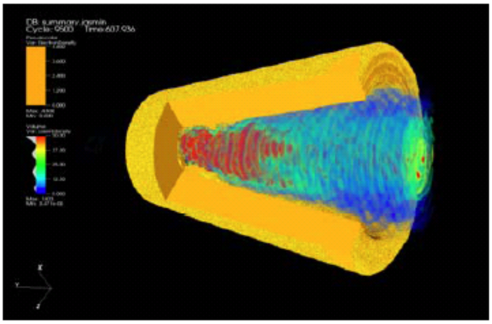


Fig. 4. Parallel visualization result for the laser plasma interaction simulation using LARED-P program

Figure 4 shows the distribution of particles in a snapshot and the related volume rendering of laser intensity. It took almost 27 second to reading data from disk to memory. The render time is about 6 second. The distribution of the particles is showed with yellow colour by using pseudo-colour method. The date field outside the particles and cone target is removed by using the iso-volumn operator. Laser energy is mainly distributed in the internal cavity of cone target. In order to make it visible, part of the cavity has been cut by using space cutting method. Meantime, the laser intensity has been drawn by using direct volume rendering method with the transparent effect. Two variables have been integrated in the same map, which demonstrated clearly the existence of plasma significantly affects the light propagation and the generation of relativistic electrons. Multi-graph coupling can effectively demonstrate the relationship between multiple variables.

2) JEMS-FDTD is a three-dimensional code named with J Electromagnetic Solver-Finite Difference Time Domain method. This program can simulate propagation, radiation, and couple of electromagnetic field by solving Maxwell equations on structured mesh explicitly with FDTD method. The time-dependent Maxwell's equations are discredited using central-difference approximations to the space and time partial derivatives.

A typical simulation on 30000 cores consists of 0.61 billions of grids. The numerical simulations output 0.468 TB time-varying datasets of 170 steps in total. When 6 processes share a raw data file, 5000 files were generated per output step.

Fig.5 shows a simulation results for high power electromagnetic pulse couples into computer box through small apertures and slots. The simulation of electromagnetic pulse coupling with computer box is important for the research of electromagnetic compatibility and defense of computer systems. Fig.6 depicts the distribution of cumulative electromagnetic field during a certain period of time inside and outside of computer box. We rendered computer box by using value extraction method along with the classification of the different parts colored. At the same time, one side of computer box has been cut by using spatial-cutting method so that scientists can effectively observe the detail distribution within computer box. The strength of cumulative electromagnetic field is rendered by using direct volume rendering method with transparent effect. Where, the red color represents high energy and the blue color represents low energy.

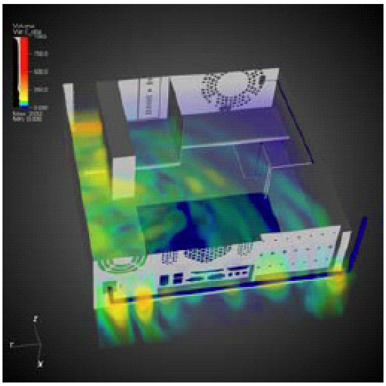


Fig. 5. Parallel visualization result for electromagnetic pulse coupling with computer box using JEMS-FDTD program

5 Conclusion

The scale of the datasets generated by large-scale simulations running on tens of thousands of cores are gigabyte or larger per output step. The terabytes time-varying datasets are providing a great challenge for efficient coupling large-scale simulations and parallel visualization in the field of scientific computing. During parallel visualization, the main bottleneck is not data rendering, but data movement between storage and rendering processors. So the datasets used for visual analysis were reorganized in the way that Visit software will efficiently read it. Numerical experiments have proved that the patch-based data reorganization method can significantly reduce I/O costs. These make it possible to analyze these datasets and understand the physical phenomena and characteristics contained in the datasets in a timely fashion.

Acknowledgments. This work was under the auspices of the National High Technology Research and Development Program of China (863 Program) (2010AA012402), the National Natural Science Foundation of China (61033009), the National Basic Key Research Special Fund (2011CB309702) .

References

1. Yang, X.J., Liao, X.K., Lu, K., et al.: The TianHe-1A supercomputer: its hardware and software. J. Computer Science and Technology 26(3), 344–351 (2001)
2. Mo, Z.Y., Zhang, A.Q., Cao, X.L., et al.: JASMIN: a parallel software infrastructure for scientific computing. Front. Compute. Sci. China. 4(4), 480–488 (2010)
3. Mo, Z.Y., Pei, W.: Scientific computing application. Physics 38(8), 552–558 (2009) (in Chinese)
4. Childs, H., Pugmire, D., Ahern, S., Whitlock, B., Howison, M.: Extreme Scaling of Production Visualization Software on Diverse Architectures. IEEE Computer Graphics and Applications 30(3), 22–31 (2010)
5. LLNL. Visit, `https://wci.llnl.gov/codes/visit` (viewed on May 2011)
6. Rubel, O., Ahern, S., Bethel, E., Biggin, M.D., Childs, H., Cormier-Michel, E., DePace, A., Eisen, M.B., Fowlkes, C.C.: Coupling visualization and data analysis for knowledge discovery from multi-dimensional scientific data. Procedia Computer Science 1(1), 1751–1758 (2010)
7. Nicolae, B., Antoniu, G., Bouge, L., Moise, D.: BlobSeer: Next generation data management for large scale infrastructure. Journal of Parallel and Distributed Computing 71(2), 169–184 (2011)
8. Chan, A., Gropp, W., Lusk, E.: An efficient format for nearly constant-time access to arbitrary time intervals in larger trace files. Scientific Programming 16, 155–165 (2008)
9. Lang, S., Carns, P.H., Latham, R., Ross, R.B., Harms, K., Allcock, W.E.: I/O performance challenges at leadership scale. In: Sc 2009: Proceedings of the 2009 ACM/IEEE Conference on Supercomputing (2009)
10. Ma, K.L.: In situ visualization at extreme: challenges and opportunities. IEEE Computer Graphics and Applications 29(6), 14–19 (2009)

11. Peterka, T., Ross, R.B., Shen, H.W., et al.: Parallel visualization on leadership computing resources. In: Journal of Physics: Conference Series SciDAC (2009)
12. Ross, R.B., Peterka, T., Shen, H.W., et al.: Parallel I/O and visualization on extreme scale. In: Journal of Physics: Conference Series SciDAC (June 2008)
13. Yu, H., Ma, K.L.: A Study of I/O Techniques for Parallel Visualization. Journal of Parallel Computing 31(2), 167–183 (2005)
14. Iskra, K., Romein, J.W., Yoshii, K., Beckman, P.: ZOID: I/O-forwarding infrastructure for petascale architectures. In: PPoPP 2008: Proc. 13th ACM SIGPLAN Symp. On Principles and Practice of Parallel Programming, pp. 153–162 (2008)
15. Mo, Z., Zhang, A. (eds.): User's guide for JASMIN, Technical Report No. T09-JMJL-01 (2009), `https://www.iapcm.ac.cn/jasmin`
16. Pei, W.: The construction of simulation algorithms for Laser Fusion. Communication in Computational Physics 2(2), 255–270 (2007)

The Boom and Bust and Boom of Educational Games

Eric Klopfer and Scot Osterweil

Massachusetts Institute of Technology
MIT Teacher Education Program and The Education Arcade,
77 Massachusetts Ave., Room 10-337, Cambridge, Mass. USA

Abstract. The history of computer-based learning games has a story arc that rises dramatically, and then plummets steeply. In the early days of personal computers, creative minds drawn to the new medium explored a variety of approaches to learning games, ranging from behaviorist drill-and-practice exercises, to open-ended environments suitable for either exploration or construction. Early practitioners were inventing new forms, and even the fundamentally limited drill-and-practice games were infused with a measure of creative energy and humor. The late 1980s and mid 1990s were a heyday for the CDROM edutainment era. However, this era came to a crashing halt as the Internet dawned and the market for edutainment dried up in the late 1990s. Despite the downfall of the edutainment era there is new energy and perspective behind the idea of learning games. While edutainment of the 1990s style has gone, there is a new take on what learning games can be. In this era, we are finding that "making a game out of learning is most certainly not the way to approach the development of learning games. However, finding the fun in that learning" and devising ways to focus on and enhance that fun as a core game dynamic is a good strategy.

Keywords: klopfer, osterweil, MIT, education, educational, technology, game, games, educational games, edutainment, computer, computer-based, computer based, learning, augmented, teaching, teacher, teacher education, augmented learning, child, children, childhood, play, playing, video game, video games, video, multimedia, nintendo, leapfrog, math, health, change, mathematics.

1 Introduction

The history of computer-based learning games has a story arc that rises dramatically, and then plummets steeply. In the early days of personal computers, creative minds drawn to the new medium explored a variety of approaches to learning games, ranging from behaviorist drill-and-practice exercises, to open-ended environments suitable for either exploration or construction. Early practitioners were inventing new forms, and even the fundamentally limited drill-and-practice games were infused with a measure of creative energy and humor. For users of these early products, each new title represented another interesting step into unknown territory.

Z. Pan et al. (Eds.): Transactions on Edutainment IX, LNCS 7544, pp. 290–296, 2013.
© Springer-Verlag Berlin Heidelberg 2013

1.1 The CD ROM Era

These products were first delivered on floppy disks and marketed alongside pure entertainment games in the few computer stores of the time. By the early 90s the adoption of compact disc (CD) drives, and improved processing speeds led to a flowering of products with increasingly rich art, animation and more sophisticated computational possibilities. As the educational titles proliferated they spawned their own industry complete with mail order catalogs for parents and teachers. The titles from this era include long-enduring lines like Where in the World is Carmen San Diego, Math Blaster, and Reader Rabbit, but notably includes many more titles that were primarily found in the bargain bins of mass retail stores. Indeed this era saw a broad infatuation with CDs as the ultimate delivery mechanism. Publishers raced to develop interactive CDs that not only occupied the computer game and educational space, but that would (as many people saw it) in theory supplant traditional reference books (encyclopedias, atlases, etc.) if not eventually all book publishing. In many ways the CD became the product, rather than merely the delivery mechanism. Parents and teachers purchased educational CD ROMs, not educational games, educational references, or multimedia guides. The CD ROM was a method for delivering educational and entertainment content of many kinds, and in fact muddied the waters of educational publishing, such that publishers focused more on the mechanism and marketing than the contents. This era came to a crashing halt with the sudden emergence of web-browsers, opening up the Internet and all of its free content to average users. The gold-rush mentality that previously animated CD publishing was abruptly redirected toward the Internet, and in that early delirium of the dot com boom, content was suddenly offered for free on-line, in the hope that money-making models would eventually emerge. Games and other products for children were a major part of this mix of web-served content, putting great stress on the market for shrink-wrapped, CD-based software. As was characteristic of many markets that emerged during the Internet boom, products were rapidly produced so that companies could be the first movers. This left little room for the development of time intensive learning games.

The delivery mechanism itself also influenced the products that were produced. A CD had to be purchased. This investment meant that most people had access to only a few titles, and subsequently could and would in turn invest time into a more in-depth experience. When things started moving at Internet speed, the way in which games had to appeal to their audience changed. They needed to be immediately and continually appealing, or the player could just leave for another site. Games became a means to attract people to a site, rather than the destination themselves.

There were other coincident pressures on the learning game industry that contributed to its decline. The rising popularity of personal computers led to their commodification. Software products that were once sold in specialty stores were now appearing on Wal-marts shelves. Big box stores required publishers to offer products at lower prices, and for lower profit margins. The need to distinguish their products in this environment led publishers to invest heavily in licensing characters with built-in market appeal (Rug Rats, Sponge Bob, etc.) that would stand out on the shelf. In an environment of diminishing profits, funds that previously went into R&D were now dedicated to these licenses. Publishers had little choice but to recycle last years product with this years hot character.

1.2 The Sinking Edutainment Ship

The final factor in this perfect storm was the rising demand by buyers (i.e. parents) for products narrowly focused on improving school performance. Drill-and-practice, led by the Math Blaster and Reader Rabbit line of products, was always a large component of the industry, but the appearance of the Jump Start series represented a turning point. The product name appealed perfectly to parents anxieties that other peoples children were getting a jump-start on education and all the future rewards that would follow in its train. Not only did these products play to the cultural moment, but they also had the kind of simplistic marketing appeal that worked in the big-box retail environment. In its two-word title, Jump Start told parents everything they needed to know about the product, what marketers called the why to buy. In response (representative of the industry at large) to the success of Jump Start, Broderbund Software, long regarded as the leader in creative software design, replaced its marketing manager with one new to the software industry, but experienced in placing packaged goods on store shelves. A diverse range of creative childrens products were suddenly marketed in the Active Minds series, with uniform packaging and lengthy checklists of educational features. Development of new product slowed, and within a few years Broderbund was swallowed up by the less adventuresome Learning Company.

The checklists on the back of software boxes that narrowly defined academic skills worthy of developing, along with the catalogs that boxed software into specific academic categories also left little room for what became known as Thinking Games. These games did not necessarily connect to academic standards, but clearly involved a diverse set of intellectual skills from problem-solving to creativity to quantitative skills. Games like The Incredible Machine, which challenged players to create various Rube Goldberg-esque contraptions, stand out in this category. In many ways these games represented the best of what the edutainment era had to offer, games that were simultaneously fun and intellectually challenging, and in fact they were fun because they were intellectually challenging. These games had a hard time neither being entirely at home in the educational or entertainment aisles.

What once had been a diverse range of products was rather quickly reduced to one very limited model. At the height of the CD boom, all childrens software productswhether designed for pure entertainment or with learning goals in mindoccupied shelf space in Toys R Us right next to the flashy entertainments designed for video-game consoles. By the end of the decade, Toys R Us had moved the video-games to the very front of the store (with increased security to protect from theft) while edutainment products were relegated to ever smaller shelves at the back of the store, and from all appearances the store was not particularly worried about customers wanting to steal the products. The ever-shrinking edutainment space was eventually replaced a decade later with the green aisle at big box toy stores, which was dominated by the early childhood Leapfrog gaming systems. In some ways Leapfrog redefined the learning games space, putting it back on the map in 21st century. And while innovative titles could be found, it quickly fell into the shelf-appeal licensing debacle that the CD ROMs had faced previously, and the new generation of learning game systems became yet another method for delivering recycled content with a fresh coat of licensed characters.

2 Why It Will Be Different This Time

Others have offered perspectives on the downfall of the edutainment era, including Mizuko Ito (2008), who chronicles not only the economic factors that contributed to this decline in childrens software, but also brings an anthropologists eye to the subject, observing the ways different products fit within long-existing approaches to childrens play and learning. A major contribution of Itos article is her categorization of three different genres of games targeted for children. She describes Educational games that are largely drill-and-practice exercises tied to the narrow curricular goals of the traditional classroom. In her formulation, Entertainment games are more exploratory, narrative based games that privilege play, and that stand decidedly apart from the institution of education. The third category, Construction includes simulation games like SimCity and Zoo Tycoon.

As solid a foundation as Ito provides, there are two respects in which her article may not be as helpful in looking forward to new approaches to Learning Games, and understanding why the new era of games represents a different model and a new opportunity. The first is the degree to which her categorization may mask other possibilities for games that break genre. Ito is clear in suggesting that such genre-bending is possible, but she largely dwells on games that define rather than violate categories. We believe there is tremendous power in blending forms, and are particularly interested in the ways in which games might follow the form of Entertainment titles but nevertheless offer intellectual challenges that contribute to academic accomplishment (much the way Macbeth, or Pride and Prejudice can be tools for fostering intellectual growth while remaining entertainment properties).

Itos article also treats the games industry as mature and suggests that there was something inevitable about the industry consolidation that stifled earlier innovations and led to the separation of her three genres into different market segments. While we would agree that there were powerful market forces that brought us to the present moment, we dont necessarily see the current situation as mature. The still-growing reach of the Internet and the proliferation of new game platforms (handhelds, web-enabled videogame consoles, cell phones) all suggest a market still ripe for innovation and creative destruction. There are new ways of finding, sampling, and buying games, and we believe over time negative factors such as big-box stores, and a limited number of dominant publishers will have a diminishing impact on the evolution of new games.

Despite the downfall of the edutainment era there is new energy and perspective behind the idea of learning games. While edutainment of the 1990s style has gone, there is a new take on what learning games can be. This perspective is primarily fueled by two ideas:

1. Video Games are not Mindless There is a new field of researchers who have shown a good deal of evidence that, despite their image in the media, commercial entertainment video games are not mindless. In fact players show a great deal of thought and learning in playing video games. This includes not only the obvious titles like SimCity and Civilization, but massively multiplayer online roleplaying games like World of Warcraft and even first person shooters. This notion was most notably popularized by Jim Gees What Video Games Have to Teach us About Learning and Literacy (2003).
2. Games Got Serious While the notion of kids learning through games fell somewhat by the wayside, other industries and interests began to capitalize on the video game model. Recognizing that video games provide an environment in which people engage in sustained challenges and enjoy it, other uses for video games have been sought by many. The Serious Games Movement, founded by Ben Sawyer and Dave Rijeski, crystallized this idea and served to advance the field. While much of the original focus was on the use of video games by the military, the uses have evolved far beyond that, encompassing many forms of training and learning. These uses have in turn spawned many offspring including Games for Health, and Games for Change.

Together, these ideas show that there is a lot of value in game play that can be meaningfully tied to learning in ways much deeper than most of the edutainment era provided. And this learning can combine both important process skills along with content learning.

2.1 Baking Educational Games

The process of designing and creating educational games can be thought of much like the process of baking. There are many attempts by a growing number of health conscious cooks to make things that are both yummy and healthy. It isn't easy to balance these two qualities. It is relatively easy to bake calorie-laden cakes filled with butter and transfats that are quite tasty, or one could make piles of oat germ laden sawdust that could reduce your cholesterol if only you could actually eat it. To get both of these factors balanced is hard, and there likely is no universal solution (other than perhaps iterative experimentation). Some recipes work really well for some groups of people, in certain contexts, with particular expectations. Similarly, in creating experiences that are both fun and filled with learning, the success of different recipes (mixes of media, immersion, styles of games, learning goals, mixtures of content, etc.) depends quite a bit on the audience, context, content, goals and facilitation.

There are two recipes that are followed quite a bit today to create a blended balance of what you want and what is good for you. One recipe takes the yummy calorie laden cake and injects beta-carotene, vitamin D and calcium right into the cake. This is no doubt a delicious cake, but its nutritional content is highly suspect. Similarly, designers of educational games that try to inject content learning into a game where it doesn't fit may create experiences that are somewhat entertaining, but their educational value is highly suspect. If your spaceship requires you to answer a math problem before you can use your blasters, chances are you'll hate the game and the math. This is the strategy taken by most of the legacy edutainment games (e.g. Math Blaster), as well as many of the new attempts to create commercially viable learning games today (e.g. the newer immersive 3D math game, Dimension M).

The other recipe simply takes all of the healthy content - wheat germ, oat bran, carrot juice, spinach leaves, etc. and bakes them into something that looks like a cake. It sure is healthy, but just because it looks like a cake, doesn't make it tasty. Many people designing educational games follow this recipe. They take educational content and do something to make it look like a game (i.e. put algebra problems in a 3D virtual world, or place the periodic table of the elements into a shooting arcade). While there may be educational potential in such an approach, these games often "suck," as Ted Castronova found and documented in a recent article in WIRED (Baker 2008). Just because it looks like a game, doesn't make it a game.

Castronova, a researcher of virtual economies, presents some tips for "Making Games That Don't Suck," which are good starters for anyone thinking about rushing into the game-making business, educational or not. Castronova created a game (a heavily modded version of Never Winter Nights) to attract players to a world in which he could study their behaviors. The world turned out to be a great looking, but primarily unfun experience, that did little to attract players. Castronova in turn advised:

- Don't Be Overly Ambitious
- Go Low Tech
- Think About Your Audience
- Get a Full-Time Staff
- Concede Screwups

These are particularly relevant for the field of educational games, which is currently taking on the appearance of a gold rush, but which frequently operates without the resources of mainstream gaming developers. Throughout academia, industry, government agencies and non-profits, many people are flocking to educational games as the silver bullet to cure our educational woes. While there is clearly power in the medium of games to teach and to learn from, educational games need to be thoughtfully designed and considered from both the educational and entertainment perspectives (as Castronova lays out) so that these games "don't suck".

So what is the magical recipe for a good educational game? While there is no silver bullet, at least the beginning of the answer lies in the framing of the problem. "Making a game out of learning will most certainly not be the way to approach the development of learning games. However, finding the fun in that learning" and devising ways to focus on and enhance that fun as a core game dynamic is a good strategy. Castronova likely has something similar in mind when he says, "consider your audience." Additionally, choosing the right technology plays into the decision, as Castronova indicates when he advises people to manage their ambitions and go low tech. We might amend that to "go with the right tech." In some cases a game may require an expensive 3D virtual world, but in other cases a text-based game may be perfect, and yet others might be appropriate for mobile devices that you can take with you anywhere. Too much time and effort has been spent on trying to capitalize on the expensive 3D virtual worlds, and not enough on the smaller, less-flashy approaches, but both offer potential for educational games.

3 Promises for Tomorrows Games

The world of learning games is currently reaping the benefits, not only of conceptual advances in the way designers and developers are linking learning and game play, but also from the rapid advances in the video games industry. For many years the advances in the video game industry were primarily reserved for higher bits and resolutions. But the last several years have seen great technical and conceptual advances in video game designs. The rise of the Nintendo DS and Wii platforms, as well as casual online and cell phone games, have shown that video games can take many forms and reach diverse audiences. This is critical for the learning games sector, in that it cannot be resigned to serving narrow privileged audiences. Instead they need to be able to reach learners of all types. These next generation ideas are already inspiring educational innovation, and demonstrating that educational games have learned a lot this time around.

References

1. Baker, C.: Trying to Design a Truly Entertaining Game Can Defeat Even a Certified Genius. Wired Magazine 16(04) (2008)
2. Ito, M.: Education vs. Entertainment: A Cultural History of Childrens Software. In: Salen, K. (ed.) The Ecology of Games: Connecting Youth, Games and Learning. The John D. and Catherine T. MacArthur Foundation Series on Digital Media and Learning, pp. 89–116. The MIT Press, Cambridge (2006)
3. Gee, J.P.: What video games have to teach us about learning and literacy. Palgrave/ St. Martins, New York (2003)

Author Index