Randall Shumaker (Ed.)

LNCS 8022

Virtual, Augmented and Mixed Reality

Systems and Applications

5th International Conference, VAMR 2013
Held as Part of HCI International 2013
Las Vegas, NV, USA, July 2013, Proceedings, Part II

2 Part II

HCI2013 INTERNATIONAL

Springer

Lecture Notes in Computer Science 8022

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

Randall Shumaker (Ed.)

Virtual, Augmented and Mixed Reality

Systems and Applications

5th International Conference, VAMR 2013
Held as Part of HCI International 2013
Las Vegas, NV, USA, July 21-26, 2013
Proceedings, Part II

Springer

Volume Editor

Randall Shumaker
University of Central Florida
Institute for Simulation and Training
3100 Technology Parkway, Orlando, FL 32826, USA
E-mail: shumaker@ist.ucf.edu

ISSN 0302-9743 e-ISSN 1611-3349
ISBN 978-3-642-39419-5 e-ISBN 978-3-642-39420-1
DOI 10.1007/978-3-642-39420-1
Springer Heidelberg Dordrecht London New York

Library of Congress Control Number: 2013942173

CR Subject Classification (1998): H.5, K.3, K.8, J.1, J.3, J.7, I.3.7, K.4.4, H.4

LNCS Sublibrary: SL 3 – Information Systems and Application, incl. Internet/Web and HCI

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)

Foreword

The 15th International Conference on Human–Computer Interaction, HCI International 2013, was held in Las Vegas, Nevada, USA, 21–26 July 2013, incorporating 12 conferences / thematic areas:

Thematic areas:

- Human–Computer Interaction
- Human Interface and the Management of Information

Affiliated conferences:

- 10th International Conference on Engineering Psychology and Cognitive Ergonomics
- 7th International Conference on Universal Access in Human–Computer Interaction
- 5th International Conference on Virtual, Augmented and Mixed Reality
- 5th International Conference on Cross-Cultural Design
- 5th International Conference on Online Communities and Social Computing
- 7th International Conference on Augmented Cognition
- 4th International Conference on Digital Human Modeling and Applications in Health, Safety, Ergonomics and Risk Management
- 2nd International Conference on Design, User Experience and Usability
- 1st International Conference on Distributed, Ambient and Pervasive Interactions
- 1st International Conference on Human Aspects of Information Security, Privacy and Trust

A total of 5210 individuals from academia, research institutes, industry and governmental agencies from 70 countries submitted contributions, and 1666 papers and 303 posters were included in the program. These papers address the latest research and development efforts and highlight the human aspects of design and use of computing systems. The papers accepted for presentation thoroughly cover the entire field of Human–Computer Interaction, addressing major advances in knowledge and effective use of computers in a variety of application areas.

This volume, edited by Randall Shumaker, contains papers focusing on the thematic area of Virtual, Augmented and Mixed Reality, and addressing the following major topics:

- Healthcare and Medical Applications
- Virtual and Augmented Environments for Learning and Education
- Business, Industrial and Military Applications
- Culture and Entertainment Applications

The remaining volumes of the HCI International 2013 proceedings are:

- Volume 1, LNCS 8004, Human–Computer Interaction: Human-Centred Design Approaches, Methods, Tools and Environments (Part I), edited by Masaaki Kurosu
- Volume 2, LNCS 8005, Human–Computer Interaction: Applications and Services (Part II), edited by Masaaki Kurosu
- Volume 3, LNCS 8006, Human–Computer Interaction: Users and Contexts of Use (Part III), edited by Masaaki Kurosu
- Volume 4, LNCS 8007, Human–Computer Interaction: Interaction Modalities and Techniques (Part IV), edited by Masaaki Kurosu
- Volume 5, LNCS 8008, Human–Computer Interaction: Towards Intelligent and Implicit Interaction (Part V), edited by Masaaki Kurosu
- Volume 6, LNCS 8009, Universal Access in Human–Computer Interaction: Design Methods, Tools and Interaction Techniques for eInclusion (Part I), edited by Constantine Stephanidis and Margherita Antona
- Volume 7, LNCS 8010, Universal Access in Human–Computer Interaction: User and Context Diversity (Part II), edited by Constantine Stephanidis and Margherita Antona
- Volume 8, LNCS 8011, Universal Access in Human–Computer Interaction: Applications and Services for Quality of Life (Part III), edited by Constantine Stephanidis and Margherita Antona
- Volume 9, LNCS 8012, Design, User Experience, and Usability: Design Philosophy, Methods and Tools (Part I), edited by Aaron Marcus
- Volume 10, LNCS 8013, Design, User Experience, and Usability: Health, Learning, Playing, Cultural, and Cross-Cultural User Experience (Part II), edited by Aaron Marcus
- Volume 11, LNCS 8014, Design, User Experience, and Usability: User Experience in Novel Technological Environments (Part III), edited by Aaron Marcus
- Volume 12, LNCS 8015, Design, User Experience, and Usability: Web, Mobile and Product Design (Part IV), edited by Aaron Marcus
- Volume 13, LNCS 8016, Human Interface and the Management of Information: Information and Interaction Design (Part I), edited by Sakae Yamamoto
- Volume 14, LNCS 8017, Human Interface and the Management of Information: Information and Interaction for Health, Safety, Mobility and Complex Environments (Part II), edited by Sakae Yamamoto
- Volume 15, LNCS 8018, Human Interface and the Management of Information: Information and Interaction for Learning, Culture, Collaboration and Business (Part III), edited by Sakae Yamamoto
- Volume 16, LNAI 8019, Engineering Psychology and Cognitive Ergonomics: Understanding Human Cognition (Part I), edited by Don Harris
- Volume 17, LNAI 8020, Engineering Psychology and Cognitive Ergonomics: Applications and Services (Part II), edited by Don Harris
- Volume 18, LNCS 8021, Virtual, Augmented and Mixed Reality: Designing and Developing Augmented and Virtual Environments (Part I), edited by Randall Shumaker

- Volume 20, LNCS 8023, Cross-Cultural Design: Methods, Practice and Case Studies (Part I), edited by P.L. Patrick Rau
- Volume 21, LNCS 8024, Cross-Cultural Design: Cultural Differences in Everyday Life (Part II), edited by P.L. Patrick Rau
- Volume 22, LNCS 8025, Digital Human Modeling and Applications in Health, Safety, Ergonomics and Risk Management: Healthcare and Safety of the Environment and Transport (Part I), edited by Vincent G. Duffy
- Volume 23, LNCS 8026, Digital Human Modeling and Applications in Health, Safety, Ergonomics and Risk Management: Human Body Modeling and Ergonomics (Part II), edited by Vincent G. Duffy
- Volume 24, LNAI 8027, Foundations of Augmented Cognition, edited by Dylan D. Schmorrow and Cali M. Fidopiastis
- Volume 25, LNCS 8028, Distributed, Ambient and Pervasive Interactions, edited by Norbert Streitz and Constantine Stephanidis
- Volume 26, LNCS 8029, Online Communities and Social Computing, edited by A. Ant Ozok and Panayiotis Zaphiris
- Volume 27, LNCS 8030, Human Aspects of Information Security, Privacy and Trust, edited by Louis Marinos and Ioannis Askoxylakis
- Volume 28, CCIS 373, HCI International 2013 Posters Proceedings (Part I), edited by Constantine Stephanidis
- Volume 29, CCIS 374, HCI International 2013 Posters Proceedings (Part II), edited by Constantine Stephanidis

I would like to thank the Program Chairs and the members of the Program Boards of all affiliated conferences and thematic areas, listed below, for their contribution to the highest scientific quality and the overall success of the HCI International 2013 conference.

This conference could not have been possible without the continuous support and advice of the Founding Chair and Conference Scientific Advisor, Prof. Gavriel Salvendy, as well as the dedicated work and outstanding efforts of the Communications Chair and Editor of HCI International News, Abbas Moallem.

I would also like to thank for their contribution towards the smooth organization of the HCI International 2013 Conference the members of the Human–Computer Interaction Laboratory of ICS-FORTH, and in particular George Paparoulis, Maria Pitsoulaki, Stavroula Ntoa, Maria Bouhli and George Kapnas.

May 2013

Constantine Stephanidis
General Chair, HCI International 2013

Organization

Human–Computer Interaction

Program Chair: Masaaki Kurosu, Japan

Jose Abdelnour-Nocera, UK
Sebastiano Bagnara, Italy
Simone Barbosa, Brazil
Tomas Berns, Sweden
Nigel Bevan, UK
Simone Borsci, UK
Apala Lahiri Chavan, India
Sherry Chen, Taiwan
Kevin Clark, USA
Torkil Clemmensen, Denmark
Xiaowen Fang, USA
Shin'ichi Fukuzumi, Japan
Vicki Hanson, UK
Ayako Hashizume, Japan
Anzai Hiroyuki, Italy
Sheue-Ling Hwang, Taiwan
Wonil Hwang, South Korea
Minna Isomursu, Finland
Yong Gu Ji, South Korea
Esther Jun, USA
Mitsuhiko Karashima, Japan
Kyungdoh Kim, South Korea
Heidi Krömker, Germany
Chen Ling, USA
Yan Liu, USA
Zhengjie Liu, P.R. China
Loïc Martínez Normand, Spain
Chang S. Nam, USA
Naoko Okuizumi, Japan
Noriko Osaka, Japan
Philippe Palanque, France
Hans Persson, Sweden
Ling Rothrock, USA
Naoki Sakakibara, Japan
Dominique Scapin, France
Guangfeng Song, USA
Sanjay Tripathi, India
Chui Yin Wong, Malaysia
Toshiki Yamaoka, Japan
Kazuhiko Yamazaki, Japan
Ryoji Yoshitake, Japan
Silvia Zimmermann, Switzerland

Human Interface and the Management of Information

Program Chair: Sakae Yamamoto, Japan

Hans-Jorg Bullinger, Germany
Alan Chan, Hong Kong
Gilsoo Cho, South Korea
Jon R. Gunderson, USA
Shin'ichi Fukuzumi, Japan
Michitaka Hirose, Japan
Jhilmil Jain, USA
Yasufumi Kume, Japan
Mark Lehto, USA
Hiroyuki Miki, Japan
Hirohiko Mori, Japan
Fiona Fui-Hoon Nah, USA
Shogo Nishida, Japan
Robert Proctor, USA
Youngho Rhee, South Korea
Katsunori Shimohara, Japan

Michale Smith, USA
Tsutomu Tabe, Japan
Hiroshi Tsuji, Japan
Kim-Phuong Vu, USA
Tomio Watanabe, Japan
Hidekazu Yoshikawa, Japan

Engineering Psychology and Cognitive Ergonomics

Program Chair: Don Harris, UK

Guy Andre Boy, USA
Joakim Dahlman, Sweden
Trevor Dobbins, UK
Mike Feary, USA
Shan Fu, P.R. China
Michaela Heese, Austria
Hung-Sying Jing, Taiwan
Wen-Chin Li, Taiwan
Mark A. Neerincx, The Netherlands
Jan M. Noyes, UK
Taezoon Park, Singapore
Paul Salmon, Australia
Axel Schulte, Germany
Siraj Shaikh, UK
Sarah C. Sharples, UK
Anthony Smoker, UK
Neville A. Stanton, UK
Alex Stedmon, UK
Xianghong Sun, P.R. China
Andrew Thatcher, South Africa
Matthew J.W. Thomas, Australia
Rolf Zon, The Netherlands

Universal Access in Human–Computer Interaction

Program Chairs: Constantine Stephanidis, Greece, and Margherita Antona, Greece

Julio Abascal, Spain
Ray Adams, UK
Gisela Susanne Bahr, USA
Margit Betke, USA
Christian Bühler, Germany
Stefan Carmien, Spain
Jerzy Charytonowicz, Poland
Carlos Duarte, Portugal
Pier Luigi Emiliani, Italy
Qin Gao, P.R. China
Andrina Granić, Croatia
Andreas Holzinger, Austria
Josette Jones, USA
Simeon Keates, UK
Georgios Kouroupetroglou, Greece
Patrick Langdon, UK
Seongil Lee, Korea
Ana Isabel B.B. Paraguay, Brazil
Helen Petrie, UK
Michael Pieper, Germany
Enrico Pontelli, USA
Jaime Sanchez, Chile
Anthony Savidis, Greece
Christian Stary, Austria
Hirotada Ueda, Japan
Gerhard Weber, Germany
Harald Weber, Germany

Virtual, Augmented and Mixed Reality

Program Chair: Randall Shumaker, USA

Waymon Armstrong, USA
Juan Cendan, USA
Rudy Darken, USA
Cali M. Fidopiastis, USA
Charles Hughes, USA
David Kaber, USA
Hirokazu Kato, Japan
Denis Laurendeau, Canada
Fotis Liarokapis, UK
Mark Livingston, USA
Michael Macedonia, USA
Gordon Mair, UK
Jose San Martin, Spain
Jacquelyn Morie, USA
Albert “Skip” Rizzo, USA
Kay Stanney, USA
Christopher Stapleton, USA
Gregory Welch, USA

Cross-Cultural Design

Program Chair: P.L. Patrick Rau, P.R. China

Pilsung Choe, P.R. China
Henry Been-Lirn Duh, Singapore
Vanessa Evers, The Netherlands
Paul Fu, USA
Zhiyong Fu, P.R. China
Fu Guo, P.R. China
Sung H. Han, Korea
Toshikazu Kato, Japan
Dyi-Yih Michael Lin, Taiwan
Rungtai Lin, Taiwan
Sheau-Farn Max Liang, Taiwan
Liang Ma, P.R. China
Alexander Mädche, Germany
Katsuhiko Ogawa, Japan
Tom Plocher, USA
Kerstin Röse, Germany
Supriya Singh, Australia
Hsiu-Ping Yueh, Taiwan
Liang (Leon) Zeng, USA
Chen Zhao, USA

Online Communities and Social Computing

Program Chairs: A. Ant Ozok, USA, and Panayiotis Zaphiris, Cyprus

Areej Al-Wabil, Saudi Arabia
Leonelo Almeida, Brazil
Bjørn Andersen, Norway
Chee Siang Ang, UK
Aneesha Bakharia, Australia
Ania Bobrowicz, UK
Paul Cairns, UK
Farzin Deravi, UK
Andri Ioannou, Cyprus
Slava Kisilevich, Germany
Niki Lambropoulos, Greece
Effie Law, Switzerland
Soo Ling Lim, UK
Fernando Loizides, Cyprus
Gabriele Meiselwitz, USA
Anthony Norcio, USA
Elaine Raybourn, USA
Panote Siriaraya, UK
David Stuart, UK
June Wei, USA

Augmented Cognition

Program Chairs: Dylan D. Schmorrow, USA, and Cali M. Fidopiastis, USA

Robert Arrabito, Canada
Richard Backs, USA
Chris Berka, USA
Joseph Cohn, USA
Martha E. Crosby, USA
Julie Drexler, USA
Ivy Estabrooke, USA
Chris Forsythe, USA
Wai Tat Fu, USA
Rodolphe Gentili, USA
Marc Grootjen, The Netherlands
Jefferson Grubb, USA
Ming Hou, Canada
Santosh Mathan, USA
Rob Matthews, Australia
Dennis McBride, USA
Jeff Morrison, USA
Mark A. Neerincx, The Netherlands
Denise Nicholson, USA
Banu Onaral, USA
Lee Sciarini, USA
Kay Stanney, USA
Roy Stripling, USA
Rob Taylor, UK
Karl van Orden, USA

Digital Human Modeling and Applications in Health, Safety, Ergonomics and Risk Management

Program Chair: Vincent G. Duffy, USA and Russia

Karim Abdel-Malek, USA
Giuseppe Andreoni, Italy
Daniel Carruth, USA
Eliza Yingzi Du, USA
Enda Fallon, Ireland
Afzal Godil, USA
Ravindra Goonetilleke, Hong Kong
Bo Hoege, Germany
Waldemar Karwowski, USA
Zhizhong Li, P.R. China
Kang Li, USA
Tim Marler, USA
Michelle Robertson, USA
Matthias Rötting, Germany
Peter Vink, The Netherlands
Mao-Jiun Wang, Taiwan
Xuguang Wang, France
Jingzhou (James) Yang, USA
Xiugan Yuan, P.R. China
Gülcin Yücel Hoge, Germany

Design, User Experience, and Usability

Program Chair: Aaron Marcus, USA

Sisira Adikari, Australia
Ronald Baecker, Canada
Arne Berger, Germany
Jamie Blustein, Canada
Ana Boa-Ventura, USA
Jan Brejcha, Czech Republic
Lorenzo Cantoni, Switzerland
Maximilian Eibl, Germany

Anthony Faiola, USA
Emilie Gould, USA
Zelda Harrison, USA
Rüdiger Heimgärtner, Germany
Brigitte Herrmann, Germany
Steffen Hess, Germany
Kaleem Khan, Canada
Jennifer McGinn, USA
Francisco Rebelo, Portugal
Michael Renner, Switzerland
Kerem Rızvanoğlu, Turkey
Marcelo Soares, Brazil
Christian Sturm, Germany
Michele Visciola, Italy

Distributed, Ambient and Pervasive Interactions

Program Chairs: Norbert Streitz, Germany, and Constantine Stephanidis, Greece

Emile Aarts, The Netherlands
Adnan Abu-Dayya, Qatar
Juan Carlos Augusto, UK
Boris de Ruyter, The Netherlands
Anind Dey, USA
Dimitris Grammenos, Greece
Nuno M. Guimaraes, Portugal
Shin'ichi Konomi, Japan
Carsten Magerkurth, Switzerland
Christian Müller-Tomfelde, Australia
Fabio Paternó, Italy
Gilles Privat, France
Harald Reiterer, Germany
Carsten Röcker, Germany
Reiner Wichert, Germany
Woontack Woo, South Korea
Xenophon Zabulis, Greece

Human Aspects of Information Security, Privacy and Trust

Program Chairs: Louis Marinos, ENISA EU, and Ioannis Askoxylakis, Greece

Claudio Agostino Ardagna, Italy
Zinaida Benenson, Germany
Daniele Catteddu, Italy
Raoul Chiesa, Italy
Bryan Cline, USA
Sadie Creese, UK
Jorge Cuellar, Germany
Marc Dacier, USA
Dieter Gollmann, Germany
Kirstie Hawkey, Canada
Jaap-Henk Hoepman, The Netherlands
Cagatay Karabat, Turkey
Angelos Keromytis, USA
Ayako Komatsu, Japan
Ronald Leenes, The Netherlands
Javier Lopez, Spain
Steve Marsh, Canada
Gregorio Martinez, Spain
Emilio Mordini, Italy
Yuko Murayama, Japan
Masakatsu Nishigaki, Japan
Aljosa Pasic, Spain
Milan Petković, The Netherlands
Joachim Posegga, Germany
Jean-Jacques Quisquater, Belgium
Damien Sauveron, France
George Spanoudakis, UK
Kerry-Lynn Thomson, South Africa

Julien Touzeau, France
Theo Tryfonas, UK
João Vilela, Portugal
Claire Vishik, UK
Melanie Volkamer, Germany

External Reviewers

Maysoon Abulkhair, Saudi Arabia
Ilia Adami, Greece
Vishal Barot, UK
Stephan Böhm, Germany
Vassilis Charissis, UK
Francisco Cipolla-Ficarra, Spain
Maria De Marsico, Italy
Marc Fabri, UK
David Fonseca, Spain
Linda Harley, USA
Yasushi Ikei, Japan
Wei Ji, USA
Nouf Khashman, Canada
John Killilea, USA
Iosif Klironomos, Greece
Ute Klotz, Switzerland
Maria Korozi, Greece
Kentaro Kotani, Japan
Vassilis Kouroumalis, Greece
Stephanie Lackey, USA
Janelle LaMarche, USA
Asterios Leonidis, Greece
Nickolas Macchiarella, USA
George Margetis, Greece
Matthew Marraffino, USA
Joseph Mercado, USA
Claudia Mont'Alvão, Brazil
Yoichi Motomura, Japan
Karsten Nebe, Germany
Stavroula Ntoa, Greece
Martin Osen, Austria
Stephen Prior, UK
Farid Shirazi, Canada
Jan Stelovsky, USA
Sarah Swierenga, USA

HCI International 2014

The 16th International Conference on Human–Computer Interaction, HCI International 2014, will be held jointly with the affiliated conferences in the summer of 2014. It will cover a broad spectrum of themes related to Human–Computer Interaction, including theoretical issues, methods, tools, processes and case studies in HCI design, as well as novel interaction techniques, interfaces and applications. The proceedings will be published by Springer. More information about the topics, as well as the venue and dates of the conference, will be announced through the HCI International Conference series website: http://www.hci-international.org/

General Chair
Professor Constantine Stephanidis
University of Crete and ICS-FORTH
Heraklion, Crete, Greece
Email: cs@ics.forth.gr

Table of Contents – Part II

Virtual and Augmented Environments for Learning and Education

Business, Industrial and Military Applications

Culture and Entertainment Applications

Table of Contents – Part I

Developing Augmented and Virtual Environments

Interaction in Augmented and Virtual Environments

Human-Robot Interaction in Virtual Environments

Presence and Tele-presence

Part I
Healthcare and Medical Applications

Gait Analysis Management and Diagnosis in a Prototype Virtual Reality Environment

Salsabeel F.M. Alfalah, David K. Harrison, and Vassilis Charissis

Glasgow Caledonian University
Cowcaddens Road
Glasgow, G4 0BA, UK
Salsabeel.Alfalah@gcu.ac.uk

Abstract. Current medical data derived from gait analysis and diagnosis of various musculoskeletal pathologies offer a plethora of text based and imaging data. The large volume and complexity of the particular data present a number of issues during the collection, storage, searching and visualisation process for gait analysis management and diagnosis. Adhering to above it is evident that a simplified, holistic and user-friendly system is required in order to improve the acquisition and comparison of medical data in a timely manner. Further consultation with health professionals suggested that the proposed prototype should entail an automated system that can extract, save and visualise the data from different sources, in order to enhance medical data visualisation, increase efficiency and thus improve quality of service and management. This work presents the development stages of a new prototype system for managing medical data for gait analysis, which additionally offer simulation capacity in a Virtual Reality environment in order to assist the medical practitioners towards a faster an better informed evaluation of each condition. In particular this paper investigates various methods of displaying medical data in a single application with a view to managing and sharing multimedia data and to employing a VR to enhance user interaction with medical data. Findings of a promising preliminary evaluation through user trials are also presented. Concluding, the paper presents future plans to incorporate a bespoke 3D human- computer interface with a view to provide the health professionals with customisable information and enhancing the interface functionalities. Finally, as the system is web-based there is scope for expansion of the application to other areas of medical assessments involving complicated datasets.

1 Introduction

Gait analysis is an assessment tool for individuals with conditions affecting their ability to walk, and is used for the formulation of medical diagnosis and future treatment improvements as well as for research purposes [1]. Some gait analysis assessments depend not only on the standard physical examination, but require a complete description of the complex pathology of an abnormal human gait pattern. Several techniques have been developed for gait analysis, with different types of

R. Shumaker (Ed.): VAMR/HCII 2013, Part II, LNCS 8022, pp. 3–11, 2013.

information datasets offered. Gait analysis applications used currently by health professionals produce multiple sets of data that often need to be investigated simultaneously; there is currently no available system to support the multitasking process simulated in 3 Dimension (3D), and thus medical data is fragmented along multiple repositories with no integrated view or filtering criteria for required information, and with limitation of data view in 2 Dimension (2D).

The aforementioned limitations of the current system hinders the concurrent diagnosis route, thus it was considered essential to transfer the medical data from a randomly distributed system to an organised system that combines all the relevant data used for gait analysis under one application package. This paper presents the methodology used to integrate the various types of diagnostic medical data by developing interactive information-visualisation web based software that provides an authenticated access to clinical information over the internet in a Virtual Reality (VR) environment with an integrated view for all the various types of medical data. In turn, this enhances interaction with gait analysis data by simulation in 3D, enables report generation based on predetermined criteria for quality control and enhances data retrieval by filtering criteria.

Three dimensional representation either in a virtual environment or a typical computer screen offers a number of advantages over the current paper/digital media based method; VR enhances the interaction and understanding as well as allowing collaboration to occur by giving the user the ability to explore structures from several viewpoints [2,3]. Storing different and scattered gait analysis data sets in an Electronic Medical Record (EMR) system offers multiple advantages such as cost benefit, confidential data storing, flexible and quick data retrieval, enhanced medical data visualisation and easier data sharing and access from remote sites [4]. The proposed approach acts for health professionals as a diagnostic and research aid for improving their 3D mental mapping, increasing productivity, facilitating medical decision-making and ultimately contributes to a better quality of service and time management process as presented in previous studies [2,5].

Evaluation and feedback on the proposed work was initially informed by a survey of 30 health professionals as to the effectiveness, usefulness, and users' satisfaction of the proposed application. The questionnaire targeted areas in the users perception of the application relating to issues such as applicability of VR in general, users' overall satisfaction with the proposed system, medical data storage in a single repository, interaction with medical data, authenticated users' access to the system, motion capture in 3D, wireless manipulation of the data, better decision making and improved quality, and time management. Preliminary evaluation for the system indicated that the proposed system enhances the gait analysis diagnostic process and improves time management.

This paper is organised as follows: Section 2 reviews the related work and the background of this work. Section 3 demonstrates the research synthesis, system evaluation is presented in Section 4 and finally Section 5 concludes the paper, with a succinct summary of the project and a future tentative plan of work.

2 Background

Gait analysis is applied to assess human gait with the required data that can distinguish between normal and pathological gait, estimate the problem and determine the therapy needed. Gait studies are developed using various techniques depending on the required information, methodology, cost and applicability (research purposes, direct treatment) [7, 8]. To facilitate both ease and accuracy of the gait analysis process medical data needs to be managed with regard to storage and visualisation regardless of the assessment tool used, or the purpose of the analysis [there are medical data produced from the analysis process with various media types]. EMR supports the coordinated and authenticated access over the internet to the specific medical information presented to the requester using appropriate HCI which enhances the medical decision-making [9, 10, and 11]. There are various techniques used to enhance medical visualisation above managing medical data in an integrated view; allowing the viewer to switch from 2D to 3D in a VR environment will be of great benefit for both the health professional and in other cases for the patient as VR systems proved their efficiency when used for rehabilitation and telerehabilitation [12, 13, and 14]. Furthermore, providing interactive navigation through 3D models using HCI will facilitate more efficient understanding of the complex models thus enhancing medical education [15, 16].

3 Proposed System

This section presents the current method that health professionals use to view, store, extract, and manage medical data produced as an output from gait analysis process. In turn, we present the proposed system which aims to enhance the gait analysis process. The particular system shows the solution provided regarding each issue in the current way i.e. storage, visualisation, time management, quality control, and interactivity.

3.1 Current Method of Gait Analysis Process

The proposed system overcomes most of the disadvantages that occur in the current way of dealing with medical data produced from the gait analysis process. Figure 1 show the current way of the gait analysis process described as follows:

- Health professionals contact participants or patients for data collection through phoning or direct contact. In both ways health professionals will describe for the patients/participants the process of data collection and give them the required details about the location and time. Once the patient arrives the HP takes the patient/participant's information (text file), filling in and signing required forms (hard copy).

  This process is time consuming for both the health professional and the participant/patient and it is repetitive; the same description has to be provided for each patient/ participant, in addition, it is not clear as much as providing examples to be viewed by a participant with no background in computer graphics and data visualisation. When the patient/participant arrives there is no defined way to store the information provided providing a non-standardised process. As such each practitioner is able to store the derived information in a different manner and format.

- Data collection: the data collection process varies from one case to another, depending on the information required and the software used for collecting data. Currently, in Glasgow Caledonian University (GCU) health professionals are employing a 3D Motion camera system (Qualysis - 16 cameras OPUS 3/5) to measure the kinematics and a Kistler Force Plate (Kistler - 9286B) to measure the kinetics of gait. These systems are synchronised using the Qualysis Track Manager software (version 2.5) during data collection. The output from this process is functional 3D models (.C3D files).

  The output file is also saved without direct linking with the patient/participant's information entered in previous sessions and the view of the file is hindered by the 2D view.

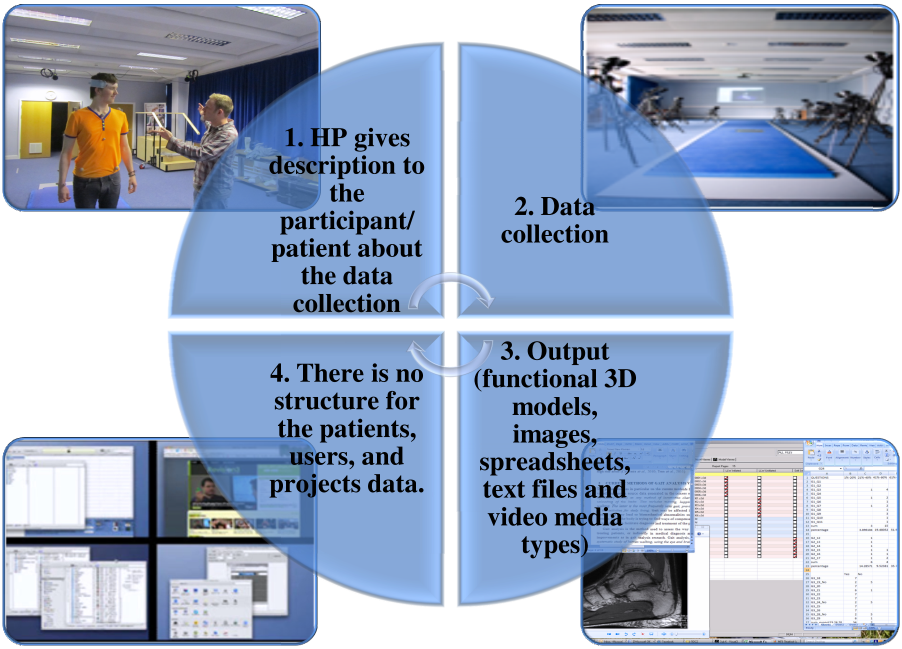


Fig. 1. Shows the Current way of Gait Analysis Process

- Data collected in track manager is exported in C3D format. In turn is imported into Visual 3D (C-Motion – version 4.75) for performing the analysis, and the output is also in C3D format (.C3D files).

  Evidently more output files are added to the previous ones without management.

- A patient/participant’s information is saved on a document in addition to the output files from the software used (Scattered Files: Text files, Spreadsheets, Images, Videos, Functional 3D models). Each health professional has his own way of saving his/her patients’ information. There is no search or filtering criteria to view data in single interface.

3.2 Proposed Gait Analysis Management and Diagnosis System

- Health professionals typically contact participants or patients for data collection through email, phone or direct contact. Due to time constrains and the volume of cases that managed and diagnosed in daily basis, health professionals do not have sufficient time to provide an analytical description for all the patients/participants about the process of data collection. Travelling and location information are also omitted due to same time limitations. As such the patient care is suffering. Notably if the participants/ patients have an internet access the required information could be provided on the website (i.e. directions, maps and parking, what will happen during the visit, multimedia demonstrations of techniques (to reassure patients), looking for information (pages which provides information on 'who we are', 'what we do', 'how we are funded', and summaries of current and active research).
  This additional non-medical information aim to save time and provide detail information and answers for the patients prior to their arrival.
- On the patient/ participant’s arrival if it is the first visit, the patient/ participant’s information will be added only once, and in each new visit the visit details will be added to the existing information.
  The user logs in to the account to add new patient (for first visit) or adds new visit (for current patient). There is static information entered once and dynamic ones are entered on each visit to the EMR.
- The patients and projects data is added to each user’s (health professional) account. The same process is repeated by each user. Search and query are provided to save time looking for a particular piece of information. For statistical analysis or producing reports filtering criteria is provided on the database and the results are shown according to the chosen view (i.e. view all/each patients’ results/files).
- Viewing medical data is not hindered by the 2D limitation as it provides a VE Real-Time 3D visualisation of patients gait analysis (figure 2) with the use of the proposed multimedia application in the Virtual Reality and Simulation Laboratory (VRS Lab) at GCU. The laboratory facilities offer state of the art dedicated space for 3D stereo visualisation, 3D surround audio and spatial gesture recognition experimentation, enabling the research team and other groups to experiment, develop, test and demonstrate novel systems and applications. Although the laboratory offers bespoke and costly equipment in this particular case we opted for off-the-shelf equipment in order to replicate reliably and in cost efficient manner the hardware required for management and VR visualisation musculoskeletal pathologies and gait analysis.

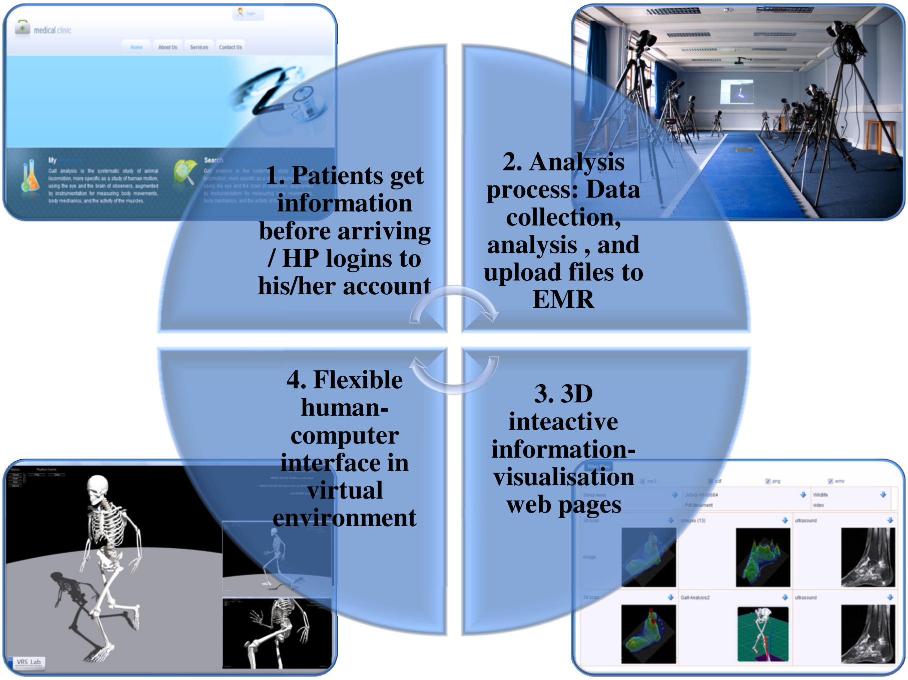


Fig. 2. Proposed Gait Analysis Management and Diagnosis in a Prototype Virtual Reality Environment

4 System Evaluation

A questionnaire was distributed to in-house health professionals for system evaluation after presenting a Real-Time 3D visualisation of patients gait analysis in the University's Virtual Reality and Simulation Laboratory (VRS Lab) to measure the HP's acceptance of the proposed work. The questionnaire was measuring multiple aspects, applicability of VR, users' satisfaction of the proposed system, data storage (EMR), interactive medical data, motion capture in 3D, wireless manipulation of data, enhancing the process of decision making and quality management, and the management of the system users.

Health professionals were first asked to supply background information, such as name, academic discipline, and academic level. They were then asked to rate their satisfaction by answering the three groups of questions as follows:

– The first group of questions measures the users' satisfaction about the proposed system and its ability to combine all the related existing software currently used for gait analysis and diagnosis under one, user- friendly package. Participants were asked to rate their percentage satisfaction. 76.6% of the participants state that the system under development satisfies 60-100% of their requirements.

- The second group of questions measures participants' satisfaction relating to the 3D representation of the derived data in an interactive VR environment; 76.2% of the participants state that the application is promising and satisfies most (60-100%) of their expected requirements.
- The third and final group of questions evaluates the participants experience about the system's navigation features and functionality, design and colour scheme, and organisation and layout of the screens; 96.4 % of the participants found that the navigation system was intuitive and helpful, 71.4 % of the participants were completely satisfied with the application's HCI, and 85.7% of the users estimated that the system would improve work flow by reducing time spent on information retrieval and archiving by a factor of 50%.

The analysis of the results derived by the survey was encouraging. Feedback shows that users are accepting the system and found that the VR environment and the HCI less distracting in contrast to viewing data in 2D and using different software packages and interfaces to look for a piece of information. In addition to the potential benefits desired from applying this system regarding reduced time spent on saving, searching, and extracting data, quality control reflected upon medical data, data analysis efficiency, and improved data visualisation.

5 Concluding Remarks

This paper describes the work carried out which examines the database and the visualisation of medical data by designing a system to manage patients' data storage in a medical record and presenting medical data in a single interface in 3D in a VR environment which simplifies the process of storing, searching and enhancing visualising gait analysis data. Apparently, from the system evaluation and participants' experience, the system managed to some degree to gain participants' satisfaction and meet most of their requirements and expectations, by providing an easy and simple way of interaction with simple tools to extract medical data, and a natural way of delivering the virtual environment and immersing the user to the virtual environment. In summary, the medical practitioners who experienced the proposed system found the subject matter exciting and interesting in both presentation methods (i.e. traditional and VR), but felt that the traditional method was more tiresome than the VR visualisation and data description.

Users' feedback raised more requirements which are in the future plans for the work; to enhance interface functionalities and provide users with customisable information. As the system is web-based, additional groups of health professionals in different areas may be added to the system which makes the data sharing and cooperation easier.

References

1. Whittle, M.: Gait Analysis an Introduction, 4th edn. Elsevier, Philadelphia (2007) ISBN: 978-0-7506-8883-3
2. Sakellariou, S., Charissis, V., Grant, S., Turner, J., Kelly, D., Christomanos, C.: Virtual Reality as Knowledge Enhancement Tool for Musculoskeletal Pathology. In: Shumaker, R. (ed.) Virtual and Mixed Reality, Part II, HCII 2011. LNCS, vol. 6774, pp. 54–63. Springer, Heidelberg (2011)
3. Wu, P., Xie, K., Yu, H., Zheng, Y., Wu, C.: A 3D Clipping Simulation Based on Virtual Surgery. In: Jin, D., Lin, S. (eds.) Advances in CSIE, Vol. 2. AISC, vol. 169, pp. 387–392. Springer, Heidelberg (2012)
4. Wang, S.J., Middleton, B., Prosser, L.A., Bardon, C.G., Spurr, C.D., Carchidi, P.J., Kittler, A.F., Goldszer, R.C., Fairchild, D.G., Sussman, A.J., Kuperman, G.J., Bates, D.W.: A Cost-Benefit Analysis of Electronic Medical Records in Primary Care. Am. J. Med. 114(5), 397–403 (2003)
5. Moschos, G., Nikolaidis, N., Pitas, I., Lyroudia, K.: A Virtual Anatomical 3D Head, Oral Cavity and Teeth Model for Dental and Medical Applications. In: Czachórski, T., Kozielski, S., Stańczyk, U. (eds.) Man-Machine Interactions 2. AISC, vol. 103, pp. 197–206. Springer, Heidelberg (2011)
6. Charissis, V., Ward, B.M., Naef, M., Rowley, D., Brady, L., Anderson, P.: An Enquiry into VR Interface Design for Medical Training: VR Augmented Anatomy Tutorials for Breast Cancer. In: McDowall, I.E., Dolinsky, M. (eds.) Proceedings of the: International Annual Symposium of IS&T/SPIE, The Engineering Reality of Virtual Reality, San Jose, California, vol. 6804, pp. 19–28 (2008) ISBN: 9780819469762
7. Kyriazis, V.: Gait analysis techniques. J. Orthopaed Traumatol 1, 1–6 (2001)
8. Gage, J.R., DeLuca, P.A., Renshaw, T.S.: Gait analysis: principles and applications: emphasis on its use in cerebral palsy. J. Bone Joint Surg (Am) 77, 1607–1623 (1995)
9. Katehakis, D.G., Kostomanolakis, S., Tsiknakis, M., Orphanoudakis, S.C.: Image Management in an Integrated Electronic Health Record Environment. In: Proc. 20th Int. EuroPACS Conf. (EuroPACS 2002), Oulou, Finland, pp. 87–92 (2002)
10. Katehakis, D.G., Lelis, P., Karabela, E., Tsiknakis, M., Orphanoudakis, S.C.: An Environment for the Creation of an Integrated Electronic Health Record in HYGEIAnet, the Regional Health Telematics Network of Crete. In: TEPR 2000, May 9-11, vol. 1, pp. 89–98. Your Connection to Electronic Healthcare, San Francisco (2000)
11. Katehakis, D.G., Kostomanolakis, S., Tsiknakis, M., Orphanoudakis, S.C.: An Open, Component-based Information Infrastructure to Support Integrated Regional Healthcare Networks. In: MEDINFO 2001, September 2-5, pp. 18–22. Tenth World Congress on Health and Medical In-formatics, London (2001)
12. Deutsch, J.E., Lewis, J.A., Burdea, G.: Technical and patient performance using a virtual reality-integrated telerehabilitation system: Preliminary finding. IEEE Transactions on Neural Systems and Rehabilitation Engineering 15, 30–35 (2007)
13. Halton, J.: Virtual rehabilitation with video games: A new frontier for occupational therapy. Occup. Ther. Now. 9(6), 12–14 (2008)
14. Deutsch, J.E., Merians, A.S., Adamovich, S., Poizner, H., Burdea, G.C.: Development and Application of Virtual Reality Technology to Improve Hand Use and Gait of Individuals Post-Stroke. Restor. Neurol. Neurosci. 22, 371–386 (2004)

15. Sonar, A., Kuxhaus, L., Carroll, J.: Simulation of Subject Specific Bone Remodelling and Virtual Reality Visualization. In: Kim, J.-J. (ed.) Virtual Reality, InTech (2010), http://www.intechopen.com/books/virtual-reality/simulation-of-subject-specific-bone-remodeling-and-virtual-reality-visualization ISBN: 978-953-307-518-1, doi:10.5772/13181
16. Seng, W.C., Seyed, H.M.: A Simple Approach for Education Using Virtual Human Dataset. International Arab Journal of e-Technology 2(1), 29–35 (2011)

Theory-Guided Virtual Reality Psychotherapies: Going beyond CBT-Based Approaches

Sheryl Brahnam

Department of Computer Information Systems,
Missouri State University,
901 S. National, Springfield, MO 65804, USA
sbrahnam@missouristate.edu

Abstract. Most VR applications in mental health care have focused on cognitive behavioral therapy. This paper is a call to expand research into other theory-guided psychotherapy practices. Evidence is presented that supports the so-called *dodo bird effect* that contends that all bona fide psychotherapies are equally effective. Two avenues for expanding research are suggested that focus on VR strengths: creating VR *playspaces* (virtual environments where therapist and client can engage playfully) and VR drama therapy.

Keywords: virtual reality, drama therapy, creative expression therapy, playspace, psychotherapy, dodo bird effect.

1 Introduction

VR as a tool for psychotherapy has come a long way since the first literature review of a few case studies in 1998 [1]. There is little doubt now that VR has made tremendous inroads into the treatment, and even prevention, of many psychological disorders. Much of this work has been based on best practices in cognitive behavioral therapy (CBT) transposed to VR. CBT describes a group of psychological treatments that are evidence based and that work on the premise that changes in thinking produce emotional and behavioral changes. CBT focuses on teaching clients specific skills to handle specific symptoms and are, as a result, usually short-term.

VR psychotherapy applications arose in the treatment of phobias and post-traumatic stress disorder (PTSD), showing great promise as a substitute for imaginative, as well as actual, exposure to fearful stimuli [2]. Psychosocial interventions, particularly exposure therapy (ET), are generally considered the empirically-supported treatments of choice [3]. Since VR uses multiple sensory modalities, providing "a perceptual illusion of non-mediation" [4] with fully adjustable parameters, VR implementations of ET were a logical first step, and VR exposure therapy (VRET) [5] has proven successful in treating a variety of phobias, acute stress disorder, and PTSD [6]. Using similar techniques, studies have shown the effectiveness of VR cue exposure therapy (VRCUE) in treating substance abuse [7]

R. Shumaker (Ed.): VAMR/HCII 2013, Part II, LNCS 8022, pp. 12–21, 2013.

and of VR food exposure therapy in treating eating disorders [8, 9]. In addition to exposing patients to stimuli, VR applications have also been developed that distract patients from unwanted stimuli, as in the case of pain management [10]. VR has also proven an effective tool for stress reduction and inoculation training [2]. For a comprehensive review of VR studies from a CBT perspective (VRCBT), see [11]; they report over 140 articles on VR employments of CBT techniques. In general, the consensus is that VRCBT provides effective tools for managing pain and for treating a variety of psychological disorders.

Given these successes it makes sense that VR research in psychotherapy has almost exclusively focused on CBT treatments and continues to do so [6, 11, 12]. Over the last 20 years, evidence-based practice (EBP), such as CBT, and empirically supported treatments (ESTs) have gained prominence in psychotherapy [13], adding yet more incentive to transpose ESTs to VR environments. It comes as no surprise then that little research has explored VR as a tool or practice modality for theory-guided and insight-based practices in psychotherapy, such as psychoanalysis, psychodynamic psychotherapy, analytical or depth (Jungian) psychotherapy, humanistic, existential, and relational psychotherapy, to name but a few. In fact, a recent search of the literature came up empty on VR studies involving the theory-guided psychotherapy practices listed above.

The intention of this paper is first to provide arguments justifying research into VR as a tool for theory-guided psychotherapies (VR-TGP). I do this in section 2 by reviewing some of the literature showing a preponderance of evidence that no therapeutic modality or bona fide psychotherapeutic treatment is better than any other. The second intention is to offer some ideas for expanding VR research in psychotherapy to include other orientations. In section 3 I briefly look at the possibility of creating VR playspaces and VR applications for drama therapy. I conclude in section 3 by calling for the development of VR systems that can be used by clinicians to develop new therapies that arise from the unique characteristics and strengths offered by VR.

2 Empirical Grounds for Exploring Theory-Guided Psychotherapies

Despite the recent emphasis on EBP, a number of meta-analyses comparing treatments for different psychological disorders show that no bona fide technique is more effective than another, thus confirming Rosenzweig's 1936 *dodo bird* conjecture [14]; empirically [15], to quote Dodo in *Alice and Wonderland*, "Everyone has won and all must have prizes." For example, both a meta-analysis of all studies published between 1980 and 2006 comparing treatments for children suffering from anxiety, depression, conduct disorder, and attention deficit hyperactivity disorder [16] and a meta-analysis focused on all studies published between 1989 and 2006 comparing treatments for PTSD [17] showed no differences in outcome between treatments. In the meta-analysis of PTSD studies [17], a broad range of treatments were represented: hypnotherapy, stress inoculation, exposure therapy, CBT, prolonged exposure, imaginal exposure, and eye movement desensitization and reprocessing. As Elkins [18] notes, "after 40 years of specificity research and millions of research dollars, there is still no scientific basis for privileging one modality and set of techniques over other modalities and techniques. Instead, scientific findings confirm that all *bona fide* psychotherapies are robustly effective, and equally so" (p. 451).

Particularly noteworthy are the results of one of the best experimentally designed studies that focused on CBT for depression. When Jacobson et al. [19] systematically removed the critical cognitive components in treatment, they found no reduction in outcomes. A more recent CBT study for PTSD also showed no attenuation of outcome when critical cognitive components were removed [20].

What research does show unequivocally is that psychotherapy is effective,[1] but how and why it works are still unanswered questions. Many researchers, however, are suggesting that certain common factors are the primary determinants of effectiveness, with the personal and interpersonal, i.e., the characteristics of the client, therapist, their relationship and alliance, accounting for most outcome variance. Andrews [21] has argued that therapist qualities, such as warmth, empathy, and respect may account for as much as 80% of the variance (for a discussion of research findings, see [18, 22]).

Frank and Frank [23] were among the first researchers to identify common factors in psychotherapy and to offer a theoretical framework (that also accounts for the dodo bird effect) based on psychotherapy as a cultural practice. They identified the following four common factors in psychotherapy: 1) an emotionally charged and confiding relationship with a helping person; 2) a healing setting that engages the client's expectations that the helper will be of assistance; 3) a rationale, conceptual scheme, or myth (it need not be true, but both client and healer must believe it) that explains the client's problems and how the client can heal; and 4) a procedure or ritual that requires the active participation of both client and therapist and that follows the rationale underlying the therapy. More recently, Tracey [24], using multidimensional scaling and cluster analysis, conceptualized common factors into three clusters: bond, information, and role. As Wampold [25] has shown, these models, as well as other explanations of psychotherapy as a cultural healing practice are consistent with research both in psychology and in medicine.

Despite the evidence suggesting that specific theoretical orientations and techniques may have little to do with outcome variance, both are essential components in Frank and Frank's model: theories provide #3, the rationale behind the therapy, and techniques are included in #4, the ritual needed for healing. According to Wampold [26], theory plays a critical, guiding role in psychotherapy. He writes, "These components are vacuous without theory–simply, there is no therapy without theory. Every client wants an explanation for what ails him or her and a set of therapeutic actions that the client believes will improve his or her condition. The last two components–the rationale and the treatment–emanate necessarily from theory" (p. 43).

---

[1] Wampold states that psychotherapy has an NNT (number needed to treat) of 3. The ideal NNT is 1, where everyone improves with treatment and no one improves in the control group. As a comparison, aspirin as a prophylaxis for heart attacks has an NNT of 129.

There are, of course, well-conducted trials that cast doubts on the dodo bird effect. Dimidjian et al. [27], for example, found that behavioral activation (BA) was superior to CBT for severely depressed patients. However, a later study, performed by the same group, that explored the enduring effects of CBT, BA, and antidepressant medication found no statistically significant differences between exposure to CBT and BA in a second year follow-up [28].

What are some of the implications of this research to VR research in psychotherapy? First, we should not be surprised to find that VR mental health applications fare no better than traditional therapies (for a survey of results, see, for instance, [29]). Second, there are no empirically valid grounds to restrict research in VR mental health care to CBT. The scientific evidence in support of the dodo bird effect justifies opening the door of VR research in mental health care to other established theory-guided practices. Third, even if the dodo bird conjecture continues to prove true, this does not obviate the need to focus much VR research in mental health care on EBP. Empirical studies supporting the efficacy of many established theory-guided practices have been published and should be utilized to establish which practices and techniques are indeed bona fide and worthy of exploration and which are not. But a word of caution here: some well-established practices, as noted in the next section, are based on the expressive arts (play therapy, drama therapy, creative arts therapy, etc.), and there is an ongoing debate about what constitutes appropriate evidence in creative arts therapy. Be that as it may, there is, nonetheless, a growing body of evidence supporting the effectiveness of creative arts therapy (see [30]).

3 VR-TGP (Playing with Possibilities)

Creative arts expression and play therapy have long been established practices in the offices of clinicians working from many orientations. Almost every major theorist has privileged creativity, the imagination, and play and has advocated methods for fostering these qualities. Freud [31] described play as the child's mechanism for working out trauma, and Jung asserted that "Not the artist alone, but every creative individual whatsoever owes all that is greatest in his life to fantasy" [32, p. 93]. Jung developed a therapeutic method that actively engages fantasy called *active imagination*, which leads people to a deeper understanding of themselves and furthers the process of individuation. Winnicott, a leading object relations theorist, has written "that it is in playing and only in playing that the individual child or adult is able to be creative and to use the whole personality, and it is only in being creative that the individual discovers the self" [33, p. 54].

That people are using computer games and VR worlds today to satisfy their needs to play and to be creative has been provocatively argued by McGonigal [34]. Therapists are taking note of this trend and are increasingly adding computer games and the Internet in their work with clients [35], especially with children and adolescents. In this section I suggest two areas where play and creative expression can be enhanced by VR for therapeutic purposes: VR healing environments, or *playspaces* [36], and VR drama therapy.

3.1 VR Playspaces

The idea of a therapeutic playspace comes from drama therapy and is defined by Johnson as “a state of playfulness that exists between the client and therapist” [37, p. 170]. Future research could explore creating virtual environments that foster play between client and therapist (thereby possibly strengthening the alliance) by providing virtual creative studios, with a host of expressive tools, and a virtual museum space to collect creative expressions (paintings, songs, poems) and important relics (photographs and videos) of the client for review and sharing with the therapist. All objects in the virtual environment could be used as containers to hold difficult emotions and traumatic memories, as well as creative expressions and emotion word clouds. Virtual spaces could thus facilitate in new ways such games for adolescents and children as “Feelings Hide-and-Seek" [38], a therapeutic version of the childhood game hide-and-seek, where feelings, written on cards, are hidden and then through the process of seeking out are retrieved and discussed.

Another possibility for play activity in a VR playspace would be applications inspired by one of the first and most widely used play therapy techniques, sandplay. As it is known today, sandplay was developed by Kalff [39] out of Lowenfeld’s “World Technique” [40][2], Eastern ideas, Neumann’s stage theory of ego development [41], and Jung’s concept of individualization [42]. Kalff’s method of sandplay therapy is divided into two stages. In the first stage, the client makes a picture in a sandtray by grouping objects and sculpturing the sand. Clients are typically instructed to look at the available materials and pick out objects that speak to them or to close their eyes and visualize a world, which they are then asked to recreate in the sand. In the second stage the client tells a story about the picture. For Kalff, a Jungian analyst, sandplay encourages communication and collaboration between conscious awareness and unconscious fantasies and images, and “serves as a bridge between inner and outer worlds” [43, p. 304].

Although I am unaware of any attempt to make a computer version of sandplay, at least one counselor, Skigen, has exploited the popular computer game Sims in a way she claims resembles sandplay therapy [44]. In her practice with children, “Simsplay,” like sandplay, takes place in her office in a protected and tolerant setting and provides an imaginative activity that is a “mirror for the individuation process” [45, p. 17]. Following Skigen’s lead, perhaps other games could be developed with therapists for VR extensions and modifications of sandplay.

Some advantages of VR sandplay would address concerns discussed by Dale and Lyddon [46], who apply a constructivist approach to sandplay. They stress the importance of the therapist’s role in recording and dismantling the sandworld. Most therapists photograph or sketch the sandworld and then videotape or audio record the stage two discussions. One advantage of sandplay games in virtual environments

[2] The first sandplay therapy technique, developed in 1929.

would be the fact that sandplay sessions could ordinarily and unobtrusively be recorded and played back when needed. Dale and Lyddon also note how taking the sandworld apart before the client leaves the office is known to be difficult for some clients. They spend some time discussing various ways of overcoming this problem. Unlike the therapists' offices of today, where the creative expressions of clients often have to be put away after each session, VR playspaces could be customized for each client and resumed for each session.

3.2 VR Drama Therapy

There are many definitions of drama therapy. Most definitions mention role-playing, dramatization, verbal and nonverbal communication, and reflection taking place within a dyad or group setting and employed for the purpose of psychological healing [37]. Although mostly originating in the ideas of Moreno, considered the father of Drama therapy, many important theorists have had a hand in the development of drama therapy, most notably, Jung, Murray, Perls, Kelly, Wolpe, Lazarus, and Winnicott (for an historical overview see [37]). Perhaps one of the most famous and commonly practiced drama therapy techniques is the empty chair dialogue intervention, where the client is asked to engage in dialogue with an imaginary person of significance sitting in a nearby empty chair [47].

It is very possible that drama therapy would work well in VR worlds. Incorporating drama and role-playing into VR games and simulations has been the focus of recent research (see [48] for a survey). The close relationship between VR and improvisational theater has been pointed out by many and explicated in terms of narrative by Aylett and Louchart [49]. They find that an important characteristic of VR is that the user "to whom a narrative is communicated is 'active' in the unfolding of the narrative as opposed to its 'passive' role in most other classical narrative media (i.e., spectator)" (p. 3). The authors also discuss the role of the "drama manager" and "game master" in ways that resonate with drama therapy and the notion of the therapist as guide, "a figure that . . . both witnesses and leads the client on the therapeutic journey" [37, p. 240].

Despite recent successes using dramatization and role-playing in games and simulation, few researchers have explored them in VR mental health applications. A notable exception is the work of Park and Ku [50], who used virtual role-playing for social skills training (SST) for patients with schizophrenia. They found that the SST-VR group showed increased communications skills compared to those using traditional SST (SST-TR). However, those using SST-VR showed less improvement in nonverbal skills compared to those using SST-TR. This latter outcome could be explained by the fact that the study used a virtual human in the role-playing sessions. As noted in section 2, considerable evidence suggests that interpersonal factors (items 1 and 2 in Frank and Frank's model) are important in therapeutic outcomes. New VR applications in mental health care that make use of role-playing might explore the benefits of bringing remotely located therapists into VR environments and work on representing as realistically as possible the nonverbal behaviors generated by the therapist in his or her conversational interactions with the patients.

Drama therapy offers a rich storehouse of ideas for new VR applications in mental health care: role rehearsal, role expansion, role reversal, narradrama, fixed-role therapy, the empty chair dialogue intervention, and playback theatre [37], to name a few. Take as one avenue of exploration, Kelly's fixed-role therapy [51], where the client is asked to take on a fictional role to play throughout the week. In follow-up sessions the client and therapist discuss the effectiveness of the new role. Fixed role therapy gives a clinician the means of understanding the client's unique frame of reference and patterns of meaning making, as well as a method for suggesting alternative patterns. VR renditions of fixed role therapy could take several forms. For example, clients could practice playing new roles with a therapist and group of virtual or real actors in a VE before moving the role out into the everyday world. Clients could also dramatically recreate significant moments they experienced the previous week as part of the discussion and working-through process.

Playback theatre (PT) is another interesting possibility for VR drama therapy. PT is a form of improvisational threatre where members of an audience tell their stories, one at a time, to a troupe of actors who then go about dramatizing the story [37]. A special member of the troupe, called "the conductor," interviews a selected member of the audience, called "the teller." Once the interview is finished, the conductor asks the teller to pick out actors to represent the characters in his or her story. The actors, along with a group of musicians, then improvise the story, with the conductor checking in with the teller to make sure that the actors have depicted the story accurately. If not, the actors are asked to replay these scenes. VR PT could unite disparate groups of actors in remote locations for the purpose of performing PT in a variety of virtual therapeutic settings.

4 Conclusion

Most VR applications in mental health care have focused on best practices in CBT. In this paper, I reviewed evidence in support of the dodo bird effect that contends that all bona fide psychotherapies are equally effective. Several implications for VR research in mental health care were drawn. In particular, it was noted that there are no empirically valid grounds to restrict research in VR mental health care to CBT. The scientific evidence in support of the dodo bird effect justifies opening the door of VR research to other established theory-guided practices.

Given the importance of common factors in psychotherapy effectiveness, another implication is that VR research may need to shift so that it focuses more on what does matter. Frank and Frank's common factor model [23] was presented as a way of understanding psychotherapy as a cultural practice. Frank and Frank identified the following four common factors in psychotherapy: 1) an emotionally charged and confiding relationship with a helping person; 2) a healing setting that engages the client's expectations that the helper will be of assistance; 3) a rationale, conceptual scheme, or myth that explains the client's problems and how the client can heal; and 4) a procedure or ritual that requires the active participation of both client and therapist.

I believe that all four factors in Frank and Frank's model can be augmented by VR. Two avenues for expanding research were suggested that focus on VR strengths: creating VR playspaces, i.e., virtual environments where therapist and client could engage playfully, and VR drama therapy, where roles could be tried out and discussed and personal stories could be dramatized. My focus in this discussion was on including real therapists and, in general, real actors (rather than virtual characters) (#1) in virtual healing settings (#2) that allowed for the virtual enactment of rituals (#4) based on established psychotherapy theories (#3). Although some of these ideas transposed current practices into VR worlds, ideally VR systems in the future, such as VR playspaces, will be developed that could be used creatively by clinicians to develop new therapies that arise from the unique characteristics offered by VR.

References

1. Wiederhold, B.K., Wiederhold, M.D.: A Review of Virtual Reality as a Psychotherapeutic Tool. Cyber Psychology & Behavior 1, 45–52 (1998)
2. Wiederhold, B.K., Wiederhold, M.D.: Virtual Reality for Posttraumatic Stress Disorder and Stress Innoculation Training. Journal of CyberTherapy & Rehabilitation 1, 23–35 (2008)
3. Wolitzky-Taylor, K.B., Horowitz, J.D., Powers, M.B., et al.: Psychological Approaches in the Treatment of Specific Phobias: A Meta-Analysis. Clinical Psychology Review 28, 1021–1037 (2008)
4. Lombard, M., Ditton, T.: At the Heart of It All: The Concept of Presence. Journal of Computer-Mediated Communication 3(2) (1997)
5. Meyerbröker, K., Emmelkamp, P.M.G.: Virtual Reality Exposure Therapy for Anxiety Disorders: The State of the Art. In: Brahnam, S., Jain, L.C. (eds.) Advanced Computational Intelligence Paradigms in Healthcare 6. SCI, vol. 337, pp. 47–62. Springer, Heidelberg (2011)
6. Safir, M.P., Wallach, H.S.: Current Trends and Future Directions for Virtual Reality Enhanced Psychotherapy. In: Brahnam, S., Jain, L.C. (eds.) Advanced Computational Intelligence Paradigms in Healthcare 6. SCI, vol. 337, pp. 31–45. Springer, Heidelberg (2011)
7. Lee, S.H., Han, D.H., Oh, S., et al.: Quantitative Electroencephalographic (Qeeg) Correlates of Craving During Virtual Reality Therapy in Alcohol-Dependent Patients. Pharmacology, Biochemistry and Behavior 91, 393–397 (2009)
8. Riva, G., Bacchetta, M., Cesa, G., et al.: Six-Month Follow-up of In-Patient Experiential Cognitive Therapy for Binge Eating Disorders. CyberPsychology & Behavior 6, 251–258 (2003)
9. Gorini, A., Griez, E., Petrova, A., et al.: Assesment of the Emotional Responses Produced by Exposure to Real Food, Virtual Food and Photographs of Food in Patients Affected by Eating Disorders. Annals of General Psychiatry 9(30), 1–10 (2010)
10. Hoffman, H.G., Doctor, J.N., Patterson, D.R., et al.: Virtual Reality as an Adjunctive Pain Control During Burn Wound Care in Adolescent Patients. Pain 85, 305–309 (2000)
11. Scozzari, S., Gamberini, L.: Virtual Reality as a Tool for Cognitive Behavioral Therapy: A Review. In: Brahnam, S., Jain, L.C. (eds.) Advanced Computational Intelligence Paradigms in Healthcare 6. SCI, vol. 337, pp. 63–108. Springer, Heidelberg (2011)

12. Riva, G.: Virtual Reality in Psychotherapy: Review. CyberPsychology & Behavior 8(3), 220–230 (2005)
13. American Psychological Association Presidential Task Force on Evidence-Based Practice Washington DC US: Evidence-Based Practice in Psychology. American Psychologist 61(4), 271–285 (2006)
14. Rosenzweig, S.: Some Implicit Common Factors in Diverse Methods of Psychotherapy. American Journal of Orthopsychiatry 6, 412–415 (1936)
15. Stiles, W.B., Barkham, M., Mellor-Clark, J., et al.: Effectiveness of Cognitive-Behavioural, Person-Centred, and Psychodynamic Therapies in UK Primary-Care Routine Practice: Replication in a Larger Sample. Psychological Medicine 38, 677–688 (2008)
16. Miller, S.D., Wampold, B.E., Varhely, K.: Direct Comparisons of Treatment Modalities for Youth Disorders: A Meta-Analysis. Psychotherapy Research 18, 5–14 (2008)
17. Benish, S., Imel, Z.E., Wampold, B.E.: The Relative Efficacy of Bona Fide Psychotherapies for Treating Post-Traumatic Stress Disorder: A Meta-Analysis of Direct Comparisons. Clinical Psychology Review 28, 746–759 (2007)
18. Elkins, D.N.: Toward a Common Focus in Psychotherapy Research. Psychotherapy 49, 450–454 (2012)
19. Jacobson, N.S., Dobson, K.S., Truax, P.A., et al.: A Component Analysis of Cognitive-Behavioral Treatment for Depression. Journal of Consulting and Clinical Psychology 64, 295–304 (1996)
20. Foa, E.B., Hembree, E.A., Cahill, S.P., et al.: Randomized Trial of Prolonged Exposure for Posttraumatic Stress Disorder with and without Cognitive Restructuring: Outcome at Academic and Community Clinics. Journal of Consulting and Clinical Psychology Review 73(5), 953–964 (2005)
21. Andrews, H.B.: The Myth of the Scientist-Practitioner: A Reply to R. King and N. King and Ollendick. Australian Psychologist (1998) 35, 60–63 (1998)
22. Wampold, B.E.: Psychotherapy: The Humanistic (and Effective) Treatment. American Psychologist 62(8), 855–873 (2007)
23. Frank, J.D., Frank, J.B.: Persuasion and Healing: A Comparative Study of Psychotherapy, 2nd edn. Johns Hopkins University, Baltimore (1991)
24. Tracey, T.J.G.: Concept Mapping of Therapeutic Common Factors. Psychotherapy Research 13(4), 401–413 (2003)
25. Wampold, B.E.: The Great Psychotherapy Debate: Model, Methods, and Findings, Erlbaum, Mahwah, NJ (2001)
26. Wampold, B.E.: The Basics of Psychotherapy: An Introduction to Theory and Practice. The American Psychological Association, Washington, DC (2010)
27. Dimidjian, S., Hollon, S.D., Dobson, K.S., et al.: Randomized Trial of Behavioral Activation, Cognitive Therapy, and Antidepressant Medication in the Acute Treatment of Adults with Major Depression. Journal of Consulting and Clinical Psychology 74, 658–670 (2006)
28. Dobson, K.S., Hollon, S.D., Dimidjian, S., et al.: Randomized Trial of Behavioral Activation, Cognitive Therapy, and Antidepressant Medication in the Prevention of Relapse and Recurrence in Major Depression. Journal of Consulting and Clinical Psychology 76(3), 468–477 (2008)
29. Gregg, L., Tarrier, N.: Virtual Reality in Mental Health: A Review of the Literature. Social Psychiatry and Psychiatric Epidemiology 42(5), 343–354 (2007)
30. Gilroy, A.: Art Therapy, Research and Evidence-Based Practice. SAGE Publications Ltd., London (2006)

31. Freud, S.: Beyond the Pleasure Principle. The International Psycho-Analytical Press, London (1922)
32. Jung, C.: Psychological Types. Princeton University Press, Princeton (1921)
33. Winnicott, D.W.: Playing and Reality. Routledge, London (1971)
34. McGonigal, J.: Reality Is Broken: Why Games Make Us Better and How They Can Change the World. The Penguin Press, New York (2011)
35. Rubin, D.B.: Multiple Imputation for Nonresponse in Surveys. J. Wiley & Sons, New York (1987)
36. Johnson, D.R., Smith, A., James, M.: Developmental Transformations in Group Therapy with the Elderly. In: Schaefer, C.E. (ed.) Play Therapy with Adults, pp. 78–108. John Wiley and Sons, Inc., Hoboken (2003)
37. Landy, R.J.: The Couch and the Stage: Integrating Words and Action in Psychotherapy. Jason Aronson, New York (2008)
38. Kenney-Noziska, S.: Techniques-Techniques-Techniques: Play-Based Activities for Children. Infinity Publishing, West Conshohocken (2008)
39. Kalff, D.M.: Sandplay: A Psychotherapeutic Approach to the Psyche. Sigo Press, Santa Monica (1980)
40. Mitchell, R.R., Friedman, H.S.: Sandplay: Past, Present and Future. Routledge, London (1994)
41. Neumann, E.: The Child. Harper Colophon, New York (1976)
42. Stewart, L.H.: Sandplay and Jungian Analysis. In: Analysis, J., Stein, M. (eds.) Stein, ed, pp. 372–390. Open Court, Chicago (1995)
43. Greenhalgh, P.: Emotional Growth and Learning. Routledge, London (1994)
44. Skigen, D.: Taking the Sand Tray High Tech: Using the Sims. In: Rubin, L.C. (ed.) Popular Culture in Counseling, Psychotherapy, and Play-Based Interventions, pp. 165–178. Springer Publishing Company, LLC, New York (2008)
45. Boik, B., Goodwin, E.: Sandplay Therapy: A Step-by-Step Manual for Psychotherapists of Diverse Orientations. Norton, New York (2000)
46. Dale, M.A., Lyddon, W.J.: Sandplay: A Constructivist Strategy for Assessment and Change. Journal of Constructivist Psychology 13(2), 135–154 (2000)
47. Daldrup, R.J., Beutler, L.E., Engle, D., et al.: Focused Expressive Psychotherapy: Freeing the Overcontrolled Patient. Guilford Press, New York (1988)
48. El-Nasr, M.S., Vasilakos, T., Robinson, J.: Process Drama in the Virtual World: A Survey. International Journal of Arts and Technology (IJART) 1(1), 13–33 (2008)
49. Aylett, R., Louchart, S.: Towards a Narrative Theory of Virtual Reality. Virtual Reality 7(1), 2–9 (2003)
50. Park, K.-M., Ku, J., Choi, S.-H., et al.: A Virtual Reality Application in Role-Plays of Social Skills Training for Schizophrenia: A Randomized, Controlled Trial. Psychiatry Research 18, 166–172 (2011)
51. Kelly, G.A.: The Psychology of Personal Constructs. Norton, New York (1955)

Development of the Home Arm Movement Stroke Training Environment for Rehabilitation (HAMSTER) and Evaluation by Clinicians

Elizabeth B. Brokaw and Bambi R. Brewer

Department of Rehabilitation Science and Technology,
University of Pittsburgh Pittsburgh, PA USA

Abstract. Stroke commonly results in severe impairment of upper extremity function, which limits independence in activities of daily living. Continued and frequent use of the affected limb can result in increased function. However, long term access to therapy is frequently limited, and home exercise compliance is low. The following paper presents the design and clinician evaluation of a Kinect based home therapy system called Home Arm Movement Stroke Training Environment (HAMSTER). The development, which focused on reducing commonly observed impairments after stroke, is discussed. Additionally the system was evaluated by twelve clinicians (occupational and physical therapists) with an average of 18 years of clinical experience with individuals with chronic stroke. The clinicians were asked about commonly prescribed home exercises, and for feedback about the HAMSTER system. Although only two of the clinicians had used the Kinect previously, the clinicians reported good usability and general satisfaction with the system. All of the clinicians felt that HAMSTER would be beneficial for individuals with chronic stroke.

1 Introduction

Stroke is the leading cause of long lasting adult disability [1]. Upper extremity function is commonly impaired after stroke and negatively affects quality of life. Therapy and intense use of the affected arm is effective for improving function [2]. However, transportation and cost are significant barriers to physical activity after stroke [3]. Nonuse leads to muscle deconditioning, which furthers the effects of stroke related impairment. In a US national survey of 312 individuals with chronic stroke only 31% reported that they exercise regularly and 27% reported rare or no exercise [4]. 86% of these individuals reported that they would be interested in an exercise program if one were available to them. However, compliance with home exercise programs is frequently low. In another study, lack of motivation was the largest reported reason for noncompliance with home exercise by individuals with chronic stroke related disability [5].

Home rehabilitation programs offer increased therapy hours at a potentially reduced cost and increased convenience to individuals with chronic stroke [6].

R. Shumaker (Ed.): VAMR/HCII 2013, Part II, LNCS 8022, pp. 22–31, 2013.

Game based systems are motivational and provide users with a sense of accomplishment that encourages continued use. A small study showed that the use of games increased therapy compliance in individuals with cerebral palsy that had previously abandoned therapy [7]. Additionally, computer games can offer accountability by recording system usage and potentially providing this data to clinicians. Several virtual reality systems have been used for therapy for individuals with stroke related disability. These systems typically involved wearable [8] or handheld sensors like the Wii [9]. Wii therapy studies have shown an increase in physical activity with a potentially more motivational intervention then general exercise [10]. However, wearable and handheld sensors can be difficult for individuals with commonly occurring hand impairments to hold or don independently. Additionally, these systems allow compensation strategies, such as flicking the wrist instead of swinging the whole arm in the Wii bowling game. These compensation methods will likely limit the functional recovery from the intervention [11] [12]. Furthermore, many systems use commercially available games, which do not sufficiently address the impairments of the user [10]. Stroke can result hemiparesis, spasticity, loss of normal joint coordination, and abnormal muscle synergies in the affected arm [13]. These conditions require thoughtful consideration during system design for this population. Conventional games may encourage abnormal or undesirable movement patterns in individuals with stroke. It is vital that the therapy be appropriately directed toward the user's unique impairments, encourage use of the affected arm without frustration, and avoid compensation strategies, which have been shown to limit recovery [11] [12].

The recent development of the Microsoft Kinect has opened up the field of virtual reality for rehabilitation. The Kinect is a low cost, \$100, vision based sensor that is used for skeleton tracking in gaming environments. The Kinect does not require sensors to be placed on the user, which allows for quick and easy setup. While the Kinect currently cannot obtain research grade motion tracking accuracy, the low cost and minimal setup time make the Kinect preferable in many situations. There have been a few relevant studies looking into the use of the Kinect for rehabilitation. One group examined the use of three commercially available Kinect games for children with degenerative ataxia and observed significant reductions in ataxia symptoms and improvements in balance [14]. Researchers have shown increased client motivation to do rehabilitation exercises with the Kinect and increased correctness of movement during training exercises [15].

The following paper presents the development of the Home Arm Movement Stroke Training Environment for Rehabilitation (HAMSTER) system for home use by individuals with chronic stroke. The system design focused on providing clinically relevant tasks with good general usability. This increases the likelihood that the system will improve function in individuals with stroke and potentially increase exercise compliance. Additionally, the system evaluation by clinicians discussed in this paper supports the potential success of the system.

2 System Development

Four sample therapeutic games were created for the Kinect. The games were designed using the Windows SDK v1.5, Microsoft XNA, and Farseer Physics Engine. All of the games were developed for use in the seated posture, due to potential safety concerns when the system is implemented in the home. The seated mode in the Windows SDK was not utilized since this would remove feedback about the subject's trunk motion. Instead standing mode was used and tracking of the lower limbs, which are difficult to distinguish from the chair, was simply ignored.

Simple graphics were developed for each game to minimize distraction and possible confusion. The visual interface for each of the games is shown in Figure 1 along with an image of a clinician user. The targeted arm movement varies widely between tasks to encompass a variety of potential exercises.

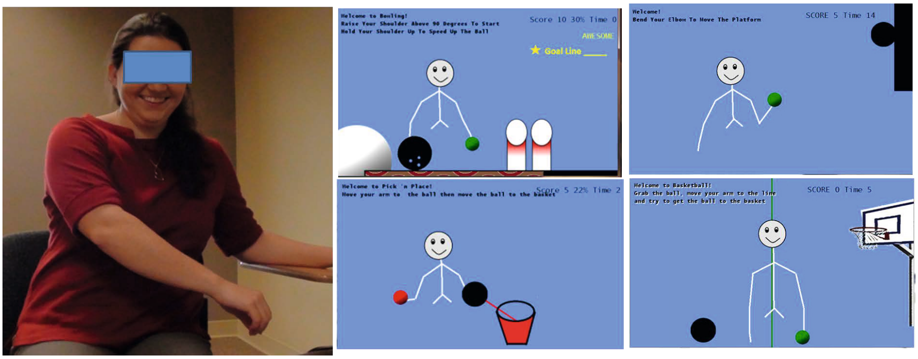


Fig. 1. On left a clinician plays the Pick and Place game. The Pick and Place graphics are shown in the lower adjacent image. The Bowling is shown above Pick and Place. The third column is Ping Pong, on top, and BasketBall below.

2.1 Bowling

The movement pattern of isolated shoulder movement with elbow extension was identified for the first task. The process of shoulder flexion with a straight elbow is often difficult after stroke due to the commonly occurring abnormal flexor synergy, which results in elbow flexion with shoulder movement. This game may help individuals break out of the abnormal flexor synergy pattern and potentially improve functional reach. In the Bowling game, the user is asked to hold their arm up to a certain height with their elbow extended. When the user achieves the goal position an impulse is applied to a bowling ball. The user receives points when the bowling ball knocks down the pins. Since elbow flexion is commonly elicited unintentionally after stroke, the game focuses on maintaining elbow extension with shoulder flexion. An error message appears and the ball will slow if the user's shoulder is above the goal line but their elbow is too bent or if the user's shoulder is horizontally abducted. If the ball has not started moving, no movement will be initiated with the elbow bent or shoulder horizontally abducted. If the user scores more than 6 points within 12 seconds the game automatically increases the difficulty by increasing the required shoulder flexion angle.

2.2 Pick and Place

The second task builds off the skills of Bowling for coordinated elbow and shoulder movement while crossing midline. This movement is more analogous to functional reach. During the Pick and Place game, the user is asked to retrieve the ball and move it into the basket. After a point is scored the ball reappears in a new target position with a different horizontal position on the screen. As the game progresses the user will have to also raise their arm to grab the balls, thus increasing the difficulty of the game. This change in difficulty is done between training sessions based on the previous session's score. The system displays an error message and does not allow for task completion if the user tries to use the unaffected hand to assist with the movement.

2.3 Ping Pong

The clinical focus of Ping Pong is isolated movement of the elbow. Control of the elbow is often limited after stroke due to spasticity. The velocity dependent aspect of this game could help increase the control of the joint at different speeds. Range of motion can also be increased as the subject progresses. It is recognized that the velocity dependent nature of this task may result in increased spasticity and that a delicate balance of speed and range of motion will need to be reached for each individual. During the Ping Pong game the user is asked to flex and extend their elbow to move a paddle to hit a ball that is moving across the screen. The ball gets progressively faster with successful hits and slows down when the ball is missed. The user gets a point for each successful ball hit.

2.4 Basketball

The clinical focus of this game is coordinated movement of the shoulder and elbow with an additional cognitive aspect of determining where the user needs to release the ball. To complete the task the individual will use a modified proprioceptive neuromuscular facilitation (PNF) movement pattern, which encourages maximal muscle activation. During the Basketball game the user is asked to retrieve the ball from across their body and then move it over the line in a position that will make the ball hit the basket. After the user successfully completes the task four times, the trajectory line changes to require more shoulder flexion.

3 System Testing with Clinicians

The protocol was approved by the University of Pittsburgh Institutional Review Board. Twelve clinicians were recruited for initial testing of the HAMSTER system via focused interviews. All subjects provided informed consent. During testing the users were always in sitting position at a distance of approximately four feet from the Kinect and computer. The screen was a 13 inch screen of a laptop. This was chosen for portability since, for the clinician's convenience, testing frequently occurred at the clinician's workplace.

Initially the clinicians completed a form asking general demographic information including number of years in clinical practice, type of home exercises usually prescribed, and perceived level of client compliance with the exercises. The procedure for each game was described prior to testing using a script guideline. The researcher's clinical goal for the game was not described. Clinicians played each game for two minutes. After one minute of the Pick and Place game, the difficulty level was increased as would be done automatically between therapy sessions. The user then played the game for an additional minute with targets requiring more shoulder flexion.

After each two minute game trial the subject was given time to jot down notes and give a general usability score from 1-4 where 1 was poor and 4 was excellent. Video recording was also used and clinicians were informed that they could also provide feedback orally. Following completion of the study, the clinician filled out a questionnaire. In the questionnaire the clinicians were asked about various aspects of the system, which covered usability, graphics, and how well HAMSTER would fulfill its role as a home therapy system.

4 Results

On average, the twelve recruited clinicians were 44 years old with a standard deviation of 15 years. The clinicians had an average of 18.8 years, with a standard deviation of 14.1 years, of clinical experience primarily with individuals with stroke. All of the convenience-sampled clinicians were female. The sample was evenly divided between occupational therapists and physical therapists. The clinicians were asked about the type of exercises that they prescribe for home exercises. Figure 2 shows the reported incident of prescribed home exercise themes, which primarily focused on functional activities and strengthening. On average the clinicians estimated that 50% of their patients are compliant with home exercises.

4.1 Game Feedback

Only two of the twelve clinicians had previously used the Kinect, and each of these clinicians had only used the Kinect once. The clinicians generally felt that the programs had good usability as shown in Figure 3. All of the clinicians felt that this system would be helpful for home exercise of individuals with

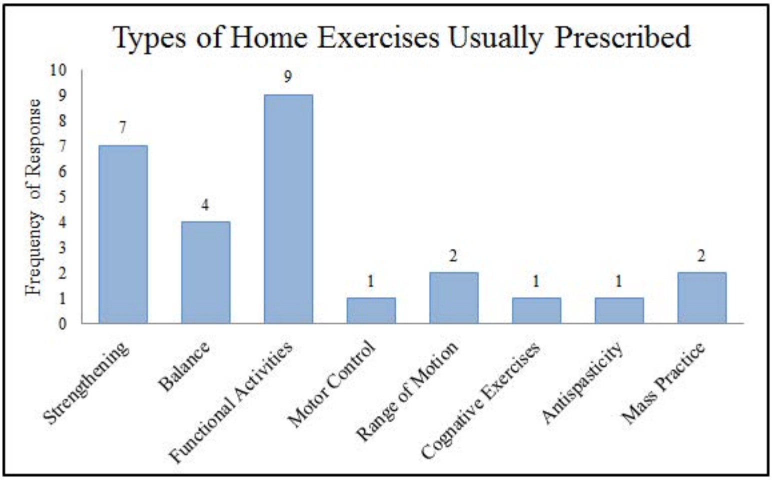


Fig. 2. Clinician prescribed home exercises by frequency reported. Upper extremity was not specified in the question. The functional activities category includes gait, and mass practice includes constraint induce movement therapy.

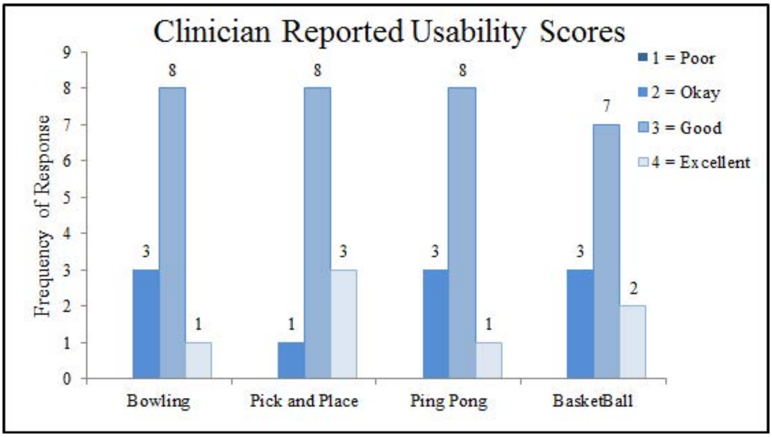


Fig. 3. Frequency chart of clinician reported usability showing the median score of good(3) for all of the tasks

stroke. Eleven of the clinicians found the physics based games (Ping Pong and Basketball) to be more entertaining than range of motion tasks like Bowling. The clinicians were divided about the need for the games to be functional task focused. Four desired games modeled after functional tasks, another four desired a mix of game and functional tasks, and the final four said game based. Two supporters of the game based design mentioned that the games should still have functional movement goals. The majority of clinicians (10/12) felt that task difficulty should be varied within each game instead of between sessions. Ten clinicians felt a mix of seated and standing exercises would be best. Several clinicians had a caveat to their answer, that this would only be true in situations where standing safety was not a concern. One clinician mentioned that special consideration of sitting balance should be made when choosing game reaching targets.

In Ping Pong two users were confused by the avatar's inability to interact with the ball. Another user found the variable timing aspect of the game, the ball speeds up with a hit and slows down with a miss, confusing. In general questioning, only two of the users reported that they found the avatar distracting or confusing. Having the avatar present, even when it isn't interacting with the virtual world, is helpful to show users what the Kinect sees, since occasionally the Kinect can have trouble with skeleton tracking.

4.2 Commonly Reported Limitations

Although the Kinect sensor has occasional limitations in tracking ability, the clinicians were generally very tolerant of the minor tracking errors that occurred during testing. The main reported concerns are shown in the Figure 4. The majority of these concerns are relatively self-explanatory. The readability was the main concern reported by the clinicians. A few individuals reported concerns about the graphics simplicity, including the "amputation" of the lower extremities of the avatar. The lower extremities were removed because the Kinect cannot accurately track their position while the user is seated. Clinicians also reported the desire to include tasks with wrist supination/pronation (6 users), grasp motion (4 users), and isolated shoulder horizontal abduction/adduction (3 users). Unfortunately, the Kinect is unable to detect supination/pronation of the wrist, nor can the Kinect cannot detect finger movements with the current skeleton tracking method, so additional sensors would be needed for these tasks.

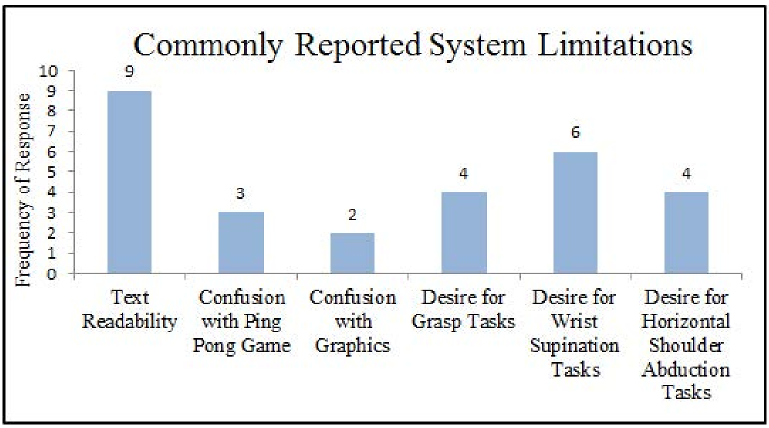


Fig. 4. Frequency of the reported limitations of the HAMSTER system

4.3 Compensation Strategies and Measurement of Kinematics

As stated previously, the use of compensation strategies limits the active use of the affected arm and thus reduces recovery from therapeutic interventions. When clinicians were asked about common compensation strategies that should be minimized during training, trunk movement was listed seven times, shoulder hiking three times, shoulder internal rotation two times, and shoulder abduction with elbow flexion two times. Unfortunately, the Kinect does not have the resolution to detect shoulder hiking. However, the Kinect is able to monitor trunk position, shoulder internal rotation, and shoulder abduction. The Bowling game already provides feedback to limit horizontal shoulder abduction and elbow flexion during the task.

It was originally thought that the kinematics recorded from the clinicians would serve as control data. However, many clinicians simulated movement patterns common to individuals with stroke during portions of the system testing. The Bowling game data is discussed here to show how the system will be able to track the movement of individuals with stroke during training. The researcher observed that the clinicians commonly tested trunk compensation patterns during Bowling. Although trunk deviation was higher in Bowling than in the other games, the subjects' average trunk angle was only 5.3 degrees from vertical with an average standard deviation of 4.7 degrees, showing that, in general, the individuals had relatively good posture when completing the task. The goal of the Bowling task is shoulder elevation with elbow extension. As a result, the maximum coordinated shoulder and elbow position that the user is able to obtain during Bowling is a measure of the user's ability to move out of the flexor synergy, which commonly impedes movement after stroke. The clinicians' average maximum coordinated position (highest shoulder angle with the elbow at less than forty degrees of flexion) was 17 degrees of elbow flexion and 138 degrees of shoulder flexion. Since the elbow error was set at 40 degrees of flexion, 17 degrees is reasonable, but shows that the clinicians did not maintain full elbow extension during the movement. This kind of quantitative data will be useful in tracking the therapy progress of individuals with chronic stroke.

5 Design Modifications and Future Work

Many simple changes to the HAMSTER system have been implemented based on the clinician feedback. The words on the screen have been reduced in number, and modified for size and clarity. Based on feedback from clinicians, HAMSTER now provides visual feedback to attempt to reduce user's trunk movement. This feedback does not limit task progression, but could for more advanced users. It is difficult to provide meaningful feedback about shoulder internal rotation since it may be used in normal movements. Feedback from testing with individuals with chronic stroke will determine how and if this kind of feedback is required for the HAMSTER games. In general the graphics were kept simplistic. The avatar has been modified to include a chair with the lower extremities to eliminate the confusion reported by one clinician. The Pick and Place task was modified to better represent a function task (feeding a fish) since integrating functional goals was encouraged by the clinicians. The Basketball game was renamed Darts, and the graphics modified to better describe the actual physics of the game and reduce confusion. An additional game was created focusing on shoulder abduction with a physics based design since clinicians found that game model more entertaining. The modified graphics for Fish, Darts, and Bird are shown in Figure 5. System usability and viability testing with individuals with chronic stroke related disability of the upper extremity is pending.

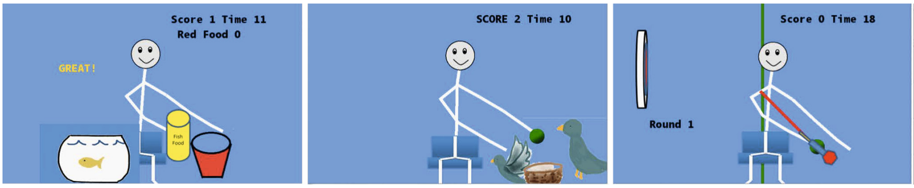


Fig. 5. The images show the modified graphics including the minimized and enlarged words, and the thicker lined skeleton with a chair. The far left image shows the modified Pick and Place task now named Fish. The middle image is the new Bird game. The user adducts/abducts their shoulder to move the mother bird to catch the baby bird. The third image is the modified Basketball game, which is now called Darts.

6 Conclusion

The results from these focused interviews have shown promise that HAMSTER will be usable and therapeutic for individuals with stroke. All of the clinicians felt that the system would be beneficial to individuals with chronic stroke. In general, the Kinect system was fairly intuitive, as noted by the good usability scores. This is especially impressive considering the users' very limited previous exposure to the Kinect. The kinematic data provided continuous and reasonable measures of arm and trunk movement. The changes discussed in future work will help improve the system effectiveness.

Acknowledgment. This work was supported in part by the Quality of Life Engineering Research Center (National Science Foundation EEC-0540865) and through the Advanced Rehabilitation Research Training Grant (National Institute on Disability and Rehabilitation Research 84.133P-1).

References

1. Roger, V.L., et al.: Heart disease and stroke statistics–2012 update: a report from the American Heart Association. Circulation 125(1), e2–e220 (2012)
2. Taub, E., Uswatte, G., Elbert, T.: New treatments in neurorehabilitation founded on basic research. Nature Reviews. Neuroscience 3(3), 228–236 (2002)

3. Zackowski, K.M., Dromerick, A.W., Sahrmann, S.A., Thach, W.T., Bastian, A.J.: How do strength, sensation, spasticity and joint individuation relate to the reaching deficits of people with chronic hemiparesis? Brain: a Journal of Neurology 127(Pt. 5), 1035–1046 (2004)
4. Shaughnessy, M., Resnick, B., Macko, R.: Testing a model of post stroke exercise behavior. Rehabilitation Nursing 31(1), 15–20 (2006)
5. Jurkiewicz, M.T., Marzolini, S., Oh, P.: Adherence to a home-based exercise program for individuals after stroke. Topics in Stroke Rehabilitation 18(3), 277–284 (2011)
6. Kairy, D., Lehoux, P., Vincent, C., Visintin, M.: A systematic review of clinical outcomes, clinical process, healthcare utilization and costs associated with telerehabilitation 31(6), 427–447 (2009)
7. Moyà Alcover, B., Jaume-i Capó, A., Varona, J., Martinez-Bueso, P., Mesejo Chiong, A.: Use of serious games for motivational balance rehabilitation of cerebral palsy patients. In: The Proceedings of the 13th International ACM SIGACCESS Conference on Computers and Accessibility, ASSETS 2011, p. 297 (2011)
8. Alankus, G., Lazar, A., May, M., Kelleher, C.: Towards customizable games for stroke rehabilitation. In: Proceedings of the 28th International Conference on Human Factors in Computing Systems, CHI 2010, pp. 2113–2122 (2010)
9. Deutsch, J.E., Brettler, A., Smith, C., Welsh, J., John, R., Guarrera-Bowlby, P., Kafri, M.: Nintendo wii sports and wii fit game analysis, validation, and application to stroke rehabilitation. Topics in Stroke Rehabilitation 18(6), 701–719 (2011)
10. Taylor, M.J.D., McCormick, D., Shawis, T., Impson, R., Griffin, M.: Activity-promoting gaming systems in exercise and rehabilitation. The Journal of Rehabilitation Research and Development 48(10), 1171 (2011)
11. Krakauer, J.W.: Motor learning: its relevance to stroke recovery and neurorehabilitation. Current Opinion in Neurology 19(1), 84–90 (2006)
12. Ellis, M.D., Holubar, B.G., Acosta, A.M., Beer, R.F., Dewald, J.P.: Modifiability of abnormal isometric elbow and shoulder joint torque coupling after stroke. Muscle Nerve. 32(2), 84–90 (2005)
13. Dobkin, B.H.: Strategies for stroke rehabilitation. The Lancet 3(9), 528–536 (2004)
14. Ilg, W., Schatton, C., Schicks, J., Giese, M.A., Schöls, L., Synofzik, M.: Video game-based coordinative training improves ataxia in children with degenerative ataxia. Neurology 79(20), 2056–2060 (2012)
15. Chang, Y.J., Chen, S.F., Huang, J.D.: A Kinect-based system for physical rehabilitation: a pilot study for young adults with motor disabilities. Research in Developmental Disabilities 32(6), 2566–2570 (2011)

A Low Cost Virtual Reality System for Rehabilitation of Upper Limb

Paweł Budziszewski

Central Institute for Labour Protection – National Research Institute,
ul. Czerniakowska 16, 00-701 Warszawa, Poland
pabud@ciop.pl

Abstract. The paper describes an on-going research aimed at creating the low cost virtual reality based system for physical rehabilitation of upper limb. The system is designed to assist in rehabilitation involving various kinds of limb movement, including precise hand movements and movement of the whole extremity. It can be used at patient's home as a telerehabilitation device. It was decided to use the system with a motion tracking (Razer Hydra) and two alternative display devices: head mounted displays (Sony HMZ-T1) and a LCD display with stereovision glasses (nVidia 3DVision). The custom software was developed to create the virtual reality environment and perform rehabilitation exercises. Three sample rehabilitation games were created to perform assessment of the rehabilitation system. In the preliminary research the usability of the system was assessed by one patient. He was able to use the system for rehabilitation exercises, however some problems with Sony HMZ-T1 usability were spotted. During the next stages of the research extended assessment of the system's usability and assessment of system's efficiency are planned.

Keywords: physical rehabilitation, virtual reality, serious games, home based rehabilitation, HMD, Razer Hydra, Sony HMZ-T1.

1 Introduction

The efficiency of physical rehabilitation is closely related to the intensity of rehabilitation exercises [1]. Various researches (e.g. [2, 3]), however, indicate that there is a large group of patients who fail to perform their exercises regularly or completely discontinue the rehabilitation process. It becomes especially noticeable when considering patients discharged from hospital. According to some estimates, this problem affects as many as 60-80% of such patients. 10% of such patients completely discontinue therapy. One of the main reasons for this fact is insufficient motivation of patients [4]. Motivation is indicated as a factor of high significance for the efficiency of rehabilitation [5]. The virtual reality (VR) based system seems to be a good solution for the physical therapy [6]. Patient's immersion in the virtual environment allows to shape the rehabilitation program in the form of a computer game (e.g. [7 - 9]). Attractive form combined with a competition (scoring, moving to a higher level of difficulty) results in focusing patient's attention on the performed task instead of upper limb

R. Shumaker (Ed.): VAMR/HCII 2013, Part II, LNCS 8022, pp. 32–39, 2013.

impairments. What is more, the virtual reality system can be successfully used for home-based rehabilitation (e.g. [10, 11]). With the internet connection the rehabilitation progress can be remotely monitored by a therapist.

The paper describes an on-going research aimed at creating the low cost virtual reality based system for physical rehabilitation of upper limb. Its goal is to create a system which would:

- assist in rehabilitation involving various kinds of limb movements, including precise hand movements and movement of the whole extremity,
- have an attractive form of a computer game using virtual reality for user interaction, which should result in increase of patients' motivation thus influencing the rehabilitation efficiency,
- be build using inexpensive, off-the-shelf components to assure the lowest possible price,
- be designed to work at patient's home as a telerehabilitation system providing contact with a therapist and enabling remote monitoring of the rehabilitation progress.

2 Methods

2.1 Overview

For the purpose of the physical rehabilitation it was decided to use the system with a motion tracking and two alternative display devices: head mounted displays (HMD) and a LCD display with stereovision glasses. The motion tracking system is used to record movements of patient's upper limb and to allow interaction with virtual environment. Both display devices are used to visualize the virtual environment and provide stereoscopic vision.

2.2 Hardware

The most important component of the rehabilitation system is the motion tracking system. It allows to record patient's movement during rehabilitation exercises. It is also the main instrument of user interaction. A lot of different devices can be used for this task, depending on desired price, accuracy and range of movement. On the one hand there are a lot of VR dedicated "professional" motion tracking devices, on the other there is a fast growing group of game controllers. Among different VR dedicated equipment one can find systems with a very good performance, tailored to the needs of rehabilitation [12]. However, this equipment is usually expensive. On the other hand, over the past few years a few interesting game controllers appeared on the market. These devices are powerful enough for many serious tasks, though sometimes with limited functionality and offering worse performance than the "professional" equipment. Considering low price and good availability, the game controllers seem ideal for the low-cost home-based rehabilitation system. One of the most interesting devices in this category is Microsoft Kinect controller. As a markerless optical tracking system it is relatively easy to use, providing tracking of the full body motion.

It can be successfully used in various aspects of the physical rehabilitation [13, 14]. Its accuracy is however too limited for rehabilitation exercises involving precise hand movements. At the same time, Kinect requires around 5 meters of empty room for proper tracking of the whole body. When considering home-based rehabilitation this is a potential problem. Another game controller often considered as a tracking device for the physical rehabilitation is Nintendo Wii Remote controller (Wiimote) [15, 16]. It is a wireless, inertial device with limited capabilities of optical system. It can provide only limited information on rotation and position of sensor, thus it is not directly suited to track spatial movements of upper limb.

In the described research it was decided to use a Razer Hydra game controller as a motion tracking device (Fig. 1). It is an inexpensive magnetic tracking system with central antenna and two corded pads. The system can track 6 DOF movement of each pad. The pads are also equipped with 7 buttons and a joystick. The controller is connected to a computer via the USB interface. Power is also provided by this interface, so no additional power supply is needed. The Hydra controller has some big advantages over other inexpensive tracking system. As a magnetic system it provides direct 6 DOF recording of sensor movements. This is not directly possible in case of inertial controllers, such as a Wiimote and is much more problematic in case of a Kinect controller. The Hydra controller also offers better accuracy than the Kinect and is free of problems with skeleton recognition when limbs are close to human body. It also requires less space for proper operation than the Kinect. There are also drawbacks of the Hydra controller. Its useful range is limited to around 1.2 m around the central antenna. However, if the central antenna is placed on the table top, it covers the whole range of upper limb movement of a person who is standing or sitting nearby. Another problem is that the controller is equipped with only two sensors and this number cannot be increased.

Fig. 1. The Razer Hydra game controller – central antenna and two game pads

For the rehabilitation purposes it is important to track position of one segment of upper limb in relation to another segment, depending on performed exercise. For these purposes two sensors of the controller are fixed to selected body segments using rubber straps. This allows to track the positions of upper limb segments in relation to one another, depending on the exercise performed.

As a visual interface two options have been chosen: HMD and computer or TV display. HMD provides stereoscopic view of the virtual environment, facilitates the ability to perceive the distance. This is particularly desirable when considering spatial rehabilitation movements [11, 17]. Combining HMD with motion tracking also enhances the level of immersion in the virtual environment, making the simulation even more attractive. Sony HMZ-T1 equipment has been chosen. After consultation with therapists it was decided to apply an alternative visual interface: display with stereoscopic glasses (nVidia 3DVision). This solution also facilitates stereoscopic vision, however immersion level is much lower. This solution is suitable for situations where wearing the HMD may be difficult, or when the budget designated for the system is lower.

2.3 Software

For the purpose of rehabilitation the custom software was created. This ensures that the system is well tailored to the needs. The software was created on the basis of open source libraries, minimizing the total cost of the system.

Following software was used:

- Ogre 3D graphic engine for visualization of the virtual environment
- Bullet Physics engine for physics simulation
- OpenAL providing sounds
- Python scripting language to set custom game parameters

The communication between therapist and patient Skype software is used.

The system provides a module for preparing the rehabilitation program, thus allowing the therapist to choose a rehabilitation game and set such parameters as the desired range of movement and exercise duration. All important data describing performed exercises and the rehabilitation progress are stored and may be accessed remotely through the internet.

3 The Study

The virtual reality system for rehabilitation of upper limb will be assessed by five therapists and three patients. For this purpose three sample rehabilitation games have been created. The goal of the first game is performing wide range movements of the whole upper limb. Patient's task is to reach small objects (apples, pears and other fruits) appearing in various places of the reach area. Their placement forces patient to perform specified exercises. The goal of the second game is to perform pronation/supination movements. Movement of patients hand is represented as movement of a fish (manta ray) swimming between various obstacles (Fig. 2). To avoid collisions with these obstacles patient must perform specified movements, in particular pronation/supination rotation. The goal of the third game is to perform precise movements of hand in various places of the reach area. In the game patient holds a key in the hand. Small boxes with a key hole appears in various points of the reach area. Patient's task is to insert the key in each key hole and perform rotation to unlock the box. All parameters of each game, such as the desired speed and placement of obstacles or other important objects, are predefined by a therapist.

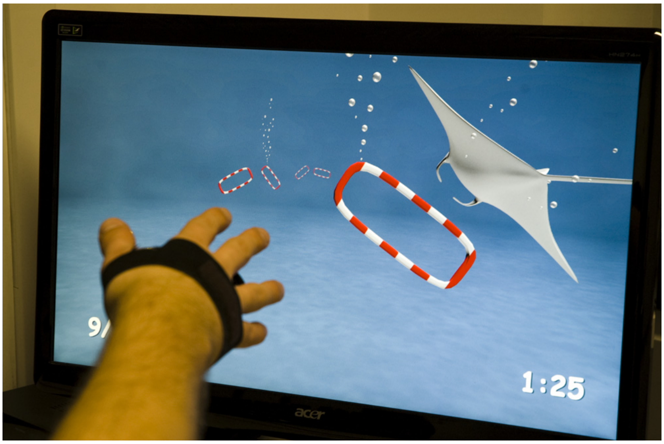


Fig. 2. Person playing the second game during system tests

These three games are suitable for rehabilitation for the following diseases:

- painful shoulder syndrome,
- radial fracture,
- olecranon fracture.

For all of these diseases at least two of the three games will be used.

4 Preliminary Results

In the preliminary phase of the research one patient (with painful shoulder syndrome) was asked to play all three games and assess system usability. Three variants of the system setup were used:

1. stereoscopic setup with HMD,
2. stereoscopic setup with LCD screen and 3DVision glasses,
3. LCD screen without stereoscopic features.

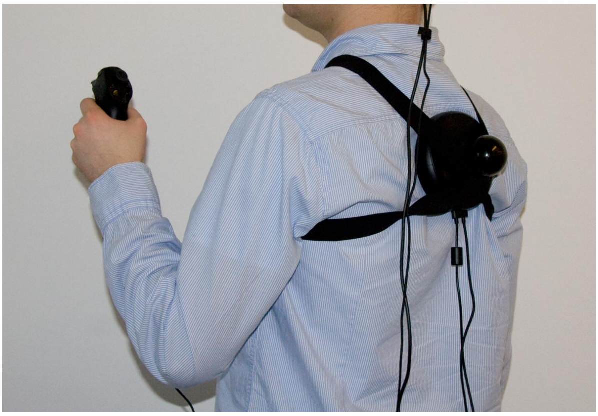


Fig. 3. The Razer Hydra setup prepared for the preliminary research

In all three cases Razer Hydra tracking system was used. As said before, this system is equipped with only two sensors. In the research central antenna was placed on patient's back using rubber strap (Fig. 3). This way it was possible to track movements of upper limb and HMD in the same time. One of sensors was held in patient's hand, second sensor was attached to the HMD (Fig 4). In the future, if HMD is not used second sensor may be attached to patient's arm for more precise results. Parameters of all three games were set by a therapist. The patient spent a few minutes playing each game in all three setups. Than he was asked to comment on usability.

Fig. 4. The Razer Hydra sensor attached to the HMZ-T1 HMD

Results of the preliminary research were as follows. The patient was able to play all three games in all system setups. However, he complained about the Sony HMZ-T1 HMD. He had problems with wearing it and with its correct adjustment. This task requires the use of both hands, which may be difficult or even impossible in case of patients with upper limb impairment. In this case assistance of another person may be needed. The patient also said, that the HMD was not comfortable to wear and he had problems with achieving stereovision. Both these problems may be caused by improper adjustment of the device. The third variant of the setup provides no stereoscopic vision, thus depth assessment is very limited. As a result, in some cases the patient had problems when trying to reach a desired point in space. In both stereoscopic setups these problems were less noticeable. The patient was asked to rank all three variants of the setup in terms of usability. He gave the highest score to the variant with no stereovision, the lowest to the variant with HMD. However, as he said, difference in scoring between both LCD-based variants is minimal.

5 Conclusion

Advanced virtual reality technology was used to create an upper limbs' rehabilitation system. Placing the rehabilitation process in the virtual environment and shaping it in the form of a computer game is supposed to have a positive effect on patients' motivation, thus influencing the rehabilitation efficiency.

In the preliminary phase of the research, system usability was assessed. As a result there was some controversy on using the HMD device, however it was decided to consider all three variants in the further research. During the next stage of the study, the system will be assessed by three patients. They will use the system for few days, checking it's usability and possibility to perform rehabilitation exercises. Results of this research will allow to prepare the system for the trial aimed to assess its efficiency, which is planned for the future.

Acknowledgements. This paper has been based on the results of a research task carried out within the scope of the second stage of the National Programme "Improvement of safety and working conditions" partly supported in 2011-2013 - within the scope of research and development — by the Ministry of Science and Higher Education/National Centre for Research and Development. The Central Institute for Labour Protection – National Research Institute is the Programme's main coordinator.

References

1. van der Lee, H., Snels, I.A., Beckerman, H., Lankhorst, G.J., Wagenaar, R.C., Bouter, L.M.: Exercise therapy for arm function in stroke patients: a systematic review of randomized controlled trials. Clinical Rehabilitation 15(1), 20–31 (2001)
2. Bassett, S.F.: The Assessment of Patient Adherence to Physiotherapy Rehabilitation. New Zealand Journal of Physiotherapy 31(2), 60–66 (2003)
3. Engström, L.O., Öberg, B.: Patient adherence in an individualized rehabilitation programme: A clinical follow-up. Scandinavian Journal of Public Health 33, 11–18 (2005)
4. Butterworth, S.W.: Influencing Patient Adherence to Treatment Guidelines. Journal of Managed Care Pharmacy 14(6)(suppl. S-b), S21–S25 (2008)
5. Maclean, N., Pound, P., Wolfe, C., Rudd, A.: Qualitative analysis of stroke patients' motivation for rehabilitation. British Medical Journal 321, 1051–1054 (2000)

6. Holden, M.K.: Virtual Environments for Motor Rehabilitation: Review. Cyber. Psychol. Behav. 8(3), 187–211 (2005)
7. Ma, M., McNeill, M., Charles, D., McDonough, S., Crosbie, J., Oliver, L., McGoldrick, C.: Adaptive Virtual Reality Games for Rehabilitation of Motor Disorders. In: Stephanidis, C. (ed.) UAHCI 2007 (Part II). LNCS, vol. 4555, pp. 681–690. Springer, Heidelberg (2007)
8. Flynn, S.M., Lange, B.S.: Games for rehabilitation: the voice of the players. In: Sharkey, P.M., Sánchez, J. (eds.) Proc. 8th Intl. Conf. on Disability, Virtual Reality and Assoc. Technologies, Augusct 31-September 2, pp. 185–194. Vina del Mar/Valparaíso, Chile (2010)
9. Burke, J.W., McNeill, M.D.J., Charles, D.K., Morrow, P.J., Crosbie, J.H., McDonough, S.M.: Designing engaging, playable games for rehabilitation. In: Sharkey, P.M., Sánchez, J. (eds.) Proc. 8th Intl. Conf. on Disability, Virtual Reality and Assoc. Technologies, August 31-September 2, pp. 195–201. Vina del Mar/Valparaíso, Chile (2010)
10. Broeren, J., Pareto, L., Johansson, B., Ljungberg, C., Rydmark, M.: Stroke rehabilitation using m-Health Care and 3D virtual environments – work in progress. In: Sharkey, P.M., Sánchez, J. (eds.) Proc. 8th Intl. Conf. on Disability, Virtual Reality and Assoc. Technologies, August 31-September 2, pp. 115–122. Vina del Mar/Valparaíso, Chile (2010)
11. Rizzo, A.A., Strickland, D., Bouchard, S.: The Challenge of Using Virtual Reality in Telerehabilitation. Telemedicine Journal and e-Health 10(2) (2004)
12. Zhou, H., Hu, H.: Human motion tracking for rehabilitation—A survey. Biomedical Signal Processing and Control 3(1), 1–18 (2008)
13. Chang, Y.-J., Chen, S.-F., Huang, J.-D.: A Kinect-based system for physical rehabilitation: A pilot study for young adults with motor disabilities. Research in Developmental Disabilities 32(6), 2566–2570 (2011)
14. Clark, R.A., Pua, Y.-H., Fortin, K., Ritchie, C., Webster, K.E., Denehy, L., Bryant, A.L.: Validity of the Microsoft Kinect for assessment of postural control. Gait & Posture 36(3), 372–377 (2012)
15. Standen, P.J., Brown, D.J., Battersby, S., Walker, M., Connell, L., Richardson, A.: Study to evaluate a low cost virtual reality system for home based rehabilitation of the upper limb following stroke. In: Sharkey, P.M., Sánchez, J. (eds.) Proc. 8th Intl. Conf. on Disability, Virtual Reality and Assoc. Technologies, August 31-September 2, pp. 139–146. Vina del Mar/Valparaíso, Chile (2010)
16. Shih, C.-H., Chang, M.-L., Shih, C.-T.: A limb action detector enabling people with multiple disabilities to control environmental stimulation through limb action with a Nintendo Wii Remote Controller. Research in Developmental Disabilities 31, 1047–1053 (2010)
17. Fidopiastis, C.M., Rizzo, A.A., Rolland, J.P.: User-centered virtual environment design for virtual rehabilitation. Journal of NeuroEngineering and Rehabilitation 2010 7, 11 (2010)

Super Pop VR^{TM}: An Adaptable Virtual Reality Game for Upper-Body Rehabilitation

Sergio García-Vergara[1], Yu-Ping Chen[2], and Ayanna M. Howard[1]

[1] Georgia Institute of Technology, School of Electrical and Computer Engineering, 85 5th Street NW, Atlanta, GA 30332
sergio.garcia@gatech.edu, ayanna.howard@ece.gatech.edu
[2] Georgia State University, Byrdine F. Lewis School of Nursing and Health Professions, Department of Physical Therapy, Atlanta, GA 30302
ypchen@gsu.edu

Abstract. Therapists and researchers have studied the importance of virtual reality (VR) environments in physical therapy interventions for people with different conditions such as stroke, Parkinson's disease, and cerebral palsy. Most of these VR systems do not integrate clinical assessment of outcome measures as an automated objective of the system. Moreover, these systems do not allow real-time adjustment of the system characteristics that is necessary to individualize the intervention. We discuss a new VR game designed to improve upper-arm motor function through repetitive arm exercises. An automated method is used to extract outcome measures of upper extremity movements using the Fugl-Meyer assessment methodology. The accuracy of the system was validated based on trials with eighteen adult subjects. With a corresponding average assessment error of less than 5%, the developed system shows to be a promising tool for therapists to use in individualizing the intervention for individuals with upper-body motor impairments.

Keywords: Cerebral Palsy, virtual reality gaming environment, Fugl-Meyer assessment, physical therapy and rehabilitation.

1 Introduction

Therapists and researchers have studied the use of virtual reality environments as part of corresponding physical therapy interventions for various patient demographics. Virtual reality (VR) refers to a computer technology that creates a three-dimensional (3D) virtual context and virtual objects that allow for interactions by the user [1]. Previous research has shown that VR environments present many benefits in the rehabilitation of individuals with motor skill disorders. Not only do they improve compliance for individuals working with their exercises [2], but they also enhance exercise effectiveness [3]. Based on this evidence, this research presents a new in-home VR game designed to improve patients' upper-body mobility through the repetition of movements associated with an individualized intervention protocol. The novelty of the system resides in the use of a 3D depth camera to capture and store

R. Shumaker (Ed.): VAMR/HCII 2013, Part II, LNCS 8022, pp. 40–49, 2013.

depth images while analyzing the subject's arm movements in real-time using the Fugl-Meyer assessment methodology [4],[11]. In addition, VR game parameters can be modified by the clinician in order to match the intervention protocol to the capabilities of the user.

This research focuses on rehabilitation activities for children with cerebral palsy. The term *cerebral palsy* (CP) describes a group of disorders of the development of movement and posture causing activity limitations that are attributed to non-progressive disturbances in the developing fetal or infant brain [5]. The Center for Disease Control and Prevention reports that CP is the most common motor disability in childhood. Often, people with CP have the inability to control some of their muscles resulting in poor movement coordination. Approximately half of the children with CP may sustain dysfunctions in upper extremity activities such as reaching, grasping, and/or manipulating objects [1]. For patients with CP, especially if detected at an early age, therapists recommend participation in physical therapy interventions in order to reduce further development of the effects of the disorder.

2 Background

Virtual reality systems have been developed to aid rehabilitation specialists in physical therapy treatments for children with cerebral palsy. Reid et. al [6] made a pilot study to show the benefits of a VR system for children with CP. Her studies suggest that a virtual environment allows for increased play engagement and the opportunity for children to practice control over their movements.

Golomb et. al [7] investigated whether an in-home remotely monitored VR videogame can help improve hand function and forearm bone health in an adolescent with hemiplegic CP. This pilot study showed that the system is prone to poor efficiency because of possible faulty Internet connections thus discarding the possibility of assessing patients remotely.

Commercially available gaming consoles have also been used in the rehabilitation process of individuals with CP in order to provide low-cost options. Deutsch et. al [8] used a Wii console to augment the rehabilitation of an adolescent with CP. The subject used a Wii controller to manipulate objects in the virtual environment. Although the participant in this study showed improvement both in performance and in learning, the research showed that the system is limited only to patients who are able to grasp the controller. Jannink et. al [9] used a motion capturing product for the PlayStation 2 platform called the *EyeToy*. Although this system presents a relatively low-cost in-home virtual environment, it lacked the ability to adjust the level of difficulty of the game to the child's capacity. To make this system more suitable for rehabilitation purposes, the option to select the desired game settings needs to be added and feedback about performance should be incorporated to increase motor learning.

Other motion capture systems require the use of additional equipment, like cyber gloves [10]. These types of systems implement the use of the IREX (Interactive Rehabilitation EX-ercise) platform. It tracks the motion of the gloves worn by the user and it maps the movements into the virtual environment. Although a novel system, it doesn't allow for a low-cost in-home rehabilitation tool.

3 Methodology

3.1 Objective

While there have been many VR systems developed for use as part of physical therapy interventions for children with CP, none incorporate a formal method of evaluating the user's upper-body motor skills in real-time and in the comfort of their own home. The goal of this research is to present a low-cost VR gaming system designed to overcome these limitations. An adaptable user interface allows the therapist or caregiver to select the desired game settings based on the user's capabilities. Users are engaged in repetitive movements during game play and are assessed in real-time by the system. Outcome measures are evaluated by quantizing one of the physical therapeutic metrics from the Fugl-Meyer assessment: range of motion (ROM).

In this paper, we focus on providing proof-of-concept evidence that: the VR system can accurately output the results of one of the metrics from the Fugl-Meyer assessment in real-time as part of the design of an in-home rehabilitation tool.

3.2 Description of Overall System

The VR game was developed to work on any general-purpose computer system running a Windows 64-bit operating system. A 3D depth camera is used to capture and store depth images from the user's arm movements. This research used the Microsoft Kinect 3D camera because of its widespread adoption and the availability of an open source software development kit. No additional equipment is needed.

The developed virtual reality application is called SuperPop VR™. When playing, the user is immersed in a virtual world where virtual bubbles appear on the screen surrounding the user. The goal is to pop as many bubbles as possible in a certain amount of time by moving a hand over the center of the bubble. The 3D depth camera is used to track the position of the user's hands (Fig. 1).

There is a set of green bubbles called Super Bubbles (SBs). Based on the user's intervention protocol established by the therapist, there is a point in time where all yellow and red bubbles on screen get erased and a set of two or three SBs appear one at a time. Each set of SBs highlights the trajectory that the therapist will use to evaluate the user's rehabilitation outcome metrics. For example, if the experimental protocol is designed to improve the user's maximum ROM, the therapist would position three SBs such that they are spaced with a slightly greater angle than the user's effective ROM (90^{o} trajectory example shown in Fig. 2). This way, through practicing the specified repetitive motion that will appear throughout the game, the user will progressively increase his/her ROM given that he/she will be forced/motivated to reach the next super bubble. When popping these bubbles, the 3D camera captures and saves depth images of the assessment arm to be run through the metrics code for evaluation in real-time (see Section IV: **Kinematic Assessment**).

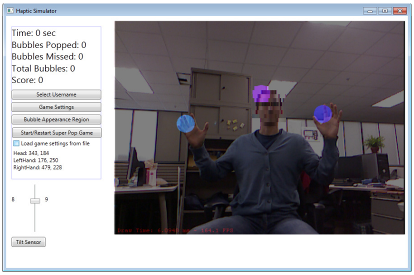


Fig. 1. Main Graphical User interface of the *Super Pop* Game

Fig. 1 also shows the main graphical user interface (GUI), which shows the virtual environment and depicts the user's progress during game play. In addition, four main buttons are located at the left side of the GUI. The first three buttons access secondary GUIs that provide the therapist options for customizing the intervention protocol of the game. The 'Bubble Appearance Region' GUI allows the therapist to select the workable region in which regular bubbles will appear and the position of the SBs. This interface allows for personalized sessions accommodating the different body structures of the users. The combination of options and features provided by the different interfaces give the therapist the freedom to match the level of difficulty of the game to the user's capacity.

3.3 Game Parameters

The difficulty of each game can be set by selecting different combinations of the following parameters: game duration in seconds, total number of levels, game speed in bubbles per second, bad bubble ratio, bubble size, good bubble score, and bad bubble score. These parameters serve different purposes in the rehabilitation protocols. The size of the bubbles and the bad bubble ratio are linked to the user's accuracy and fine motor skills. Intervention protocols designed for users with poor accuracy and/or poor fine motor skills will include larger bubbles and a lower bad bubble ratio such that the user doesn't have to worry about avoiding bubbles. The speed of the bubbles is linked to the speed of the user's movements. Intervention protocols for users with slower movements will include games with bubbles that appear at a lower rate. A scoring system is added to the game in order to motivate the users to pop some bubbles while avoiding others. Given that the SBs mark the point where the camera will capture depth images; these are worth twice as much as the good bubbles in order to increase the motivation to pop them.

All the game levels have equally distributed durations determined by dividing the total game duration by the total amount of levels. At each passing level, the game

Fig. 2. Example of a 90° trajectory created by the position of the three Super Bubbles

increases its difficulty by: increasing the game speed, increasing the bad bubble ratio, and/or decreasing the bubble radius.

4 Kinematic Assessment

The goal of the VR system is to autonomously evaluate the user's performance during game-play using the Fugl-Meyer assessment methodology. This method is a cumulative numerical scoring system for measurement of motor recovery, balance, sensation, and joint ROM [4]. Given that this research focuses only on non-touch upper-arm rehabilitation, we limit these preliminary experiments to measuring the ROM of the user's movements. This process is accomplished with the following algorithms.

4.1 Assessment of Arm Movements

A computational method for assessing upper body movements was developed in [12]. This method uses computer vision techniques to determine the user's outcome metrics through a non-touch scenario during game play.

First, a methodology called Motion History Imaging (MHI) is used to represent the user's movements using temporal templates [13]. The result is a scalar-valued image where more recently moving pixels are brighter in intensity. Once the user's movements are recognized, a contour representing the shape of the movements is extracted. A canny edge detection algorithm is used to extract the edges of the contour representing the upper-arm movement. Given that there are always unwanted contours in the image, a convex hull of the edge detected image is calculated. Afterwards, the RANdom SAmple Consensus (RANSAC) determines the best possible line fit by iteratively selecting a random subset of the original input data and returns points from the original input data that are inliers [14]. Once the points that create the upper and lower lines are recognized, the slopes of each line are used to calculate the angle between the two lines. That is,

$$ROM = \left|\arctan\left[(m_2 - m_1)/(1 + m_2 * m_1)\right] * (180/pi)\right| \tag{1}$$

where m_1 and m_2 are the slopes of each line respectively. The maximum angle found over the total amount of saved images gives the ROM of the user's movements.

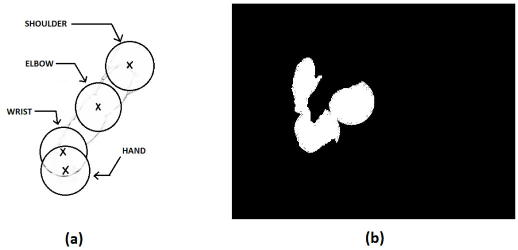


Fig. 3. (a) Imaginary circles drawn around the coordinates of the four arm joints with centers in the corresponding joints, and (b) Silhouette of a left arm isolated from the rest of an image (arm pixels in white for better visualization).

Although the aforementioned algorithm was successfully shown to output Fugl-Meyer assessment values that correlate with ground-truth data [12], it relies on the assumptions that: 1) the user's sagittal plane is perpendicular to the optical axis of the camera, 2) the user's elbow must be straight, and 3) only the arm under evaluation can move during an experiment such that the MHI doesn't take into consideration the movement of other body limbs. If these assumptions aren't met, the accuracy of the results is reduced. Given that the standard VR game play violates these assumptions - since it requires users to face the camera and bend the elbows in order to pop the bubbles - we modified the aforementioned algorithm to not only increase the accuracy of the assessment results but also enable the assessment methodology to run in real-time.

4.2 Extracting Upper Body Limbs

One of the assumptions of the previous approach was that only the assessment arm can be moved in order for the MHI to work. To improve the accuracy of the results, our algorithm isolates the arm from the rest of the image. We first identify the user's body contour in the image using the depth data obtained from the camera. As such, let

$$\begin{aligned} &\Phi = \{\phi_0, \phi_1, ..., \phi_{\Sigma-1}\} \\ &\Omega \subset \Phi \end{aligned} \tag{2}$$

where Φ is the depth data array, Σ is the total number of elements in Φ, and Ω is a subset of the depth data array that contains the elements belonging to the user's body contour. Each element of Φ, ϕ_i for i = 0, 1, …, Σ-1, corresponds to each of the pixels of the captured image. Each element is a 16 bit integer where the highest 13 bits represent the measured depth of the corresponding pixel in millimeters, and the lowest 3 bits represent the player index. We can define the lowest 3 bits of ϕ_i as:

$$L = \phi_i \quad AND \quad 0x0007 \tag{3}$$

An element of Φ is also an element of Ω if the equivalent decimal value of the lowest three bits of the integer is different than zero. Otherwise, the pixel is not included in the subset meaning that it's a part of the background of the captured image. That is, based on the definition of Equation (3):

$$\phi_i \in \Omega \quad if \quad L \neq 0x0000 \quad \forall i = 0,1,...,\Sigma - 1 \tag{4}$$

After separating the user's body contour from the rest of the image background (identifying the members of subset Ω), the next step is to separate the user's assessment arm from the body contour. Using the depth data, we can extract the coordinates of the user's arm joints (hand, wrist, elbow, and shoulder) during game play. Four imaginary circles are drawn around the four arm joints of the assessment arm (Fig. 3a). The radii of the circles are empirically determined such that the circles cover most of the assessment arm. These circles are used to determine if a given pixel in the image belongs to the arm or not using the following procedure.

Based on Equations (2), (3), and (4), Equation (5) calculates the (x,y) coordinates of the elements of Ω. That is

$$\begin{aligned} y &= floor(i / width) \\ x &= i - y * width \end{aligned} \qquad \forall \phi_i \in \Omega \tag{5}$$

where i is the index of element ϕ_i, width is equal to the number of pixel columns the images have (640), and *floor(m)* is a function that rounds the value of m to the lowest integer. Images obtained from the Kinect are 640 by 480 pixels, which means the depth data array contains a total of 307,200 elements. The resulting (x,y) coordinates are used to calculate the Euclidean distance between the pixels corresponding to the ϕ_i members of Ω and all four arm joints. If at least one of the calculated distances is less than the selected radii of the imaginary circles surrounding the arm joints, then the pixel corresponding to element ϕ_i is part of the user's assessment arm and it is assigned an RGB value that is slightly darker than the rest of the image. Otherwise, the pixel is ignored. This is an iterative process that is repeated for all elements in the depth data array every time the algorithm is called.

Fig. 3b shows the silhouette of the isolated arm resulting from this process. A set of images like this one is used as input for the assessment algorithm. Since all irrelevant information is removed from the images, the previous assumption where users can only move the assessment arm doesn't have to be met. Thus, the assessment algorithm in [12] can make more accurate calculations which allows for real-time assessment of the user's arm during game play.

5 Testing Environment

Ten male and eight female abled bodied adults within the ages of 19–27 years old were recruited for this study. Each participant was asked to play the game twice in order to provide assessment metrics for both arms. The SBs were placed such that they created a 90^0 trajectory in an adduction motion from start to end. A set of SBs would appear every 20 seconds and each SB would stay on screen for a maximum of 5 seconds. Some individuals with cerebral palsy may have difficulty standing up for

Table 1. Average and standard Deviation of the Measurements and Measurement Error

	Range of Motion (Degrees)		Error from 90^0 (Degrees)	
	Left Arm	Right Arm	Left Arm	Right Arm
Average	93.1765	92.8772	4.0954	4.0123
Standard Deviation	3.6894	4.1319	2.6323	3.0418

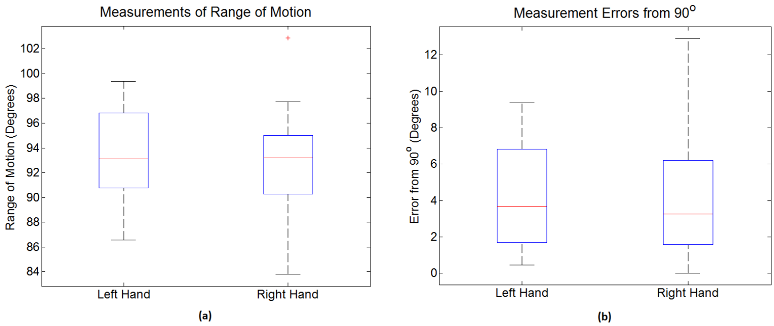


Fig. 4. (a) Boxplot showing the distribution of the range of motion measurements, and (b) Boxplot showing the error of the measurements from the 90° motion

extensive periods of time due to limitations in their functional motor abilities [5],[15]. As such, to better simulate realistic use, subjects were asked to play the game sitting down in front of the depth camera. Before each game, the following instructions were given to the users: *Pop the yellow and green bubbles (these are worth 5 and 10 points respectively), and avoid the red bubbles (these are worth -5 points each).* Afterwards, the users were asked to play for the duration of two minutes for each game.

6 Results and Analysis

Fig. 4a and 4b show boxplots for the range of motion (ROM) measurements and the error of the measurements from the 90° motion respectively. All the ROM measurements were calculated in real-time. Table 1 shows the average error and standard deviation of error. Results show that the system is able to calculate the ROM with an average error of **4.0954° ± 2.6323°** and **4.0123° ± 3.0418°** for the left and right arm respectively. Given that the overall error is less than 5%, we believe that the presented VR system provides a viable approach to performing real-time assessment of rehabilitation metrics. This fact, coupled with the ability to customize the game play settings, enables the ability to individualize the intervention protocol. We believe this is a necessary step for creating a system that can serve the therapy needs of individuals with upper-body motor impairments, such as children with cerebral palsy.

Another observation throughout the experimental sessions is that all the users were focused during game-play. Based on a study made by [16], 69% of people with motor impairments don't perform the recommended exercises which can further develop their symptoms. In order to encourage patients to work on their therapy exercises, we developed the VR system such that users can see their therapy sessions as a game and not just a repetitive set of movements which can be tedious.

7 Conclusion and Future Work

The ***Super Pop VR Game™*** is a virtual reality system designed to assist individuals with limited upper-body movement in achieving their rehabilitation goals. The VR system presented in this work allows for individuals to use it in the comfort of their homes without the need for additional equipment. This enables therapy interventions to be accessible to a larger demographic of patients with disorders that affect their motor skills. Most importantly, the system allows the therapist to select the parameters of any game such that they match with the user's needs. The experiments showed that the presented VR system is able to accurately calculate the range of motion as part of the implementation of the Fugl-Meyer assessment methodology in real-time. This is a necessary first step in designing an in-home rehabilitation tool for individuals with neurological movement disorders such as cerebral palsy.

Future work for this research is to incorporate additional tests from the Fugl-Meyer methodology as well as other assessment methodologies [17]. This will provide more comprehensive assessment of the outcome metrics of a user. Moreover, the results from the outcome metrics will later be used as feedback for the VR gaming system in order for it to automatically adapt to the user's limitations and needs. This control and adaptation approach will play an important role in rehabilitation given that the games will be able to adjust to the users in real-time providing more individualized therapy sessions targeted to the specific needs of each individual.

Acknowledgements. This research was partially supported by the National Science Foundation Graduate Research Fellowship under Grant No. DGE-1148903 (and 1208287), and the Grammy Foundation Grant Program. Any opinions, findings, and conclusions or recommendations expressed in this material are those of the author's and do not necessarily reflect the views of the National Science Foundation or the Grammy Foundation.

References

1. Chen, Y., Kang, L., Chuang, T., Doong, J., Lee, S., Sai, M.T., Jeng, S., Sung, W.: Use of Virtual Reality to Improve Upper-Extremity Control in Children with Cerebral Palsy: A Single-Subject Design. Physical Therapy: Journal of the American Physical Therapy Association 87(11), 1440–1454 (2007)
2. Reid, D.: The influence of virtual reality on playfulness in children with cerebral palsy: A pilot study. Occupational Therapy International 11(3), 131–144 (2004)

3. Bryanton, C., Bossé, J., Brien, M., McLean, J., Mccormick, A., Sveistrup, H.: Feasibility, Motivation, and selective Motor Control: Virtual Reality Compared to Conventional Home Exercise in Children with Cerebral Palsy. CyberPsycology & Behavior 9(2), 123–128 (2006)
4. Fugl-Meyer, A., Jääskö, L., Leyman, I., Olsson, S., Steglind, S.: The post-stroke hemiplegic patient: A method for evaluation of physical performance. Scandinavian Journal of Rehabilitation Medicine 7, 13–31 (1975)
5. Bax, M., Goldstein, M., Rosenbaum, P., Leviton, A., Paneth, N.: Proposed definition and classification of cerebral palsy. Developmental Medicine & Child Neurology 47(8), 571–576 (2005)
6. Reid, D.: Benefits of a virtual play rehabilitation environment for children with cerebral palsy on perceptions of self-efficacy: a pilot study. Pediatric Rehabilitation 5(3), 141–148 (2002)
7. Golomb, M., McDonald, B., Warden, S., Yonkman, J., Saykin, A., Shirley, B., Huber, M., Rabin, B., AbdelBaky, M., Nwosu, M., Barkat-Masth, M., Burdea, G.: In-Home Virtual Reality Videogame Telerehabilitation in Adolescents With Hemiplegic Cerebral Palsy. Archives of Physical Medicine and Rehabilitation 91(1), 1–8 (2010)
8. Deutsch, J., Borbely, M., Filler, J., Huhn, K., Guarrera-Bowlby, P.: Use of a Low-Cost, Commercially Available Gaming Console (Wii) for Rehablitation of an Adolescent With Cerebral Palsy. Physical Therapy: Journal of the American Physical Therapy Association 88(10), 1196–1207 (2008)
9. Jannink, M., Van Der Wilden, G., Navis, D., Visser, G., Gussinklo, J., Ijzerman, M.: A Low-Cost Video Game Applied for Training of Upper Extremity Function in Children with Cerebral Palsy: A Pilot Study. Cyber. Psychology & Behavior 11(1), 27–32 (2008)
10. You, A., Jang, S., Kim, Y., Kwon, Y., Barrow, I., Hallet, M.: Cortical reorganization induced by virtual reality therapy in a child with hemiparetic cerebral palsy. Developmental Medicine & Child Neurology 47(9), 628–635 (2005)
11. Duncan, P., Propst, M., Nelson, S.: Reliability of the Fugl-Meyer Assessment of Sensorimotor Recovery Following Cerebrovascular Accident. Physical Therapy: Journal of the American Physical Therapy Association 63(10), 1606–1610 (1983)
12. Brooks, D., Howard, A.: A Computational Method for Physical Rehabilitation Assessment. In: IEEE International Conference on Biomedical Robotics and Biomechatronics (BioRob), pp. 442–447 (2010)
13. Bobick, A., Davis, J.: The Recognition of Human Movement using Temporal Templates. IEEE Transactions on Pattern Analysis and Machine Intelligence 23(3), 257–267 (2001)
14. Fischler, M., Bolles, R.: Random Sample Consensus: A paradigm for model fitting with application to image analysis and automated cartography. Communications of the ACM 15(6), 381–395 (1981)
15. Palisano, R., Rosenbaum, P., Walter, S., Russell, D., Wood, E., Galuppi, B.: Development and reliability of a system to classify gross motor function in children with cerebral palsy. Developmental Medicine and Child Neurology 39(4), 214–223 (2008)
16. Shaughnessy, M., Resnik, B., Macko, R.: Testing a model of post-stroke exercise behavior. Rehabilitation Nursing 31(1), 15–21 (2006)
17. Camerirão, M., Bermúdez, S., Badia, I., Duarte, E., Frisoli, A., Vershure, P.: The Combined Impact of Virtual Reality Neurorehabilitation and Its Interfaces on Upper Extremity Functional Recovery in Patients With Chronic Stroke. Stroke: A Journal of Cerebral Circulation 43(10), 2720–2728 (2012)

Asynchronous Telemedicine Diagnosis of Musculoskeletal Injuries through a Prototype Interface in Virtual Reality Environment

Soheeb Khan[1], Vassilis Charissis[1], David Harrison[1], Sophia Sakellariou[2], and Warren Chan[1]

[1] Glasgow Caledonian University
School of Engineering and Built Environment,
Department of Computer, Communications and Interactive Systems,
Glasgow, UK
[2] Royal Infirmary,
Department of Radiology,
Glasgow, UK
soheebkhan@googlemail.com

Abstract. Telehealth provides a much needed option for remote diagnosis and monitoring of various pathologies and patients. Remote provision of health care can offer a two fold support for the medical system and the patients. Primarily it could serve isolated locations and secondly it could monitor a large number of outpatient cases directly on their homes instead of the hospital premises. However in specific cases direct communication and visual data acquisition can be a major obstacle. To this end we have developed a prototype system that could enable the medical practitioners to have real-time diagnosis through 3D captured visual and motion data. This data are recreated in a Virtual Reality environment in the hospital facilities offering a unique system for remote diagnosis. This paper presents the design considerations and development process of the system and discusses the preliminary results from the system evaluation. The paper concludes with a tentative plan of future work which aims to offer the medical practitioners and the patient with a complete interface which can acquire gait data and thus analyse a large variety of musculoskeletal pathologies.

Keywords: Virtual Reality, HCI, 3D Visualization, Asynchronous Diagnosis, Telemedicine, Motion Capture.

1 Introduction

The demand and utilisation of telemedicine-based care is on the increase due to the highly amplified number of populations striving towards country life and remote living locations. Additionally the financial climate, limited resources and the constant growing population around the globe have prompted an interest from various governing bodies to reform and seek alternative methods for delivering high quality health care [1]. The technological advancements and the increase of communication innova-

R. Shumaker (Ed.): VAMR/HCII 2013, Part II, LNCS 8022, pp. 50–59, 2013.

tions have made telemedicine a promising solution for many issues faced by the current health care systems [2, 3]. Patients with musculoskeletal (MSK) issues and injuries comprise the largest number of presentations to General Practitioners (GPs), and continue to require input even post surgical intervention or during the rehabilitation process [4]. Yet the gait analysis of each patient is time-consuming and costly if the patients are located away from the medical and city centres.

The technological innovations in motion-capture (Mo-Cap) systems have made it possible to acquire and collect complex motion data for the biomechanics of musculoskeletal (MSK) structures. Mo-Cap techniques are being deployed and implemented across a range of disciplines. In the past, Mo-Cap setups have been expensive ventures, as they required specialised costly equipment, professional setup, training and allocation of a dedicated large space. This has limited the amount of gait analysis and MSK diagnosis laboratories confining them to cities and highly populated areas. Such laboratories are most commonly situated at a designated facility which requires the patients to travel to and from routinely. Due to the limitation of such facilities, people living in isolated and rural areas have limited or no access to this triage. Markerless motion capture techniques are new and only in the last decade the technology has surfaced as a useable product for the mass consumers. Contemporary technological breakthroughs, related to cameras, projectors and videogames fuelled the development of cost-efficient, consumer-based peripherals such as the Microsoft "Kinect" that could be utilised off-the-shelf for a fraction of the typical Mo-Cap suites [5]. Despite the significantly reduced price, only a minor impact to the final quality of the derived motion data is noted, the latter not constituting a significant drawback for the majority of the applications.

The current capacity in computer processing and accessibility of variety of 3D development packages has enabled visualisation of complex anatomy in real-time virtual environments. Mass developments of videogames, virtual reality (VR) and 3D programs have made possible for 3D engines to be utilised across various industries as common platforms for real-time visualisation purposes. This has prompted medical information to be presented in a much-improved manner by the use of photorealistic 3D models and user-friendly interfaces.

The main objective of our proposed system is to employ inexpensive hardware such as the 'Kinect' in conjunction with 3D VR environments, and a user-friendly interface to obtain locomotion data that can be of diagnostic value for clinicians. The proposed work presents an asynchronous Telemedicine based Musculoskeletal Diagnosis Service, which enables the health professionals to asynchronously perform gait analysis and diagnosis. The telemedicine based system utilises 3D medical visualisation, Virtual Reality and Motion Capture systems. Furthermore the system supports active learning and development for healthcare professionals through the use of a user-friendly interface and appropriately designed Human Computer Interaction (HCI) in conjunction with photorealistic 3D medical visualisation. The system was evaluated both with qualitative and quantitative methods by ten medical practitioners offering promising results with regards to the HCI usability, and data manipulation through online real-time communication. The initial trials utilised a default MSK 3D human model which was attached to various motion capture data produced in real-time in a laboratory environment and transmitted to the medical group's location.

2 Innovative Motion Capturing Solutions

Contemporary technological breakthroughs, related to the videogames industry and associated peripheral devices have fuelled researchers to begin development of cost-efficient, home based telemedicine systems. Innovative inexpensive optical sensor technology designed to enhance the video game experiences have become an easy obtainable solution for multiple complex issues and desires for human computer interaction (HCI) [6,7]. Clinicians and researchers have begun to experiment and develop tools with such devices to assist motion driven tasks [8,9,10]. A sufficient amount of interest has been to use such tools as an aid for rehabilitation purposes [10,11]. Attygalle developed a tool for a home based rehabilitation setup, which incorporated the standard Biofeedback rehabilitation program. Using a Nintendo Wii remote, Attygalle successfully recorded data concerning rehabilitation for stroke patients. The data was asynchronously analysed by therapists to track the patient's progression over time.

Concurrent studies developed a system that utilised the Kinect's hand gesture capabilities to be employed in live surgery [11]. This system facilitated health professionals in a sterile environment to navigate, manipulate and control MRI and CT images through Natural User Interface (NUI). The system proposed by Zöllner successfully employed the Kinect as a navigation tool for the visually impaired [12]. Similarly there are many projects that are currently experimenting with this technology in various other disciplines [10,11,12]. Other studies demonstrated Kinect's tracking capabilities which delivered moderate results when compared to the Vicon system [13,14].

IPI Soft is a motion capturing software which utilises such videogame peripherals to achieve low budget motion tracking. Various other tools are available that also utilise this technology for tracking purposes but in comparison IPI Soft excels. Although used mostly by the entertainment industry, this research showed that this tool has not yet been utilised for gait analysis. It is yet to be demonstrated if such devices and software could be utilised to obtain data that can be of diagnostic value to the clinician and comparative studies of its performance against the traditional marker based motion capture system (vicon).

3 System Rationale

The proposed system design rationale consists of three phases: Data Collection, Data Processing and Diagnosis. Phase one consists of gathering and submitting the essential data required for motion analysis and clinical musculoskeletal diagnosis. The proposed system requires this process to be carried out remotely and independently by patients or health professionals.

The submitted data will be transmitted to a designated processing centre. Phase 2 requires the data to be processed into a viewable and readable medium for the health professional. Phase 3 enables the specialist/ health professionals to view and analyse the data to make a diagnosis. Further courses of action could be delivered to the patient via email or phone. These phases will be asynchronous and independent from one another, which will not require any instant reaction or live contact between them. A detailed description of each phase is illustrated in the architecture of the proposed system. See Fig 1.

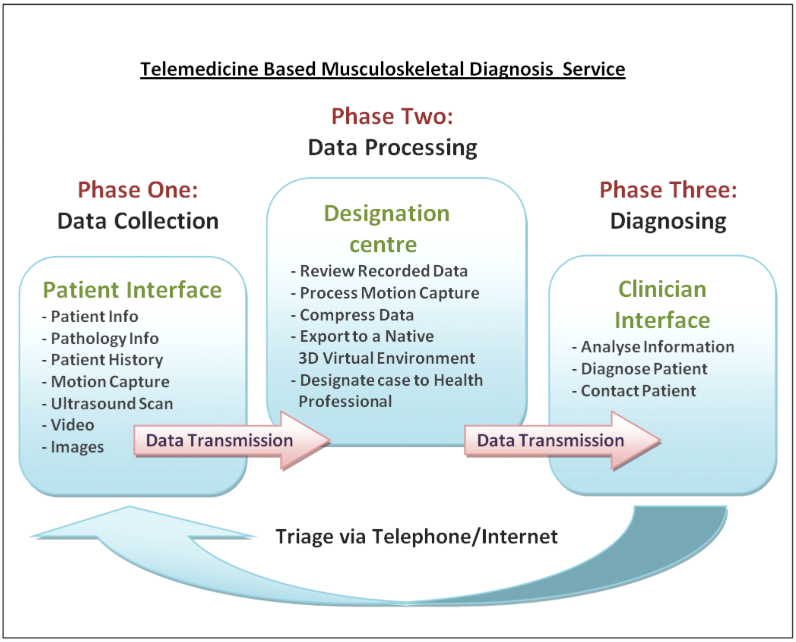


Fig. 1. Architecture of the Proposed System

4 Comparative Analysis of Optical Sensors

A thorough analysis of existing off-the-shelf sensor equipment was deemed necessary so as to define the best possible combination of hardware. As software compatibility is a major issue with all new technologies analysed, we concluded that the best option currently is somehow dictated by the designated software used for multiple tracking sensors, namely ipiSoft. Consequently, the peripherals that were taken in consideration were chosen as compatible with the specific software and utilised by the ipi Desktop motion capture system.

The three main systems that are most compatible with ipiSoft are Microsoft Kinect Sensor, Asus Xtion sensor and PlayStation Eye Digital camera. By analysing the user guides for each of the systems provided by ipiSoft, the different optical sensors in regards to the proposed work were compared. The systems were compared on price, portability, calibration, space utilisation, complexity and time of setup as well as the size and accuracy of the recorded data. This allowed the selection of the most appropriate kit for the proposed system. The paper furthermore presents an analysis of benefits and drawbacks of three different setups recommended by ipiSoft in relation to the proposed system as presented in Table 1 below.

Table 1. Affordable Optical Sensors Comparison

Optical Sensors	Positive	Negatives
Microsoft Kinect (Available from £99 per unit)	• Simple to set up(plug play) • Multiple device compatibility (maximum 2) • Simple to calibrate • Compact setup - Minimum required space: 3m by 3m (10 by 10 feet) • Suitable for living rooms or limited space • Portable setup works in most indoor environments • Popular device, regular support for updates and high driver quality (Microsoft) • Compatible with various desktops and laptops • Remote adjustment of device with iPi recorder application • Captures at resolution of 640x480 • Good Quality results • Subject is not required to dress in an applicable way	• Prone to Occlusion Problem whilst using single device • Require ACDC power supply • Only 30 fps frame rate • Maximum capture volume7 by 7 feet (approx. 2 by 2 meters); • Only 30° of coverage while using multiple devices • require systems with multiple USB controllers
ASUS Xtion (Available from £130 per unit)	• Simple to set up(plug play)Multiple Device compatibility • Compact setup - Minimum required space: 3m by 3m (10 by 10 feet • Does not require power supply except USB • Allow 60 fps frame rate (at 320 x 240 resolution only) • Decent Quality Results • Subject is not required to dress in an applicable way	• Recently introduced (prone to errors) • Less popular device with no regular updates and lower drivers quality • limited compatibility to certain desktops and laptops • positioning and adjustment, only manual • Maximum capture volume7 by 7 feet (approx. 2 by 2 meters) • The More Expensive Device
PlayStation Eye Camera(Available from £15 per unit)	• Multiple device compatibility (maximum 6 cameras) • Range of setups -4 cameras 320 by 240 resolution - 4 cameras at 640 by 480 resolution - 6 cameras at 640 by 480 resolution • 60 fps frame rate • Maximum capture volume 20 feet by 20 feet (7 meters by 7 meters) • more performance space • Does not work with some USB controllers • Compatible with various desktops and laptops • Cheaply Available Devices • Can achieve some high quality results	• Complicated to set up and calibrates • Needs large space, minimum space required: 4 meters by 4 meters • additional tripods are required for an effective setup • Subject is required to wear certain clothing • A powerful computing system is required to record the data • Limited Portability • Can be effected by shadows • not suitable for compact or living room environments • require systems with multiple USB controllers • No motor, allows only manual positioning

After analysing the gathered information the most appropriate optical sensor for the proposed system is the Microsoft Kinect. In comparison to the 'ASUS Xtione', the 'Kinect' is the observable choice. Both of the depth cameras are very alike in technology and deliver similar results. The Kinect excels as the more popular device due to the regular updates and high quality driver provided by Microsoft. The Kinect is compatible with most laptops and desktops which has prompted both professionals and enthusiasts to develop features using this device to be utilised for a variety of purposes. These features could potentially be incorporated and utilised by the proposed system. In contrast ASUS Xtione is less popular and is only available to purchase from specific shops. The system driver's and updates are not as frequent as the updates provided by Microsoft. This could make the setup unstable and prone to error. ASUS Xtione is unsuitable due to its limitation and restrictions in compatibility as it works only with certain laptops and desktops. The Asus device allows recording at 60 frames per second (fps) in comparison to the 30fps provided by the Kinect, this is only possible though when recording at a capture resolution of 320x240. Both devices record at 30fps when recording at capture resolution of 640x480. Although the recorded frame rate is faster at a lower resolution this has an effect on the quality of the data which will eventually be used to track the motion resulting in less accurate results and thus negating the advantage of a higher frame rate.

Although the six-camera based PlayStation Eye setup has been suggested as the recommended choice by ipiSoft for the best results, there are certain limitations and attributes that make this setup less applicable and practical for the proposed system. As it is essential to utilise minimum 3 PlayStation Eye cameras for any motion tracking data, this results in a complex, time consuming installation requiring significant fine tuning. There are various procedures and tasks that need to be carried out in order to obtain functional and efficient calibration. The calibration process is prone to errors and requires the user to patiently adjust and calibrate the cameras until the desired results are achieved. In a remote location, however, the system would need to follow a fast and errors free installation. Maintenance is also required to be minimum with easily accessible and regular software and hardware updates.

Portability is also essential as the proposed system would require a remote setup which will enable the user to capture their movement independently from their homes or remote facilities. As mentioned before, the proposed system is also targeted at people who may have limited or no knowledge of using computers. Calibration and installation of the setup needs to be very simple so that it could be carried out independently without having to carry out complicated procedures. Comparatively to the PlayStation Eye setup, the Kinect setup is substantially easier to install and does not require any additional equipment for recording other than the laptop and cables. The minimum required space for the Kinect setup is 3m by 3m (10 by 10 feet) which is ideal for environments such as living rooms and facilities with limited space capacity. This setup is portable as it does not require special calibration unlike the Sony PlayStation Eye setup.

Overall the three tested systems were further compared with other different Mo-Cap technologies used in prior research studies [15,16]. Particular focus was given to alternative inexpensive but effective solutions. These included the optical sensors predominantly utilised for video gaming purposes and Mocap tool’s like ipiSoft that are utilised for low budget projects. By experimenting with these alternative solutions the best device for the proposed system was chosen as presented above.

5 Evaluation

Based on the aforementioned observation our system entailed 2 Kinects and a laptop in order to retrieve the motion data and compile them with the use of ipiSoft.

The three-dimensional visualisation was further exported and combined through Unity 3D to a default three-dimensional model of a human body. The latter serves the purpose of presenting the potential injury location rather than presenting a detailed model of the patient.

The system was evaluated with 10 users in the Virtual Reality and Simulation laboratory (VRS lab) of Glasgow Caledonian University (Glasgow, UK). The laboratory is equipped with state of the art HD/3D projector, 65inch 3DTV, and surround audio equipment. The aforementioned devices run through a set of custom-built workstations powered by dual Hexacore Xeon CPUs and NVidia Quadro 4000 Fermi graphics cards.

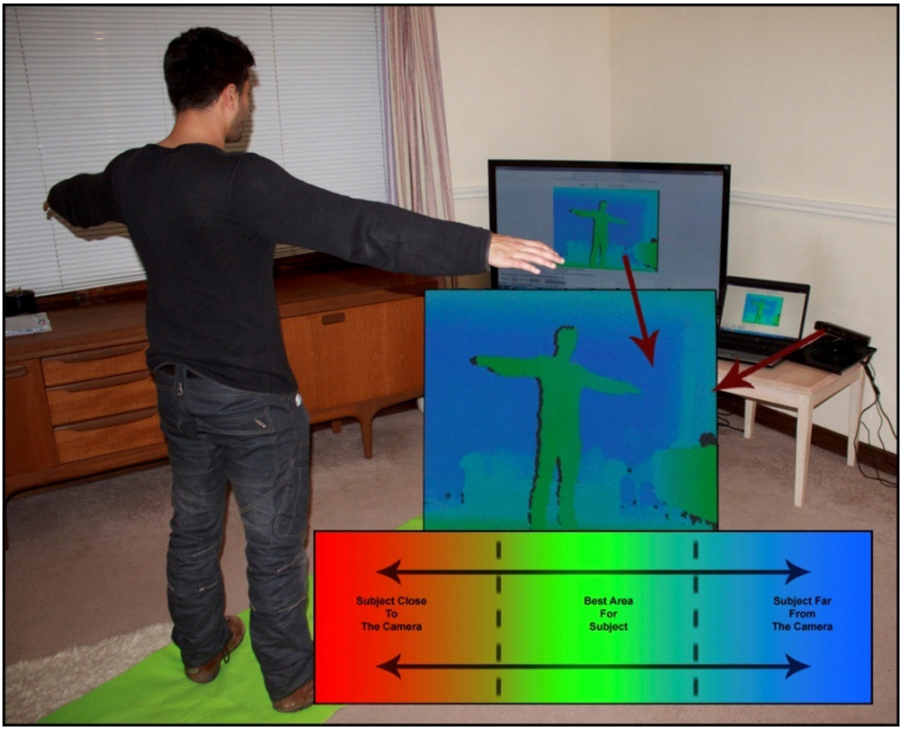


Fig. 2. The subject is being recorded using a Kinect sensor

Data from this experiment was successfully retrieved which was processed and mapped on to a 3D model to replicate the movement in to a virtual environment. This process has been demonstrated in Fig 3. Further work is required to determine if the retrieved data could be utilised for gait analysis and MSK diagnosis.

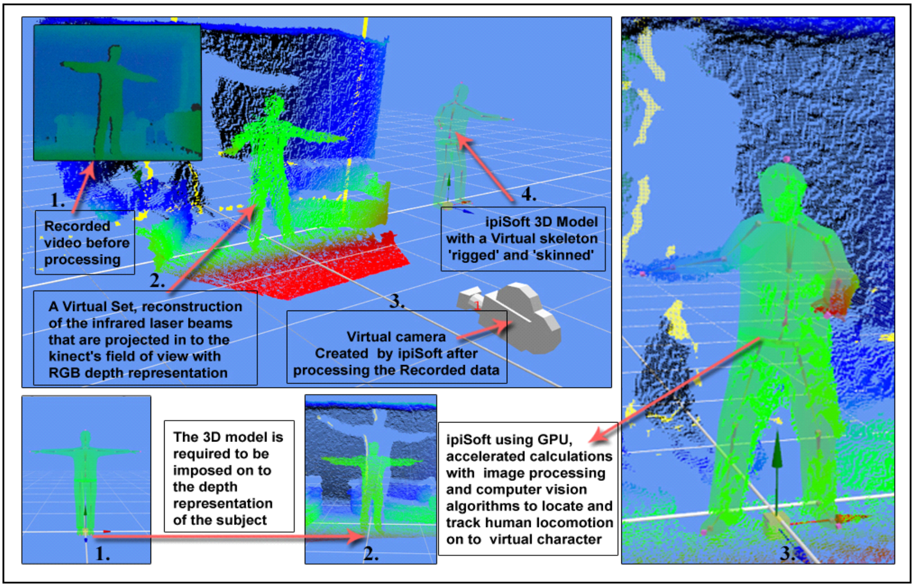


Fig. 3. Motion mapping data on to the virtual character in ipiStudio

6 Results and Discussion

Through the initial trials the system performs sufficiently in order to capture and transmit motion information for a preliminary diagnostic appraisal. This observation was unanimously noted by the participants. As this is the intention of the proposed system the rough estimation of movement offers clear information that could assist a medical practitioner to estimate the musculoskeletal injury in a long distance surgery. In turn the medical doctor can provide a therapy and through the specific system can monitor the progress of the patient remotely. Some users suggested also that such system could extend its usage from the remote locations to even the typical follow up sessions in order to minimize the crowding effect of visiting patients in the hospital premises.

The system's main drawback is the quality of the derived data that could not be used, in this raw format directly, for precise gait analysis. Such action requires very specific measurements, which need to be captured by multiple very precise cameras and through an iterative process of marking different sections of the leg, foot and ankle. Interestingly the complete gait analysis of data is also falling within the asynchronous process and as such the detailed information need further exploration before any patient consultation or treatment. The precision required in this case affects significantly the cost of the related hardware and software that is required to capture the data. Furthermore complicated and time-consuming calibration and maintenance as well as specialist operators are other major issues for deploying such services in multiple points across the country and particularly in remote locations.

On our proposed system we are particularly keen to increase quality to acceptable levels for generic diagnosis and monitoring as the required equipment is accessible, cost effective and simple to run and maintain. The users during the trial commented in favour of this approach. The required time of set up of the system was within a 5 minutes margin and the data transmission was in all the trials well within an hour (without any major data optimization).

7 Conclusions

This paper presented the development process and system rationale of our proposed, prototype telemedicine system. The particular system is designed for remote diagnosis of musculoskeletal injuries and monitoring patient rehabilitation through gait analysis motion capture. The main objective of the system if to offer a clear, simple and cost effective method for both medics and patients. As such the system employs a variety of off-the-shelf equipment and a custom made software application captures and visualise the information for both groups through a virtual environment. The initial trials presented successful transfer and visualisation of motion capture and anatomical information.

Our tentative plan for future work will aim to identify the potential accuracy issues in comparison to a full motion capture system designed for precise gait analysis which is currently used by the medical and allied health practitioners. Additional information might be incorporated and explored through simplified ultrasound data acquisition. Finally we aim to install the system in satellite GP practises in remote locations across Scotland and evaluate the speed, usability and efficiency of the system in real life conditions.

References

1. Bower, P., Cartwright, M., Hirani, S., Barlow, J., Hendy, J., Knapp, M., Henderson, C., Rogers, A., Sanders, C., Bardsley, M., Steventon, A., Fitzpatrick, R., Doll, H., Newman, S.: A comprehensive evaluation of the impact of telemonitoring in patients with long-term conditions and social care needs: protocol for the whole systems demonstrator cluster randomised trial. Studies in BMC Health Services Research 11, 184 (2011)
2. Blank, E., Lappan, C., Belmont, P.J., Machen, M.S., Ficke, J., Pope, R., Owens, B.D.: Early Analysis of the United States Army's Telemedicine Orthopaedic Consultation Program. Studies in Journal of Surgical Orthopaedic Advances 20, 50–55 (2011)
3. Syst Kuo, C.H., Liu, L.: Development of a Web-Based Telemedicine System for Remote ENT Diagnoses. In: Studies in 2010 IEEE International Conference on System Science and Engineering (ICSSE), pp. 565–570 (2010)
4. Sakellariou, S., Charissis, V., Grant, S., Turner, J., Kelly, D., Christomanos, C.: Virtual Reality as Knowledge Enhancement Tool for Musculoskeletal Pathology. In: Shumaker, R. (ed.) Virtual and Mixed Reality, Part II, HCII 2011. LNCS, vol. 6774, pp. 54–63. Springer, Heidelberg (2011)
5. Stone, E.E., Skubic, M.: Passive In-Home Measurement of Stride-to-Stride Gait Variability Comparing Vision and Kinect Sensing. In: IEEE EMBS, Massachusetts USA (2011)

6. Mitchell, G., Clarke, A.: Capturing and Visualising Playground Games and Performance: A Wii and Kinect based motion capture system. In: The Proceedings of the 2011 International Conference on Electronic Visualisation and the Arts (EVA 2011), pp. 218–225. British Computer Society, UK (2011)
7. Charissis, V., Ward, B.M., Naef, M., Rowley, D., Brady, L., Anderson, P.: An Enquiry into VR Interface Design for Medical Training: VR Augmented Anatomy Tutorials for Breast Cancer. In: McDowall, I.E., Dolinsky, M. (eds.) Proceedings of the: International Annual Symposium of IS&T/SPIE, The Engineering Reality of Virtual Reality, San Jose, California, vol. 6804, pp. 19–28 (2008) ISBN: 9780819469762
8. Desloover, K., Wong, P., Swings, L., Callewaert, B., Vandenneucker, H., Leardini, A.: Range of motion and repeatability of knee kinematics for 11 clinically relevant motor tasks. Studies in Gait & Posture 32, 597–602 (2010)
9. Foxlin, E., Harrington, M., Pfeifer, G.: Constellation: A wide range wireless motion-tracking system for augmented Reality and Virtual set applications. In: Studies in Proceedings of Siggraph 1998 Computer Graphics Proceedings. ACM SIGGRAPH Conference Series (1998)
10. Zhou, H., Hu, H.: Human motion tracking for rehabilitation—A survey. Studies in Biomedical Signal Processing and Control, Elsevier LTD 3, 1–18 (2008); Perry, J., Burnfield, J.M.: Gait Cycle. In: Gait Analysis Normal and Pathological Function, vol. 2, pp. 3–19. Slack Incorporated (2010)
11. Noonan, D., Mountney, P., Elson, D., Darzi, Yang, G.: A Stereoscopic Fibroscope for Camera Motion and 3D Depth Recovery During Minimally Invasive Surgery. In: Proc. ICRA 2009, pp. 4463–4468 (2009)
12. Zöllner, M., Huber, S., Jetter, H., Reiterer, H.: NAVI – A Proof-of-Concept of a Mobile Navigational Aid for Visually Impaired Based on the Microsoft Kinect. Studies in University of Konstanz, Germany (2012)
13. Dutta, T.: Evaluation of the Kinect sensor for 3-D kinematic measurement in the workplace. Studies in Applied Ergonomics, 1–5 (2011), doi:10.1016
14. Stone, E.E., Skubic, M.: Passive In-Home Measurement of Stride-to-Stride Gait Variability Comparing Vision and Kinect Sensing. IEEE EMBS, Boston (2011)
15. Henry, P., et al.: RGB-D mapping: Using Kinect-style depth cameras for dense 3D modeling of indoor environments. The International Journal of Robotics Research, SAGE Journals (2012), doi:10.1177/0278364911434148
16. Izadi, S., et al.: KinectFusion: real-time 3D reconstruction and interaction using a moving depth camera. In: Proc. of 24th UIST annual ACM symposium on User Interface Software and Technology, pp. 559–568 (2011)

Developing a Theory-Informed Interactive Animation to Increase Physical Activity among Young People with Asthma

Jennifer Murray[1], Brian Williams[2], Gaylor Hoskins[2], John McGhee[3], Dylan Gauld[3], and Gordon Brown[4]

[1] Edinburgh Napier University, School of Life, Sport and Social Sciences, Edinburgh, Scotland
j.murray2@napier.ac.uk
[2] Nursing, Midwifery and Allied Health Professions Research Unit, University of Stirling, Stirling, Scotland
{brian.williams,Gaylor.hoskins}@stir.ac.uk
[3] Duncan of Jordanstone College of Art and Design, University of Dundee, Dundee, Scotland
{j.mcghee,d.gauld}@dundee.ac.uk
[4] Asthma UK Scotland, Edinburgh, Scotland
gbrown@asthma.co.uk

Abstract. The current paper describes the development of a theory-informed interactive animation and which aims to increase levels of physical activity in young people with asthma. The project adopts a multi-disciplinary theoretical perspective, applying knowledge from applied health research, human centred design and psychology in order to best approach and develop a meaningful and effective health intervention.

Keywords: Asthma, interactive animation, multidisciplinary, theory-informed.

1 Introduction

To begin, an introduction to the issues surrounding physical activity and young people with asthma will be discussed, to provide a context for the need for the development of the theory-informed interactive animation. The key theoretical bases informing the development of the animation are then discussed, as are the key HCI principles that apply to the animation's development. This is then followed by a broad description of initial assessments of the animation and future plans to assess and refine it.

2 Background

Within the UK, the number of new diagnoses of asthma has increased significantly over the past 20 years [1]; one in five children is now affected by asthma within the UK [2,3]. One in five of all child GP consultations relate to asthma [4], and within England and Wales in 1999 alone there were over 30,000 asthma related hospital

R. Shumaker (Ed.): VAMR/HCII 2013, Part II, LNCS 8022, pp. 60–65, 2013.

admissions and 25 related child deaths [5]. Improving asthma control is a necessity for both the health of young people and for reducing the burden on health services. One simple and cost effective way to do this is through physical activity; which can lead to improvements in aerobic fitness [6, 7] and asthma related benefits such as reduced hospital admissions, reduced absences at school, reduced medication use [8] and improved ability to cope with asthma [7]. However, young people with asthma are actually less likely to be physically active than their peers [9-13], and that they attribute this to their asthma. For example, in a survey of young people in Los Angeles between 1999-2000, 53% of those with a diagnosis of asthma reported their activities were limited because of their asthma [14]. The overwhelming majority of studies, however, indicate that young people with asthma can exercise safely when appropriately treated [15].

3 Theoretical Basis

We drew on the research team's extensive background in asthma and in developing innovative theoretically-informed visualisations to encourage behaviour change to develop the interactive animation. In depth qualitative work in the area of asthma and physical activity carried out previously by members [16] of the research team identified important beliefs that act as the principle reasons for low activity among young people with asthma. The study also found that young people were unable to distinguish when their asthma was the cause of shortness of breath, or when they were simply short of breath due to normal physical processes, and this then impacted on concerns for their safety and beliefs about their ability to take part in physical activity [16]. Subsequently, motivation to be physically active was diminished, and issues around self-efficacy (i.e., the level of belief that a person has about their capabilities) arose.

Thus the animation that we have developed aims to address these issues. First, by using the psychological literature on motivation, self-efficacy and learning theory (how young people learn most effectively) it is believed that the optimal learning environment can be created to enhance the user's experience of the animation and the messages given, encouraging them to take control (thus increasing self efficacy) and increasing motivation to alter behaviours and beliefs. There is insufficient scope within the context of the current paper to discuss these important elements in greater detail; however the authors plan to publish this information in future articles.

Second, previous work by the team in developing novel theory-informed animations has shown some success in changing health related behaviours through transforming 'abstract' concepts into 'concrete' ones. For example, an animation was developed using behavioural theory to create a 3D animation to illustrate the link between obesity and atherosclerosis, showing the user what was actually happening to the heart and blood vessels when weight was increased and decreased. We therefore applied this concrete-to-abstract method to the current project's animation, demonstrating what is happening inside the lungs during physical activity [17]. Williams and colleagues [17] found that through showing patients an animation detailing the impact of weight on the health of blood vessels and the heart, patients were able to better conceptualise what was happening within their own bodies, leading to increased motivation to alter behaviour and increased self-efficacy (i.e., the belief that they can change their behaviour).

We build on this research using the same key principle of transforming the 'abstract' concept of visualising the internal effects on the lungs and bronchioles during breathlessness due to asthma compared to normal breathlessness to a 'concrete' visualisation within the animation. This progression from an abstract concept to a concrete representation is thought to not only aid understanding and conceptualisation [17, 18], but also aid retention in memory [19]. Thus, through strengthening the concreteness of the visualisation, young people with asthma may be more likely to: 1) recall the lessons given within the animation; 2) improve their conceptualisation, thus understanding of the issue and; 3) through their increased understanding, begin to develop a better awareness of when they are breathless due to asthma versus when they are breathless simply due to normal physical exertion.

A key difficulty in current health based interventions which aim to alter behaviours lies in the translation of intentions to become more physically active to actually becoming more physically active: it is well known within the psychological literature that intentions do not always guarantee behaviour change. Put simply, the mere statement of intention to alter a behaviour does not always mean that the behaviour is altered in real life. We identified two key theories of behaviour change which were highly relevant to our topic area and applied these while designing our animation: the Theory of Planned Behaviour [20] and Leventhal's Common Sense Model of behaviour [21].

These models aim to explain why people behave in certain ways and what barriers and facilitators exist between the intention to act and the actual behaviour. Once again, the full discussion of these theories in relation to the development of the current animation is outwith the scope of the current paper, but will be discussed in future papers.

Addressing these psychological elements and applying theoretical models within our animation was therefore of paramount importance in developing a useful, evidence based intervention.

4 Developing and Assessing the Animation

Using the theoretical bases described above, we have shaped the animation in such a way as to maintain a strong narrative, interactive elements to enhance engagement and immersion, and the animation that we have developed allows the user to select a character (choice of six, with initial testing piloting a choice of two, who will acts as a 'buddy') and a type of activity for the character to participate in (running, soccer or dance, with initial testing only piloting running). The characters introduce themselves to the user and discuss in a 'chatty', friendly tone their asthma and their inhaler use. The user is prompted to, at set points, decide whether their character will use inhalers or not; and they then see the consequences of this action. Both external views of the characters and internal views of the lungs are shown, to aid the user in establishing a concrete conceptualisation of what happens within the lungs during breathlessness caused by asthma compared to normal breathlessness, and what happens when inhalers are used correctly.

Based on lessons learned previously, we sought not only to embed psychological and health theory to the current animation, but to also embed human centred design principles to a greater degree and involve an online consultative user-group feedback mechanism throughout the design process at each key decision making stage.

The consultative user-group consisted of a representative sample 23 people, including young people with asthma (N=5), parents (N=4), teachers (N=6), adults with asthma (N=3) and health professionals (N=5) who met online to discuss general issues surrounding asthma and physical activities and more project-specific aspects (e.g., design preferences). General discussions with the user-group were found to be very fruitful in shaping our initial ideas and thoughts on the project; more so than expected. These discussions began with a prompt question or discussion point (e.g., which forms of physical activity do you enjoy most and least' and 'are there any types of physical activity that you feel your asthma stops you doing?'), and the users would discuss this in a typical online forum format, with members of the research team contributing additional prompts for follow-up information. These types of questions allowed us to form rapport with our user-group prior to asking for feedback on specific elements of the animation and allowed bonding between the members of the user group to allow an open and comfortable feedback and discussion platform for the project. It also allowed us to identify whether there were common themes in terms of likes, dislikes and general attitudes, helping us to form initial storyboards and concepts for the animation.

More specific feedback was sought from the user group through the use of closed or short-text surveys on specific components of the animation. In some cases these surveys were expanded and advertised on asthma-related forums and social networking pages to maximise feedback. This form of feedback informed development was particularly useful when designing the characters to be used within the animation. The user is able to select one of six characters. At the initial concept stage we canvassed opinions on likability, realism and general attitudes towards the characters. All but one character had some alterations made as a direct result of the feedback and two were substantially revised due to lack of popularity on multiple character-elements by the sample.

This was of particular importance, as one of the key design principles that we sought to embed within the animation was that of empathising with the characters. Through creating an empathic bond with the character, users are more likely to pay more attention and become engaged in [22] and immersed in [23] the animation's narrative. Having an empathic bond with the character and increased engagement and immersion within the animation should, therefore, lead to optimal learning and conceptualisation of the target information (i.e., the impact of physical activity and inhaler use on the lungs) and encourage users to utilise this in their future understandings and decisions about their physical activity in relation to their asthma.

The final key design principles adhered to included: the use of visual and audio narratives to enhance immersion and engagement [24]; the use of interactivity to allow choice, user engagement, and immersion [23]; allowing customisation of characters and physical activity scenarios to enhance user engagement and immersion [23]; the use of minimal additional text to enhance learning and comprehension but not detract from the visual narrative; and ensuring cognitive load was not too high when multiple messages were given to enhance comprehension.

As previously discussed, the 'look' of the animation was shaped by user-group feedback. We plan to carry out a more detailed and rigorous analysis of the effectiveness of the animation and its acceptability to key stakeholders (young people with asthma, parents, teachers, health professionals) via in-depth qualitative interviews and through an online interactive modelling experiment.

5 Conclusion

We hypothesise that through this rigorous blend of multidisciplinary theory and through the use of high quality design, the final animation will be highly attractive for use by health professionals as an intervention and will appeal to young people, their parents and to other key stakeholders. Through developing health (and potentially other) interventions in this way, it is hypothesised that behaviour and attitudinal change can be achieved to the benefit of target users.

References

1. Anandan, C., Nurmatov, U., Schayck, O., Sheikh, A.: Is the prevalence of asthma declining? Systematic review of epidemiological studies. Allergy Asthma Proc. (2009), doi:10.1111/j.1398-9995.2009.02244.x
2. Asher, M.I., Montefort, S., Bjorksten, B., Lai, C.K.W., Strachan, D.P., Weiland, S.K., et al.: Worldwide time trends in the prevalence of symptoms of asthma, allergic rhinoconjunctivitis, and eczema in childhood: ISAAC Phases One and Three repeat multicountry cross-sectional surveys. The Lancet 368(9537), 733–743 (2006)
3. Beasley, R.: The International Study of Asthma and Allergies in Childhood (ISAAC) Steering Committee. Worldwide variation in prevalence of symptoms of asthma, allergic rhinoconjunctivitis, and atopic eczema: ISAAC. The Lancet 351(9111), 1225–1232 (1998)
4. National Asthma Campaign. An audit of children's asthma in the UK. The Asthma Journal 8(2), 3–11 (2002)
5. National Asthma Campaign. Out in the open. A true picture of asthma in the United Kingdom today. The Asthma Journal 6(3), 3–14 (2001)
6. Ram, F.S.F., Robinson, S.M., Black, P.N.: Effects of physical training in asthma: a systematic review. Br. J. Sports Med. 34, 162–167 (2000)
7. van Veldhoven, N.H.M.J., Vermeer, A., Bogaard, J.M., Hessels, M.G.P., Wijnroks, L., Colland, V.T., et al.: Children with asthma and physical exercise: effects of an exercise programme. Clin. Rehabil. 15, 360–370 (2001)

8. Welsh, L., Kemp, J.G., Roberts, R.G.D.: Effects of physical conditioning on children and adolescents with asthma. Sports Med. 35(2), 127–141 (2005)
9. Glazebrook, C., McPherson, A.C., Macdonald, I.A., Swift, J.A., Ramsay, C., Newbould, R., Smyth, A.: Asthma as a barrier to children's physical activity: implications for body mass index and mental health. Pediatrics 118(6), 2443–2449 (2006)
10. Lang, D.M., Butz, A.M., Duggan, A.K., Serwint, J.R.: Physical activity in urban school-aged children with asthma. Pediatrics 113, 341–346 (2004)
11. Meyer, A., Machnick, M.A., Behnke, W., Braumann, K.M.: Participation of asthmaic children in gymnastic lessons at school. Pneumologie 56(8), 486–492 (2002)
12. Sawyer, S.M., Fardy, H.J.: Bridging the gap between doctors' and patients' expectations of asthma management. J. Asthma 40(2), 131–138 (2003)
13. Yeatts, K., Sotir, M., Music, S., Herget, C.: Health consequences for children with undiagnosed asthma-like symptoms. Arch. Pediatr. Adolesc. Med. 157, 540–544 (2003)
14. Los Angeles County Department of Health Services. Los Angeles County health survey. Los Angeles: County of Los Angeles Department of Health Services Public Health (2001), `http://publichealth.lacounty.gov/ha/reports/habriefs/v3i6_asthma/asthm.pdf`
15. Lucas, S.R., Platts-Mills, T.A.E.: Physical activity and exercise in asthma: relevance to etiology and treatment. J. Allergy Clin. Immunol. 115(5), 928–934 (2005)
16. Williams, B., Hoskins, G., Pow, J., Mukhoadhyay, S., Coyle, J.: Low exercise among children with asthma: A culture of over protection? A qualitative study of experiences and beliefs. British Journal of General Practice 60(577), e319–e326 (2010)
17. Williams, B., Anderson, A., Barton, K., McGhee, J.: Can theory be embedded in visual interventions to promote self-management? A proposed model and worked example. International Journal of Nursing Studies 49(12), 1598–1609 (2012)
18. Rich, M., Chalfen, R.: Showing and telling asthma: Children teaching physicians with visual narrative. Visual Sociology 14(1), 51–71 (1999)
19. Walker, I., Hulme, C.: Concrete words are easier to recall than abstract words: Evidence for a semantic contribution to short-term serial recall. Journal of Experimental Psychology: Learning, Memory, and Cognition 25(5), 1256–1271 (1999)
20. Ajzen, I.: The theory of planned behaviour. Organisational Behavior and Decision Processes 50, 179–211 (1991)
21. Leventhal, H., Brisette, I., Leventhal, E.A.: The common-sense model of self-regulation of health and illness. In: Cameron, L.D., Leventhal, H. (eds.) The Self Regulation of Health and Illness Behaviour, pp. 42–65. Routelidge, London (2003)
22. Kennedy, D.M., McNaught, C.: Design elements for interactive multimedia. Australian Journal of Educational Technology 13(1), 1–22 (1997)
23. Keen, S.: Fast tracks to narrative empathy: Anthropomorphism and dehumanisation in graphic narratives. SubStance 40(124), 135–155 (2011)
24. Plowman, L.: Narrative, linearity and interactivity: making sense of interactive multimedia. British Journal of Educational Technology 27(2), 92–105 (1996)

The Design Considerations of a Virtual Reality Application for Heart Anatomy and Pathology Education

Victor Nyamse, Vassilis Charissis, J. David Moore, Caroline Parker, Soheeb Khan, and Warren Chan

Glasgow Caledonian University
School of Engineering and Built Environment,
Department of Computer, Communications and Interactive Systems,
Glasgow, UK
v.charissis@gmail.com

Abstract. Anatomy and pathology of the human body are complex subjects that cannot be elucidated easily to the medical students through traditional description and illustration methods. The proposed interactive system aims to present clear information on demand. For enhancing further the three-dimensional understanding of the anatomical information, a virtual reality environment was developed in order to accommodate different 3D models of the human body. In this case we opted for the heart model as it presents a unique section of the body that can produce motion and sound. The produced model was further simplified for use by patients who wish to understand better the generic anatomy and typical pathologies of the heart. Additionally the paper presents the data results of the system evaluation performed by ten users. The derived results although promising, highlighted some benefits and drawbacks of the proposed system that we aim, to improve in the near future. Finally the paper concludes with a plan of future work which will entail further interactivity through audio incorporation and gesture recognition.

Keywords: Virtual Reality, HCI, 3D Visualization, Heart Disease, Anatomy, Pathology.

1 Introduction

The transfer of anatomical information and knowledge has been of great importance to medical practice for centuries. This has however been sidelined over the past decades and the decline was more predominant in medical schools where, changes to teaching methods, reduction in curricular teaching time and reduction in staff have led to limited teaching of anatomy. Concerns have been raised on the future of anatomical knowledge provided at an undergraduate level and unfortunately, these have been reinforced by reports of errors due to low levels of anatomical knowledge [1].

Initial dismissal of detailed human anatomy teaching was primarily attributed to problems with the educational methods employed in such teaching. The role of traditional methods such as didactic teaching and cadaveric dissection, in a modern

R. Shumaker (Ed.): VAMR/HCII 2013, Part II, LNCS 8022, pp. 66–73, 2013.

medical curriculum has been contested. These methods are obsolete to the current teaching practices due to practical and ethical issues. Dissection, viewed by traditional medical educators as the most effective method of learning anatomy [2], is currently affected by major shortages of cadaveric material as well as a variety of ethical issues involved in the practice. Furthermore, didactic teaching is mainly teacher-centered, not interactive enough and thus unable to stimulate students interests and encourage active student participation in learning [13]. Some institutions have removed these teaching methods from their curricula and are thus seeking more innovative approaches to teaching and learning anatomy. The reason most commonly cited for the reform in practices is that methods were outdated and not in keeping with current student learning needs and demands. In seeking to contribute and develop an innovative product, a virtual reality system was designed to effectively present anatomical information to a variety of audiences.

This paper details the design approach, the method of evaluation and the initial results for a virtual learning environment developed to aid the learning of human anatomy. It also introduces the interface and explains the rationale behind the design, which placed emphasis on a simplified interface and incorporated multi-sensory interactions. The virtual system provides users with the interactive environment to explore the anatomy of the heart using photorealistic three dimensional (3D) representations. It offers real time visualization in an immersive stereoscopic environment, providing anatomical information in a novel manner. The heart was chosen as the initial anatomical model for the virtual reality (VR) system due to the high rate of mortality from cardiovascular disease, especially in Scotland where the study takes place [14]. This renders the anatomy and pathology of the heart a high priority theme of the medical curriculum. The system was designed to suit various users including patient who could visualize the possible effect on the heart of certain risky behavior and in the process prevent future detrimental activities being taken, thus aiding in reducing patient risk factors.

2 Virtual Reality and Anatomical Information

The present reality of medical education and representation of anatomical information makes it ripe for innovation. This necessitates the adoption of a new approach to presenting anatomical information. To this end we created a novel design of an anatomy tutoring application using present advances in technology to augment the effect of the application tool.

Computerized tools have become increasingly important in most institutions that teach anatomy and are widely believed to be the alternative medium of choice. These tools have the ability to result in interactive learning, which is preferred to traditional resources of plastic models, or 2D illustrations that abound in most practitioners' workplaces. This has resulted in the creation of various interactive learning resources [11]. Different computerized tools have been developed, implementing various forms of technology including virtual, augmented and mixed reality. Currently augmented reality is becoming increasingly used in medical in-situ visualization [3]. Sakellariou et al implemented an augmented reality environment for training medics on the complex anatomy of the inguinal canal and had positive feedback in terms of knowledge enhancement and learning experience [4]. Results show that 3D virtual reality has the potential to reduce the learning curve [7]. In assessing the effect of a virtual reality environment, most studies showed the importance of a simplified interface and concluded that future works should include more intuitive interaction [3,4]. Also, these tools achieve increased active student participation in the learning process.

3 System Interface Design

To achieve the set objectives of enhancing knowledge and the learning experience the interface was simplified, as findings from previous research show the importance of a simplified interface [11]. Furthermore the design decisions were made based on their ability to reduce cognitive load and improve knowledge. This theory based design approach mainly relied on research from cognitive psychology.

The interaction was designed to offer the best anatomy learning experience, by providing real-time rotation and intelligent structural identification. This is a deviation from the limitations imposed by pseudo 3D, which is commonly used by most anatomy tutoring application, as pseudo 3D has been found to constrain the learners' ability go learn new anatomical information and impedes the exploration of structures [11]. The identification of individual anatomical structures, an integral part of the learning action, was simplified and made less cognitively stressful, by making it clearly visible. This was implemented with labels that were close to the structure being identified (Figure 1), in line with current understanding on the effect of proximity between images and

Fig. 1. Screenshot of the VR interface presenting a testing low-polygon heart model and the selection of the aorta

associating text [8]. To reduce cognitive load and augment knowledge storage, visual texts were also replaced with auditory cues. This reduced the burden of excess text, which as observed from previous trials disrupts the learning experience [11].

In creating the heart prototype, 3D models were developed by manipulating CT and MRI scan data with further enhancements utilizing 3D Max, a modeling software. This was aimed at increasing photorealism in the models whilst optimizing their polygon number to allow incorporation into Unity 3D gaming engine.

The use of a gaming engine offered unrestricted manipulation of models and liberates the application from limitations associated with pseudo 3D. Previous tutoring system utilized the pseudo 3D technology that entailed a combination of still images put together. Providing real time rotation and interaction improves the learning and enhances the exploration of individual structures.

The interactions, discussed above, were developed with Unity3D. The particular game engine was deemed ideal for such system as it offers great flexibility and compatibility with a vast number of other software packages and applications. To create the ideal immersive environment for the iterative development process, we utilised the Virtual Reality and Simulation Laboratory (VRS Lab) at Glasgow Caledonian University which offers large stereoscopic projection (3x2.5m) in combination with surround sound system and gesture recognition capacity.

Notably during the design and development the system was tested in the specific facility it was our intention to produce a final application that could perform equally well even with the use of a laptop. The latter hardware was able to reproduce the stereoscopic effect through the appropriate graphics card and monitor. Due to cost efficient requirements the system portability was one of the main objectives of the proposed system.

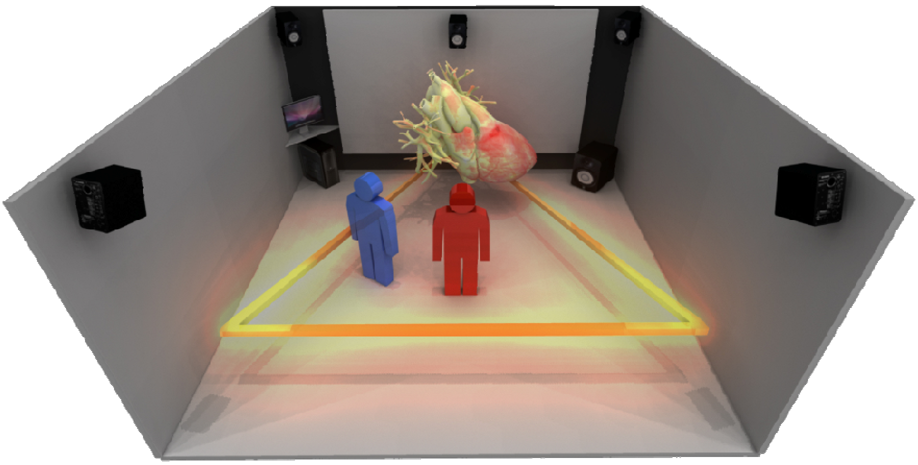


Fig. 2. Virtual Reality and Simulation Laboratory (VRS Lab) configuration which supports large scale stereoscopic projection and 3D surround sound system

4 Case Study Evaluation

This prototype application was designed primarily to suit the group of medical student users. Yet it was made clear that such as system could be utilised routinely for also enhancing patient information on related heart disease issues.

Although the learning requirement of these two groups differ, the application was designed so that depth of transferable anatomy knowledge could be adjusted; A detailed representation of the anatomy of the heart was presented to the medical students whilst a simplified version was presented to patients. Previous studies have shown that good anatomy grounding is beneficial to understanding pathology[11]. Patients' knowledge of anatomy is generally insufficient to fully comprehend the information provided to them by medical practitioners. This is further hindered by the inability of the average patient to reconstruct a mental map of verbally communicated anatomical terms and conditions by their medical practitioner. Studies have indicated that verbally presented information in the context of a medical consultation is neither well retained nor adequately comprehended by patients [15]. Real-time 3D medical visualisation offers a unique platform of communication between doctors and patients which could be employed easily in any computer or mobile device in a timely manner in order to aid information communication.

This creates a shift from traditional paper based material and aims to improve the ability of doctors and other health practitioners to convey simplified yet accurate anatomical information to patients. This could potentially help them make appropriate decisions, from life-style changes to coping with their diagnosis.

For the first stage of the evaluation of this application, a heart anatomical model was employed for an audience of lay public i.e patients. The heart model was initially used for evaluation as cardiovascular pathologies contribute to the highest mortality and morbidity statistics in Britain, and comparatively more in Scotland [14]. A photorealistic model of the heart was created from CT data. The developed model was accessed through a simple and clear interface which entail all the typical three-dimensional interactions such as rotation, zoom, pan and selection of individual sections. The latter are annotated in analogous fashion to medical anatomy books.

5 Results and Discussion

For the initial evaluation of the system we requested from ten users to test the system and provide us with feedback through a Likert scale (1 to 5) questionnaire and one to one interviews. The individuals chosen were members of the public with no specific medical background. Thus they had little prior knowledge on the anatomy of the heart. The questions presented to them after using the application are noted below:

1. The system helped me create a 3D mental map of the heart
2. I can remember most of the structures of the heart.
3. This learning experience was enjoyable.
4. The placement of labels helped me identify structures
5. I have a better understanding of the heart's anatomy

6. I can understand various cardiac pathologies after using this system.
7. Unrestricted rotation improved my learning experience
8. I can use this system on my own time without required support
9. Medical terminology was demystified
10. I would like to see similar applications for other vital organs

This preliminary evaluation offered succinct yet crucial information with regards to the system performance and acceptability from the users as presented in Figure 3. More analytically the system scored over 80% in all except one (question 6) of the questions. Question 6 scored 78% as some users felt that they can have a better understanding of particular pathologies than previously, although they were not sure of how much in-depth information they should acquire. Some of the users commented that even the simplified system presented some information probably more relevant to medical practitioners.

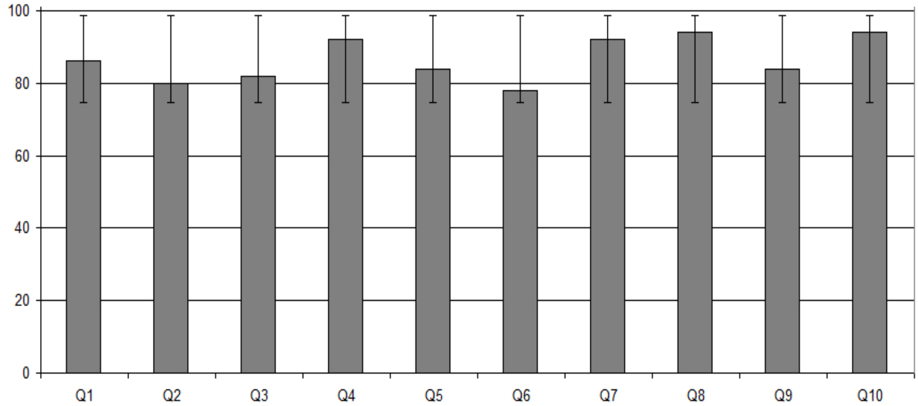


Fig. 3. Scores in percentage for each question derived from the 10 users

From a visualisation point of view the current 3D model of the heart was intentionally reduced to very low polygon count so as to optimise the system performance and test it in a typical laptop. The users noted that although the refresh rate of the stereoscopic visualisation was adequate, the 3D model requires additional smoothing of the surfaces and further optimisation.

The interface was considered user-friendly without any unconventional and unexpected manipulation tools. However we intent to improve significantly the system interactivity with the use of gesture recognition and audio tools. To this point, the users were in favour of annotations although a voice over description could further assist with the pronunciation of some complex terminologies.

Overall the system offered a good experience for the users and delivered to its promise to enhance the medical understanding of the users. Interestingly the users suggested that the particular system could be very useful for other organs and pathologies of the human body.

6 Conclusions

This work presented the design process of a prototype medical education system. For experimentation purposes we choose to develop a detailed 3D model of a heart. The proposed interface was developed with a view to support both medical students and patients. With this in mind the system offers a two fold approach which can facilitate both requirements with minor system alterations.

The first version of the system was evaluated by ten users and with the use of Likert questionnaire and post-trial user interviews. The preliminary results were promising and the subjective feedback highlighted some benefits but also drawbacks of the application that are planned for amendment in future versions.

For future work we plan to extend the visual information and blend it with 3D audio which will enhance the understanding of different heart diseases. We anticipate that the correlation of visuals and audio will offer a more realistic representation of heart pathologies and convey the subject matter in a clearer and more intuitive manner.

References

1. Ellis, H.: Medico-legal Litigation and its Links with Surgical Anatomy. Surgery 20(8), i–ii (2002)
2. Mclachlan, J.C., Pattern, D.: Anatomy teaching: ghosts of the past, present and future. Medical Education 40(3), 243–253 (2006)
3. Sritharan, K.: The rise and fall of anatomy. Student BMJ 331, 309–352 (2005)
4. Chien-Huan, C., Chien-Hsu, C., Tay-Sheng, J.: An Interactive Augmented Reality System for Learning Anatomy Structure. In: Proceedings of the International MultiConference of Engineeers and Computer Scientists, IMECS 2010, Hong Kong (2010)
5. Blum, T., Kleeberger, V., Bichlmeier, C., Navab, N.: Mirracle: An augmented reality magic mirror system for anatomy education. 2012 IEEE Virtual Reality Workshops (VR), 115–116 (2012)
6. Charissis, V., Zimmer, C., Sakellariou, S., Chan, W.: Exploring the simulation requirements for virtual regional anesthesia training. In: Society of Photo-Optical Instrumentation Engineers (SPIE) Conference Series, p. 4 (2010)
7. Leahy, W., Sweller, J.: Cognitive load theory, modality of presentation and the transient information effect. Applied Cognitive Psychology 25(6), 943–951 (2011)
8. Wickens, C., Andre, A.: Proximity compatibility and information display: Effects of color, space, and objectness on information integration. Human Factors 32, 61–77 (1990)
9. Champney, T.: A proposal for a policy on the ethical care and use of cadavers and their tissues. Anatomical Science Education 4, 49–52 (2011)
10. Charissis, V., Ward, B.M., Naef, M., Rowley, D., Brady, L., Anderson, P.: An Enquiry into VR Interface Design for Medical Training: VR Augmented Anatomy Tutorials for Breast Cancer. In: McDowall, I.E., Dolinsky, M. (eds.) Proceedings of the: International Annual Symposium of IS&T/SPIE, The Engineering Reality of Virtual Reality, San Jose, California, vol. 6804, pp. 19–28 (2008) ISBN: 9780819469762

11. Sakellariou, S., Charissis, V., Grant, S., Turner, J., Kelly, D., Christomanos, C.: Virtual Reality as Knowledge Enhancement Tool for Musculoskeletal Pathology. In: Shumaker, R. (ed.) Virtual and Mixed Reality, Part II, HCII 2011. LNCS, vol. 6774, pp. 54–63. Springer, Heidelberg (2011)
12. Nielsen, J., Landauer, T.K.: A mathematical model of the finding of usability problems. In: Proceedings of ACM INTERCHI 1993 Conference, Amsterdam, The Netherlands, April 24-29, pp. 206–213 (1993)
13. Older, J.: Anatomy: A must for teaching the next generation. The Surgeon 2(2), 79–90 (2004)
14. Allender, S., Peto, V., Scarborough, P., Kaur, A., Rayner, M.: Coronary heart disease statistics (2008)
15. Miller, R.: Approaches to Learning Spatial Relationships in Gross Anatomy: Perspective from Wider Principles of Learning. Clinical Anatomy 13, 439–443 (2000)
16. Leahy, W., Sweller, J.: Cognitive load theory, modality of presentation and the transient information effect. Applied Cognitive Psychology 25(6), 943–951 (2011)

Human-Computer Confluence for Rehabilitation Purposes after Stroke

Rupert Ortner[1,*], David Ram[2], Alexander Kollreider[2], Harald Pitsch[2], Joanna Wojtowicz[3], and Günter Edlinger[1]

[1] Guger Technologies OG, Graz, Austria
{ortner,edlinger}@gtec.at
[2] Tyromotion GmbH, Graz, Austria
{David.Ram,Alexander.Kollreider,Harald.Pitsch}@tyromotion.com
[3] Akademia Górniczo-Hutnicza im. Stanisława Staszica w Krakowie
jo.wojtowicz@gmail.com

Abstract. In this publication, we present a Motor Imagery (MI) based Brain-Computer Interface (BCI) for neurologic rehabilitation. The BCI is able to control two different feedback devices. The first one is a rehabilitation robot, moving the fingers of the affected hand according to the detected MI. The second one presents feedback via virtual reality (VR) to the subject. The latter one visualizes two hands that the user sees in a first perspective view, which open and close according to the detected MI. Four healthy users participated in tests with the rehabilitation robot, and eleven post stroke patients and eleven healthy users participated to tests with the VR system. We present all subjects' control accuracy, including a comparison between healthy users and people who suffered stroke. Five of the stroke patients also agreed to participate in further sessions, and we explored possible improvements in accuracy due to training effects.

Keywords: Medical and healthcare, Applications: Rehabilitation.

1 Introduction

Brain-computer interface (BCI) technology has been used widely for communication and device control in a closed loop system [1]. The choice of the BCI approach depends on which device the user wants to control. The P300 and steady state visual evoked potentials (SSVEP) approaches are based on evoked potentials. Hence, they need an external stimulation device, and are not useful for people suffering from visual impairments if visual stimulation is used. The approach based on changes in sensorimotor rhythms (SMR) is the third popular one. These BCIs rely on power changes in the mu- (8Hz-12Hz) and beta bands (18Hz-26Hz) over regions active during motor imagery (MI). These rhythms are associated with the cortical areas most directly connected to the brain's normal neuromuscular outputs [1]. The MI based

* Corresponding author.

R. Shumaker (Ed.): VAMR/HCII 2013, Part II, LNCS 8022, pp. 74–82, 2013.

BCI was already successfully used for helping people suffering motor impairments. Pfurtscheller et al. demonstrated the MI-BCI based control of functional electrical stimulation for restoring hand grasp in a patient with tetraplegia [2], Millan et al. controlled an intelligent wheelchair via executing three different mental tasks [3], and other systems have been described (e.g. [4], [5]). Recently, the idea of utilizing the MI for neurological rehabilitation became popular. The idea is to use the BCI not to replace lost motor function, but to improve motor functions in patients.

Prior research showed that mentally rehearsing movements (that is, performing MI) could be used as an effective therapy in stroke rehabilitation [6] even if no feedback about the performance is given to the user. MI may be a method to overcome learned nonuse in chronic stroke patients, and could also be practiced by patients with poor motor performance, which otherwise excludes four out of five patients from active movement therapies [7]. A review, comparing the effects of conventional therapy plus MI to those of only conventional therapy proved the positive effects of MI interventions [8]. Zimmermann-Schlatter et al. identified four studies performed in Asia and North America. Two of them found significant effects on the Fugl-Meyer Assessment (FMA) score and in the Action Research Arm Test. One study only found significant effects in the task related outcomes.

The additional advantages of not only performing MI, but also tracking MI with a BCI and presenting online feedback to the user, seem clear: (i) the feedback helps and motivates the patient to perform accurate MI, (ii) the therapist gets feedback about the performance of MI and can track changes over time, and (iii) real-time feedback may increase Hebbian plasticity, which is likely to increase cortical activity [9]. To test this approach, Ang et al. compared rehabilitation success across 54 hemiparetic stroke patients who received either standard robotic rehabilitation or rehabilitation with a MI-BCI and robotic feedback [10]. They showed that significant gains in FMA scores were observed in both groups at post-rehabilitation and 2-month post-rehabilitation, but no significant differences were observed between groups. Furthermore, they proved that hemiparetic stroke patients can operate EEG-based MI-BCI, and that EEG-based MI-BCI with robotic feedback neurorehabilitation is effective in restoring upper extremity motor function after stroke.

This manuscript presents a MI based Brain-Computer Interface (BCI) that can control different feedback devices. The BCI was connected either to an upper limb rehabilitation robot (Amadeo, Tyromotion GmbH, Austria) or a Virtual Reality (VR) system (gVRsys, g.tec medical engineering GmbH, Austria). Both the VR system and the rehabilitation robot provide online feedback to the user about the detected MI. A total of eleven post-stroke patients and a control group of eleven healthy people took part in the VR based experiment. First results from 4 healthy users performing the experiment with sensory feedback with the rehabilitation robot are presented.

2 Methods

2.1 Detection and Classification of MI

For better classification of MI via a Linear Discriminant Analysis (LDA), the EEG channels are spatially filtered with Common Spatial Patterns (CSP). This method yields a set of spatial filters designed to minimize the variance of one class while maximizing variance for the other class. For proper classification, it is sufficient to choose only the four best discriminating filters. These are the two filters leading to the highest variance of class one, and the two filters generating the highest variance of class two (while each of the four filters minimizes the variance of the other class). Given N channels of EEG for each left and right trial, the CSP method provides an N x N projection matrix. This matrix is a set of subject-dependent spatial patterns, which reflect the specific activation of cortical areas during hand movement imagination. With the projection matrix W, the decomposition of a trial X is described by

$$Z = WX \tag{1}$$

This transformation projects the variance of X onto the rows of Z and results in N new time series. The columns of W^{-1} are a set of CSPs and can be considered as time-invariant EEG source distributions. Due to the definition of W, the variance for a left hand movement imagination is largest in the first row of Z and decreases as the number of subsequent rows increases. The opposite occurs for a trial with right hand motor imagery. For classification of the left and right trials, the variances have to be extracted as reliable features of the newly designed N time series. However, it is not necessary to calculate the variances of all N time series. The method provides a dimensionality reduction of the EEG. Mueller-Gerking et al. [11] showed that the optimal number of CSPs is four. Following their results, after building the projection matrix W from an artifact corrected training set X_T, only the first and last two rows (p=4) of W are used to process new input data X. Then the variance (VAR_p) of the resulting four time series is calculated for a time window T. These values are normalized and log transformed according to the formula:

$$f_p = \log_{10}\left(\frac{VAR_p}{\sum_{p=1}^{4} VAR_p}\right) \tag{2}$$

Where f_p (p=1..4) are the normalized feature vectors and VAR_p is the variance of the p-th spatially filtered signal. These four features can be classified with a linear

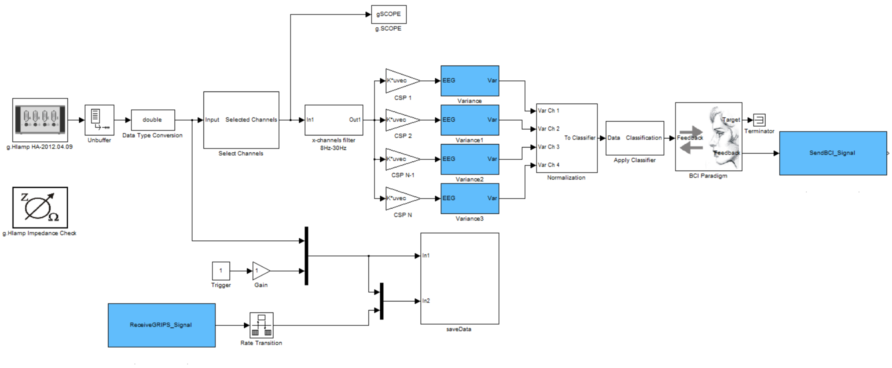


Fig. 1.

discriminant analysis (LDA) classifier. For a very good overview of the CSP method, please see [12], [13].

2.2 Experimental Workflow

The BCI experiment was set up with g.BCIsys, as shown in the Simulink model in Fig. 1. The data were recorded over 64 positions (see Fig. 2A) distributed over the cortex and sampled at 256 Hz. Active EEG electrodes (g.LADYbird) were used to make the preparation procedure faster and easier and to increase data quality. A g.HIamp biosignal amplifier (g.tec medical engineering GmbH, Austria) was used for data recording. The unit has 256 ADCs with 24 bit precision and performs oversampling to increase the signal to noise ratio. Before applying the spatial filters, the EEG data were converted to double precision and bandpass filtered between 8 and 30 Hz. Then, the variance was calculated within a time-window of 1.5s length. These features were normalized, log transformed and classified with the LDA. The LDA classification result drives the BCI Paradigm feedback block. This block controls the paradigm timing and sends the feedback commands to the external feedback device, either the rehabilitation robot or the VR system. The block ReceiveGRIPS_Signals tracks the actions of the rehabilitation robot and saves this data in synchrony with the EEG data for offline analysis.

2.3 Session Timing

One experimental run lasted about six minutes and contained 40 randomized commands of either left-hand or right-hand MI. Fig. 1B shows the trial timing. One trial lasted eight seconds. A random intertrial interval between 0.5 and 1.5 seconds was included between each trial. The cue (command) was presented at 3 seconds. The feedback phase lasted from three seconds until the end of the trial.

2.4 Robotic Feedback Paradigm

Four healthy subjects (mean age 24 ± 5.2 years, 2 left-handed, 2 right-handed) participated in the tests with the rehabilitation robot Amadeo (see Fig. 3B). The Amadeo is a mechatronic finger rehabilitation device that allows each individual finger, including the thumb, to move independently and separately. The positions, as well as the forces of each finger, were measured constantly during the paradigm and saved with the EEG data, allowing detailed offline analysis. One experimental paradigm consisted of 4 runs. With the data of the first 3 training runs, a specific classifier for detecting the MI was generated. No feedback was presented during the training runs. In the following run, this classifier was tested and the online error rate was calculated with the novel run. The robot gave feedback to only one of the two hands. In a real rehabilitation session, this would be the affected side. For the healthy subjects, we selected the dominant hand to receive robotic feedback. The cue was given via a red arrow, pointing either to the left side or the right side of a computer

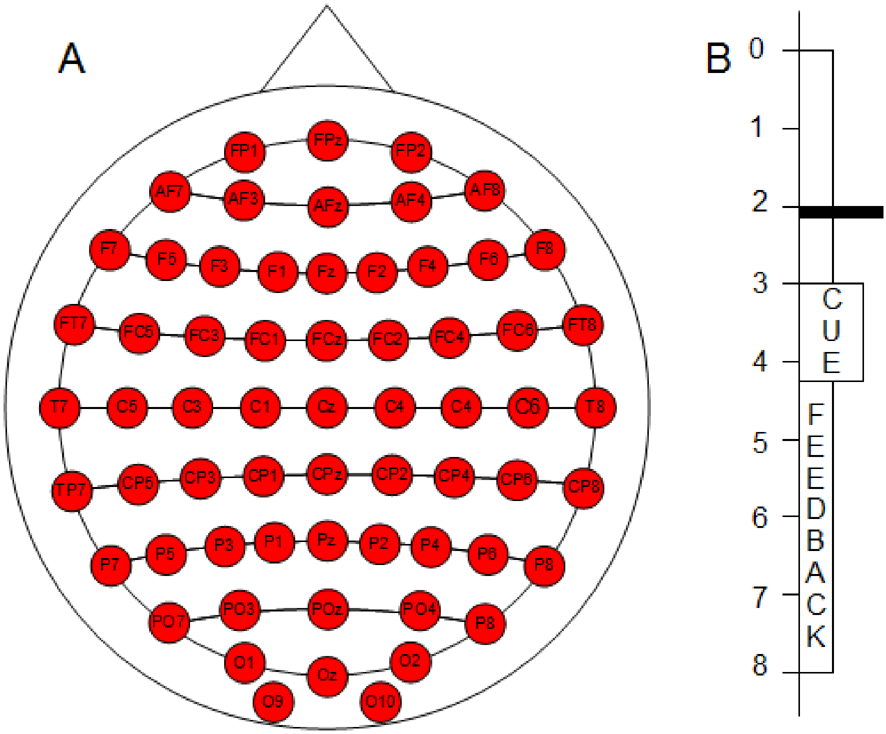


Fig. 2. A: Positions of the 64 active EEG electrodes. The ground electrode was placed on the forehead (near FPz) and the reference on the right earlobe. B: Timing of one feedback trial for either the robotic feedback paradigm or VR feedback paradigm.

screen. If the cue pointed to the side that was not fixed in the rehabilitation robot, the subject was asked to perform a real (not imagined) full flexion and extension of his/her fingers. If the cue pointed to the other hand, the subject was asked to instead imagine the same movement. When the correct MI was detected, the robot provided feedback by performing a flexion and extension of the five fingers. Within one trial, only one full flexion and extension was done. If no correct MI was detected during the feedback phase of the trial, then the robot performed no movement.

2.5 VR Paradigm

Eleven post-stroke patients (mean age 67.5 ± 10.3 years) and eleven healthy subjects (mean age 22.3 ± 4.2 years) participated in the tests with the VR system. The measurements with stroke patients were performed at the Krzeszowice Rehabilitation Center, Poland. The measurements with the healthy users were performed at Guger Technologies OG, Austria. Because the patients were not able to participate in longer sessions, the number of runs was reduced to either three, or sometimes two, runs per session. Two runs were conducted if the user was very tired or did not feel good.

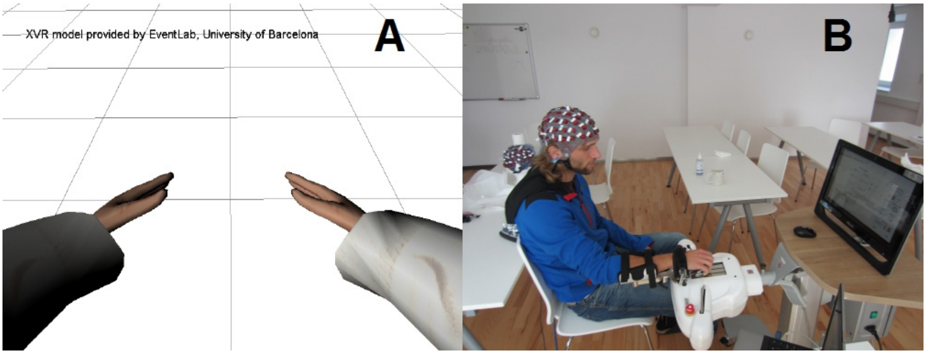


Fig. 3. The two feedback strategies. (A) VR paradigm. (B) Robotic feedback paradigm.

Hence, training data recorded during such sessions were not sufficient to set up a subject specific classifier. Therefore a generic classifier (generated from a large pool of previously recorded MI sessions of other users) was used. For comparison, the same procedure was tested on the group of eleven healthy users, using the identical generic classifier and always with three runs. Feedback was presented for both hands, visualizing the user's hand in VR as they seen in the user's first perspective (see Fig.3A). The cue was presented by flexion and extension of the left or right hand. After the cue phase, the user had to imagine the same flexion and extension as seen during the cue phase. A beep indicated the start of the cue phase. A second beep indicated the end of it and the beginning of the feedback phase. The feedback was then presented as flexion and extension of the detected hand side of MI, thus presenting real-time online feedback to the user. If during the feedback phase the detected hand side changed, then the feedback also flipped from one hand to the other.

3 Results

Table 1 shows the mean accuracy of the group of eleven stroke patients and the control group of eleven healthy users. Five of the stroke patients participated in four further sessions. For this extra comparison, the results of this group after the first session and after the fourth session are depicted in the last two columns.

Table 2 shows the results of the healthy users performing the robotic feedback sessions. The mean accuracy across the four users was 86.55%. The accuracy level is averaged over 40 trials.

Table 1. Mean accuracy rates of the two groups participating in the VR paradigm

	Healthy	Stroke		
Session #	1	1	1	4
Participants	11	11	5	5
Mean Acc.	63.77	60.67	59.7	72.48
SD	16.52	13.05	6.08	8.45

Table 2. Accuracy rates of the healthy users participating in the robotic feedback paradigm

Session #	Accuracy (%)
1	92.50
2	95.00
3	68.70
4	90,00
mean Acc.	86.55
SD	13.97

4 Discussion

The aim of this study was to evaluate a novel rehabilitation strategy, which can present feedback either via a rehabilitation robot or a VR system. The error rate during online control was calculated. The robotic feedback was tested only on healthy users with a specific classifier for each session. The groups testing the VR feedback used a generic classifier. This is the reason why the classification result of the latter groups is lower than for the robotic feedback group. The difference in control accuracy between healthy users and stroke patients is only about 3% on average, although the mean age of the stroke patients (67.5 years) is much higher than that of the healthy control group (22.3 years). One very important finding of the study is the improvement of control accuracy of the stroke patients during only 4 training sessions. As could be seen in Table 1, they improved from 59.7% to 72.48%. The motivation of the user and the advances in the rehabilitation process due to the BCI approach depends on the accuracy of the BCI, hence these improvements seem very promising.

One difference between the two feedback approaches was the delay in presenting the feedback. The VR feedback gave feedback in real-time and in synchrony to the MI. If the MI changed during the feedback phase, then also the feedback changed. For the robotic feedback, the user had to first perform the MI, then a full flexion and extension was performed by the robot, regardless of what the user did while the robot moved. In a recent publication, Ramos-Murguialday et al. called this approach discrete proprioceptive feedback, and stated that the feedback contingency is of vital importance to enable neuro-motor-rehabilitation [14]. Gomez-Rodriguez et al. also wrote that synchronization is likely to increase cortical plasticity due to Hebbian-type learning, and could improve the functional recovery [9]. For future studies, we aim to adapt the robotic feedback that way to deliver synchronized online feedback, similar to the VR feedback approach.

The advantage of the robot is that it delivers both visual and proprioceptive feedback, which can stimulate the afferent pathways even more than the VR based feedback and thereby could be more effective. Another future goal will be to investigate the combination of the two rehabilitation strategies.

The BCI communicates to the VR system and the robot via an interface that is based on UDP (g.UDPinterface, g.tec medical engineering GmbH, Austria). With this generic interface, it is easy to create other feedback devices, and we will also evaluate functional electrical stimulation.

Acknowledgements. This research is supported by the European Union FP7 Integrated Projects BETTER (No. 247935) and CSI (120209).

References

1. Wolpaw, J.R., Birbaumer, N., McFarland, D.J., Pfurtscheller, G., Vaughan, T.M.: Brain-computer interfaces for communication and control. Clin. Neurophysiol. 113, 767–791 (2002)
2. Pfurtscheller, G., Müller, G.R., Pfurtscheller, J., Gerner, H.J., Rupp, R.: "Thought"-control of functional electrical stimulation to restore handgrasp in a patient with tetraplegia. Neuroscience Letters 351, 33–36 (2003)
3. del Millan, J.R., Galan, F., Vanhooydonck, D., Lew, E., Philips, J., Nuttin, M.: Asynchronous Non-Invasive Brain-Actuated Control of an Intelligent Wheelchair. In: Conf. Proc. IEEE Eng. Med. Biol. Soc., vol. 2009, pp. 3361–3364 (2009)
4. Pfurtscheller, G., Guger, C., Müller, G., Krausz, G., Neuper, C.: Brain oscillations control hand orthosis in a tetraplegic. Neurosci. Lett. 292, 211–214 (2000)
5. King, C.E., Wang, P.T., Mizuta, M., Reinkensmeyer, D.J., Do, A.H., Moromugi, S., Nenadic, Z.: Noninvasive Brain-Computer Interface Driven Hand Orthosis. In: 33rd Annual International Conference of the IEEE EMBS Boston, Massachusetts USA, August 30-September 3 (2011)
6. Sharma, N., Pomeroy, V.M., Baron, J.-C.: Motor imagery: a backdoor to the motor system after stroke? Stroke 37, 1941–1952 (2006)
7. Grotta, J., Noser, E., Ro, T., Boake, C., Levin, H., Aronowski, J., Schallert, T.: Constraint-induced movement therapy. Stroke 35, 2699–2701 (2004)
8. Zimmermann-Schlatter, A., Schuster, C., Puhan, M.A., Siekierka, E., Steurer, J.: Efficacy of motor imagery in post-stroke rehabilitation: a systematic review. J. Neuroeng. Rehabil. 5, 8 (2008)
9. Gomez-Rodriguez, M., Grosse-Wentrup, M., Hill, J., Gharabaghi, A., Schölkopf, B., Peters, J.: Towards Brain-Robot Interfaces in Stroke Rehabilitation. In: Conf Proc. IEEE International Conference on Rehabilitation Robotics 2011 (2011)
10. Ang, K.K., Guan, C., Ang, K.S.G.B.T., Kuah, C., Wang, C., Phua, K.S., Chin, Z.Y., Zhang, H.: Clinical study of neurorehabilitation in stroke using EEG-based motor imagery brain-computer interface with robotic feedback. Conf. Proc. IEEE Eng. Med. Biol. Soc. 2010, 5549–5552 (2010)
11. Müller-Gerking, J., Pfurtscheller, G., Flyvbjerg, H.: Designing optimal spatial filters for single-trial EEG classification in a movement task. Clinical Neurophysiology 110, 787–798 (1999)
12. Guger, C., Ramoser, H., Pfurtscheller, G.: Real-Time EEG Analysis with Subject-Specific Spatial Patterns for a Brain–Computer Interface (BCI). IEEE Trans. Rehab. Eng. 8, 447–456 (2000)

13. Blankertz, B., Tomioka, R., Lemm, S., Kawanabe, M., Müller, K.-R.: Optimizing Spatial Filters for Robust EEG Single-Trial Analysis. IEEE Signal Processing Magazine 25(1), 41–56 (2008)
14. Ramos-Murguialday, A., Schürholz, M., Caggiano, V., Wildgruber, M., Caria, A., Hammer, E.M., Halder, S., Birbaumer, N.: Proprioceptive Feedback and Brain Computer Interface (BCI) Based Neuroprostheses. PLoS ONE 7(10), e47048 (2012)
15. Wolpaw, J.R., Boulay, C.B.: Brain-Computer Interfaces. In: Graimann, B., Allison, B., Pfurtscheller, G. (eds.), pp. 29–46. Springer (2010)

Projected AR-Based Interactive CPR Simulator

Nohyoung Park, Yeram Kwon, Sungwon Lee, Woontack Woo, and Jihoon Jeong

Department of GSCT, Korea Advanced Institute of Science and Technology,
305-701 Daejeon, Republic of Korea
{nypark,yeram,plaming,wwoo,jjeong}@kaist.ac.kr

Abstract. In this paper, we propose a new approach of a cardiopulmonary resuscitation (CPR) simulation system that exploits both AR-based visualization and embedded hardware sensing techniques. The proposed system provides real-time interactive visual feedback to the CPR trainee with the projected AR indicator plane that visualizes results of an interlocking signal of the trainee's actions using embedded sensors. This system also provides proper guidelines about the CPR trainee's posture by detecting a user's articular pose from a RGB-D camera in real-time. As implementation results, our system provides interactive feedback, that enabling more accurate and effective training experience to the trainee and more cost-effective rather than traditional CPR education training systems.

Keywords: CPR simulation, AR-based CPR simulator, augmented reality training system, augmented reality simulation, projected augmented reality.

1 Introduction

Although it is seriously related to human life and death, CPR implementation rate in real-life emergency situations by those who have CPR education experience is relatively low [1–3] because current CPR education depends mainly on a theoretical and limitative approach; for example, practicing on a nonresponsive mannequin model and watching a recorded video or slide presentation. Use of a passive CPR mannequin model does not provide real-time feedback even when trainees' procedure is correct. Some educational CPR mannequin tools for common users show a hardware-based reaction, but it is not providing real time or accurate feedback with respect to trainees' behavior. However, a professional medical CPR mannequin or a system, which can offer proper reaction and feedback, is highly expensive. Therefore, with traditional CPR training education processes, it is difficult to apply adequate behavior when responders encounter an emergency situation.

To solve this problem, we consider with two different approaches: The first approach is a hardware-based sensing technique. There are some current studies related to CPR mannequin models, mainly focused on hardware improvement of sensor itself and hardware design for simulation control [4]. They both consider sensing accuracy for artificial respiration and cardiopulmonary press by a trainee. However, improvement of hardware accuracy itself does not makes it possible to

R. Shumaker (Ed.): VAMR/HCII 2013, Part II, LNCS 8022, pp. 83–89, 2013.

experience real-time feedback to a CPR trainee; further, it has educational limits as training models with respect to the trainees' experience in various situations. Meanwhile, some research about virtual reality and augmented reality show that the VR or AR environmental experienced by the users enhances their spatial presence [5–7] or coexistence [8–10] about indirect experience of various situations. In addition, [11-14] researches showed that VR or AR-based real-time simulation enables direct visual feedback to users.

Thus, in this paper, we propose a new approach of training simulation system that combines both hardware-based sensing and projected AR-based interactive visualization techniques, especially on a CPR simulation system.

2 Proposed Projected AR-Based CPR Simulator

This system consists of hardware-based sensor module, pose control module using camera, and projected AR-based real-time visualization module. The hardware sensor module installed inside the CPR mannequin model detects and gives digital output of the cardiopulmonary press, artificial respiration rate given by the trainee, and the pose control module calculates the accuracy of the trainee's pose using the RGB-D sensor. Using these data, the visualization module provides real-time feedback to the trainee by visualizing computed information. Figure 1 shows the overall architecture of the system.

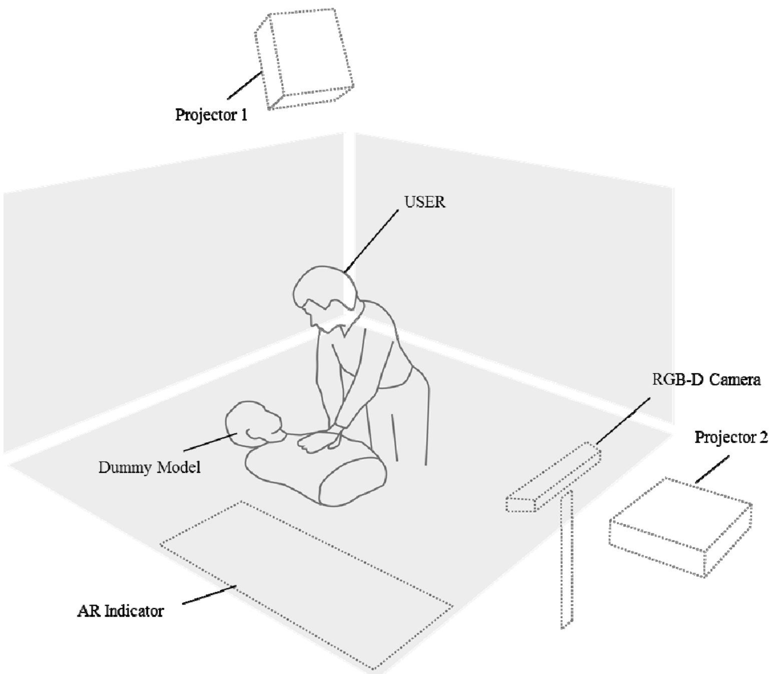


Fig. 1. The concept diagram of the projected AR-based interactive CPR simulator

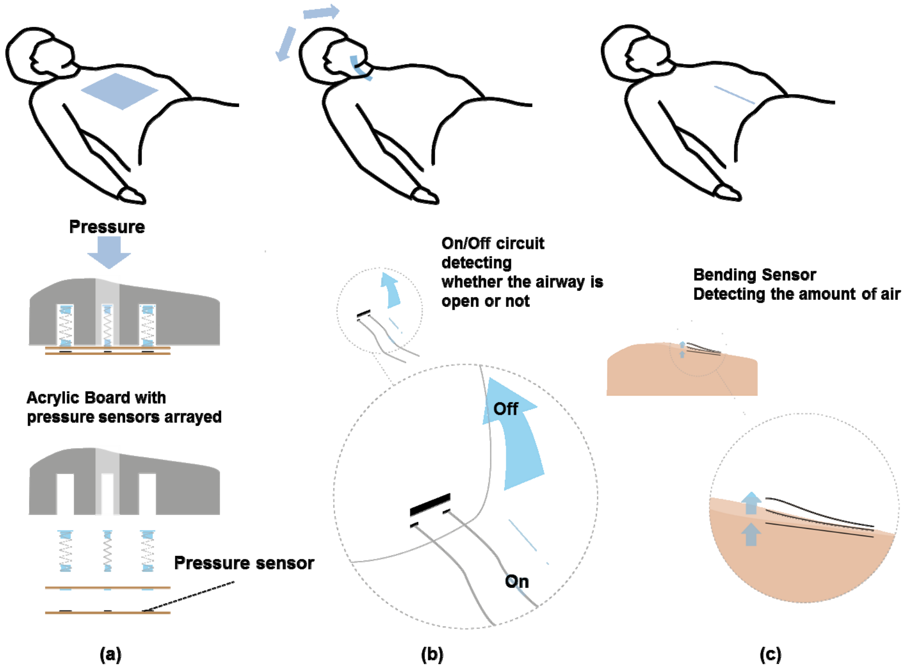


Fig. 2. The concept diagram of embedded hardware. (a) Pressure sensor; (b) On-off circuit; (c) Bending sensor.

2.1 Hardware Sensor

To obtain precise strength of cardiopulmonary press and artificial respiration data, sensors controlled by Arduino [15] are installed in the model. A pressure sensor and a bending sensor are connected directly to Arduino. The analog signals from embedded sensors translate as digital signals through Arduino. Arduino then sends the digital signal from the sensors to the visualization module through a serial port or Bluetooth connection.

In Figure 2(a), four cross-arrayed pressure sensors, wedged between acrylic boards, detect the cardiopulmonary pressure, rate, and pressurized position determined by the user's action on the mannequin model. Through the simple circuit attached on the model's airway, Figure 2(b), the system is able to perceive whether the model's airway is open when artificial respiration is needed. The bending sensor attached on the model's stomach as shown in Figure 2(c), which is inflated when the air comes into the airway, right beneath the model's skin, to detect if the user is giving proper amount and pressure of air.

2.2 User Pose Control

The user pose control module provides proper guidelines to the CPR trainee by extracting his or her articular pose data and comparing it with the desired data. In this module, an RGB image is used to extract 2D image coordinates of four color markers attached on the side of the trainee's shoulder, arm, and waist. The system also obtains the 3D point data of colored markers from corresponding depth image. Using those 3D point data, the module computes information – such as whether the angle between the arm and waist is proper and if the elbow is straight as desired – and provides real-time visual information to the trainee. Figure 4 shows the visual effect that is shown to the trainee.

2.3 Projected AR-Based Visualization

The visualization module's goal is properly visualizing the sensor and user's articular pose data driven from 2.1 and 2.2, considering the environment and circumstances. Therefore, projected AR-based technique is used in this system to maximize the sense of realism and visualize certain possible situations.

As shown in Figure 3, in the real environment where the CPR simulation is executed, the computed graphical information based on the sensing and pose information is shown in real time. Through this, the trainee is able to get real-time visual feedback of artificial respiration and cardiopulmonary.

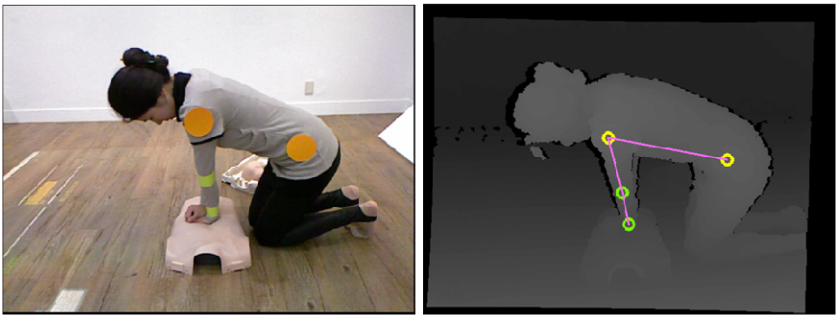


Fig. 3. Articular pose estimation with RGB-D sensor

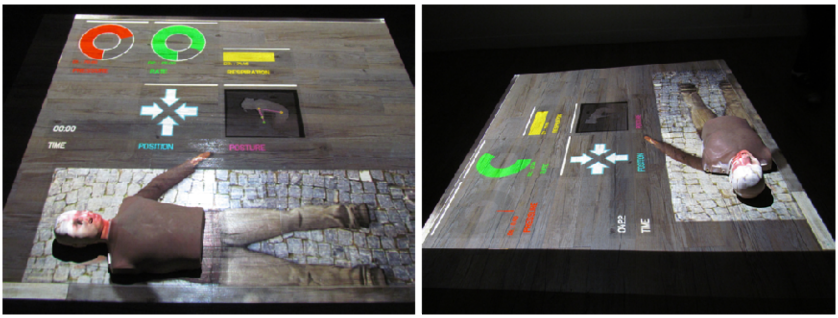


Fig. 4. Projected AR-based visualization

3 Implementation Result

As shown in Figure 5, we preceded the CPR training course from beginning to end with the proposed system. The result showed that the system successfully received the data from the 8 trainees and, each time, displayed accurate artificial respiration and cardiopulmonary press and rate information through the projected AR-based graphic indicators, augmented in real space during the procedure. As shown in Figure 6, three kinds of well-designed hardware were mounted properly. During the eight simulations, the hardware module accurately sensed the user's action on the mannequin, and it provided detected information to the visualization module stably.

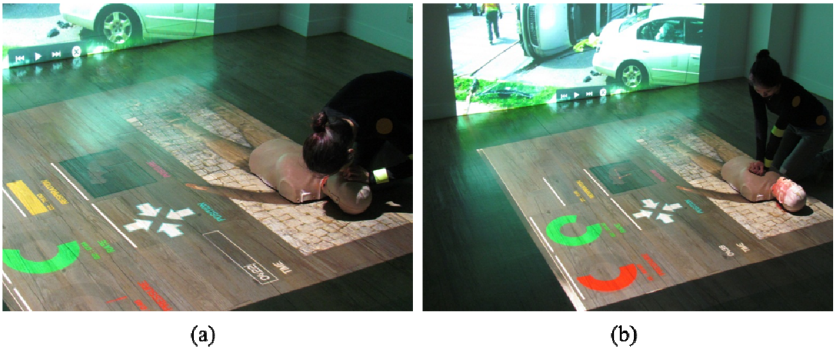


Fig. 5. CPR training process with the proposed system. a) Artificial respiration; (b) Cardiopulmonary press.

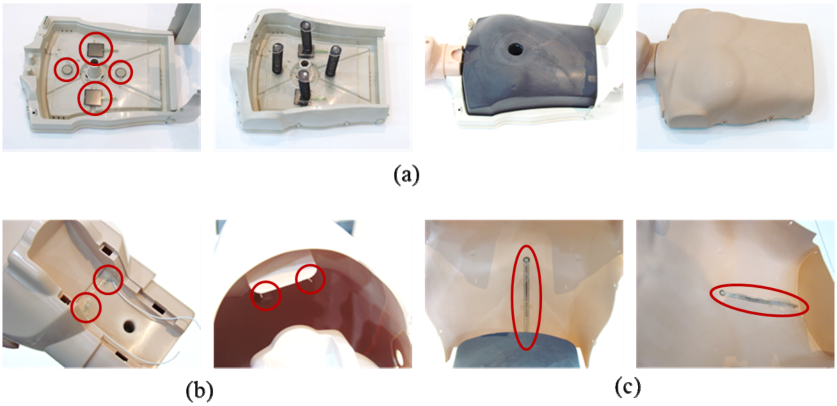


Fig. 6. Embedded hardware sensors. (a) Array of pressure sensor; (b) On-off circuit; (c) Bending sensor.

On the other hand, a problem occurred in the user pose control module. Because of the projected AR-based visualization, we turned the lights low. Therefore, the RGB image-based color extraction algorithm failed frequently.

4 Future Work

We expect that the proposed AR-based interactive CPR simulation system will have more effective educative experiences than are currently gained in a traditional CPR training process. To prove effective educative experiences, we will precede user-case studies with information including an experiment questionnaire and scenario-based user behavior analysis.

Acknowledgments. This work was supported by the DigiLog Miniature Augmented Reality Research Program funded by KAIST Research Foundation.

References

1. Swor, R., Khan, I., Domeier, R., Honeycutt, L., Chu, K., Compton, S.: CPR Training and CPR Performance: Do CPR - trained Bystanders Perform CPR? Academic Emergency Medicine 6 (2006)
2. Todd, K.H., Braslow, A., Brennan, R.T., Lowery, D.W., Cox, R.J., Lipscomb, L.E., Kellermann, A.L.: Randomized, controlled trial of video self-instruction versus traditional CPR training. Annals of Emergency Medicine, 364–369 (1998)
3. Gombeski Jr, W.R., Effron, D.M., Ramirez, A.G., Moore, T.J.: Impact on retention: comparison of two CPR training programs. American Journal of Public Health, 849–852 (1982)
4. Li, H.N., Wang, Z.T., Li, X.B.: Hardware design of CPR Simulation Control System based on SCM. In: Proceedings of the IEEE Electrical and Control Engineering (ICECE), pp. 4802–4805 (2011)
5. Slater, M., Usoh, M., Steed, A.: Depth of presence in virtual environments. Presence-Teleoperators and Virtual Environments, 130–144 (1994)
6. Cho, D., Park, J., Kim, G.J., Hong, S., Han, S., Lee, S.: The dichotomy of presence elements: The where and what. In: Proceedings of the IEEE Virtual Reality, pp. 273–274 (2003)
7. Witmer, B.G., Singer, M.J.: Measuring presence in virtual environments: A presence questionnaire. Presence, 225–240 (1998)
8. Meehan, M., Insko, B., Whitton, M., Brooks Jr, F.P.: Physiological measures of presence in stressful virtual environments. In: ACM Transactions on Graphics (TOG), pp. 645–652. ACM (2002)
9. Klinger, E., Légeron, P., Roy, S., Chemin, I., Lauer, F., Nugues, P.: Virtual reality exposure in the treatment of social phobia. Studies in Health Technology and Informatics (2004)
10. Boulos, M.N.K., Hetherington, L., Wheeler, S.: Second Life: an overview of the potential of 3 - D virtual worlds in medical and health education. Health Information & Libraries Journal, 233–245 (2007)

11. Sliney, A., Murphy, D.: JDoc: A serious game for medical learning. In: 2008 First International Conference on IEEE Advances in Computer-Human Interaction, pp. 131–136 (2008)
12. Kron, F.W., Gjerde, C.L., Sen, A., Fetters, M.D.: Medical student attitudes toward video games and related new media technologies in medical education. BMC Medical Education, 50 (2010)
13. Ruiz, J.G., Mintzer, M.J., Leipzig, R.M.: The impact of e-learning in medical education. Academic Medicine, 207 (2006)
14. Mantovani, F., Castelnuovo, G., Gaggioli, A., Riva, G.: Virtual reality training for health-care professionals. CyberPsychology & Behavior, 389–395 (2003)
15. Arduino, `http://www.arduino.cc`

Affecting Our Perception of Satiety by Changing the Size of Virtual Dishes Displayed with a Tabletop Display

Sho Sakurai[1], Takuji Narumi[2], Yuki Ban[2], Tomohiro Tanikawa[2], and Michitaka Hirose[2]

[1] Graduate School of Engineering, The University of Tokyo
[2] Graduate School of Information Science and Technology, The University of Tokyo
{Sho,narumi,ban,tani,hirose}@cyber.t.u-tokyo.ac.jp

Abstract. In this paper, we propose a tabletop system for affecting our perception of satiety and controlling energy intakes by controlling a size of a projected image around the food. We hypothesized that ambiguous perception of satiety can be applied to control our food intake. Given that estimating portion size is often a relative judgment, apparent food volume is assessed according to the size of neighboring objects such as many cutleries. Especially, the effect of the size of dish on food intake has been debated. Based on the knowledge, we constructed a tabletop system which projects virtual dishes around the food on it, in order to change the assessed apparent food volume interactively. Our results suggest that the size of virtual dish change the perception of satiety and the amount of food consumption.

Keywords: Augmented Satiety, Human Food Interaction, Cross-modal Interaction, Augmented reality, Food Consumption.

1 Introduction

Obesity has become a serious public health concern all over the world [1]. To decrease rates of obesity, many researchers have developed systems and services in the field of engineering [2, 3]. Many of these promote physical activity and bring users to the attention of the amount of the food eating.

However, Sustaining highly conscious effort to control amount of food consumed adequately is difficult. One of the reasons is that humans cannot accurately assess the volume or nutrition value of the food they consume. Therefore, humans estimate the volume of food eaten by using indirect cues: distension in the stomach and bowels, elevated blood-glucose levels, and apparent size of food. This estimation is inaccurate because some kinds of cues are evaluated relative to an individual's surroundings.

Recent studies in psychology have revealed the apparent volume of food can influence eating behavior. Given that estimating portion size is often a relative judgment, apparent food volume is assessed according to the size of neighbouring objects such as dishes and cutlery [4, 5].

R. Shumaker (Ed.): VAMR/HCII 2013, Part II, LNCS 8022, pp. 90–99, 2013.

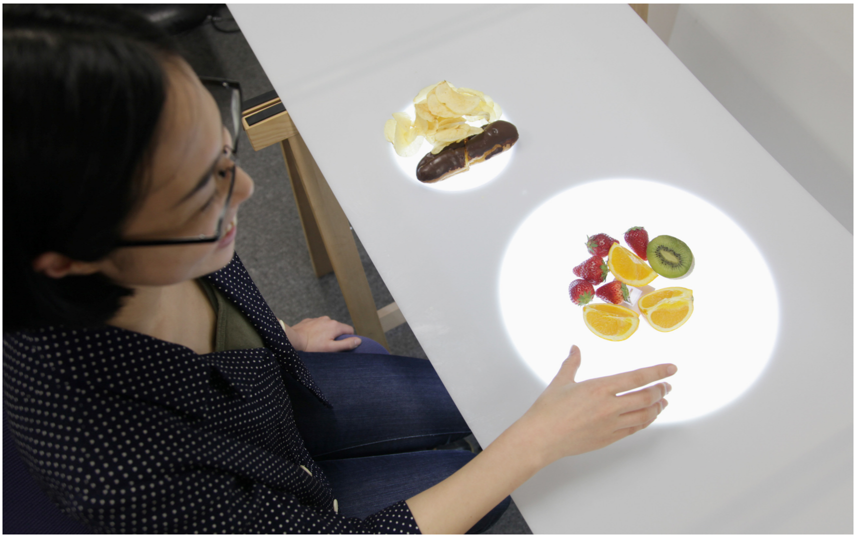


Fig. 1. Tabletop system which projects virtual dishes around the food for affecting the perception of satiety

Our early research focuses on changing the apparent size of food for affecting the perception of satiety as the cue. We have proposed a method for food-volume augmentation using shape deformation processing in real-time [6]. We have shown that our system could change the consumption volume from about -10% to about 15% by changing only its apparent size without changing perceived fullness.

Meanwhile, the previous system based on augmented reality paradigm has some limitations. First, it has strong limitation in target food and environment. It is useful only for food eaten with the fingers and cannot be used with arbitrary backgrounds. Second, the change in the apparent size of the food is presented to users by using a head-mounted display (HMD) in the system. However, use of wearable devices such as a HMD in every eating is unnatural and unrealistic.

To solve these problems, we propose a tabletop system for controlling perception of satiety (Fig. 1). The proposed system displays visual stimuli for affecting our eating behavior around the food on it using marker-based food tracking techniques. This enables us to use the system with any kind of food except liquids. It also eliminates the need to use wearable devices.

In this paper, we first explain the implement of tabletop system which displays virtual dish around the food for controlling perception of satiety. Next, we discuss its validity through an exploratory study.

2 The Effect of the Size of Dishes on the Estimation of Volume of Food and Food Consumption

Among some eating utensils, the effect of the size of dishes on the estimation of volume of food and food consumption has been vigorously debated [5, 7]. However, recently, a possibility is indicated that the effect of the size of dishes on amount of estimation and consumption of food based on principle of the Delboeuf illusion [8].

Delboeuf illusion is phenomenon that is concentric circles surrounded by a largecircle around seems smaller, and concentric circles surrounded by almost same size circle appear larger (Fig 2) [9]. This is known that the former effect is most effective when the ratio of external diameter of the circle to inner circle is 1 to 3, the latter effect is most effective when the ratio of external diameter of the circle to inner circle is 2 to 3. Also this has no effect when external diameter to the circle and inner circle is 1 to 2.

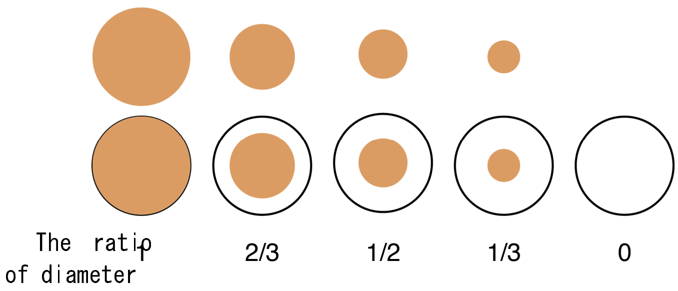


Fig. 2. Delboeuf illusion

Ittersum et al. showed that the size of dishes affect on the estimation of amount of food in the range about -10% to 12% even in the same amount of food by Delboeuf illusion (Fig 3) [8]. Through the experiments, Ittersum et al. have confirmed that the above-mentioned paradoxical results between [5] and [7] can be explained by the effect of this illusion.

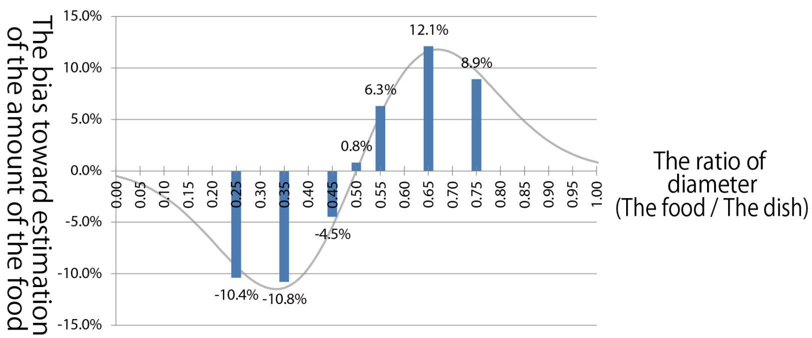


Fig. 3. The bias on estimation of food volume based on Delboeuf illusion (created based on [8])

While this experiment showed that estimation of the amount of food change depending on the size of dishes, whether amount of food consumption varies or not depending on its. Considering the result, amount of food intake that is influenced by the size of the dish will be similar to reversing plus and minus of the estimation of amount of food in above experiment. On the basis of above knowledge, we control our satiety and amount of food intake by changing the size of virtual dish displayed around the food using the system we propose.

3 A Tabletop Display for Affecting Our Perception of Satiety

We propose a method for controlling the perception of satiety by changing a size of projected image around the food using a tabletop display. We aim to affect the perception of satiety by changing the size of projected dishes and the relative size of the food on it. This gives the users the impression that there is a difference in food-volume on a virtual dish although amount of food is same.

Figure 4 shows the setup of our system. The system consists of a laptop PC, transparent plastic plates with AR markers, a semi-transparent lacteous acrylic board, a WEB camera, a mirror and a projector. The users put food on the transparent plastic plate. Each transparent plastic plate is attached an AR marker for detecting the position of the plate and food on it. The AR markers can be seen through the semi-transparent lacteous acrylic board from under the table when the plates are put face up on the board. A WEB camera put under the table captures the AR tags and the system calculates their position. This enables the tabletop system to aware position of each plates and food on it, and projected virtual dishes around each plate.

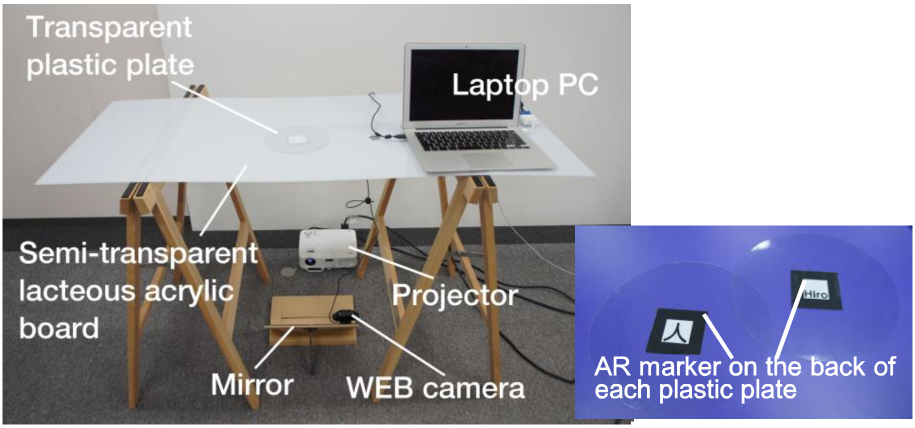


Fig. 4. System configuration

4 Exploratory Study

We conducted an exploratory study to investigate how the perception of satiety and food consumption would change by changing the size of projected dishes. In the current study, we examined whether our system would change consumption volume even though users are not consciously paying attention to a change in volume. Participants were not made aware of the true purpose of the experiment.

4.1 Subjects

The study consisted of 9 subjects and they were all males. The subjects' ages ranged from 21 to 25 and the average age was 22.6 years-old. The subjects were screened to determine that they were in good health, had no food allergies/restrictions, were not currently dieting for weight- loss or trying to gain weight, were not depressed, were not using medication known to affect appetite.

4.2 Experimental Design

The exploratory study used a within-subjects design. On 3 separate days, subjects came to our laboratory to eat potato chips. The area where is put the potato chips were confined to area on transparent plastic plate. The size of the plate acts as a benchmark of comparison between the size of food and virtual dishes. The diameter of the plate is 17.0 [cm]. The amount of the food that is put on the plate at once is 50.0 [g] and its energy density was 5.6 [kcal/g].

During each testing session, subjects were presented with one of three apparent size of the virtual dish changes by our tabletop display system: small dish condition (× 1.0, 17.0 [cm]), medium dish condition (× 1.5, 25.5 [cm]) and large dish condition (× 2.0, 34.0 [cm]). The appearance of the virtual dish and potato chips in each condition is shown in Figure 5. To eliminate any effect of presentation order of the size, the presentation order was randomly assigned and balanced across subjects. The test days were separated by at least 2 days in order to prevent any satiation effects.

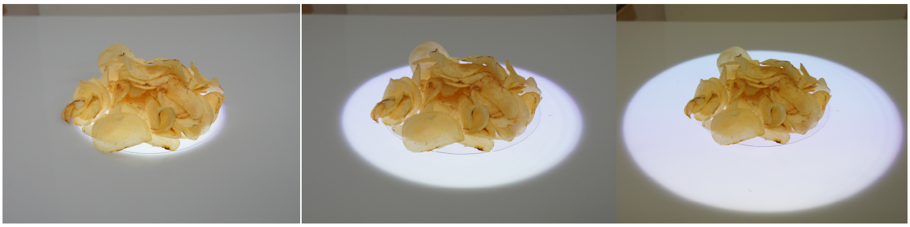


Fig. 5. Projecting a virtual dish around the food on plate (Left: small dish, Middle: medium dish, Right: large dish)

4.3 Procedures

Subjects were asked to keep their activity level as similar as possible on the day before each testing session. They were also asked to eat a similar meal as their last meal prior to the experimental session. During the day of each testing session, subjects were instructed not to consume any food for 2 hours prior to the exploratory study. Subjects kept a brief record of their activity patterns and food intake on the day prior to, and each day of, the testing session. Prior to each testing session, subjects recorded their life cycle and we made sure that subjects had followed the prescribed protocol prior to each session: at least two hours had passed since their last meal, there was no significant change in diet since the previous testing session and that their activity level had been similar as previous days prior to the testing session (e.g., no all-night work or hard exercise). If activity patterns and menu differed from that of the day prior to the previous testing session, we postponed the testing session for at least one day.

First, subjects were asked questions regarding their hunger conditions. We describe these questions below. After this test, subjects were instructed to eat as much potato chips and drink as much water as they desired (Fig 6.) Subjects were told that they did not have to eat the entire potato chips on the dish and we allowed them to have leftovers. In the beginning, 50.0 [g] of the potato chips is put on the dish. Subject were also told that they can order additional potato chips or water, if they ate up the potato chips or drank down a cup of water and desire much further them. In this case, 50.0 [g] of the potato chips were put on the plate per an addition. In order not to be affected perception of satiety and food intake by other external cues, we asked subjects to focus on the food and the table before they ate it. After the subject stopped eating, we asked questions identical to those asked prior to the exploratory study.

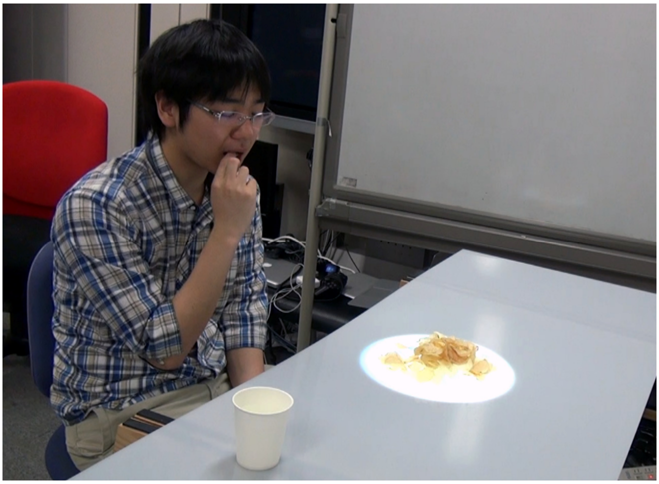


Fig. 6. A Subject eating potato chips using our system

4.4 Visual Analogue Scale Ratings

Subjects completed ratings of hunger and satiety immediately before and after the exploratory study. Subjects rated their hunger, thirst, nausea and fullness when completing the analogue scales. All questions are shown in Figure 8.

For example, subjects answered the question “How hungry are you now?” by rating hunger on a 100 mm line anchored by “not hungry at all” on the left and "extremely hungry" on the right. Other anchors in other questions consisted of the phrases "not at all" and "extremely" combined with the adjectives "thirsty," "nauseated," and "full." If a rating during the current testing session had changed more than 30 [mm] compared with the rating from his/her prior testing session, we postponed his/her testing session.

The subjects were asked to rate the palatability of the potato chips on a 100 [mm] visual analogue scale prior to the experiment. We originally planned that subjects whose palatability rating was less than 40 [mm] would be excluded from this experiment in order to avoid inflicting them. However, all subjects found the potato chips highly palatable; thus, no subjects were excluded based on this criteria.

4.5 Results

Average potato chips consumption volume and standard error of the mean in each experimental condition are as follows: small dish condition: 72.1 ± 7.9 g; medium dish condition: 59.7 ± 7.7 g; large dish condition: 66.9 ± 9.9 g (Fig. 7). Figure 7 illustrates the normalized results based on the consumption volume under small dish condition by subject. The curve in figure 7 illusrates expected amount of food consumed from biases which affects on estimation of food volume by the results of Ittersum's experiments [8].

The amount of food consumed was expected to be going todecrease by about 12% under medium dish condition than small dish condition. Actual food consumption under medium dish condition decreased by an average of -13.4%. than small dish condition. Also amount of food consumed under large dish condition increased by average of 2.7%. this is a result that agree with the expectation that the amount under large dish condition is almost same as under small. Though clear significant difference is not shown, this result gives close agreement with expected changing amount of food intake from [8].

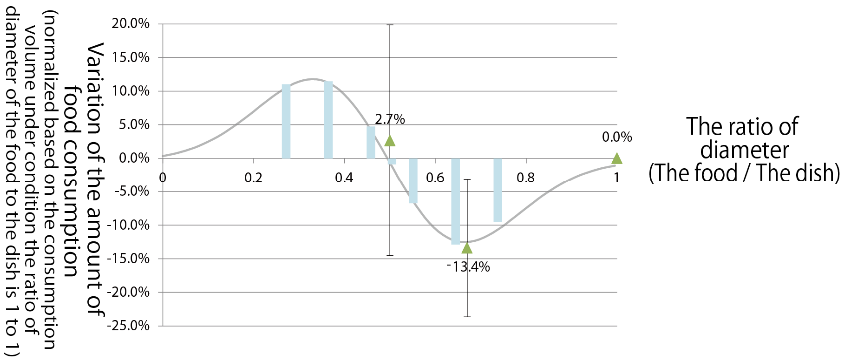


Fig. 7. Amount of potato chips consumed (± standard error of the mean) by the size of the projected dish

This shows possibility that amount of food consumption is predictable by effect of Delboeuf illusion when ratio of size of virtual dish to the food is changed using a tabletop display. Therefore, this result suggests a possibility that the ratio between diameter of dishes and food should be decided based on the effect size of Delboeuf illusion to change the amount of the food intake as it is intended.

Figure 8 shows the ratings of hunger and satiety before and after the testing session by experimental condition. Across all conditions of the size of the projected image, no significant differences were found prior to the testing session on ratings of hunger, thirst, nausea and fullness. There were also no significant differences on these ratings after each testing session.

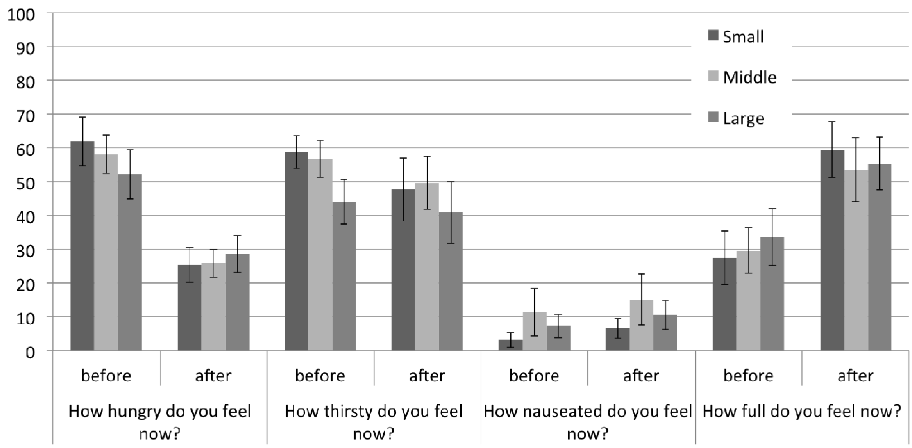


Fig. 8. Ratings of hunger and satiety before and after the testing session in each experimental condition. (Average ± standard error of the mean).

4.6 Discussion

Our results suggest that the size of a projected virtual dish affects food intake and the perception of satiety; this is accomplished from the same amount of food during a single meal. These results also suggested the validity of our experimental design, because there were also no significant differences on these evaluation values on survey item before and after each testing session. These results also showed that our system maintains perceived fullness even after subjects ate the food, since there were no significant differences about perceived satiety after each test session, this result suggested a possibility that our system changed amount of food consumed of subjects with remaining satiety of after each eating, or can change perceived satiety derived same amount of potato chips.

According to questionnaires after experiment, more than half of subjects were aware of different of the size of virtual dishes. Nevertheless average of amount of food consumption changed mostly like the result of [8]. This shows our system affects on perception of satiety regardless of being conscious or not of subjects.

It was considered about the evoking of strangeness to the users when they put food on the plate and eating it. However, subjects answered they did not have feelings of resistance toward eating the food on the plate in their free description after all session. Therefore, it is thought that putting the food on the plate did not affect the experimental design.

5 Conclusion

In this paper, we proposed a tabletop system which displays virtual dishes around the food in order to control consumption volume of the food. It displays virtual dishes as visual stimuli for affecting our eating behavior around the food using marked-based food tracking techniques. This is based on previous research in the field of psychology. With this method, we can create the illusion of apparent food size and the perception of satiety. We developed a prototype system for implementing the proposed method. Also we conducted the exploratory study to evaluate the effectiveness of the proposed system.

Our study examined how the size of virtual dishes affects our "consumption volume" of food on these. As the result, we suggested that the ratio between size of food volume and virtual dish should be determined based on the effect of Delboeuf illusion to change amount of food intake.

We will test validation of the system through further sophisticated experiment with more subjects and expansion in the range of the ratio between the size of the virtual dishes and the food. Also we will discuss validation of cutleries, food and chromaticity contrast between dishes and tablecloth described in the reference of Ittersum et al. Through these assessments, we aim to realize a more effective system.

References

1. Maitland, J., Chalmers, M.: Designing for peer involvement in weight management. In: Proceedings of CHI 2011, pp. 315–324 (2011)
2. Grimes, A., Harper, R.: Celebratory technology: new directions for food research in HCI. Proceedings of CHI 2008, 467–476 (2008)
3. Mankoff, J., Hsieh, G., Hung, H.C., Nitao, E.: Using low-cost sensing to support nutritional awareness. In: Borriello, G., Holmquist, L.E. (eds.) UbiComp 2002. LNCS, vol. 2498, p. 371. Springer, Heidelberg (2002)
4. Wansink, B., Cheney, M.M.: Super Bowls: serving bowl size and food consumption. The Journal of the American Medical Association 13, 293(14), 1727–1728 (2005)
5. Wansink, B., Ittersum, K.V., Painter, J.E.: Ice cream illusions bowls, spoons, and self-served portion sizes. American Journal of Preventive Medicine 31(3), 240–243 (2006)
6. Narumi, T., Ban, Y., Kajinami, T., Tanikawa, T., Hirose, M.: Augmented Perception of Satiety: Controlling Food Consumption by Changing Apparent Size of Food with Augmented Reality. Proceedings of CHI 2011, 109–118 (May 2012)

7. Rolls, B.J., Liane, S.R., Kitti, H.H., Jennifer, S.M.: Using a Smaller Plate Did Not Reduce Energy Intake at Meals. Appetite 49(3), 652–660 (2007)
8. Ittersum, K.V., Wansink, B.: Plate Size and Color Suggestibity: The Delboeuf Illusion's Bias on Serving and Eating Behaviors. Journal of Consumer Research (2012) (Online Published in November 2011)
9. Nicolas, S.: Joseph Delboeuf on Visual Illusions: A Historical Sketch. American Journal of Psychology 104, 563–574 (Winter 1995)

Part II

Virtual and Augmented Environments for Learning and Education

An Experience on Natural Sciences Augmented Reality Contents for Preschoolers

Antonia Cascales[1], Isabel Laguna[2], David Pérez-López[3], Pascual Perona[3], and Manuel Contero[3]

[1] Universidad de Murcia, Avda. Teniente Flomesta 5, 30003 Murcia, Spain
[2] Universidad de Alicante, Cra. San Vicente del Raspeig s/n, 03690 San Vicente del Raspeig, Spain
[3] Instituto de Investigación en Bioingeniería y Tecnología Orientada al Ser Humano (I3BH), Universitat Politècnica de València, Camino de Vera s/n, 46022 Valencia, Spain
antonia.cascales@um.es, isabel.laguna@ua.es, {dapelo,pperona,mcontero}@i3bh.es

Abstract. Early education is a key element for the future success of students in the education system. This work analyzes the feasibility of using augmented reality contents with preschool students (four and five years old) as a tool for improving their learning process. A quasi experimental design based on a none-quivalent groups posttest-only design was used. A didactic unit has been developed around the topic "animals" by the participant teachers. The control group followed all the didactic activities defined in the developed didactic materials, while the experimental group was provided in addition with some augmented reality contents. Results show improved learning outcomes in the experimental group with respect to the control group.

Keywords: augmented reality, preschool, teaching/learning process.

1 Introduction

Early educational intervention [1] has been proposed as an effective tool to fight against poverty and inadequate learning environments, promoting child development and school success. The OECD reports that those students that attended preschool for one year or more scored more than 30 points higher in reading by age 15 in the PISA assessment than those who did not [2].

Educational professionals must be ready for tomorrow's school, facing the constant challenges of our society [3]. In the path of achieving this aim, it is necessary to motivate students, since "learning is more effective when the apprentice voluntarily engages in the process" [3]. Therefore there is a necessity for novel attractive technologies [4] where teachers can provide students efficient and interesting environments to learn [5]. In this context, this work tries to contribute to the improvement of early childhood education by means of the following objectives:

R. Shumaker (Ed.): VAMR/HCII 2013, Part II, LNCS 8022, pp. 103–112, 2013.

1. Promote educational innovation by a gradual change in teaching methodology in order to utilize the advantages provided by Augmented Reality (AR).
2. Analyze the possibilities that AR can have on early childhood education.
3. Restructure the classroom environment to incorporate AR
4. Assess improvements that can promote the learning of the students.

Therefore, we have tried to answer the following questions:

- What happens in the teaching and learning process from the use of AR in the classroom?
- How can AR help us to achieve the educational objectives?
- What kinds of interactions are produced when this technology is implemented in the classroom?

In order to answer the previous questions we have developed some augmented reality educational contents and implemented a teaching/learning strategy around them that has been tested on a real preschool scenario.

1.1 Augmented Reality in Preschool and Kindergarten

Augmented Reality is a technology which introduces virtual contents such as 3D computer-generated objects, texts and sounds, onto real images and video all in live time. There are different definitions and classifications of AR: Azuma [6] describes AR as a variation of Virtual Reality (VR), a technology which consists of the complete immersion of a user inside a synthetic environment. In VR the user is not able to perceive the surrounding real world. However, in AR it is possible. In this fact AR differs from VR, because AR adds artificial information to reality while it does not hide the surrounding real world.

Augmented Reality has been touted as one of the most interesting emergent technologies for education, being a powerful and motivating tool which can involve several senses of the student by means of the proper combination of sound, sight and touch. Application of AR technology in education is just beginning to be explored, especially when using it with preschoolers.

One of the first experiences who understood this potential in kindergarten and preschool education was developing a mixed reality platform to provide a tool for collaborative learning [7]. Other studies were conducted in language learning in which AR was used for improving pronunciation and memorization of the Chinese language [5]. Other authors introduced AR educative games to study English as a foreign language [8, 9]. In these works virtual models and texts in Chinese and English were shown to children in an interactive way controlled by them. Besides, the field of Mathematics was also explored using a serious game to learn Mathematics [10]. Other researchers did not study specifically a subject, instead of it, they studied the learning process. For example, a multisensory mixed reality system was designed for children between 3 and 5 years old that used it in a learning environment formed by images, sounds, and haptic objects [4]. This interactive platform was really helpful to study the advantages of learning through sensorial experiences. Another AR system [11] was developed using AR and robots to deliver positive stimuli to students. In that work, researchers took advantage of an immersive environment to familiarize children with a dramatic activity. Finally, other authors provided an AR tool for an interactive storyteller to enhance children reading [12].

As it could be deducted in the last paragraphs, AR contributes in many ways to support the teaching/learning process: students' senses are involved in interactive activities by using manipulative material [7]. Besides, self-learning is promoted by enjoyable edutainment in friendly interfaces [5, 8]. And from the point of view of technology approval, in previous studies it is shown that students and their parents have made positive valuations about AR [13].

In the next sections, a detailed description of the didactic contents is provided, after a short introduction about AR in educational contexts. Then, the experimental design is presented, followed by the results, discussion and conclusions.

2 Materials and Methods

2.1 Didactic Materials

The augmented reality application consists of a launcher, a camera configuration tool and a content installation tool. The launcher is designed to manage a collection of AR contents and it allows the user to launch a specific AR application, selecting its academic year, subject and language through pull-down menus as seen in figure 1-a. It also provides a preview of the content; a handbook about it and the AR marker which has to be printed in order to use the AR application (see figure 1-b).

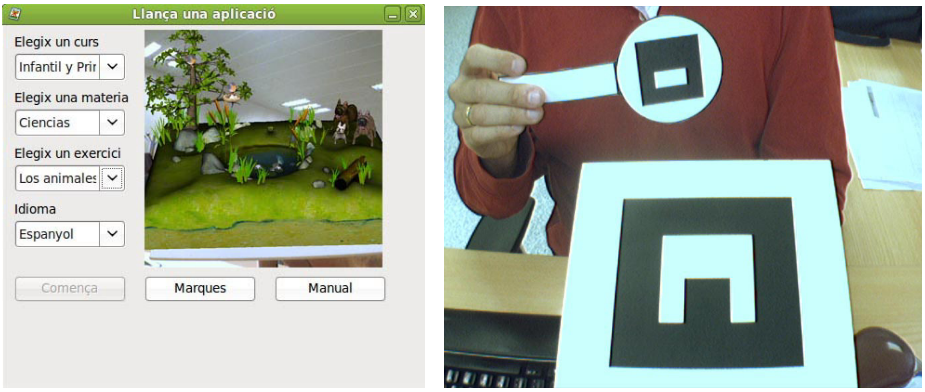


Fig. 1. (a) Launcher to run the AR animal application; (b) AR marks used by the application

The camera configuration tool allows the user to configure video modes, exposition, color and tracking parameters. Finally, the installation tool allows installing and uninstalling AR contents from a local drive or from a URL. All these elements are expected to be used by the teacher, who launches the selected content for the preschoolers.

2.2 Application Structure

The application, like other AR ones, follows the “magic mirror” paradigm, where a computer monitor shows a live video stream from a webcam, and renders some 3D models over the AR markers. The application shows a set of icons at the top right side of the screen, providing some common functionality as seen in figure 2. The AR marker has been provided with a special handle to be used more comfortably by pre-schoolers, as can be seen in figure 3-a. The AR application incorporates some interface elements to access much more information, and to control certain simulation parameters. They are common tools, such as changing the zoom level of the 3D model that appears on the AR mark, which can be accessed using the mouse. Other common tools are the audio switch on/off, to show the application user guide, to quit the application and to display a context-sensitive help on the action being performed at any moment. In addition, it is worth mentioning a functionality that has proved to be very useful for the teachers. It is the possibility of stopping the tracking of the AR marker. This functionality allows freezing the position of the 3D model on the screen, without holding the AR marker. This feature is very useful when a teacher wants to point to a part of the model with their hands, for example, if the AR application is visualized on a digital whiteboard.

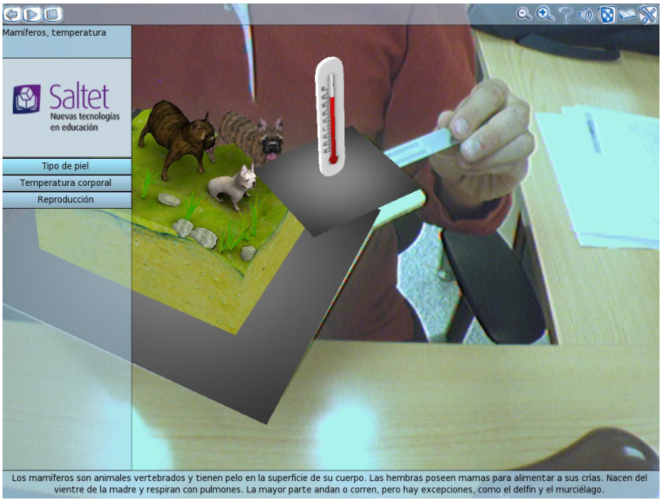


Fig. 2. Interface characteristics. Example of secondary mark acting as a thermometer

The system also includes buttons to control an auditory narrative, (play, pause and stop) at the top left part of the screen. And finally, close to those buttons, there is a button used to return to the previous screen. Moreover, the left side of the screen is reserved to include menus that dynamically change. In this way, these menus can give access to different activities, “Presentation” and “Lesson”, as presented below. Also, under these menus, there is a rectangular area reserved to show informative texts relative to the visualized scene. Under this area, there is a small square which is used to show 2D images which can produce a better content visualization. Finally, in the lower side of the screen, there is also another rectangle reserved to show more detailed explanations in text format. Note that those explanations are also shown as an auditory narrative. And also note that this text box incorporates an automatic scroll to support long texts.

It should be noted that the system can be run in several languages by using XML configuration files, so the system loads these files which contain links to all the digital resources such as 3D models, texts and audio. These resources are organized in language folders, so, the way to change the language is to load a different set of digital resources from the right language folder. The system provides that option by a drop-down menu in the launcher.

2.3 AR Animals Application

This AR application provides two activities, "Presentation" and "Lesson", and it includes an innovative solution in AR. That solution is a special magnifying glass, which will be described below. The main objective of this application is to help the teachers to show the students the vertebrate animal classification and to show in detail these animals' coats, kind of reproduction and corporal temperature. Thus, "Presentation" shows, over the AR marker, a kind of park where five animated couples of animals appear. This mode only allows the user to observe the scene; there is no more interaction than moving the AR marker and to observe the animals from different points of view.

Fig. 3. (a) Preschooler using the AR application (b) Two marks interaction: magnifying glass

The second activity, "Lesson", shows the same park over the AR marker and explains the vertebrate animal classification using different texts and auditory explanations. Thus, this section is composed of five sub-activities which can be accessed by a set of buttons located at the left of the screen: "Mammals", "Birds", "Fishes", "Amphibians" and "Reptiles". Each sub-activity presents a male and a female of each kind of animal over the AR marker, and three buttons; "Coat", "Corporal temperature" and "Reproduction". These buttons activate an operating mode where two AR marks can be used at the same time. A big mark is used to represent 3D models, and a small one to represent a magnifying glass or a thermometer, depending on the user selection. So, the thermometer shows hot, cold or variable temperature when the user moves the small mark closer to the big one, that is, when the thermometer is over the animal as seen in Figure 2.

In order to control the user interaction with the two marks, the system calculates the distance between the 3D models and the small mark, and defines some distance thresholds, so, when these thresholds are exceeded an event occurs. The magnifying glass works as is expected, that is, it applies a zoom to the animals, hence, when the user looks at the virtual glass, he obtains an augmented vision through that glass, like a real magnifying glass, as can be seen in figure 2. In this way, the user can observe coat details and how the animals reproduce. In order to do this, a 3D transformation between the markers is calculated by the system and it is applied to a shader, which is also applied to the glass. The idea is to look at the big mark from the small one's point of view, and to apply some zoom.

It is important to mention that those augmented reality resources were initially designed for primary school students, but after a preliminary evaluation with preschoolers, it was observed that they were able to access the basic functionality of the application without any problem.

2.4 Libraries

This system was conceived and developed to be run under Ubuntu Lucid Lynx (10.04), however, it can be run under Windows. The system was developed using OGRE as the graphics engine and a set of public libraries like OIS to manage input devices, Audiere to manage sound streams, tinyXML to manage XML configuration files, v4l2 to capture video from a webcam under Linux, MyGUI to develop the graphical user interface and OpenCV to develop our own AR engine under Linux.

In order to support the development of AR applications, our research group has developed its own software library [14]. Although there are several public libraries with AR capabilities, we decided to develop it in order to overcome some drawbacks present in some public libraries (jitter, bad performance under illumination variations, lack of support of infrared markers, etc.).

2.5 Participants

The research involved two groups of eighteen preschoolers in each group, with ages between four and five years from the public school Virgen de los Desamparados in Orihuela (Spain). One group was taken as the control group, while the other was taken as the experimental group. Both groups were composed of third graders from the second cycle of pre-primary education, according to the Spanish education system. Both groups had the same teachers.

The school is located in a rural area. It is one of the seventeen pilot technological schools in the province of Alicante (Spain). The school is fully equipped with technology and also has formed a team of teachers experienced in ICT, which works hard to improve the use of ICT in the classroom. Regarding the students participating in our research, they have been using ICT in the school since they were three years old.

Table 1. Students' categorical estimation scale

Pupil:	NA- Not Achieved IP- In progress A- Achieved		
Item		**Criteria**	
Animals			
• Child recognizes vertebrate animals: dog	NA	IP	A
• Child recognizes invertebrate animals: caterpillar	NA	IP	A
• Child recognizes mammals	NA	IP	A
• Child recognizes birds	NA	IP	A
• Child recognizes fish	NA	IP	A
• Child recognizes amphibians	NA	IP	A
Knowledge about mammals:			
• Mammals are born from the womb	NA	IP	A
• Mammals nurse their young with milk	NA	IP	A
• Mammals have hair	NA	IP	A
• Some mammals are terrestrial	NA	IP	A
• Some mammals are aquatic	NA	IP	A
• Some mammals can fly	NA	IP	A
Knowledge about birds:			
• Birds are born from eggs	NA	IP	A
• Birds do not nurse their young with milk	NA	IP	A
• Birds have feathers	NA	IP	A
• Usually birds can fly	NA	IP	A
• Some birds live on earth	NA	IP	A
Knowledge about fish:			
• Fish hatch from eggs	NA	IP	A
• Fish have scales	NA	IP	A
• Fish breathe through gills	NA	IP	A
Knowledge about amphibians:			
• Amphibians hatch from eggs	NA	IP	A
• When amphibians are immature they breathe through gills	NA	IP	A
• When amphibians are adults they live on land and water	NA	IP	A

Students often work with technology in their own classrooms, where they have several computers and an interactive whiteboard.

2.6 Experimental Design and Method

In this research a nonequivalent group posttest-only [15] design has been chosen. Under this scheme, one group (the experimental group) received the intervention (augmented reality contents), while the other group (the control group) does not use AR. The intervention was done in a natural situation, without a random selection of groups [16].

Initial conditions for both groups were similar: each group was composed by 18 children between 4 and 5 years old. Moreover, both groups had had the same teachers in the previous year and they had also studied the same contents.

Besides, this experience has been developed using an active and communicative methodology. On the one hand, teachers were deeply implied providing feedback data about student experiences. On the other hand, preschool students worked properly following the didactic guides developed by participant teachers. The chosen didactic unit for the two groups involved was “animals”. Two versions of these didactic materials were created. The only difference between them was that the “experimental unit” provided the augmented reality resources described in the previous section. In this way both units have the same educational curriculum content, one with AR and one without it. Therefore the independent variable of this research was the presence of Augmented Reality as a didactic tool.

The assessment of experimental and control groups was performed using an evaluative categorical scale completed by teachers (see Table 1). This scale consisted of 22 items, where each item was checked according to the following categories: A (Achieved), IP (In Progress) and NA (Not Achieved).

3 Results

Table 2 summarizes the results obtained from applying the categorical estimation scale presented in Table 1. Experimental group reflects a slight improvement with respect to control group although there is no statistical significant difference.

Table 2. Results of learning outcomes in didactic unit “animals” (absolute frequencies)

	Control Group			Experimental Group		
	A	IP	NA	A	IP	NA
Classification of animals:						
Mammals	11	5	2	13	4	1
Birds	11	5	2	13	3	1
Amphibians	9	5	4	13	4	1
Fish	10	5	3	12	5	1
Characteristics of animals	13	4	1	15	3	-
Body parts of animals	12	5	1	15	3	-

4 Discussion and Conclusions

The slight improvement that results reflect must be put in context. Augmented reality contents are one of the ingredients of the didactic activities designed to support this learning unit. The learning process was organized around team work, and AR contents were used by the preschoolers at their own pace. AR contents served as a catalyst providing a real motivation and stimulus for the children, and teachers observed a very positive impact on students.

Participant teachers had no previous exposure to this technology, but it was easily integrated in the class dynamics. The perception of participant teachers was that augmented reality improved learning activities, supporting students’ constructive learning approach. Teachers considered AR to be a resource that is tailored to the characteristics of their students and thus is useful for learning.

All study participants considered that the use of AR is a good tool in the teaching-learning process. The conclusions we reached in our experience of inclusion of AR as a part of the teaching-learning process are the following:

- The use of AR promotes active behavior in the student.
- The work of teaching improves with the use of AR. Daily work is more playful and fun for both students and teachers.
- Students learn more when they are using AR and they achieve more learning goals than if they are not using AR.
- AR also promotes communication skills, promoting all kinds of interactions in the classroom between teacher and students, students and students, students and families, families and families and teachers and teachers.

The experience with the use of AR has been very positive for teachers and students. After the results obtained with the use of AR, participant teachers will motivate their colleagues to use AR in their classrooms. Finally, teachers are in full agreement with the resource and its viability of application. They considered feasibility, time spent on the implementation, distribution of content and resources and infrastructure available, appropriate and sufficient.

Acknowledgements. The Spanish Ministry Economy and Competitiveness partially supported this work (Project ref. TIN2010-21296-C02-01).

References

1. Barnett, W.S.: Effectiveness of Early Educational Intervention. Science 333(6045), 975–978 (2011)
2. OECD: Investing in high-quality early childhood education and care (ECEC). OECD Publishing, `http://www.oecd.org/dataoecd/0/28/48980282.pdf` (retrieved)
3. Campos, P., Pessanha, S.: Designing Augmented Reality Tangible Interfaces for Kindergarten Children. In: Shumaker, R. (ed.) Virtual and Mixed Reality, HCII 2011, Part I. LNCS, vol. 6773, pp. 12–19. Springer, Heidelberg (2011)
4. Lim, J., Kim, S.: A Study on Markerless AR-based Infant Education System using CBIR. Communications in Computer and Information Science 78, 52–58 (2010)
5. Chen, C.H., Su, C.C., Lee, P.Y., Wu, F.G.: Augmented Interface for Children Chinese Learning Technologies. In: 7th IEEE International Conference on Advanced Learning Technologies, pp. 268–270. IEEE Press, New York (2007)
6. Azuma, R.: A Survey of Augmented Reality. Presence: Teleoperators and Virtual Environments 6(1), 355–385 (1997)
7. Winkler, T., Herczeg, M., Kritzenberger, H.: Mixed Reality Environments as Collaborative and Constructive Learning Spaces for Elementary School Children. In: Barker, P., Rebelsky, S. (eds.) Proceedings of World Conference on Educational Multimedia, Hypermedia and Telecommunications 2002, pp. 1034–1039. AACE, Chesapeake (2002)

8. Hsieh, M.C., Lee, J.S.: AR Marker Capacity Increasing for Kindergarten English Learning. In: International Multi Conference of Engineers and Computer Scientists, vol. 1, pp. 663–666 (2008)
9. Hsieh, M.C., Lin, H.C.K.: Interaction Design Based on Augmented Reality Technologies for English Vocabulary Learning. In: Wong, S.L., et al. (eds.) Proceedings of the 18th International Conference on Computers in Education, vol. 1, pp. 663–666. Asia-Pacific Society for Computers in Education (2010)
10. Lee, H., Lee, J.: Mathematical Education Game Based on Augmented Reality. Technologies for E-Learning and Digital Entertainment, 442–450 (2008)
11. Hyun, E., Choi, K., Kim, G.J., Han, J., Jo, M., Kim, N.: Delphi Survey on the Use of Robot Projector based Augmented Reality in Dramatic Activity for Young Children. International Journal of Digital Content Technology and its Applications 5(11), 272–282 (2011)
12. Kim, H.M., Song, T.H., Jung, S.M., Kwon, K.H., Jeon, J.W.: Virtual Storyteller Using Marker Based AR and FPGA. In: IEEE 54th International Midwest Symposium on Circuits and Systems, pp. 1–4. IEEE Press, New York (2011)
13. Dunleavy, M., Dede, C., Mitchell, R.: Affordances and Limitations of Immersive Participatory Augmented Reality Simulations for Teaching and Learning. Journal of Science Education and Technology 18, 7–22 (2009)
14. Martín-Gutiérrez, J., Saorín, J.L., Contero, M., Alcañiz, M., Pérez-López, D., Ortega, M.: Design and validation of an augmented book for spatial abilities development in engineering students. Computers & Graphics 34(1), 77–91 (2010)
15. Cook, T.D., Campbell, D.T., Day, A.: Quasi-experimentation: Design and Analysis Issues for Field Settings, pp. 19–21. Houghton Mifflin, Boston (1979)
16. Buendía, L., Colás, P., Hernández-Pina, F.: Métodos de Investigación en Psicopedagogía. McGraw Hill, Madrid (1997)

Teaching 3D Arts Using Game Engines for Engineering and Architecture

Jaume Duran[1] and Sergi Villagrasa[2]

[1] Universitat de Barcelona, Barcelona, Spain
[2] La Salle - Universitat Ramon Llull, Barcelona, Spain
jaumeduran@ub.edu, sergiv@salle.url.edu

Abstract. The main objective of this paper is to evaluate the application of 3D virtual worlds for teaching different subjects mainly oriented to architectural visualization and creating 3D models for multimedia. The use of 3D technologies, multi-user virtual environments and avatars are new methodologies for the student to have a much richer experience and therefore more motivating for a deeper understanding of the assessment and help understand more collaborative the projects. In this paper we work on the concept e-learning and blended learning technologies related to interactive 3D spaces such as OpenSim, Activeworlds, Secondlife, Unity and others. The students' participation in these virtual 3D environments will help to understand the concept of an architectural project and 3D creation, improving collaboration between students and teacher, and dramatically increase in a greater understanding of the project and a high degree of their involvement with design develop. The paper describes the method of teaching 3D arts using Game Engines like Unity.

Keywords: Virtual reality, Game engines, Visual learning.

1 Introduction

Although many educators have created outstanding learning environments, the belief persists that the classroom is the domain of PowerPoint presentations and similar tools. This is a non-effective use of electronic capabilities in education. Meanwhile, creating more effective learning environments using virtual environments is a highly intensive labor. The learning curve is steep for both student and educator.

It is necessary to evaluate the degree of satisfaction, need [1], and interest that the use of these technologies can provide, both in the training of the student and their perception of the technology [2][3]. In this paper, we present a proposal for appropriate use of technology, specifically the implementation of Virtual Reality (VR) with a classroom-based video game Engine.

Nowadays, 3D technology and multimedia must assist education in activating all senses in order to improve student learning [4][5]. Concretely, we propose the enhancement of learning with 3D material, from the application to the Multimedia Engineering Grade and Architecture. Our proposal represents mixed learning and blended learning [6], usable with both face-to-face and virtual classes.

R. Shumaker (Ed.): VAMR/HCII 2013, Part II, LNCS 8022, pp. 113–121, 2013.

The applied learning methodology means learning by doing [7], as well as the use of applied gamification techniques. The purpose of learning is to create a collaborative project [8][9]. The main objective for ideal blended learning is the use of interactive methods with technology such as Unity or Opensim, plus gamification for engaging students and achieving better learning outcomes.

Our hypothesis is that the use of 3D virtual environments, video game engines, or multi-verse environments will be one of the challenges in the future of education [10][11]. This challenge is due in part to the high degree of interaction and collaboration between all the stakeholders.

To enact this experiment, we will use videogame engines, which we shuffle Unreal (UDK) and Unity because of their portability with mobile devices and tablets, their high graphical quality, and their multiplatform nature. Also we will experiment with multi-verse virtual worlds where the social aspect will take on particular relevance. That is, virtual worlds such as Second Life or Open SIM will be important to other learning-system developers who are more interested in the social aspect of virtual education.

In this paper, we present a new way of teaching using the modeling, texturing, and illuminating capabilities of 3D. This represents a new way of delivering practical exercises that substantially involve the student. The purpose is to maintain and increment the motivation or engagement in different ways. In the 3D technology references in this article we propose to create a greater experience and interaction from and for the student, in the name of greater global learning.

The experiment improves how a 3D model is presented. The traditional method is composed of presentation classes in which deliveries of 3D model are made by students who have to create a determinate 2D model. After this first step the student send to the teacher some static images of the model. This kind of exercises is strictly procedural, in that the student demonstrates that he or she knows how to use the tools for the 3D model and then render the result. However, the implication curve is low because the student looks for a result that may more or less overpass the minimum and doesn't go further in his/her skills with 3D tools.

Also in the experiment, we add an improvement to the model visualization. The presentation implies a greater knowledge for videogames and real-time engines. At the same time, we look for a greater interaction and collaboration among students and teachers. To involve more implies a pupil with greater motivation. And a greater motivation makes the student propose greater and more disciplined goals.

2 3D Modeling Innovation in Ways to Learn

In this new way of presenting and learning 3D modeling, two technologies are used: the first one, it is proposed that the delivery of models be online, where the 3D model can be uploaded and visualized on the web. In this case, the methodology will be 3Dclever so that 3D models will be directly uploaded on the web in a simple and effective way. The web allows the visualization and interaction with the object on a web navigator (such as Firefox or Chrome) with HTML5 support. This type of presentation is useful for directly visualizing the model and evaluating it independently of the modeling tool used. Moreover, is possible to visualize in various hardware formats: personal computer, laptop or tablet.

To follow the evolution of the creation of a concrete model is better than a drawing. A 3D model allows for more student involvement and more interactivity.

The second proposed is use a game engine called Unity for an enhanced interactive and user experience.

To exemplify the last methodology proposed, we shall proceed to explain a real exercise applied in a Multimedia degree on the subject of “Computer Animation” at La Salle, Ramon Llull University, a six-ECTS-credit course that is taught annually. This course has a total of 62 students.

3 Implementation of the Proposal

The objects represent a model creation that includes a geographic area inhabited with roads, wheels, a water tower, etc. [Fig. 1]. Also depicted is a typical house and structures that could be banks, saloons, food shops, police headquarters, etc. [Fig. 2]

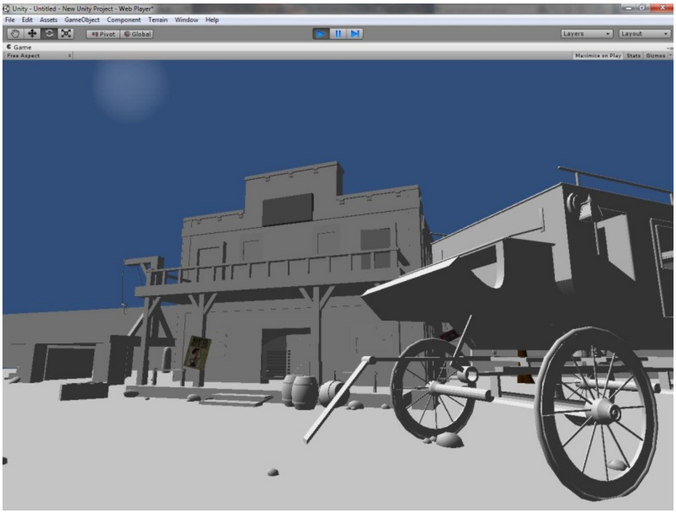


Fig. 1. 3D model of a caravan and saloon in Unity

The implementation that is proposed is a graphic-quality motor in which students' models can be easily incorporated in a virtual world using web navigation in mobile devices and personal computers. Students would be able to interact and visit the models virtually [12]. Also, to potentiate the collaboration among all the class, we propose not only to create independent models but to create a far-west town using the input of all the students, along with teacher supervision. This project would call upon greater collaboration and interactivity.

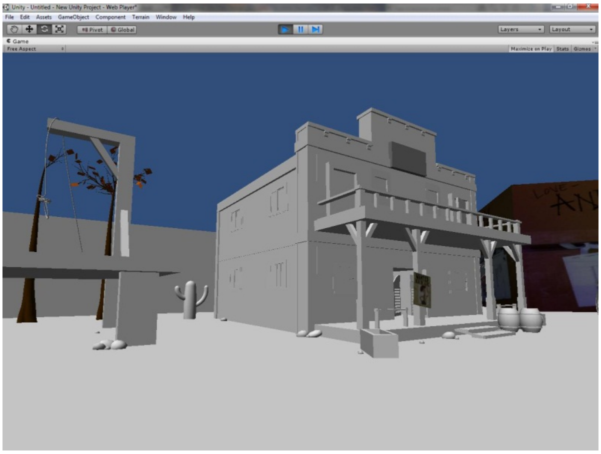


Fig. 2. 3D model of a gallow and a saloon in Unity

So, once all the models are presented, the teacher will introduce the models in the engine and from a tablet on the web. All the students will be able to use and "play" in the streets of the city created by themselves, visiting the shops of their peers' creations [Fig. 3]. The term "play" will have a certain importance in this method, because even the use of videogames is fun, it is also educational.

Fig. 3. 3D textured model of a saloon and caravan

To create a bit of competition and attempt to be cool to peers and the rest of the class, the engagement part of the motivation process improves notably. These processes, applied in the gamification, are taken into consideration when executing the process of this methodology of 3D impartation [13].

To evaluate correctly the progress and determine if the objectives in the hypothesis represent an improvement in student involvement and greater learning due to the increment in motivation, we will proceed to distribute two surveys, at the beginning of the experiment to obtain the student's profile "see Fig. 4" [14] and at the end of the experiment to evaluate the user experience "see Fig. 5" [15]. With these types of surveys we obtained a subjective motivation, efficiency and satisfaction that the student has perceived using this new methodology, and basic data about the elements to improve.

NEW TECHNOLOGIES				
How much interest do you have for the computers and technological advances in general?				
Nothing	Little	Something	Quite	Much
What technologies are you using from the list below?				
Mobile	Cam	MP3-MP4	Computer	
Ladtop	Console	Smartphone	Tablet	
Do you have any of these technologies? Indicates which:				
How many hours a day do you use the computer?				
<1	1-2	2-4	4-8	>8
You use the computer to:				
Estudy	Work	Leisure	Other	
INTERNET, SOCIAL NETWORK AND OTHER TOOLS				
Which device you use to connect to Internet?				
Mobile	Ladtop	PC	Smartph.	Tablet
How many hours a day you connect to Internet? (regardless of the device)				
<1	1-2	2-4	4-8	>8
Where you usually connect to Internet?				
Home	Univ.	Work	Ciber	
WIFI públic	Mobile	Other		
What type of connection you usually use? (regardless of the device)				
WI-FI	ADSL	3G	TV	Other
Which services from Internet you usually usually use?				
E-mail	Chats	Browser	Games	
Architecture	Blogs	Sports	News	Others
Do you use social networks?				
To what do you use the social network?				
Professional	Estudies	Friendship	Others	

Fig. 4. Extract of the profile test

Profile test incorporates specific question about the technology that students will use to know the degree of knowledge and expectations about its use. Some questions related are:

- Do you know what VR is?
- Do you think that is useful for your studies?
- Do you think that applicating VR would improve your?
- Do you thing that VR will be difficult in its application?
- Do you thing that VR technology can be a limitation for the final?

After completion the experiment, we will define the post-test “Fig. 5”. The main purpose of this test is to evaluate student assessments of the course content and the support technology (VR in this case). Through the structured test, based on International Organization of Standardization (ISO) 9241-11, it will be possible to evaluate the feasibility of using VR technology in educational environments while focusing on the usability guidelines of Effectiveness, Efficiency, and Satisfaction.

EFFICIENCY
The use of game engine Unity is robust and stable?
The use of game engine Unity is easy and fast for create a 3D environment?
The interface of Unity is clear and intuitive?
The interaction of the 3D world is more interactive than other tools that you known?
SATISFACTION
The quality of the 3D content models has a good definition?
The models created for the engine are difficult to achieve the minimun polygon resolution?
Is the Virtual Reality useful for 3D learning process?
Is the Virtual Reality good enough for 3D content creation?
The use of Virtual reality of the instructor is useful?
LEARNING
What is your perception of the interactive use of the Unity?
Playing with the virtual world has a improvement for understanding 3D space?
What is your perception of the utility of Unity for 3D creation?
What is your perception of the interactivity of the use of 3D worls with the professor?
What is your perception of the interactivity of the use of 3D worls with the other students?
Playing with the virtual world created with Unity is a collaborative way of learning 3D content?
What is your perception of the use of Virtual reality engine is more engaging that other 3D tools?
What is your perception of the self interactivity for use Unity engine?

Fig. 5. Usability post test

In the experimentation and investigation of the scientific hypothesis that serves as the basis of the study, a correctly designed and used “user test” will be created to allow the extraction of data to study.

In the teaching field, the typology of any test used has as a principal objective to value the usability of any new processes of the training project. This focus means that the kind of questions must be oriented to the teaching methodology and not the project itself, so that the project evaluation is realized with specific questionnaires relative to same. So that depending on the function of the training method and the obtained results, it will be possible to validate the initial hypothesis and to review a more effective implementation of how teaching methods can be incorporated into the new technologies favorably.

Surveys are designed to model the answer of the implementation of a technology or kinds of technologies. In the university, teaching resources depend on the user profile, highlighting those that are focused on measuring the effectiveness and efficiency of the course, as well as the opinion and degree of satisfaction and pupils’ preferences.

The survey will be a questionnaire that will be presented to the participants in paper format. The questions of efficacy and efficiency have been created using a Likert scale. Each question will be assigned a numerical value. The value assigned will indicate the degree of accordance or disagreement with the question one a five-point scale, so that the questionnaire is answered with accuracy in terms of the degree of accordance over the affirmations.

The Likert scale is the most-used scale in investigations where each option is valued and the answers of each person are summed to obtain a unique punctuation of a theme [15]. Consisting of a set of items presented according to affirmations or prejudices in which reactions are sought from the participants, who choose one of the five criteria indicated by the following table "Table 1".

Table 1. Likert scale

Value	Equivalence
1	Strongly disagree
2	Disagree
3	Fine
4	Agree
5	Totally agree

Technically it is chosen the Unity engine to create a virtual world for:

- Notable graphic quality.
- Upgrade plug-in architecture.
- Free except exportation to IOS and android.
- Easy scenario creation without having to program. Fast implementation.
- Is possible to export to multi-platform web and mobile devices such as android and IOS even if an expense is involved.
- Multi-player future possibility (with the smartphone plug-in, for example) so that all the class can be at the same time in the virtual town, creating a game.

The possibility exists that, once implemented, the multi-player becomes interdisciplinary. This would allow for programming and videogames by the students, adding interactivity, actors, and even mini-games within the town. For example, if a student has created a bank, in the videogame subject they may implement a banking game. Therefore all students apart from visualizing the models created in an interactive world, also will use them for executing their own routines of programming and videogames. Creation is interactive, and 100% unique. The models must have certain minimum model conditions so that unity will work.

4 Future Lines

It is also been experienced that the possibility exists to create multi-diverse worlds for avatars and to interact in the virtual world. As in second life, the alternative that it is being shuffled is OpenSim.

OpenSim adds respect to Unity including databases, avatars, positioning a house in a town inside a global world, server scripts, etc. In other words, a persistent world that is open to everyone. OpenSim is making headway as a viable alternative to Second Life. About 98% of the functionality of Second Life is present in OpenSim. The remaining 2% primarily deals with vehicle physics.

For avatar creation, existing plug-ins as EVOLVER where one can make their own avatars can be used in OpenSim and Unity where they can also be implemented in the multi-player world.

Unity and OpenSim have the potential to be engaging, interactive ways to deliver educational experiences to students. They are more productive, because they include students rather than solely relying on teachers to create content.

5 Conclusions

In this paper we have reviewed and conceptualized the research of teaching 3D arts with virtual reality technologies in a university setting, always considering the model sample selection and type of analysis.

The methodology designed will begin deployment in the academic year 2012-2013, beginning in the second half when results will be obtained in the course of “Computer Animation” at La Salle, Ramon Llull University.

In the first qualitative samples taken, it has been observed that the proposed survey design is consistent and allows analysis without any problems. It seeks to implement technology in teaching through the study of the user profile of students. The survey has been implemented with the support of digital Moodle Intranet, allowing fast and accurate results.

The expected results for students are that their academic progress will be faster and more satisfying through the use of a virtual reality environment.

This technology applies learning 3D virtual creation, 3D modeling, 3D animation, and creating spaces for architecture, which should lead to a paradigm shift in the presentation, visualization, and understanding of media and architectural projects.

For these reasons we can say that we are in front of teacher and students’ great improvement of how learn 3D arts for multimedia and architecture with virtual reality, because it pushes for collaboration among students and further achievement levels for learning 3D.

References

1. Pozo, J.I.: Teorías cognitivas del aprendizaje. Morata, Madrid (1993)
2. Piaget, J.: Inteligencia y afectividad. Aiqué, Madrid (2001)
3. Vygostsy, L.S.: Obras escogidas, vol. I-V. Antonio Machado, Madrid (1991)
4. Carr, W., Kemmis, J.: Teoría crítica de la enseñanza. Martínez-Roca, Barcelona (1988)
5. Trilla, J. (coord.): El legado pedagógico del siglo XX para la escuela del siglo XXI. Graó, Barcelona (2001)

6. Siemens, G.: Connectivism. ElearnSpace (2004), http://www.elearnspace.org/Articles/connectivism.htm
7. Downes, S.: E-learning 2.0 (2005), http://www.downes.ca/post/31741
8. Cobo, C., Pardo, H.: Planeta web 2.0 o medios fast food (2006)
9. Lankshear, C., Knoble, M.: Nuevos alfabetismos. Su práctica cotidiana y el aprendizaje en el aula. Morata, Madrid (2008)
10. Prieto, L. (coord.): La enseñanza universitaria centrada en el aprendizaje. Octaedro/ICE UB, Barcelona (2008)
11. Rheingold, H.: Multitudes inteligentes: la próxima revolución social. Gedisa, Barcelona (2004)
12. Piscitelli, A.: Internet, la imprenta del siglo XXI. Gedisa, Barcelona (2005)
13. Kemp, J.E., Smellie, D.C.: Planning, Producing and Using Instructional Media. Harper & Row, New York (1989)
14. International Journal of Information Technology and Application in Education (JITAE). Augmented Reality in architecture degree. New approaches in scene ilumination and user evaluation. Ernest Redondo, David Fonseca, Albert Sánchez, Isidro Navarro. 1(1), 19–27 (2012)
15. Barberà, E., Mauri, T., Onrubia, J. (coords.): Cómo valorar la calidad de la enseñanza basada en las TIC. Graó, Barcelona (2008)
16. Gardner, H.: Aproximaciones múltiples a la comprensión. In: Reigeluth, C.M. (coord.) (ed.) Diseño de la instrucción: teorías y modelos: un nuevo paradigma de la teoría de la instrucción, vol. 1, pp. 77–98. Santillana, Madrid (2000)

The Characterisation of a Virtual Reality System to Improve the Quality and to Reduce the Gap between Information Technology and Medical Education

Jannat Falah, David K. Harrison, Vassilis Charissis, and Bruce M. Wood

Glasgow Caledonian University
Cowcaddens Road
Glasgow, G4 0BA, UK
Jannat.Falah@gcu.ac.uk

Abstract. Contemporary medical training is hindered by an excessive amount of information provided to students through mainly traditional teaching methods yet the younger generations are accustomed to digital data and information on-demand. As such they have developed a fully customised manner of learning, which in turn requires a new, innovative and equally customised teaching method. This inherited customisation and accelerated manner of learning stems from contemporary lifestyle trends. As such, a reduced learning curve requires innovative and efficient teaching methods, which comply with existing curriculums, yet facilitate the contemporary learning mantra. In particular medical education requires a plethora of information related to the understanding of spatial relations and the three-dimensionality of the human body. Previous studies successfully employed Virtual Reality (VR) and high fidelity patient simulation in order to improve and enhance the medical education and clinical training. The benefits of this technological adoption in the teaching field offered safer experimentation environments, reduced time and cost. Furthermore the Virtual Reality facilities and systems can be extensively customised with relatively low cost and be re-used for various applications. The purpose of this paper is to identify the differences between current education methods and the proposed technology. This research will exploit current teaching trends and attempt to provide recommendations based on a University of Jordan case study. Overall the paper describes the design process of the survey questionnaire that was used for this evaluation and provides valuable insights to both academics and practitioners regarding the potential benefits and drawbacks of adopting such a system.

Keywords: Applications: Education, Virtual Reality, system characterisation, medical education, Middle East.

1 Introduction

Virtual Reality is gradually spreading as a teaching aid due to a plethora of benefits (Onyesolu, 2009). In medicine this technology facilitates several teaching and

R. Shumaker (Ed.): VAMR/HCII 2013, Part II, LNCS 8022, pp. 122–131, 2013.

diagnostic activities. In turn, VR teaching methods enable the students to investigate the human body and create a three-dimensional mental picture of the human body structures and relationships. Real-time VR applications offer a rich, interactive, and highly engaging educational context, thus supporting experimental learning-by-doing. Notably, it can contribute to increasing interest and inspiration in students and to effectively support skills (Mantovani et al, 2003).

The study presented, took place in the medical school of the University of Jordan and involved interaction and consultation with staff and students in the form of questionnaires and interviews. The obtained data were subsequently analysed according to standard tools. The findings of this study will inform the decision making process that will determine the efficiency of applying the idea of a new service in the particular University environment. The recommendations derived by this work will define the framework for the development of an exclusive facility, namely the Virtual Reality and Simulation laboratory, which will aim to enhance the traditional ways of learning in the medical field. The study develops a conceptual model that integrates the SERVQUAL Gap model (Parasuraman et al, 1988), to help the University to discover the shortfalls and evaluate student satisfaction. This will guide their improvement efforts by highlighting the weaker attributes that must be strengthened. The first stage of the empirical work has been to design, administer and analyse a questionnaire that is based on the SERVQUAL framework adapted to the University of Jordan.

2 Background

Through extensive review of existing teaching areas that require technological enhancement it was deemed ideal to engage with the medical disciplines as they involve direct and critical interaction with human beings, which requires a high level of knowledge and dexterities. This can only be acquired through practical training, and is not without potential risk to the patient as with traditional teaching methods (see one, do one, and teach one). One of the basics of the medical profession is to do no harm, and human error is unacceptable. For this reason providing a virtual environment that simulates an identical situation to what the doctor or student may encounter as part of their duties towards patients, allows them to gain confidence in dealing with these situations skillfully before proceeding to the stage of real doctor-patient interaction.

2.1 Simulation in Medical Education and Training

Medical education throughout the past decade has observed a major growth in the use of simulation technology (Scalese et al, 2007), Simulation basically reconstructs the ‘‘experience’’ of patient care, the participants can have the opportunity to practice variety of skills in a safe environment, and this kind of practice will reflect on their performance (Bond et al, 2007). Medical simulation techniques have shown great ability in other areas especially in medical disciplines.

2.2 Virtual Reality and Medical Education

Over the past few years Virtual Reality has been used in many areas and has proven to be a powerful teaching tool in several fields; one of the most important fields is medical education and there are many applications which have been designed for various fields within medicine (Vozenilk et al, 2004).

There are multiple benefits of using VR technology in medicine such as having an opportunity for repeating a training task (Riener, 2012), in addition to the chance of practice without taking a risk for the patient. Furthermore VR aims to enhance the quality of the education, raise safety, and allow extensive and effective training.

2.3 Education in Jordan

Jordan places huge emphasis on education and it is one of the most educated Arab countries. Out of a population of 6,249,000 million (Department of statistics, 2011), higher education in Jordan has a major role in the procedure of comprehensive development at numerous levels and areas. Throughout the last ten years, higher education in Jordan has observed a vast growth in terms of the diversity of study programs, techniques of teaching and learning that control both the quality and quantity and growth of higher education institutions (Higher education in Jordan, 2012).

Despite the limited human and financial resources in Jordan, higher education is within the priorities of the state, for the important role it plays on the structure of a knowledge based society.

2.4 Service Quality (SERVQUAL)

The SERVQUAL model was developed by Parasurman et al (1988) with a 22 item instrument. The instrument items perform five dimensions; Reliability, Responsiveness, Tangibles, Assurance, Empathy (Jabnoun and Al Rasasi, 2005).

The particular model offers primarily a tool for measuring and managing services. In addition, many organizations have adopted the SERVQUAL model to improve their quality. As an example the Midland and Abbey National banks used this model (Buttle, 1995). SERVQUAL is a vastly used model to rate and evaluate the service quality provided to the end users (Pawitra and Tan, 2003). Furthermore this model is used to measure the satisfaction of the customers, students, end-users for any system or for any service. As such it has been employed as the main evaluation method in a diverse set of fields across different industries, including dental services, hotels, higher education, business schools, hospitals and banking (Buttle, 1995).

The aim of this study is to measure the student experience and needs. One of the advantages of the model is measuring students experience rather than just measuring the experience of teaching (Cuthbert, 1996b). Furthermore, focusing on the learning process is much more important than focusing on the end results for the students. Cuthbert (1996a) pointed out that SERVQUAL gives attention to the service delivery against the final result and for this reason is appropriate for measuring the learning process.

3 Methodology

The objective of this research is to identify and comprehend the lecturing needs of the students within the University of Jordan and to determine ways to improve the current medical training within the University.

The first part of the research involved designing a questionnaire to explore differences between the perceived training and the student expectations. The survey instrument was adapted from the SERVQUAL model. Focus groups of friends and family medical doctors were used to pilot the questionnaire. The final questionnaire, as shown in Appendix section, contains 22 statements that reflected the five different SERVQUAL dimensions. Each dimension contains four to six questions, as shown in Table 1. Students were first asked to supply some demographic information, such as gender and age. They were then asked to rate their general expectations from the medical school on a 5-point Likert scale ranging from not expected (1) to essential (5). On the reverse side of the questionnaire students were then asked to rate their perceived qualities in the medical school using 5-point Likert scale ranging from strongly disagree(1) to strongly agree(5). Final year medical students were asked to answer a total 44 questions (22 expectations against 22 experiences), within each of five categories, as shown in Table 1.

Table 1. Definition of SERVQUAL Dimensions

Categories	Definition	Number of items in questionnaire
Tangibles	The appearance of physical facilities in the university, laboratories, physical models.	4 items
Reliability	Ability to perform the promised knowledge regularly and truthfully	6 items
Responsiveness	Willingness to help students	4 items
Assurance	Knowledge and Courtesy of doctors and their ability to convey trust and confidence to the students	4 items
Empathy	Caring, respect, attention, and friendliness	4 items

The questionnaire was distributed to 30 medical students of the University of Jordan Hospital in their last taught year, consisted of 17 males and 13 females. This group was of particular interest as its members typically have more experience in the clinical training environment than the earlier years. The feedback was collected by hard copy in order to improve the response rate. The aim of the questionnaire is to measure students' expectations in contrast to university's provision, in order to highlight the potential gaps between knowledge offering and demand.

4 Data Analysis

The analysis of the SERVQUAL questionnaire was conducted by using the SPSS software program that calculated the frequencies and the gaps (differences), within the data. The T-Test conducted was used to calculate the frequencies and the gaps (differences between perceived and expectation) Parasurman et al (1988), within the data. Preliminary results of the data analysis showed the demographics for the final year medical students including gender.

A Paired-Sample T-test of all 22 statements T-test results using the SPSS program was used to compute and compare the means scores for students' expectation and experience statements presenting satisfying results (Appendix 1). In term of expectation, the mean ranged between 3.90 (the university training is available when required) and 4.80 (school doctors are knowledgeable about their training materials) except one question, which scored 1.30 (the school doctors are often too busy to respond to the student needs). Experiences ranged between 2.83 (the physical models are flexible enough for training) and 4.30 (the school doctors give patient centered communication skills to the students) except one question that scored 1.93. This result indicates that there are gaps and the students are not satisfied in some statements, as shown in Table 2.

Table 2. SERVQUAL Scores

	Maximum	**Minimum**	**Likert Scale Ranging**
Expectation Mean	Mean 4.80: (The school doctors are knowledgeable about their training materials).	Mean 3.90: (The university training is available when required).	3.9 Range: Between No feelings and Preferred 4.8: Between Preferred and Essential.
Experience Mean	Mean 4.30: (The school doctors give patient centred communication skills to the students).	Mean 2.83: (The physical models are flexible enough for training).	2.83 Range: Between Disagree and Moderated 4.30: Between Agree and Strongly Agree.

The SERVQUAL statements were grouped into five dimensions (in both the expectations and experience sections) each with its range of relevant statements:

- Tangibility (Statements 1-4).
- Reliability (Statements 5-10).
- Responsiveness (Statements 11-14).
- Assurance (Statements 15-18).
- Empathy (Statements 19-22).

The results presented in Appendix, show the differences between the ratings, which students assigned to expectation statements, and experience statements for the 30 questionnaire responses. For each pair of statements, the SERVQUAL score was computed as follows; Service quality = Experience - Expectation. Figure 4 shows the differences between expectations and experiences for all 22 statements. The statement number 13 was a positive gap (0.633): The School Doctors are often too busy to respond to the student's needs so that means a good thing because the students pointed out that the doctors do respond to their needs.

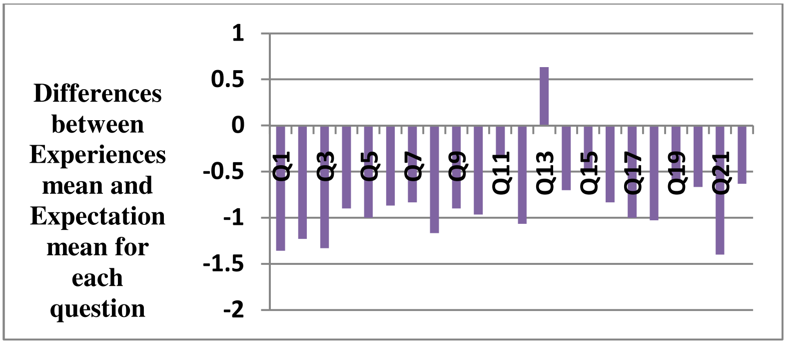


Fig. 1. SERVQUAL 22 differences

SPSS was used as an analysis tool, to calculate the differences or gaps between the student expectations and experiences over the SERVQUAL 5 dimensions. As shown in Table 2, there is a gap between expectations and experience in all dimensions, but the biggest gap was tangible and reliability which is (-1.208) (-0.955), and the smallest gap was in responsiveness which is (-0.366).

Paired-Sample T-test of the 5 dimensions the descriptive, correlation and paired-sample t-test using the SPSS program was done in order to compare the 5 mean dimension scores for expectations and experiences and their differences. The t-test was carried out to compare the mean to show the significant differences (Gaps) between the expectations and experiences of the SERVQUAL 5 dimensions.

Table 3. SERVQUAL Dimension Gaps

Dimensions	Expectations (Mean)	Experiences (Mean)	Differences (Mean)
Tangible	4.44	3.23	-1.208
Reliability	4.12	3.17	-0.955
Responsiveness	3.70	3.33	-0.366
Assurance	4.69	3.85	-0.841
Empathy	4.72	3.90	-0.825

As shown in Figure 2 there is a gap between expectations and experiences in all dimensions, but the biggest gap was in tangible, and the smallest gap was in responsiveness.

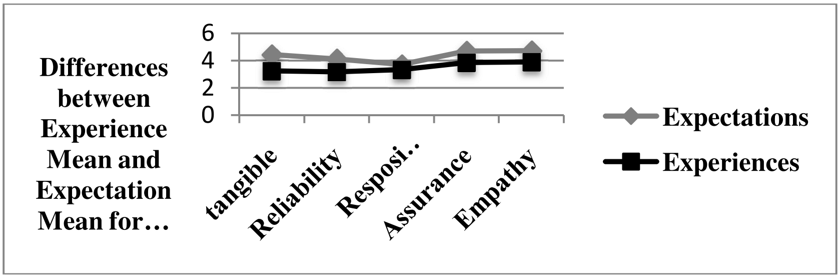


Fig. 2. SERVQUAL Dimension Gaps between Expectations and Experience

5 Conclusion

Initial Analysis of the survey of 30 respondents, who are final year medical students at University of Jordan, shows that there is a gap between student expectations and experiences in all dimensions of the SERVQUAL model, with the biggest gap in tangible attributes and the smallest gap in responsiveness. It was clear from the SERVQUAL results that there is a gap between what students expect and what the university provides. The SERVQUAL instrument identified weaknesses in the university services and the gap for all the dimensions were negative value indicators that the current methods of clinical training provided by the university do not satisfy the student expectation. The findings from this research show that there are gaps in all SERVQUAL dimensions, with the largest gap in 'Tangibles' and smallest gap in 'Responsiveness'. Adhering to the above, our future plan of work will focus further in the development of a bespoke research and teaching facility which will utilize extensively Virtual Reality and synthetic environment interaction methods. In particular the next stage for this work will have three experiments in order to make comparisons between Virtual Reality (using 3D models) vs traditional methods (using physical models). Consequently, depending on the result, the best solution will be applied to improve the clinical training within the University of Jordan.

References

1. Vozenilk, J., Huff, J.S., Reznek, M., Gordon, J.A.: See One, Do One, Teach One: Advanced Technology in Medical Education. Academic Emergency Medicine 11(11) (2004)
2. Onyesolu, M.: Virtual Reality Laboratories: An Ideal solution to the Problems Facing Laboratory Setup and Management. In: Proceedings of the World Congress on Engineering and Computer Science, San Francisco, USA, vol. 1 (2009)
3. Mantovani, F., Castelnuovo, G., Gaggioli, A., Riva, G.: Virtual Reality Training for Health-Care Professional 6(4) (2003)
4. Parasuraman, A., Zeithaml, V.A., Berry, L.L.: SERVQUAL: A Multiple Item Scale for Measuring Consumer Perception of Service Quality. Journal of Retailing 64(1), 12–40 (1988)
5. Jabnoun, N.A., AL Rasasi, A.J.: Transformational leadership and service quality in UAE hospitals. Managing Service Quality 15(1), 70–81 (2005)
6. Buttle, F.: SERVQUAL: review, critique, research agenda. European Journal of Marketing 30(1), 8–32 (1996)
7. Pawitra, T.A., Tan, K.C.: Tourist satisfaction in Singapore - a perspective from Indonesian tourists. Managing Service Quality 13(5), 399–411 (2003)
8. Cuthbert, P.F.: Managing service quality in HE: is SERVQUAL the answer? Part 1. Managing Service Quality 6(2), 11–16 (1996a)
9. Cuthbert, P.F.: Managing service quality in HE: is SERVQUAL the answer? Part 2. Managing Service Quality 6(3), 31–35 (1996b)
10. Kim, G.: Designing Virtual Reality Sytems, p. 3. Springer (2005) ISBN-13: 987-1-85233-958-6
11. Stone, J.R.: The reality of virtual reality, World Class Design to Manufacture 2(4), 11–17 (1995)
12. Riener, R.: Virtual reality in medicine, p. 1. Springer, London (2012) ISBN: 978-1-4471-4011-5
13. Department of statistics (2011): Department of statistics, `http://www.dos.gov.jo/dos_home_e/main/ehsaat/alsokan/1.pdf.pdf` (last accessed October 2012)
14. Higher education in Jordan (2012): Hashemite Kingdom of Jordan Ministry of Higher education and Scientific Research (last accessed December 2012)
15. `http://www.mohe.gov.jo/brief/briefmohe1/tabid/558/language/enus/default.aspx`
16. Scalese, R., Obeso, V., Issenberg, S.: Simulation Technology for Skills Training and Competency Assessment in Medical Education 23(1), 46–49 (2008)
17. Bond, W.F., Lammers, R.L., Spillane, L.L., Coggins, R.S., Fernandez, R., Reznek, M.A., Vozenilek, J.A., Gordon, J.A.: The Use of Simulation in Emergency Medicine (2007)
18. A Research Agenda, on Behalf of the Society for Academic Emergency Medicine Simulation Task Force 14(4), 354–362 ISSN: 1069-6563

Appendix

Appendix shows 30 questionnaire responses grouped into five categories. Expectation means and actual experience means were compared.

Tangible	**Expectations** (Mean)	**Experience** (Mean)	**Difference** (Mean)
Q1. 2D virtual models are enough to learn and understand.	4.27	2.90	-1.36
Q2. Having an opportunity for repeating a training task.	4.63	3.40	-1.23
Q3. The University provides an appropriate range of physical models for each student.	4.53	3.20	-1.33
Q4. The scenarios that are used in the lecture are relevant.	4.33	3.43	-0.90
Reliability	**Expectations (Mean)**	**Experience (Mean)**	**Difference (Mean)**
Q5. The University provides training within a controlled environment.	3.97	2.97	-1
Q6. The University training is available when required.	3.90	3.03	-0.86
Q7. The way of training in the anatomy laboratory sufficient when using the physical 3D models.	3.90	3.07	-0.83
Q8. The physical models flexible enough for training.	4.0	2.83	-1.16
Q9. The essential material is covered by the training.	4.73	3.83	-0.90
Q10. Quality of cadaver is important to avoid decay.	4.27	3.30	-.967
Responsiveness	**Expectations (Mean)**	**Experience (Mean)**	**Difference (Mean)**
Q11. The School Doctors give each student the opportunity to practice.	4.03	3.70	-.33
Q12. The School Doctors give patient centred communication skills to the students.	4.70	3.63	-1.06
Q13. The School Doctors are often too busy to respond to the student needs.	1.30	1.93	-.633
Q14. The School Doctors are always willing to help with any student problems.	4.77	4.07	-0.70
Assurance	**Expectations (Mean)**	**Experience (Mean)**	**Difference (Mean)**
Q15. School Doctors are knowledgeable about their training materials.	4.80	4.30	-0.50
Q16. School Doctors are generally able to solve any problems at the training class.	4.73	3.90	-0.833
Q17. The period of training time matches the time in the syllabus.	4.57	3.57	-1.0
Q18. The knowledge given by the School Doctors gives confidence to the students.	4.67	3.63	-1.03

Empathy	Expectations (Mean)	Experience (Mean)	Difference (Mean)
Q19. The Faculty of Medicine listens carefully to student requirements.	4.83	4.23	-0.60
Q20. The Faculty of Medicine gives each student individual attention.	4.67	4.0	-0.667
Q21. The style of training is appropriate for all students.	4.67	3.27	-1.40
Q22. The Faculty of Medicine understands a student's specific needs.	4.73	4.1	-0.633

A Mobile Personal Learning Environment Approach

Francisco José García-Peñalvo, Miguel Ángel Conde, and Alberto Del Pozo

Computer Science Department. Science Education Research Institute (IUCE), GRIAL Research Group, University of Salamanca, Salamanca, Spain
{fgarcia,mconde,delpozo1988}@usal.es

Abstract. Learning and teaching processes are not restricted to an institution or a period of time. A person can learn from experience, from the interaction with peers, because he/she has a personal interest on something, etc. Lot of such learning activities are today mediated by the Information and Communication Technologies. Such technologies allow the users decide what tools and contexts use to learn. But, in order the learning activities can be taken into account they should be visible for the institutions. In this paper a service-based framework to facilitate this is presented. It is specially focus on the communication of the mobiles devices used as a learning tool with the traditional institutional learning platforms. The framework is implemented as an Android solution and tested by students. From these tests, it can be seen that a mobile Personal Learning Environment is possible and its use motivates students' participation in learning activities.

Keywords: Mobile Learning, Mobile Devices, Personal Learning Environments, Android, Web Services, Interoperability.

1 Introduction

The way in which users interact with different systems has changed along the time. The evolution of Information and Communication Technologies (ICT) and their application in different contexts have effects in the way in which people carry out their activities [1]. One of these application areas is learning and teaching processes. ICT has influenced the way to access the contents, how the interaction is carried out among the different stakeholders, the kind of tools and resources used, etc. [2].

Nowadays, a learner uses very different applications and technologies to learn. From an institutional point of view, these tools are usually provided by the Learning Management Systems (LMS), which is a very popular and common solution in both Industry and Academic contexts [3]. However, learners not only learn in such official contexts and only with institutional applications, nor in specific periods such as academic courses, they learn along their lives, using tools that are not provided and controlled by the institution [4, 5].

To represent these other learning ways, including the personalization capabilities that belong to the current digital are we live, Personal Learning Environment (PLE) concept is defined. PLEs facilitate the users' learning process allowing them to use

R. Shumaker (Ed.): VAMR/HCII 2013, Part II, LNCS 8022, pp. 132–141, 2013.

those tools they want and not binding them to a specific institutional context or learning period such as the traditional LMS does [6]. The implementation of this concept is not an easy task because the PLE should include very different tools and they should coexist and interact with existing LMSs, especially to provide information about other learning activities that the learners may carry out. However, the learner does not use just different tools, but also different technologies such as mobile devices. Mobile technologies are quite popular nowadays and provide very sophisticated services that make them a proper platform for the definition of a Mobile PLE, mPLE.

However, regardless of the PLE underlying technology, they should interact with the LMS. The introduction of a PLE does not suppose the demise of the learning platforms [6]. LMSs have been highly successful in stimulating online engagement by teachers and learners and besides they are widely used, and large investments have been made on them [7]. The likely coexistence of LMSs and PLEs introduces a requirement for interoperation between the two contexts that should be taken into account in web and mobile platforms.

Based on this idea of mPLE there exist several solutions. Some of them consider the mobile device as a PLE by itself (with the problem about the lack of the learner's tool integration, which may mislead her) [4]. Other solutions use specific features of the mobile to adapt learning to the user (GPS, accelerometer, the camera and so on); the problem in this case is the great dependence on the mobile devices hardware and software [8]. Other initiatives use widgets to represent different functionalities in the mobile, but these have also the problems related to the diversity of software and with the low use of standards [9, 10]. Finally, there are several experiences that make easy the LMS representation in the mobile device [11, 12]; they have the problem that learners are limited to a specific set of institutional tools.

Given this context, it is necessary a solution that facilitates learners the possibility to include and use the tools they want for learning, including both institutional and non-institutional applications. In addition, it should be possible to return to the LMS the information about what is happening with these tools in the mPLE context. To do this, this contribution presents a mobile PLE based on Android technologies and an interoperability framework.

To describe this solution the paper is structured as follows. Section 2 presents the solution approach, the interoperability framework and how the mobile device is connected with the LMS. Section 3 describes the Android implementation of the solution. Section 4 poses the result of the evaluation of the implementation. Finally some conclusions are presented.

2 Proposal for a Mobile PLE

As commented above, there is a necessity to take into account the learning activities that happen in the tools used by learners in their mobile devices. To make this possible it is necessary to define communication ways between such tools and the LMSs in a way that the activities carried out in the devices can be taken into account. This is the idea of mPLE, understood as a set of mobile services, tools and communication channels that make easy the learner to carry out learning activities outside of the institutional environment with the possibility to return to it the outcomes achieved.

In this sense the authors of the paper have defined a service-based framework that, by using web services and interoperability specifications, facilitates the definition of a PLE with activities that returns information to the LMS [13]. Such framework includes a mobile device that by using web services would include into it an institutional activity. However, the framework did not consider the reality of a mobile PLE and it has been adapted (Fig. 1).

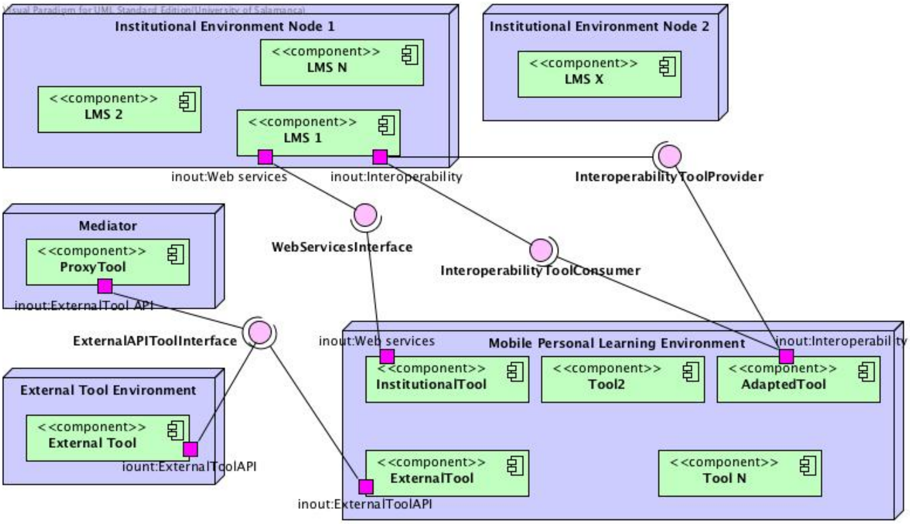


Fig. 1. Deployment diagram of the framework

The LMS is the basic tool of the institutional environment. Each node can include one or several and the framework does not define the specific LMS to use or if it is going to be use one or several. However, the LMSs should satisfy a minimum set of requirements, they should support web services and interoperability specifications. In Fig 1. is seen that the LMS implements a web service interface (WebServicesInteface) and a interface as the consumer of a interoperability specification (Interoperability-ToolConsumer). In addition, it uses the interface implemented by the tools to integrate them, the InteroperabilityToolProvider interface.

Other important elements are the tools included in the mPLE. These are the tools that the learner may use to learn in his/her personal environment defined in the mobile device. There are three types of tools:

- Tools that do not interact with the LMS. These are tools that can be employed in learning activities, but in order to take into account what the user does in them, the teacher should leave the LMS, enter into the tool and check the activity carried out. In example, the use Instagram in the mobile device.

- Tools that use the LMS web services. These kinds of tools use the web services provided by the learning platforms in order to access to information and functionalities from outside of this environments. Such tools should include a web service consumer that uses the web service interface provided by the LMS. In example, an adaptation of the forum that can be use in the mobile.
- Tools that can integrate the students' activity through the use of interoperability specifications. These tools, using interoperability specifications, can be set up and instantiated as learning activities by the teacher in the LMS. In this way, the student may use them in the mobile device and it is possible to return the learners' outcomes achieved in such context to the LMS. The teacher does not need to access to other contexts to check what the learner has done.

It should be noted that a tool can consume web service and at the same time act as a Tool Provider by using interoperability specifications.

Other kind of components, as commented above, is the mediators, also known as Proxy Tools. These components facilitate the communication among the tools and the learning environments. They have two main aims. The first one is to facilitate the integration of tools that cannot implement a Tool Provider. This means tools that cannot be accessed to modify the source code. In this case, the mediator also interacts with the tool with the interfaces that they provide (ExternalToolAPI). The other aim of the Mediator is to provide additional functionalities o pre-process data. In example, it can be used to provide an interface to evaluate learners' activity in tools that are integrated but were not thought as learning tools.

3 Implementation of the Framework and the Mobile Solution

In order to check the suitability of the framework it is implemented as a proof of concept. In consists of:

- Institutional Context. Although different LMSs could be used, several Moodle 2.1. instances will be used in the proof of concept. There are different reasons for using of Moodle in this context: apart from the fact that Moodle is one of the most popular LMSs all over the world, it also is: 1) open source; 2) developed and supported by an international community with more than 1000000 members [14]; 3) a system with more than 68000 installed servers in which there are more than 58 millions of students; 4) translated to more than 75 languages [15]; 5) has great success in different institutions [16]; and 6) it includes a web service layer that open it to new technologies and facilitates it to be integrated with service oriented architectures [17].
- Communication channels. In order to implement those channels, web services are used to exchange information and interaction with the LMS and BLTI to integrate the students' activity performed in other environments and to guarantee the portability of the framework to other contexts. The web services will be those provided by the LMS, which can be extended by following the Moodle extension protocol in case of need. However, it is not possible only to use web services because this

would mean that the framework should be adapted to the service layer of each platform to use. This is solved by using BLTI, implemented by most of LMS [18].

- The Mobile Personal Learning Environment. There are several possibilities to implement the mobile PLE.
  It can be implemented as a widget container with different widgets representing the tools. This kind of solution facilitates the representation of the applications not only in a mobile context, but also in a browser or in the computer desktop. The main problem of it is that the existing mobile run-time containers are beta products and some widgets are not working properly and such containers are linked to a specific mobile operating system (there is a version for android, other for iOS, etc.).
  Other possibility is to use web services to connect a HTML5 [19] solution. In this way, it is possible to represent the mobile device independently of the device. The problems of this solution are that it is joined to a specific LMS and which means that is not easy to include other tools. In addition, a HTML5 solution does not allow the learner to add any tool he/she uses to learn.
  Taking this into account, for the implementation it was decided to use a specific mobile system, Android. It is one of the most popular mobile devices operating system and, although this constrained the solution by a specific technology, it facilitates a free, open and scalable definition of the mobile PLE.
  In order to carry out such implementation, first a container is defined. It gives access to the user to the different tools included into the PLE (Fig. 2).

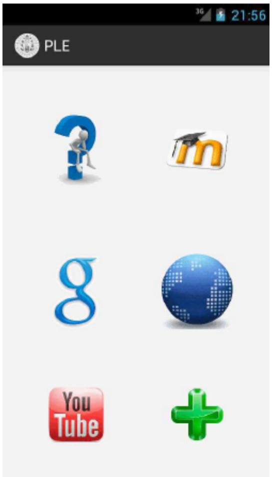


Fig. 2. Screenshot for the mobile PLE with access to all the tools

In addition, it allows leaner adding any tool installed in the phone and facilitates to download and install others. They can be institutional tools adapted to the mobile that uses the LMS web services interface; and also other tools that, by using interoperability specifications, return the outcomes of the learners' activity in the mobile environment. By default, two tools have been adapted for the mPLE. One of them is a Moodle forum. It represents how an institutional tool can be accessed through the mobile PLE and the activities that take place are reflected in the LMS (Fig. 3). With this tool the learner can read his/her forums in Moodle courses, create discussions, posts, etc.

Fig. 3. Screenshot of the forum application into the Mobile PLE

The other tool uses interoperability specifications to return learners' outcomes to the LMS. It is a quiz tool; the learner can answer quizzes and return the results to the LMS (Fig. 4). The teacher in Moodle instantiates a quiz-based activity and the learners can access to it through his/her mobile and fulfil the quizzes.

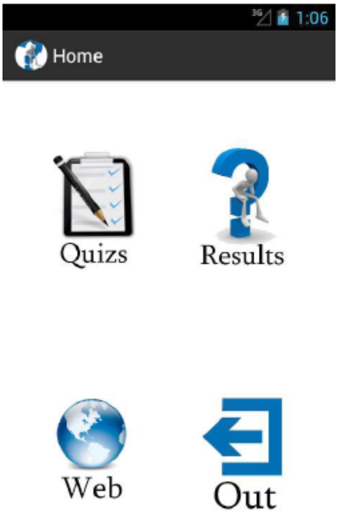


Fig. 4. Mobile Quizz tool in the Mobile PLE

Other tools installed in the phone could be also added to the application (Fig 2), YouTube, Wikipedia, Instagram, etc.

Given this implementation of the framework, different interaction scenarios are evaluated; one of them is described in the next section.

4 Pilots and Evaluation

In order to evaluate the framework and the Mobile PLE a pilot is carried out with students of the third course of Education Degree of the subject Evaluative Research in Education. The pilot involved 56 students. The methodology used to validate the system is a quasi-experimental design [20]. This methodology is used because in this experiment pre-established groups of students (class-groups) are involved, so it is not possible to have a complete randomized group of people [21]. Thus experimental design is not applicable.

Quasi-experimental design implies the definition of a hypothesis that is checked by using an experimental group and a control one (independent variable). In both groups the same tests are applied, a pre-test at the beginning of the experiment and a post-test after it. The students of the experimental group test the system (that is to say they use the forum application in the mobile Personal Learning Environment) while the people in the other group do not. After running the experiment, data is analyzed by using probabilistic techniques to validate the initial hypothesis.

The scientific hypothesis of this experiment would be "The students appreciate as a positive asset to the use of institutional functionalities in a mobile device, which helps them to learn". From such hypothesis, a dependent variable is defined: "The impact of the use of institutional functionalities through mobile devices". To operationalize this dependent variable, some asserts (also called items) have been proposed to the students and they have graded their agreement by using five value levels (1=Strongly disagree, 2=disagree, 3=indifferent, 4=agree, 5=Strongly agree).

In the pre-test:

- I.1. Sometimes I use my Smartphone to access to Moodle and its resources.
- I.2. I use my mobile device to learn through online tools and some mobile applications.

In the post-test:

- I.3. The application of online tools, mobile native applications and Moodle functionalities into the mobile help me to learn.

The scientific hypothesis is accepted if the results of the pre-test are similar in both groups (which prove that both groups are similar and have a common knowledge and background) and the results of the post-test between the persons involved in the experimental group and the control group are different (those who have tried the tool should answer in a different way). So, the following null hypothesis for both groups is proposed: **H0: $\mu E = \mu C$** (where μE is the average grade for the experimental group and μC for the control group). To check it, Student's T test was applied. The results of the first test can be seen in the Tab. 1, with a signification of a 0.05. If the signification of the item is under 0.05 the null hypothesis is accepted, if not, it is rejected.

In Tab. 1 one can see that in both pre-test items the null hypothesis is retained (that is, the experimental and control group answer more or less the same) with a bilateral signification of 0.700 and 0.449 that is greater than 0.05. In the post-test the null hypothesis is rejected (the results between the experimental and control group are different). It should be noted that in item I1 and item I2 the average for the experimental and control groups are around 2 or 3, which means that most of them do not use mobile devices to access Moodle or other learning tools. It is also interesting to consider the average of the experimental group in the post-test (4.25) which shows that the students who tested the system consider it useful for learning.

Table 1. The results of the Student's T-test

Pre-test results for Student's T test						
VD	$\overline{X_E}$	S_{X_E}	$\overline{X_C}$	S_{X_C}	t	p
I.1	2.90	1.081	2.80	1.348	-0.393	0.700
I.2	3.10	0.960	2.80	1.239	1.193	0.449
Post-test results for Student's T test						
I.3	4.25	0.749	3.27	0.978	2.315	0.021

To support this conclusion an opinion assertion about the experience was posed to the students of the experimental group. This assertion is: "After using the Moodle forum through a mobile device I consider export tools like that to mobiles make me easy to follow discussions and participate in the forum, so my learning is improved and the forums use is in my opinion more attractive". The 95% of the students agree or strongly agree with the assertion, they consider useful to export this kind of functionalities.

The conclusions obtained from these experiences allow the validation of scenarios but always from the students' perception; as a future work they should be checked in other contexts, with other kind of students, with teachers, etc. In addition other statistical tests can be applied such as non-parametric tests as the Mann-Whitney U test.

5 Conclusions

During this paper two main problems have been posed. The learners do not learn just in the institutional environments, they learn during their whole life, using very different tools and in very different contexts, and it is necessary to have such learning into account. Moreover, nowadays mobile technologies facilitate to access to that kind of learning, so it is necessary to take into account what happens in those devices. This can be done through the definition of a mobile Personal Learning Environment that allows, on the one hand, that the students can use any application of their device to learn and, on the other, to make visible such learning instances to the institutions.

In order to make this possible the paper presents a service-based framework approach. This framework, by using interoperability specifications and web services, makes possible the communication among the mobile PLEs and the institutional platforms.

Such framework has been implemented as an Android mPLE and tested with students of the University of Salamanca. From such test, it can be concluded that, from the students' perspective and in a controlled context, the opportunity to represent students' PLE in a mobile device that includes functionalities and/or information from the LMS, which could be combined with other tools they use to learn, encouraging them to participate in the subjects and helps them to learn.

As a future work, regarding the framework implementation, it is interesting to improve the client to support logging, and in this way have quantitative data about which are the most used mobile tools to learn; to define new versions of the system for other operative systems; to adapt more tools; to measure the usability of the solution; to compare the results of validation with other implementations, i.e.: HTML5 clients; etc. With regards to the validation, it is needed to consider different contexts far from the university world, with people with less knowledge about the use of mobiles, in different countries and to apply other statistical tests and qualitative techniques during the pilots.

As a final conclusion, it is possible to say that the implementation of a mobile PLE is a fact and, although the solution can be improved, the key issue is that it is possible to make visible learning activities to the institution.

Acknowledgements. This work is partially supported by Spanish Science and Innovation Ministry through the oiPLE project (ref. TIN2010-21695-C02) and also by the Regional Council of Education of Junta de Castilla y León through the projects GR47 and MPLE (ref. SA294A12-2).

References

1. García-Peñalvo, F.J.: Preface of Advances in E-Learning: Experiences and Methodologies. Information Science Reference, Hershey, PA, USA (2008)
2. Conde, M.Á., García, F.J., Rodríguez-Conde, M.J., Alier, M., Casany, M.J., Piguillem, J.: An Evolving Learning Management System for new educational environments using 2.0 tools. Interactive Learning Environments, doi:10.1080/10494820.2012.745433 (in press)
3. Prendes, M.P.: Plataformas de campus virtuales de Software Libre: Análisis compartivo de la situación actual de las Universidades Españoles., Informe del proyecto EA-2008-0257 de la Secretaría de Estado de Universidades e Investigación (2009), `http://www.um.es/campusvirtuales/informe.html`
4. Downes, S.: E-learning 2.0. Elearn magazine 2005, p. 1 (2005)
5. Wesch, M.: From Knowledgable to Knowledge-able: Learning in New Media Environments. Academic Commons (2009), `http://www.academiccommons.org/commons/essay/knowledgable-knowledge-able`
6. Adell, J., Castañeda, L.: Los Entornos Personales de Aprendizaje (PLEs): una nueva man era de entender el aprendizaje. In: Roig Vila, R., Fiorucci, M. (eds.) Claves Para la Investigación en Innovación y Calidad Educativas. La Integración de las Tecnologías de la Información y la Comunicación y la Interculturalidad en las aulas. Stumenti di Ricerca per l'innovaziones e la Qualità in Ámbito Educativo. La Tecnologie dell'informazione e della Comunicaziones e l'interculturalità Nella Scuola. Marfil – Roma TRE Universita degli studi, Alcoy (2010)

7. Sclater, N.: Web 2.0, Personal Learning Environments, and the Future of Learning Management Systems. Research Bulletin (2008)
8. Thüs, H., Akbari, M., Schroeder, U.: Mobile Personal Environment for Higher Education. In: PLE Conference 2011, Southampton, UK, pp. 1–10 (2011)
9. Aplix Web Runtine, `http://wiki.webvm.net/wrtc/` (Última vez accedido April 20, 2012)
10. Sachse, J.: The Standardization of Widget-APIs as an Approach for Overcoming Device Fragmentation. GRIN Verlag GmbH (2010)
11. Molly Project - The open source Mobile portal, `http://mollyproject.org/` (Última vez accedido April 20, 2012)
12. Jennings, T.: Technology Audit - CampusM. Ovum's Knowledge Centers (2011)
13. García-Peñalvo, F.J., Conde, M.Á., Alier, M., Casany, M.J.: Opening Learning Management Systems to Personal Learning Environments. Journal of Universal Computer Science 17, 1222–1240 (2011)
14. Moodle Stats, `http://moodle.org/stats` (Última vez accedido September 14, 2012)
15. Cole, J., Foster, H.: Using Moodle. O'Really (2007)
16. Molist, M.: Moodle llena la geografía educativa española de campus virtuales. Diario el Pais (2008), `http://www.elpais.com/articulo/portada/Moodle/llena/geografia/educativa/espanola/campus/virtuales/elpeputec/20081204elpcibpor_1/Tes`
17. Casany, M.J., Alier, M., Conde, M.Á., García, F.J.: SOA Initiatives for eLearning: A Moodle Case. In: 23rd International Conference on Advanced Information Networking and Applications, AINA, May 26-29, pp. 750–755. IEEE Computer Society, Bradford (2009)
18. Common Cartridge and Basic Learning Tools Interoperability Progress and Conformance Status, `http://www.imsglobal.org/cc/statuschart.html` (Última vez accedido August 30, 2012)
19. W3C: HTML5. A vocabulary and associated APIs for HTML and XHTML (2011), `http://www.w3.org/TR/html5/`
20. Campbell, D.T., Stanley, J.C.: Diseños experimentales y cuasiexperimentales en la investigación social. Amorrortu Editores (1970)
21. Dendaluce, I.: Diseños cuasiexperimentales. In: Hoz, V.G. (ed.) Problemas y Métodos de Investigación en Educación Personalizada. Rialp, Madrid (1994)

Perceived Presence's Role on Learning Outcomes in a Mixed Reality Classroom of Simulated Students

Aleshia T. Hayes[1], Stacey E. Hardin[2], and Charles E. Hughes[3]

[1] University of Central Florida: Synthetic Reality Lab., Institute for Simulation and Training, Orlando, FL, USA
[2] University of Central Florida - College of Education, Orlando, FL, USA
[3] University of Central Florida: Synthetic Reality Lab, Computer Science Division, Orlando, FL, USA
{Aleshia.prof,Staceyehardin}@gmail.com, ceh@cs.ucf.edu

Abstract. This research is part of an ongoing effort on the efficacy and user experience of TLE TeachLivE™, a 3D mixed reality classroom with simulated students used to facilitate virtual rehearsal of pedagogical skills by teachers. This research investigated a potential relationship between efficacy, in terms of knowledge acquisition and transfer, and user experience in regard to presence, suspension of disbelief, and immersion. The initial case studies examining user experience of presence, suspension of disbelief, and immersion were used to develop a presence questionnaire revised from the work of Witmer and Singer (1998) to address the TLE TeachLivE™ mixed reality environment. The findings suggest that targeted practice, authentic scenarios, and suspension of disbelief in virtual learning environments may impact learning.

Keywords: Mixed Reality Classroom, Simulation, Presence, Suspension of Disbelief, Immersion, Engagement, Knowledge Acquisition, Virtual Learning.

1 Introduction

1.1 The Environment

The research reported here is part of a large scale study at the University of Central Florida investigating the efficacy and user experience of TLE TeachLivE™, a 3D mixed reality classroom with five simulated students, used to facilitate virtual rehearsal of pedagogical skills in pre-service and practicing teachers. The classroom consists of five students cognitively and behaviorally modeled after research-based student archetypes whose avatars are visually modeled to be representative of diverse middle school populations. The classroom and students are displayed on a large screen high definition display and the current interface tracks user movement allowing the teacher's physical movement to be reflected by changes in the perspective position of the virtual camera. This creates a relative point of view, which is

R. Shumaker (Ed.): VAMR/HCII 2013, Part II, LNCS 8022, pp. 142–151, 2013.

reflective of his or her physical proximity and orientation in relation to the students and the classroom environment. The virtual students, through human in the loop simulation, interact with each other and with the teachers. Teachers move freely in the environment to interact with students in the simulated classroom as they deliver practice lessons. The student avatars respond to teacher's questions, behaviors, classroom strategies and lessons in ways that are authentic to each student's characteristics and learning styles as well as the pedagogical delivery of the teacher.

1.2 System Function

Teachers enter the simulator and virtually rehearse pedagogical or content driven objectives. Immersed in TLE TeachLivE™, they deliver lessons or spontaneously interact with the five virtual students' avatars. The learning objectives are extensive, ranging from classroom management to science or literacy. Iterative virtual rehearsals constitute experiential learning, enhancing pedagogical and interpersonal skills.

Fig. 1. Teacher delivering a lesson to the TLE TeachLivE™ classroom students

1.3 The Study

This formative pilot study explores user experience of TLE TeachLivE™ in regard to the experience of presence, immersion and suspension of disbelief. This research thrust is building toward an investigation of the impact of perceived presence, immersion, and suspension of disbelief on knowledge acquisition and learning transfer. Frequency and fidelity of practice can lead to effective application in the field (Dieker, Hynes, Hughes, & Smith, 2008; Mahon, Bryant, Brown, & Kim, 2010). Presence, suspension of disbelief, and immersion were chosen as constructs effective at describing and explaining interpersonal communication and human psychology involved with teaching interpersonal skills (Taylor, 2002; Wagner et. al., 2009). The interrelated nature of these constructs confounds research of them as distinct entities (Harteveld, 2011; Murray, 1997). The reciprocal relationship between presence and immersion was identified as a potential confound to this research, as many scholars erroneously use the terms interchangeably (Witmer & Singer, 1999; McMahon, 2007; Harteveld, 2011). Embracing and studying the symbiosis between presence, suspension of disbelief, and immersion can elucidate significance of these elements on learning.

Because TLE TeachLivE™ is used predominately with educators in K-12 classrooms, the learning objectives were derived from accepted knowledge, skills, and abilities for educators. These behaviors and learning objectives were synthesized from the Bill & Melinda Gates Foundation's Leveraging High Level Practices and the Measures of Effective Teaching (MET) study (Foundation, 2010).

2 User Experience

2.1 Presence

The International Society for Presence Research defines presence as, "a psychological state or subjective perception in which even though part or all of an individual's current experience is generated by and/or filtered through human-made technology, part or all of the individual's perception fails to accurately acknowledge the role of the technology in the experience" (2000). Slater and Usoh (1993) distinguish factors of presence as either exogenous or endogenous; exogenous factors are created by the generation of the virtual environment, while endogenous factors are subjective and occur within the user. This distinction calls for analysis of experiences by multiple methods to deepen and enrich understanding of the potential impacts of presence.

2.2 Suspension of Disbelief

Wirth & Saskia synthesize body of knowledge on suspension of disbelief and distill it to the "tolerance of media users towards unreal or implausible content in fictional media" (2005). For the purposes of this study, we expound on this to define suspension of disbelief as the phenomenon in which a participant is able to overlook and even forget the fact that the environment is not natural, but constructed and contrived, in order to enhance engagement, presence, and belief of the experience (Boelstorff, 2011; Dede 2009; Maynes et al. 1996; Jeffries, 2000; Kantor et al. 2000; Hindle 2002; Kushner, 2004; LeRoy et al. 2008; Park, Calvert et al. 2008; Serby, 2011.)

2.3 Immersion

Scholars explain immersion as an individual's subjective experience of virtual objects in which they seem to be authentic, which is facilitated by the user's willing suspension of disbelief (Dede, 2009; Witmer & Singer, 1999). Dede (2009) refines immersion to, "subjective impression that one is participating in a comprehensive, realistic experience" (p.1). The authors of this paper synthesized these to call immersion "an experience in which participants feel not only that they are 'there' with the virtual characters, but also that they have meaningful impact on the environment and entities in the environment" (Dede, 2009; Witmer & Singer, 1999; McMahan, 2007).

3 Measuring Learning and Transfer

Research has been conducted for years attempting to quantify the return on investment in areas of functional fidelity, physical fidelity, and other contributors to user experience (Martin, 1981; Lapkin & Levvit-Jones, 2011). Some of the methods used to distinguish this relationship include cost-utility analyses that compare the learning outcomes gained from simulators of different levels of fidelity (Lapkin & Levvit-Jones, 2011; Aldrich 2009). The researchers for this study chose to focus on exploring the relationship between the user's experience and the learning outcomes and transfer of training. In order to test this, we integrated learning objectives into scenarios that could be measured through evaluation before and after the training was delivered. Pedagogical strategies were the education constructs identified for this study: a) specific vs. general praise, b) wait time, and c) higher order questions.

3.1 Education Constructs to Be Evaluated

Specific Praise. For the purposes of this study, specific praise refers to positive statements about performance that are explicit in identifying the exact behavior, in order to reinforce and increase the occurrence of the targeted behavior (Kalis, Vannest, & Parker, 2007; Hawkins & Heflin, 2010; Feldman, 2003); for example, "Good job showing your work on every question in your assignment". It is a combination of a positive statement linked to the behavior being reinforced such as "Excellent work using a strategy to write your paragraph" (Scheeler, Bruno, Grubb, & Seavey, 2009).

Wait Time. For the purposes of this study "wait time" is defined as the elapsed time after a teacher ask students a question, and before students respond or the question is rephrased or repeated (Stahl, 1994; Tincani & Crozier, 2007; Novak, 1963).

Higher Order Questioning. For the purposes of this study "higher-level questioning" is posing questions that allow students to use past experiences, prior knowledge, and previously learned content and relate it to newly learned content in order to create an open ended and well thought out answer (Danielson, 2011; Winne, 1979). Teaching Works, in their report entitled Measures of Effective Teaching, distinguish higher order questioning as strategy of a highly effective teacher (Foundation, 2011).

4 User Experience Measures

4.1 Qualitative User Experience: Interviews

In the interview the participants elucidated their experience of presence, suspension of disbelief, and immersion through self-report. The interview questions included learning, suspension of disbelief, presence and immersion:

Overall, how successful do you feel your virtual rehearsal performance was?
How can you tell that the students are engaged or not engaged with you?
How did the virtual students compare to students you encounter in a classroom?
How did the virtual classroom compare to your experience of a physical classroom?
When you were teaching the virtual students, were you able to suspend disbelief?
When teaching the students did you feel like you were in the same physical space as them?
When teaching the students did you feel like you were in the same physical space as them?
How would you describe your use of specific praise?
How would you describe your use of higher order questioning?
How do you feel you used wait time?

4.2 Qualitative User Experience Measure: Observation

Observation was utilized by the researchers to interpret the user experience. Presence can be measured by observing reflexive responses to stimuli, such as a participant reaching to catch a ball or flinching or jumping at a stimulus (Sheridan, 1994). This reflexive response can be a physical response, but may also be a reflexive social response to measure presence (Sheridan, 1994). The social responses might include replying to a question, apologizing, or simply saying goodbye before walking away. While this measure could also be refined to an objective measure of the level at which a user has suspended disbelief, this study does not apply it as such a measure.

4.3 Quantitative User Experience: Questionnaire

The interview questions for the study were derived from the operational definitions for presence, suspension of disbelief, and immersion. After explaining to participants the meaning of each construct, the researcher then asked them to verbally evaluate the experience according to each of the three. This study utilized interviews and questionnaires that were derived from the Witmer and Singer constructs and their Presence Questionnaire (1998). The Witmer and Singer questionnaire measuring a virtual environment was modified to reflect the needs of a mixed reality environment.

The researchers anticipated relationships between suspension of disbelief and feelings that the students and environment felt real would be revealed by the questionnaire data. The hypotheses that are being tested are:

H1: There will be a relationship between suspension of disbelief and the rating of the environment feeling real.
H2: There will be a relationship between suspension of disbelief and the rating of the students feeling real.

5 Methodology: Mixed Methods Inquiry

This study began with both subjective measures and objective measures to explore the user experience and the learning outcomes. The participants were practicing and pre-service (K-12) teachers in a southeastern state. The time frame of the study was a period of three weeks that began with a baseline observation of their teaching style, followed by three ten minute sessions in TLE TeachLivE™ that concluded with after action review (AAR) of performance, followed by a final observation in their classroom.

5.1 Stage 1: Case Study Methodology

The first stage of this formative research was qualitative, utilizing open-ended user experience interview questions in two case studies investigating the constructs of presence, suspension of disbelief, and immersion. Two middle school teachers were observed in their live classrooms in order to establish a baseline of performance for the constructs (wait time, higher order questioning, and specific praise). The teachers were then immersed in TLE TeachLivE™ with session objectives of increasing wait time, higher order questioning, and specific praise. Upon completion of each of three 10 minute sessions, the teachers were given feedback in the form of After Action Review (AAR). After the last session in the mixed reality classroom, the teachers were observed again in their live classroom for a post intervention evaluation of performance. The teachers in training were asked open ended questions about their perceptions of the mixed reality classroom environment and the authenticity of the simulated student avatars after their sessions.

In both cases, responses indicated some initial apprehension with the mixed reality classroom environment and the student avatars. They also indicated that the teachers felt that the students in the mixed reality classroom were very much like students that they experience in a live classroom. The participants also indicated that that they established emotional relationships with the students such as frustration, empathy, joy, and pride when they succeeded in getting them engaged.

5.2 TLE TeachLivE™ Presence Questionnaires

The first stage of this formative phase of this efficacy research study moved to explore the modified presence questionnaire. The preliminary questionnaire was administered to a convenience sample of 24 pre-service and practicing teachers who were assigned to teach a ten minute lesson in TLE TeachLivE™ by the professor for their teaching instruction course. After teaching in TLE TeachLivE they were administered the abbreviated TLE TeachLivE™ presence questionnaire.

6 Current Findings

The preliminary case study supports the idea that virtual learning environments impact learning with targeted practice, authentic scenarios, presence, and suspension of disbelief.

6.1 Does the Experience of TeachLivE™ Effect Learning Outcomes?

The methods in the baseline research were to observe the teachers as they taught their students in their professional classroom environment. The frequency of the target behaviors, specific praise, higher order questioning, and wait time were measured by two raters, to ensure inter-rater reliability. The teachers were immersed into the environment three separate instances over a period of two weeks, in the TeachLivE™ classroom as the teacher for 10 minutes. After virtual rehearsal, teachers were given a chance to reflect and also receive feedback in an After Action Review session. Students shared what their perception of their performance was and were met with the reality of how they had actually performed. Each day, the users in the case study improved their performance of the targeted behaviors in the mixed reality classroom environment. The qualitative research revealed user's learning in the lab, as evidenced by the fact that they improved the target behaviors with each iteration of teaching the virtual students. The skills also transferred to the physical classroom with live students, as scores of each of the behaviors were considerably higher in the second measure than the baseline measure of practice in the classroom.

6.2 Do Students Using TeachLivE™ Experience Presence?

The interviews, observations and questionnaires revealed that users are experiencing presence in TLE TeachLivE™. In post intervention interviews, teachers commented:

"I was so nervous."
"I can't believe I made Sean cry; I feel so bad,"
"I couldn't get her to put her cell phone away."
"It just feels so real."

The observations exposed the characteristics of presence. The participants demonstrated behaviors that indicated presence, such as walking up to the virtual students' approximate locations in the physical space while speaking to them.

Similarly, presence is also evident in the fact that when the session time ends, teachers try to "wrap up the lesson." They say goodbye to the students, and they start planning the next session. This corresponds with the presence measure of reflexive response. These natural responses are transferring to the classroom.

6.3 Are Students Using TeachLivE™ Suspending Their Disbelief?

The case studies and observations of education students indicate that these students are able to suspend their disbelief in the mixed reality classroom. This suspension of disbelief is demonstrated by their emotional response to the virtual students' characteristics and behaviors. The responses in the interview questions asking students if they were suspending their disbelief indicated that they were. This could be observed as teachers rarely treat the simulated students differently than they would real students; which the researcher's code as an indicator of participants having suspended disbelief.

Statements such as, "Next time I will work at managing Kevin a little better," indicate the suspension of disbelief because the participants are talking about Kevin in a way that ignores the fact that he is a simulated student and not "real".

6.4 Do Students Using TeachLivE™ Experience Immersion?

The interviews revealed that participants felt immersed, according to our definition. Relationships teachers developed with the simulated students and the sense and behavior indicating they felt they could impact the students demonstrated immersion. Likewise, the questionnaire data demonstrated that a majority of users feel the students portray "living kids" (see Table 1). Participants rated the mixed reality classroom as moderately to not at all like a physical classroom. This was consistent with interview responses indicating perceived functional limitations to their practice in the system, such as inability to physically manipulate desks and the limitation of five students.

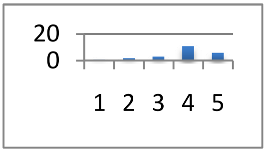


Fig. 2. Rating of the level of realism of the virtual students

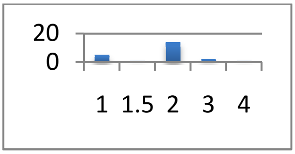


Fig. 3. Rating of the level of realism of the mixed reality classroom

6.5 Is There an Interaction between Presence, Immersion, and the Suspension of Disbelief?

This research is focusing primarily on the endogenous factors occurring within the user. Specific constructs investigated in the area of user experience included sense of immersion, presence, and suspension of disbelief. While the questionnaire data did not reveal any significant interaction, the case study interviews support the hypothesis that there is some kind of interaction between these constructs.

7 Discussion

These preliminary findings have created a foundation for an additional, larger scale quantitative research study that is currently in process to investigate the relationship between learning and the user experience; specifically suspension of disbelief and presence. The long-term research path for this research includes the exploration of the potential impact on presence derived from different interfaces through which the virtual students may manifest themselves.

7.1 Discussion: Qualitative Findings from Interviews

The interviews revealed that participants felt a sense of presence in the classroom, but wanted even greater immersion, asking for new features. Teachers indicated they would like to be able to see work samples of each student's actual progress and be able to look over their shoulders. Some of the teachers asked for higher physical fidelity, wanting desks in the room to indicate exactly where the students would be physically located as they navigated the mixed reality classroom. While these experiences were easy to discuss qualitatively, the questionnaire did not differentiate them in the same way.

7.2 Discussion: Questionnaire Findings

The questionnaire data did not prove as informative as we hoped, but it has provided a baseline for future qualitative questions. The questionnaire provides only ordinal data in a case where ratio data would be more useful. This could be addressed by changing the scale from a Likert to semantic differential, in order to give the participants shared reference criterion. For instance, would a participant rating of highest realism for the avatar represent a sense of its being indistinguishable from "living" children or the most realistic that the participant had witnessed to date?

Finally, the sample of 24 participants who completed the questionnaire was not large enough to adequately represent the population. This was further confounded by the fact that half of the participants had teaching experience in a physical classroom while the other half had no experience. This was confounding as it effectively rendered our sampling as being two samples of 12 participants in each group.

7.3 Future Research

Future research in this area is currently being conducted. The next phase of the research into the efficacy of the TeachLivE™ simulator is a large-scale study that will include 200 practicing teachers across the United States. The control group of this study will have their performance evaluated at the baseline and again after three weeks. TLE TeachLivE™ users will complete the revised presence questionnaire.

Acknowledgements. We would like to acknowledge the Bill & Melinda Gates Foundation, whose gift in part supported the TeachLivE™ team and the execution of this study, and the National Science Foundation (CNS1051067) whose support contributed to the technical infrastructure. We would like to acknowledge the other principle investigators (PIs) of TLE TeachLivE™, Drs. Lisa Dieker, and Michael Hynes. We would also like to acknowledge the work of Dr. Carrie Straub, the coordinator of the second phase of TeachLivE™ research. Finally, we would like to acknowledge the efforts of our partner universities, whose reflections facilitated development and yielded consensus to the areas of highest concern. Opinions, findings, and conclusions or expressed in this material do not necessarily reflect the views of these supporters.

References

1. Aldrich, C.: Learning Online with Games, Simulations, and Virtual Worlds: Strategies for Online Instruction. Jossey-Bass, San Francisco (2009)
2. Bailenson, J., Beall, A., Loomis, J., Blascovich, J., Turk, M.: Transformed Social Interaction: Decoupling Representation from Behavior and Form in Collaborative Virtual environments. Presence: Teleoperators& Virtual Environments 13(4), 428–441 (2004)
3. Boellstorff, T.: Virtuality. A Companion to the Anthropology of the Body and Embodiment, 504–520 (2011)
4. Dieker, L., Hynes, M., Hughes, C., Smith, E.: Implications of Mixed Reality and Simulation Technologies on Special Education and Teacher Preparation. Focus on Exceptional Children 40(6), 5 (2008)
5. Foundation, B.M.G.: Learning about Teaching: Initial Findings from the Measures of Effective Teaching Project. Bill & Melinda Gates Foundation, MET Project (2010)
6. Harteveld, C.: Triadic Game Design Balancing Reality, Meaning and Play (2011)
7. International Society for Presence Research.: The Concept of Presence: Explication Statement (2000), `http://ispr.info/` (retrieved January 2012)
8. Kalis, T., Vannest, K., Parker, R.: Praise Counts: Using Self-Monitoring to Increase Effective Teaching Practices. Preventing School Failure 5193, 20–27 (2007)
9. Mahon, J., Bryant, B., Brown, B., Kim, M.: Using Second Life to Enhance Classroom Management Practice in Teacher Education. Educational Media International 47(2) (2010)
10. McMahan, A.: Immersion, Engagement, and Presence: A Method for Analyzing 3D Videogames. In: The Video Game Theory Reader, pp. 67–86. Routledge (2003)
11. Murray, J.: Hamlet on the Holodeck: The Future of Narrative in Cyberspace, pp. 98–99. The MIT Press, Cambridge (1997)
12. Sheridan, T.: Further Musings on the Psychophysics of Presence. Paper presented at the Systems, Man, and Cybernetics. IEEE Conference on Humans, Inf. and Technology (1994)
13. Slater, M., Usoh, M.: Presence in Immersive Virtual Environments. In: IEEE Virtual Reality Annual International Symposium, Seattle, Washington, pp. 90–96 (1993)
14. Slater, M.: Measuring Presence: A Response to the Witmer and Singer Presence Questionnaire. Presence: Teleoperators& Virtual Environments 8(5), 560–565 (1999)
15. Witmer, B., Singer, M.: Measuring Presence in Virtual Environments: A Presence Questionnaire. Presence: Teleoperators& Virtual Environments 7(3), 225–240 (1998)
16. Wirth, W., Boecking, S.: Towards conceptualizing suspension of disbelief for communication research: International Communication Association (2005)
17. Nakano, Y.I., Ishii, R.: Estimating User's Engagement from Eye-Gaze Behaviors in Human-Agent Conversations. In: Proceedings of the 15th International Conference on Intelligent user Interfaces (IUI 2010), pp. 139–148. ACM, New York (2010)

The Building as the Interface: Architectural Design for Education in Virtual Worlds

Luis Antonio Hernández Ibáñez and Viviana Barneche Naya

VideaLAB. Universidade da Coruña. Spain
{luis.hernandez,viviana.barneche}@udc.es

Abstract. This paper focuses on architectural spatial design for virtual tridimensional learning environments through the lens of a case study. The work describes the design methodology of a flexible and interactive set of virtual constructions where the architecture itself acts as a dynamic interface whose spaces adapt to the activities that avatars carry out in their interior, sometimes interacting with them. This approach considers the multiple innovative parameters that have to be taken into account in the process of cyberarchitectural design.

Keywords: Metaverses, Virtual Worlds, Cyberarchitecture, V-Learning.

1 Introduction

"We are on the extreme promontory of the centuries! What is the use of looking behind at the moment when we must open the mysterious shutters of the impossible?"
Filippo Marinetti. Manifesto of Futurism (1909)

Throughout history, technological changes have contributed to generate strong cultural and social changes. The "digital revolution" that arose from ICT's produced a profound transformation in the way people live, work and learn.

The number of virtual worlds is increasing day by day. Millions of users make practical use of virtual worlds for communication and as a mean for developing a wide range of activities [1]. Following this growing trend, an increasing number of educators and organizations make use of virtual worlds to create teaching programs and learning activities [2, 3]; that employ virtual architectural spaces that emulate their real equivalents.

Usually, the architectural design of virtual facilities follows the same compositive an organizational rules that are used in the design of buildings in the real world. That is commonly done as an easy way to provide the user with a reference frame which looks familiar to him by the use of well-known building typologies. The lack of background in architectural design of the vast majority of virtual world designers also plays an important role in this direction.

The design of such virtual buildings can follow very different rules than those used in the architecture of the real world, but in spite of the large number of cases of virtual

R. Shumaker (Ed.): VAMR/HCII 2013, Part II, LNCS 8022, pp. 152–161, 2013.

constructions used for educational purposes that can be found, examples of research under the perspective of their architectural design are almost nonexistent. The literal translation of the real world architectural language to the virtual worlds is becoming a burden that delays the raising of new formal languages which should be exclusive of the virtual realm.

Those innovative aspects of the virtual medium constitute a field for research and experimentation in architectural design. Cyberarchitects can get rid of rules and limitations of the real world such as gravity, weather or even the rules of Euclidean geometry, being able at the same time to associate behaviors to the contents they create, thus allowing the interaction of the user-avatar with objects and the environment.

There are many concepts in the field of cyberarchitectural design that do not have an equivalent in the architecture of real world; some refer to new properties than can be found in the virtual building itself, such as immateriality, the possible absence of gravity or the mutability of shapes and materials. Other concepts relate to the avatar, such as anthropometry (even for non-human avatars), ability for flight, or spatial location of the avatar's camera. Those aspects are also important in the design of the virtual buildings and urban spaces intended to host interactive educational activities.

2 Objectives

This paper deals with several aspects of spatial design of virtual tridimensional learning environments by means of a case study: a virtual world, named "Isla Videa", built in Second Life which emerged as a formal response to the need of implementation of a virtual site for a master degree in the authors' university.

Two courses, with a duration of 25 hours each, on the topic of tridimensional interactive design, were given in cyberspace with both professors and students from different geographical locations being there through their avatars. This courses were part of the "Master on Digital Creation and Communication" offered by the university.

Considering the aforementioned concepts, the design of those virtual facilities sought to obtain a flexible and interactive learning environment inside a dynamic architecture, with responsive spaces capable to mutate in order to adapt to different activities. This research considered the multiple new parameters of this kind of architecture that can be now taken into account in the design process. The building experience itself represented a challenge for the authors-architects, since they had to work together in the virtual environment from two different physical locations.

The subject of the courses was "Interactive 3D design". Twenty students and four teachers took part in each experience, all of them located in different physical places. Aside from the educative goals of the project, the virtual environment served as a test bed for the interactive 3D contents created by the students.

3 Methodology

The process to implement these virtual facilities was divided in four interrelated stages: definition of parameters of spatial design; interaction design; building of the architectural model and usability test. This can be considered an interesting mix of the processes of creation of two apparently distant disciplines such as the architectural design and the creation of digital interactive contents.

3.1 Definition of Parameters of Spatial Design

Cyberspace is rich in singular concepts that can be of use in virtual architecture design, expanding the creative limits of the cyberarchitect. There were two important key concepts, immersion and presence, inherited from Virtual Reality and related to space perception [4] that played an important role in the design of Isla Videa.

The initial sketches were drawn based on the analysis of the movements and trajectories of the users inside the metaverse as a mean to define the spaces, the limits of the buildings' skins and the behavior of the different elements that compose this project. Many aspects of cyberarchitectural design utilized in this case have already been described by the authors in previous publications. [5]

The study of circulations includes criteria derived from the flight capabilities of the avatars and the use of teleporting. Hence, the different proposed pathways act in this project as helpers for orientation, defining at the same time the zoning for activities.

The absence of predetermined physical laws in the metaverse make possible to create floating, weightless architectures, without the need of structural elements such as pillars or beams. Consequently, walls do not fulfill any supporting role nor they serve as a protection against climatic agents. Walls are used just to limit spaces and the activities which take place inside them and to control visual permeability.

This medium also enables the mutability of shapes to interactively adapt to user's needs. That results in designing volumes and skins for the buildings with a flexible behavior which can also display variability in their materials like changing colors or transparency, morphing textures, etc. Text, video, sound and web content can also be used to define the materiality of a constructive element.

3.2 Building the Architectural Model

The design of Isla Videa starts from the idea of the utilization of architecture as an interface for the dynamic organization of space, where the virtual building respond to the avatar's needs as he or she carries out learning or social activities.

Based in such premises, the project unfolds from two structural axis: a water course that crosses the island from North to South defining two large zonings of activities which materialize in the form of two platforms and a belt, or pathway that runs Southeast to Northwest integrating transversally all activity zones (Fig.1).

Platforms are organized based on their main use. On one side the one corresponding to the Main Building, and the Exterior Classroom that takes the shape of a floating flat volume covering the FlexSpace located underneath, whose name derives from the shapes that it holds and its use.

Fig. 1. Overview of the project

Crossing the watercourse one arrives to the Sculptures Field, a large area designed to host exhibits. Over this area, several floating platforms hold the personal working space for students. A big sphere suspended in the air acts as Entry and Welcome Point. Annex to this platform is the Deck over the channel on one side and the Beach and the Auditorium on the other side.

The majority of the built elements that constitute Isla Videa, from the buildings to the urban environment support some class of user interaction through LSL scripting. For instance, part of the pavement of the Exterior Classroom (Fig. 2) raises on demand and swings until it is transformed into a giant display where teachers and students can drop graphic contents to share and discuss during the course activities.

The transparent box that gives shape to the Main Building is as a permanent visual link among the different spaces of the project (Fig. 3). This building is used a neutral container for multiple activities. Entering the building, the user attention focuses on two suspended elements, the Floating Classroom and the Mediatheque, both situated over a lower level dedicated to creative activities and audiovisual exhibition on screens activated by user's presence.

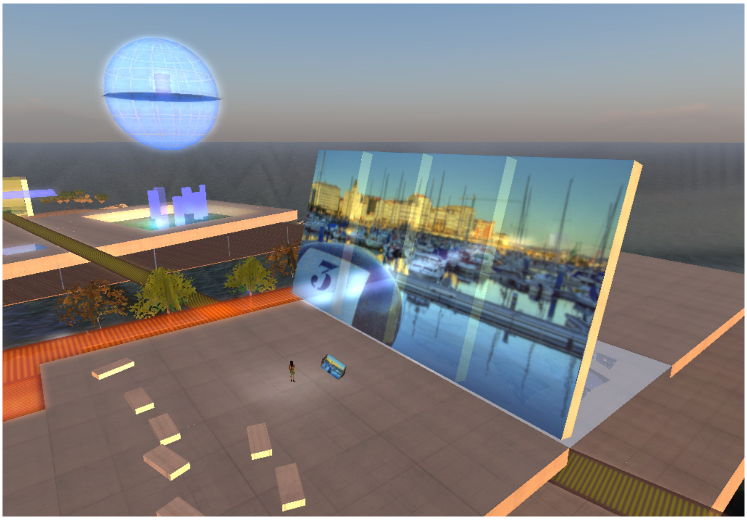


Fig. 2. Exterior Classroom

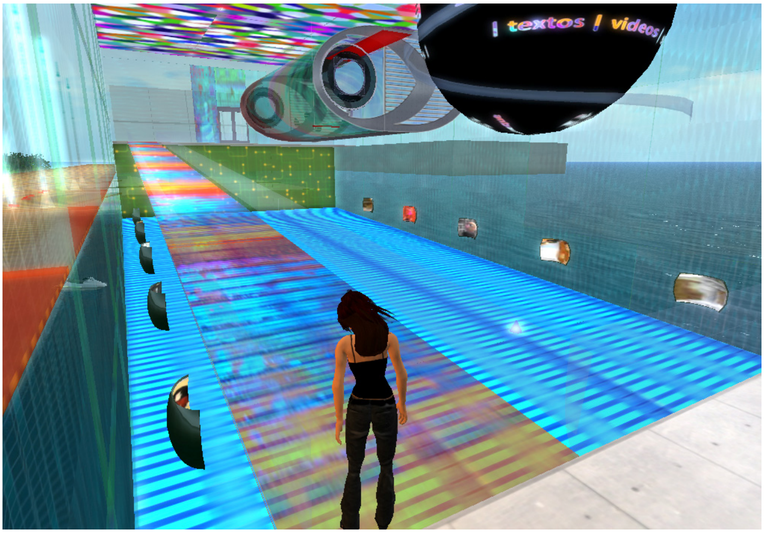


Fig. 3. Main Building. Interior space.

A network of teleporters permits the transport among the different elements, as it similarly happens throughout the island.

The Floating Classroom, shaped as an ellipsoidal cylinder, is a self-enclosed space designed either as a meeting room or as a classroom for small audiences. The element is defined by an exterior skin of constantly varying appearance while from the interior behaves as a translucent object. Lattice windows in both edges allow the users to watch the activities taking place underneath. It has no door, since access and exit are done by teleportation.

Fig. 4. Floating Classroom

On the inside, the room is equipped with a multimedia table that can be used to make presentations to all avatars seated around it. Once the activity is finished, it can be stored. For this purpose, table and seats fold together and the set transforms into a decorative lamp attached to the ceiling. (Fig.4).

The black sphere of the Mediatheque, with interior walls that seem to be made of textured light; host the information containers that hold the course notes, resources, assignments and documents of teachers and students (Fig.5).

Exiting the Main Building and using the teleporting network, the lower platform can be accessed. This zone contains several morphing spaces generated by the movement and continuous change of shape of several video screens, like cloths that twist and fold under the influence of user presence and the wind. Those FlexiSpaces are intended to integrate art, music and audiovisual works.

This area is bordered by a Reactive Garden, whose plants react to the user presence and movement changing their colors. The pavement follows the steps of the avatar illuminating its tiles on every step. The end of the garden is presided by a kinetic sculpture made of primitives and particle systems.

Crossing the watercourse along the walking belt, the pathway ends at the Auditorium over the sea, which is resolved with simple formal elements, but using the "ghost" feature available in the system to model the seats, so they can be crossed through to reach directly the seat desired instead of the classic row-column movement.

Fig. 5. Mediatheque

Fig. 6. A class in the Auditorium

The Sculpture Field completes the island. It is made by a combination of cubes of light and opaque materials forming a large open space to display academic works and host temporary exhibitions.

3.3 Interaction Design

One of the advantages of the use of metaverses like Second Life or Opensim is the possibility to program behaviors associated with every geometric element by means of a scripting language. This allows multiple levels of interaction between the user-avatar and the virtual environment. Therefore, both the container and the contents can be modified anytime to adapt to new requirements.

Three kinds of interaction were established in Isla Videa: touch, used to manually activate events, like unfolding a projection screen; proximity, used to trigger events automatically, like door opening in presence of avatars and location, used to modify the shape or orientation of objects with respect to avatar's location, like in the cloth-screens situated in the FlexiSpace.

There were other modes of interaction not related to avatars. Objects could react to solar time, i.e. turning on and off the lights of buildings and lampposts at dusk and dawn. Some objects could also communicate with others, such as the presentation devices that teachers used to control the images displayed in one or several screens along the island.

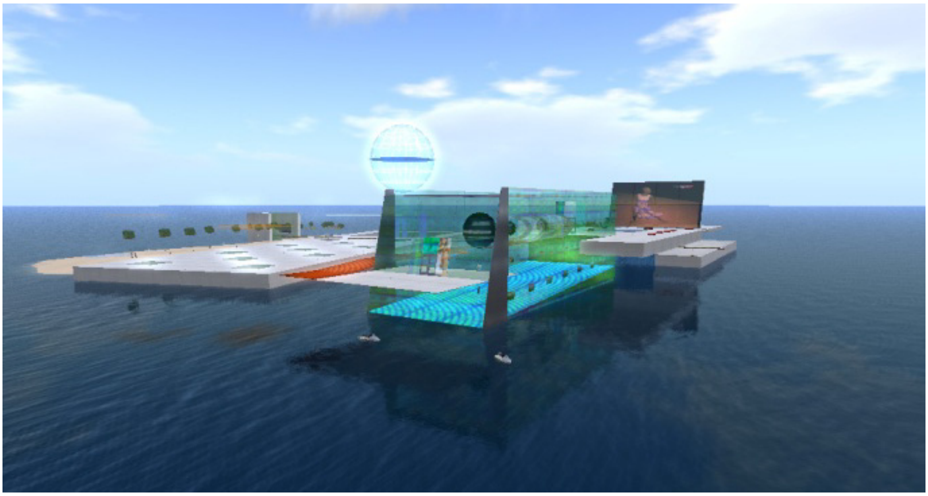


Fig. 7. Overview of Isla Videa

3.4 Usability Test and Results

The project was tested in several ways. Two complete courses were carried out successfully in the virtual world. Both students and teachers highlighted the easiness to understand the organization of the virtual facilities and remarked the intuitive, yet surprisingly new concepts present in the architectural design.

Fig. 8. Architectural 3D Award exhibition

Isla Videa was also used to run a temporal exhibition of works submitted to the Architectural 3D Award contest [6]. A set of interactive exhibitors were designed to showcase pictures and videos nominated for the prize (Fig. 8). Those displays were responsive to the presence of the avatar as well, enlarging their content and playing the videos automatically when needed. Finally, the 3D Award ceremony was broadcasted from the real world to the virtual facilities, so the participants were able to assist virtually, met and discuss despite of living in very distant places around the globe.

4 Conclusions

The exploration of new formal languages derived from the unique features of the virtual realm leads to the development of innovative architectural forms than can enhance the educational experience comparing with the use of building shapes that mimic those of the real world.

As technology advances there is every indication that in a few decades, computer simulation systems will be able to provide the user with sensations of virtual presence and immersion in an extremely vivid way that is not available today. There is little doubt the mission of Architecture and architects is to contribute their theoretical work, talent and knowledge to design the spatial experiences of future Internet users in virtual buildings.

References

1. Kzero Worldswide Consulting Co: Radar chart for Q1-2012 (First Quarter 2013), `http://www.kzero.co.uk/blog/slideshare-presentation-q1-2013-radar-chart/` (retrieved February 2013)
2. Merrick, K., Gu, N., Wang, X.: Case studies using multiuser virtual worlds as an innovative platform for collaborative design. Journal of Information Technology in Construction (ITcon) 16, 165–188 (2011)
3. Williams, D.: The Mapping Principle, and a Research Framework for Virtual Worlds. Communication Theory 20(4), 451–470 (2010), doi:10.1111/j.1468-2885.2010.01371.
4. Hernández Ibañez, L., Barneche Naya, V.: Cyberarchitecture: A Vitruvian approach. In: Proceedings of 2012 International Conference on Cyberworlds, Germany, pp. 283–289 (2012), doi:10.1109/CW.2012.48 ISBN: 978-0-7695-4814-2/12
5. Hernández Ibáñez, L., Taibo, J., Seoane, A., Jaspe, A.: Perception in Architectural Visualization through Immersive Virtual Reality. In: EGA Revista Expresión Gráfica Arquitectónica, vol. 18, pp. 252–261 (2011) ISSN: 1133-6137
6. Architectural 3D Award site at CGarchitect.com. (retrieved February 2013), `http://3dawards.cgarchitect.com`

Mixed Reality Space Travel for Physics Learning

Darin E. Hughes[1], Shabnam Sabbagh[2], Robb Lindgren[2], J. Michael Moshell[2], and Charles E.Hughes[1,2,3]

[1] Institute for Simulation and Training, University of Central Florida, Orlando, Florida, USA
dhughes@ist.ucf.edu

[2] School of Visual Arts and Design, University of Central Florida, Orlando, Florida, USA
shabnam.api@knights.ucf.edu,
{robb.lindgren,mmoshell}@ucf.edu

[3] Department of Electrical Engineering and Computer Science, University of Central Florida, Orlando, Florida, USA
ceh@cs.ucf.edu

Abstract. In this paper we describe research being conducted on a mixed reality simulation called MEteor that is designed for informal physics learning in science centers. MEteor is a 30 x 10 foot floor area where participants use their bodies to interact with projected astronomical imagery. Participants walk and run across the floor to simulate how objects move in space, and to enact basic physics principles. Key to the success of this learning environment is an interface scheme that supports the central metaphor of "child as asteroid." Using video data collected in our studies we examine the extent to which feedback mechanisms and interface conventions strengthened the metaphorical connection, and we describe ways the interaction design can be improved for future iterations.

Keywords: STEM, mixed reality, whole-body learning, informal education, physics simulation.

1 Introduction

Mixed Reality (MR) is used to describe technology environments that mix the real and the virtual. MR merges the physical with the virtual world, either by augmenting real environments with digital elements, or augmenting virtual environments physical elements [7]. Physically interacting with digital artifacts has benefits for learning [10], and several researchers have demonstrated the potential for using MR technologies to promote education across various domains [1, 2, 4, 8]. Some have even described design principles for MR environments that facilitate learning [6] such as allowing for intuitive mechanisms for controlling digital objects through direct manipulation, and supporting interactions on a "human scale." In this paper we describe an MR environment that was designed to enact these principles by engaging users in a metaphorical interface scheme. By analyzing video of participants using the system,

R. Shumaker (Ed.): VAMR/HCII 2013, Part II, LNCS 8022, pp. 162–169, 2013.

we describe the ways that the interface succeeded in supporting the central metaphor, as well as obstacles that must be addressed in future iterations of the technology.

MEteor is a MR game designed as an informal education simulation to be installed in science centers for middle school-aged children, providing a whole-body experience centered around learning Newton's laws of motion and Kepler's laws of planetary motion. Using multiple, calibrated projectors, the experience is displayed on the floor while minimizing shadows that often limit visibility and playability in floor-projected, interactive displays [4]. In this simulation, the learner is presented with a 30 foot by 10 foot display of a star field. When the learner walks onto the display, a tracking circle is projected around his or her feet – providing immediate feedback to the user that the simulation is interactive. In addition to the floor display, a separate wall display provides guidance in the form of short text instructions, a graphical display that shows the participant's performance, and additional gameplay information such as "level" and "score." In its current testing phase, a human guide is present to answer questions, provide assistance, and regulate the flow of participants.

In order to familiarize the participant with the gameplay mechanics, they first complete a series of training exercises. In these exercises the participant learns to launch their asteroid by jumping off a virtual space platform. The asteroid leaves the platform with the same velocity as the participant. Once launched, the participant must stay with the asteroid as it moves through space. The training exercises teach the participant that their tracking circle will change color (green to red) if their body gets too far from the launched asteroid's trajectory. They also learn that the graph on the wall display shows their path and how closely they adhered to the asteroid's path.

Once the practice levels are complete, the learner is taken through a series of physics games wherein they learn how to launch their asteroid past planetary objects with varying amounts of gravitational pull, how to knock a satellite planet out of orbit, and lastly how to place the asteroid into its own successful orbit. Learners are not only scored on their ability to successfully launch the asteroid on the correct trajectory, but also on their ability to follow the asteroid on its path, thus predicting the movement based on the various gravitational forces present at each level.

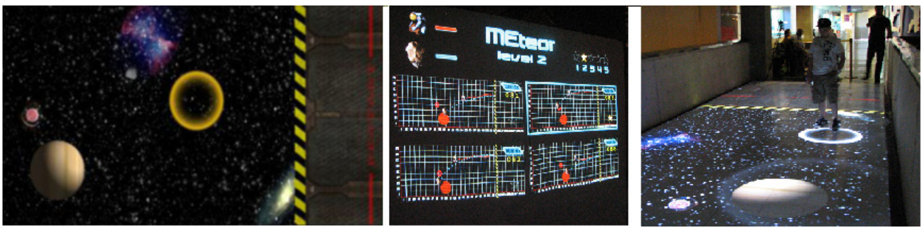


Fig. 1. The *MEteor* MR simulation game. Left: A screen shot of the floor projected imagery. Middle: Graphs and game information displayed on the wall. Right: A participant interacting using MEteor at the museum.

2 Objectives

A central challenge in teaching physics is the cognitive persistence of the "Aristotelian model." In everyday experience, moving objects tend to stop moving unless pushed, and everything falls toward the earth. Space provides an opportunity to observe the motion of objects without the influence of friction and air resistance, but learners are typically unable to experience and observe motion from this controlled perspective. The MEteor project is built around a set of experiments in which we use MR technology to provide a simulated space travel experience to young learners.

This space travel experience, however, does not involve passive ridership in simulated rockets. Learners instead embody the movement of a digital asteroid, making active predictions about how its trajectory will be affected by the presence of planets and other entities in the simulation. Our central hypothesis is that a realistic interactive simulation, experienced through whole-body motion, affords opportunities to directly confront the implicit models that we all learned from our earliest days, and replace them with new models and intuitions that can be built upon by formal instruction.

MEteor leverages the theories of embodied cognition wherein it is posited that learning is shaped by our physical interactions with the world. This understanding is aided by the use of functional metaphors.

Embodiment researchers and philosophers [3, 5] have argued that the body's form and activity provide functional metaphors that constitute the foundation of our language and our understanding. This research provides insight into the ways that computer systems and interfaces can support these body metaphors through immersive simulations and robust systems of feedback. In this simulation, the learner is encouraged to take the perspective of the asteroid and through this virtual embodiment, it is predicted that they will learn concepts in ways that they might not encounter in formal educational interventions.

3 Design and Methods

In the spring of 2012 we invited approximately 270 middle-school-aged students to use the MEteor simulation, either in our lab or at our installation at a local science center. Participants typically used the simulation for approximately 20 minutes, and while they were guided in their use by a member of the research team, students were given the flexibility to explore the interface and discover the means to achieve the goals of the game. All sessions were videorecorded, giving the research team the opportunity to observe and analyze features of the system that supported the target metaphorical connection, and those features that appeared to detract from it.

In conducting our analysis, it was useful to conceptualize the interface scheme in terms of a set of control inputs and system responses, with the overall goal being the optimal configuration of these such that learners achieved a better understanding of Newton and Kepler's laws.

3.1 Control Inputs

The control inputs in this simulation are all based on the learner's position (tracked by a laser scanner) within a space divided into the "launch deck" and "outer space." While on the launch deck the position of the digital asteroid is determined by the runner (projected right in front of her). When the asteroid crosses the edge of the deck the asteroid follows physical laws (e.g., the pull of a planet's gravity) and the runner must try to stay with the asteroid. Three degrees of freedom are determined by this launching run: (1) the lateral position where the asteroid enters space; (2) the angle of entry, and (3) the speed.

3.2 System Responses

How do the learner's actions generate feedback that maximizes learning about science, rather than about the control mechanisms of this particular simulation? Visual and auditory cues indicate the degree to which the participant stays with the asteroid in real time. Additionally, a score and a graphical replay projected on the wall provided feedback about how well they tracked the asteroid in a form of after-action review.

4 Results and Lessons Learned

Analysis of the video sessions indicated overall mixed success with maintaining the "child as asteroid" metaphor using the control inputs and system responses. In many cases, participants would not stay with the asteroid once they had determined that it was not going to hit its target. Their attention seemed focused on the task of hitting the target rather than that of predicting, following and understanding the asteroid's motion. However, participants who grasped the interface metaphor early on in their trials had greater success with the simulation. For example, one participant kept hitting the target but failing to get a good score. A member of the research team reminded the participant that, in order to get a good score, he needed to not only hit the target but also stay with the asteroid. In the next try the participant kept his eyes on the asteroid, but at one point fell behind and noticed he had lost contact. He immediately ran after it to catch up again, apparently realizing that he no longer had control over the asteroid after the launch. This "a-ha" moment allowed him to successfully control the asteroid and to more effectively traverse the subsequent levels.

In the following sections we describe a few of the interface interaction patterns that we observed in our review of the video sessions. The first two of these observations, decoupling and competing metaphors, are events that appeared to have detracted from the overall interaction and learning goals of the MEteor experience. The second two observations are instances of behaviour and insight that appeared to advance the learning and performance goals of the simulation.

4.1 Decoupling

As previously discussed, the participant is asked to perform two tasks: launch the asteroid on the correct trajectory and follow the asteroid's path. In reviewing the video recordings, it was clear that many of the participants would prioritize hitting the target over following the asteroid – despite the cost of a lower score. In many cases, hitting the target presented itself to the participant as the primary goal. This prioritization led the participants to decouple themselves from the asteroid, breaking the key interface metaphor.

The current design of the simulation perhaps reinforces this behaviour in a few ways. Successfully following the asteroid but not hitting the target does not lead to the completion of the level. Achieving a high score then becomes less essential where it might not be viewed as a success. Furthermore, hitting the target generated satisfying sound and visual effects indicating success. Another factor is based on the speed of the asteroid; in some instances the asteroid would move quite quickly and a participant would have to run to keep up with it. As one might expect, the adolescent participants in this study were not always willing to exert the necessary physical activity to meet these objectives.

Fig. 2. Decoupling: A participant not following the asteroid. The participant stops moving just after the launch line, allowing the asteroid to get away from him and the tracking circle to go red.

4.2 Competing Physical Metaphors

Whole body simulations by their very nature must be built around some pre-existing metaphors of interaction. A common "mistake" made by many of the participants is to try to kick the asteroid, or even drag it with their foot. In the real world, if an object is on the floor and the intent is to move that object, kicking it is often a valid approach – especially in the context of games (e.g. soccer, kickball, etc.). Likewise, some participants thought that the trajectory of the asteroid was based on which leg was forward when it was launched, while others tried jumping off the platform thinking that this was the required action to put the asteroid into space. One participant even attempted stomping the target when the asteroid failed to hit it.

In a traditional computer-mediated game run on a desktop, it is less likely that these kinds of mixed metaphors would occur. Not surprisingly, creating a simulation that occupies real physical space appears to invoke real-life metaphors. There is the potential for these real life metaphors to be leveraged in powerful ways with natural user interfaces. However, these metaphors can also lead to confusion and difficulty when real-life metaphors and virtual representations do not align.

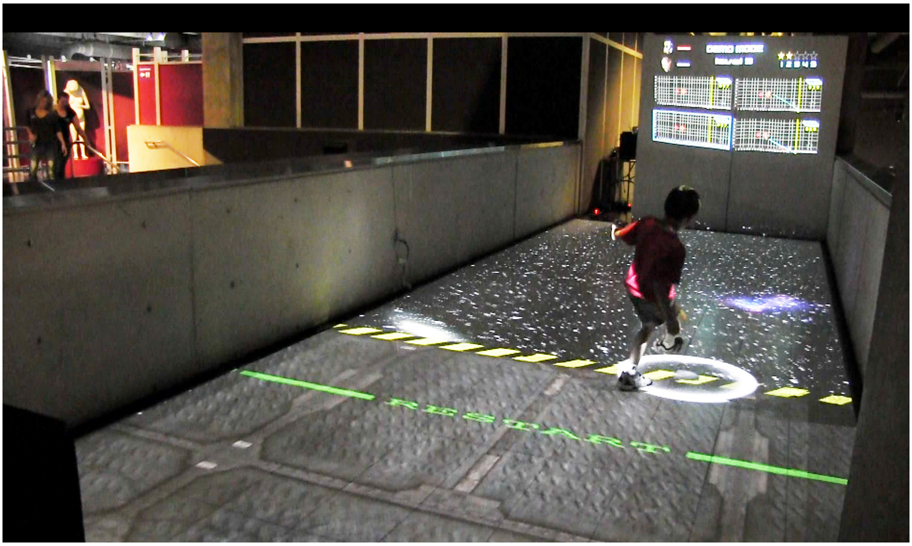


Fig. 3. Competing physical metaphors: A participant trying to kick the asteroid

4.3 Real-Time Visual Cueing

As soon as a participant enters the 30 foot by 10 foot projected game space, a green circle is projected around their feet – providing a visual cue to inform the participant that they are being tracked. To reinforce the coupling of the participant to the asteroid and to encourage the participant to follow and predict its trajectory, the circle gradually changes to red when they deviate from the asteroids path. This commonly-used metaphor proved to be quite effective in many instances.

In figure 4, we see a participant trying to catch up to the asteroid after seeing the circle change from green to red. This simple use of color-coding provided immediate and clear feedback to the participant. Not only did this convention keep participants connected to the asteroid, but it compelled them to perform an accurate enactment of how objects move in space, reinforcing important physics principles in physical way.

Fig. 4. Real-time visual cueing: A participant trying to catch up to asteroid

4.4 Post-Action Representational Supports

The primary function of the wall display is to provide a post-action review of the last-completed trial. A graph shows the path of the asteroid and the path of the participant, thus allowing the participant to make modifications or adjustments on their next trial. The display provides an abstraction of the participant's performance in order to invoke increased awareness of the process and principles while also allowing other participants to make comments and observations. In the science center environment this often led to further discussion among groups of participants as they watched a classmate play the game.

In figure 5 (left), a participant ponders his past trial while receiving input from an instructor. In figure 5 (right), the participant gestures to the place on the graph where he lost track of the asteroid and discusses with the researcher why he believes this event happened.

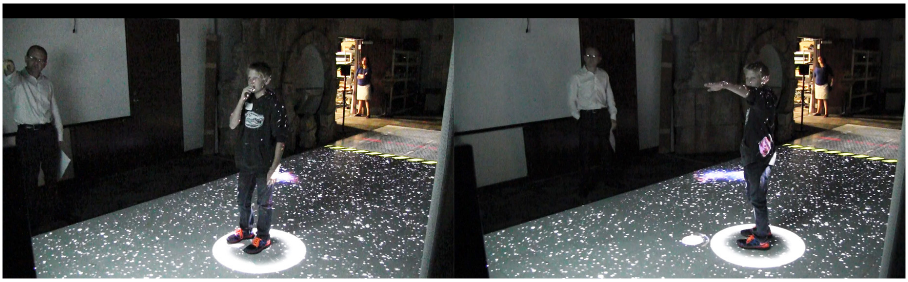


Fig. 5. Post-action representational support: A participant studying post-action graph

5 Conclusions

Whole-body, MR simulations that utilize real physical space in combination with virtual objects can provide unique educational experiences and perspectives not offered by traditional gaming technologies. Designing such a simulation requires an awareness of potential competing metaphors and other usability considerations – wherein real-life experiences with physical objects, individual physical capabilities, and task priorities must be addressed. Visual and audio cueing, consistent and tightly coupled metaphors, and post-action representational supports can be used to improve usability and comprehension.

Acknowledgements. The contents of this paper are based upon work supported by the National Science Foundation under Grant No. DRL-1114621. Any opinions, findings, and conclusions or recommendations expressed in this material are those of the authors and do not necessarily reflect the views of the National Science Foundation.

References

1. Birchfield, D., Megowan-Romanowicz, C.: Earth science learning in SMALLab: A design experiment for mixed-reality. Journal of Computer Supported Collaborative Learning 4(4), 403–421 (2009)
2. Chang, C.-W., Lee, J.-H., Wang, C.-Y., Chen, G.-D.: Improving the authentic learning experience by integrating robots into the mixed-reality environment. Computers & Education 55(4), 1572–1578 (2010)
3. Gallagher, S.: How the body shapes the mind. Oxford University Press, Oxford (2005)
4. Hughes, C.E., Stapleton, C.B., Hughes, D.E., Smith, E.: Mixed reality in education, entertainment and training: An interdisciplinary approach. IEEE Computer Graphics & Applications 26(6), 24–30 (2005)
5. Johnson, M.: The body in the mind: The bodily basis of meaning, imagination, and reason. University of Chicago Press, Chicago (1987)
6. Johnson-Glenberg, M., Koziupa, T., Birchfield, D., Li, K.: Games for learning in embodied mixed-reality environments: Principles and results. In: The Proceedings of the Games + Learning + Society Conference (GLS), pp. 129–137 (2011)
7. Milgram, P., Colquhoun, H.: A taxonomy of real and virtual world display integration. In: Tamura, Y.O.H. (ed.) Mixed Reality: Merging Real and Virtual Worlds, Springer, Verlag (1999)
8. Pan, Z., Cheok, A.D., Yang, H., Zhu, J., Shi, J.: Virtual reality and mixed reality for virtual learning environments. Computers and Graphics 30(1), 20–28 (2006)
9. Pillat, R., Nagendran, A., Lindgren, R.: Design requirements for using embodied learning and whole-body metaphors in a mixed reality simulation game. In: Proceedings of ISMAR 2012 (IEEE and ACM International Symposium on Mixed and Augmented Reality), Atlanta, GA (2012)
10. Price, S., Rogers, Y.: Let's get physical: The learning benefits of interacting in digitally augmented physical spaces. Computers & Education 43, 137–151 (2004)

Picking Up STEAM: Educational Implications for Teaching with an Augmented Reality Guitar Learning System

Joseph R. Keebler[1], Travis J. Wiltshire[2], Dustin C. Smith[1], and Stephen M. Fiore[2]

[1] Wichita State University, Wichita, Kansas
[2] University of Central Florida, Orlando, Florida
{joseph.keebler,dcsmith}@wichita.edu,
{twiltshi,sfiore}@ist.ucf.edu

Abstract. Incorporation of the arts into the current model of science, technology, engineering, and mathematics (STEAM) may have a profound impact on the future of education. In light of this, we examined a novel technology at the intersection of these disciplines. Specifically, an experiment was conducted using augmented reality to learn a musical instrument, namely the guitar. The Fretlight® guitar system uses LED lights embedded in the fretboard to give direct information to the guitarist as to where to place their fingers. This was compared to a standard scale diagram. Results indicate that the Fretlight® system led to initial significant gains in performance over a control condition using diagrams, but these effects disappeared over the course of 30 trials. Potential benefits of the augmented reality technology are discussed, and future work is outlined to better understand how embodied cognition and augmented reality can increase learning outcomes for playing musical instruments.

Keywords: STEAM, augmented reality, embodied learning, music education, Fretlight® guitar.

1 Introduction

The STEAM education movement is currently building momentum towards incorporating the arts into the science, technology, engineering, and mathematics (STEM) thrusts [e.g., 1]. To justify the benefits of the Arts, recent efforts by the Dana Foundation [2-3] have organized multi-university consortiums of leading cognitive neuroscientists in order to examine the relationship between learning, arts, and the brain. The consortiums have established that the arts play a critical role in the development of generalizable cognitive skills that can serve as a benefit to learning. One form of art related to a number of neurological and educational benefits is playing a musical instrument. More specifically, music training may lead to increases in brain plasticity across the lifespan [4], processing of both musical and non-musical auditory stimuli [5], and demonstrable improvements in reading comprehension and math skills [6].

R. Shumaker (Ed.): VAMR/HCII 2013, Part II, LNCS 8022, pp. 170–178, 2013.

Previous research has highlighted a need to incorporate quality music training more heavily into education at all levels, with early stages being particularly more beneficial [2-3], [5]. However, there are limitations to the distributed nature in which traditional music learning often occurs [7]. Recent advances in computing technology have sought to mitigate the limitations of distributed learning materials through the use of augmented reality, particularly for learning to play the guitar [8-9]. These systems were experimental and faced several shortcomings. The present research seeks to extend this line of work by examining an augmented reality guitar learning system that may foster an embodied learning environment. Embodied learning environments facilitate learning by allowing individuals to be engaged both physically and mentally leading to a stronger interconnection of auditory and visual stimuli with motor responses [10] and thus the integration of augmented reality musical learning systems should serve as a benefit to music education as well as STEM education.

1.1 STEAM and the Benefits of Arts Education

The STEAM education movement seeks to incorporate arts as an essential element to the STEM education movement [1]. More specifically, proponents of this movement claim that an integrated STEM and arts curriculum is essential to foster creativity and innovation in the STEM disciplines [11]. Over the years, a number of research efforts have tried to establish the link between learning, the arts, and the brain, though maybe none as vast as the multi-university consortiums of cognitive neuroscientists organized by the Dana Foundation [2-3]. The overarching consensus from these large research efforts was that training in the arts is correlated with generalizable educational and neurological benefits and thus the arts should be advocated for in accordance with other STEM disciplines and integrated into current educational systems at multiple levels.

Though the arts incorporate a large number of disciplines, the focus of this paper is specifically on the benefits of music education, and more specifically, guitar training. A benefit to STEM education found by Spelke [in 2] was that intensive and prolonged music training was associated with better performance in core brain systems associated with geometrical tasks. This claim is also supported by Vaughn [12] who used meta-analytic techniques to examine this association between music training in both geometrical and spatial reasoning. More generally, Jonides [in 2] found that extensive music training led to the development of enhanced memory by an increased ability to focus attention on rehearsal tasks. Likewise, Neville et al. [in 2] found that children given music training showed an increase in executive control function analogous to those given specific attention training. These findings suggest music training leads to positive benefits for learning as a function of increased attention and spatial reasoning abilities.

Additionally, Kraus, and Chandrasekaran [5] found that music training has an overarching benefit to auditory processing that includes improvements in auditory-verbal-memory and auditory attention for both music and non-music auditory signals. Further, the researchers have found that music training leads to increases in brain plasticity that are correlated with performance on executive function tasks [5], [13].

These improvements are primarily a function of practice and age during the onset of training; however, even practice at later stages in life can lead to increases in neuroplasticity [5]. In regards to the benefits of music training during early phases of learning, Skoe and Kraus [14] found that short durations of formal music training during childhood showed significant differences in auditory brain stem responses approximately seven years after formal training had commenced. Drawing from the results of this study and their prior work, Skoe and Kraus [14] conclude that there are long lasting positive benefits to music training. As such, the evidence is mounting in support of the generalizable benefits of music training, especially during early stages of education. Further, research needs to establish the educational benefits associated with certain types of music training strategies and technologies.

1.2 Augmented Reality Guitar Systems

Differential benefits likely result from the specific instructional strategies employed in the varying types of musical training that individuals receive. Rather than detail the number of differences resulting from music training types, one solution for mitigating a pervasive issue across all types of music training is proposed in the context of electric guitar training. Specifically, one barrier to music learning is the distributed nature of the materials required for learning an instrument [7]. This poses a problem during learning of new music material as it is often presented in an abstract form such as notation, guitar tablature, or diagrams that have to be memorized and translated in some form to the instrument (see Figure 1 for example diagram). A related finding to this notion stems from piano learning research where Lim and Lippman [15] found that performance was greater for participants who were able to physically practice a piece rather than attempt to memorize it. Though the piano research utilized distributed materials, it suggests that performance may be improved though the stronger coupling of the learning materials with performance on that actual instrument.

One approach to mitigating the problem with distributed learning materials may be overcome by integrating augmented reality technology with guitar training. Augmented reality involves the mixing of the physical world with digital information [16]. Liarokapis [8] and Motokawa and Saito [9] have developed experimental augmented reality guitar learning systems that represent a step towards integrating the digital representation of the material to be learned with the physical instrument; however, these systems are experimental and not easily accessible to those less technically inclined. A recent commercially available product, known as the Fretlight® guitar (Optek Systems, Inc.), includes an LED circuit board underneath the fret board. When coupled with the included computer software, the guitar is thought to help individuals at any level of musical experience to advance their performance. The company claims their guitar is useful because it displays the information guitar players need (i.e., where to place their fingers) right where they need it (see Figure 1). The benefits of this device seem clear though they have not been empirically evaluated. Utilizing the Frelight® guitar will allow for an examination of sensorimotor learning processes within the framework of embodied and enactive cognition.

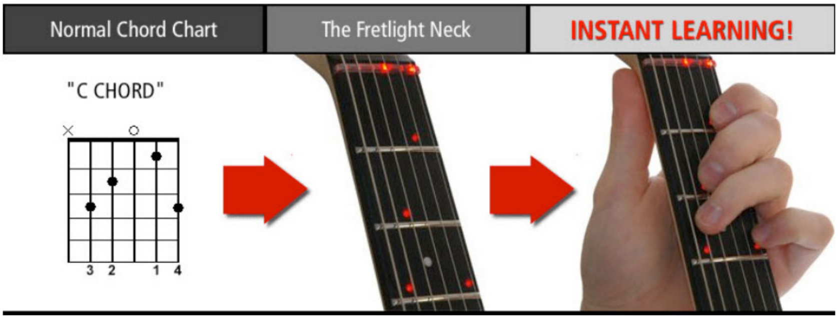


Fig. 1. The leftmost part of this image represented a standard chord diagram utilized in learning to play the guitar. The center image represents how the Fretlight® guitar would display the same information and the rightmost image represents how an individual is able to see and place their fingers in the proper location.

1.3 Embodied and Enactive Learning

Kagan [in 3] advocates for the importance of arts in education because the arts are able to multi-modally engage the mind; thus leading to the integration of motor, perceptual, and cognitive skills. This emphasis on the mind is somewhat misleading as the body plays a major role in music learning through the sensorimotor system. Specifically, sensorimotor integration is the result of both learning and practicing a musical instrument. That is, these activities lead to the association of motor actions with certain auditory signals and visual information resulting in strengthened connections between the auditory and motor regions of the brain [4]. Further, Kraus and Chandrasekaran [5] found that children under seven that were given musical training show superior sensorimotor integration than those whom received training later in their lifetimes. This notion suggests the temporal occurrence of music training makes a difference such that sensorimotor integration is stronger when music training occurs earlier in life.

Two theories that are particularly relevant for understanding the sensorimotor learning processes are embodied [e.g., 18] and enactive cognition [e.g., 17]. Recent advances in the cognitive sciences have supported the embodied and enactive theories using the exemplar of music cognition [19-21]. That is, music and more specifically, the process of learning to play a musical instrument, is not an amodal cognitive process that occurs independently of the environment. Rather, the embodied and enactive theories claim that there is an inextricable multi-modal link between the environment (e.g., the musical instrument) and the brain created through the body's sensorimotor system [20-21]. Given this notion, knowledge of music and how to play an instrument can be thought of as arising from the co-occurring perceptions and actions apparent in the learning process.

Along these lines, embodiment has been identified as a key theory for the design of augmented and mixed-reality learning environments as it supports learning by allowing individuals to be engaged both physically and mentally and thus leading to sensorimotor integration [10]. So far, researchers have not attempted to integrate the embodied approach to musical learning with mixed or augmented reality learning systems. This research represents such an attempt by utilizing the Fretlight® guitar in which digital information displayed on the guitar, rather than through distributed materials, visually indicates the location to place the fingers. As an individual places their fingers and plays the notes indicated by the Fretlights®, motor actions couple with the perceptions of the auditory signals and visual information and thus theories of embodied and enactive cognition likely underpin learning with the Fretlight® guitar.

1.4 Research Hypotheses

The goal of this research is to compare the learning processes that occur during training of the A minor pentatonic scale with the Fretlight's LED system compared with a scale diagram (a more traditional guitar learning aid). More specifically, accuracy and precision were compared between the two conditions during both the learning of the scale as well as during a test performance in which there was no learning aid. The two conditions were compared in order to evaluate learning gains. Our hypotheses were as follows:

- H1: Performance during training, in terms of accuracy and precision, for the Fretlight® guitar condition will be better than in the diagram condition.
- H2: Test performance, in terms of accuracy and precision will be better in the Fretlight® guitar condition.

2 Method

2.1 Participants

55 undergraduate students from the University of Central Florida voluntarily participated in this study. There were 30 female and 25 male participants. Participants voluntarily participated in exchange for class credit. Eligibility for this study was restricted to right-handed individuals who had no formal or informal stringed instrument training.

2.2 Experimental Design

A between-subjects design was utilized for this experiment where participants were randomly assigned to one of two levels of the independent variable (IV). The IV for this study was the learning aid that participants utilized for learning the A minor pentatonic scale. The two levels of this IV are as follows: (1) Fretlight® LED learning aid in which LED lights illuminate the scale so that participants can see the proper locations of notes embedded within the fretboard and (2) scale diagram learning aid, which is a traditional distributed method for learning the guitar consisting of a paper diagram with the notes of the scale shown on a depiction of the fretboard.

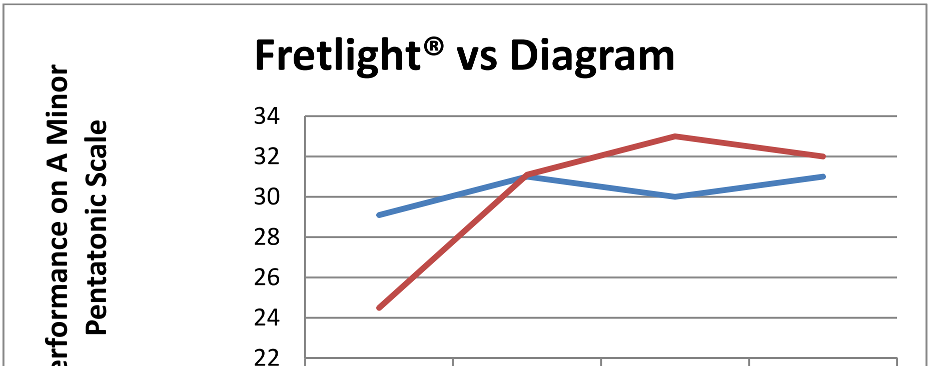
The dependent variables (DVs) of the study were participants' training and test performance ratings. In order to determine the training and test performance, a rating scale was developed to assess the accuracy and precision of each of the 12 notes of the A minor scale for each trial of the training and test performances. Specifically, the constructs of accuracy and precision were incorporated into a rating scale in which each note played by participants was assigned a value of 0 – 4 depending on whether it was a correct note and how precise that correct note was played.

2.3 Procedure

When the experiment first began, participants were given an informed consent document and randomly assigned to one of the levels of the IV. Then, all participants were introduced to several basics of the guitar utilizing a PowerPoint presentation with embedded videos. The presentation familiarized participants with how to hold the guitar, how to fret notes with the left hand, and how to use the right hand to pick the strings. The training information included in the presentation was based on guidelines included with the Fretlight® software. During the presentation, participants were given opportunities to practice the basics of the guitar. Participants were then trained to play the A minor pentatonic scale over approximately 30 training trials with one of the two learning aids (Fretlights® or diagram). All participants used the same guitar. For the training, participants were asked to try their best to play through the whole scale and to continue on to the next note if they skip or mess up a note in the scale. To determine the extent to which participants learned to play the scale, they were required to complete a test performance without the learning aid (i.e., Fretlights® or diagram were removed) for 10 test trials.

3 Results

The performance data were analyzed using a 2 x 4 mixed ANOVA model to examine performance during the first and last training trials as well as the first and last test trials. Results indicated a significant interaction effect, $F(2,53) = 6.6$, $p < .05$, partial eta^2 = .112. The Fretlight® condition led to significantly higher performance (M = 29.1, SD = 7.72) than the diagram condition (M = 24.5, SD = 11.85), but only during training trial 1. Following, there was a significant increase in the diagram condition performance from training trial 1 (M = 24.5, SD = 11.85) to training trial 2 (M = 31.1, SD = 8.1). There were no significant differences between or within conditions during either of the testing portions of the experiment (see Figure 2).



	Training Trail 1	Training Trail 2	Test Trial 1	Test Trial 10
Fretlight® Means	29.1	31	30	31
Diagram Means	24.5	31.1	33	32

Fig. 2. Comparison of performance on A minor pentatonic scale between the Fretlight® and diagram conditions

4 Conclusions

4.1 Discussion

Based on these data, it seems there is some support for H1 during the initial training trial difference. That is, participants in the Fretlight® training condition performed better than those in the diagram training condition. However, because this is followed by a convergence in performance at the end of the training, we cannot conclude full support for H1. Rather, we can develop a new hypothesis: Performance during initial training sessions with the Fretlights® will be greater than those in the diagram condition though, performance will converge as the number of training sessions increase. Further, H1 was not supported as no significant differences between the test performance scores were found between the two conditions.

4.2 Implications

These data demonstrate that 1) embodied learning through the Fretlight® guitar may create a lower "barrier to entry" for learning to play the guitar and 2) and that we must attend to training design implications. Specifically, Fretlight® training programs needs to make it explicit whether the aim is for the device to be utilized in the real-world (i.e., during a performance) or just as a training aid to make learning easier (this is likely the case). For example Tang, Owen, Biocca, and Mou [22] assessed the effectiveness of an augmented reality system for an object assembly task with other forms of media; however, they did not compare how participant performance was influenced when the training medium was removed, as was done in the present study. Further, if the aim is to eventually play the guitar without the aid of the Fretlights® then the findings from this study suggest there is a need for some form of scaffolding incorporated into training in which the individual transitions from use of the technology to non-use.

More broadly, future research needs to examine the specific tasks and components of learning a given musical instrument utilizing measures to better comprehend which aspects of learning to play a musical instrument are associated with specific cognitive benefits especially more longitudinally and at different ages. Some constructs that may be beneficial to examine include tests of spatial-temporal reasoning, reading comprehension, auditory discrimination skills, and neuroplasticity [4-5], [23].

In sum, a large body of research has amassed pointing to long term benefits of music training and support for incorporating the arts, specifically music training, into STEAM education. This research represents a unique exploration of learning through a novel augmented reality guitar learning system and as novel technologies are developed for advancing the science of learning a musical instrument, researchers must examine the extent to which benefits inherent to the traditional music training process still remain.

Disclaimer. Partial support for this research was provided by Optek Systems, Inc. in the form of an academic discount on the experimental materials (i.e., Fretlight® guitarts); however, the views and findings expressed in this manuscript are solely those of the authors and do not necessarily reflect those of Frelight ® or Optek Systems, Inc.

References

1. Pomeroy, S.R.: From STEM to STEAM: Science and Art Go Hand-in-hand. Scientific American (2012), `http://blogs.scientificamerican.com/guest-blog/2012/08/22/from-stem-to-steam-science-and-the-arts-go-hand-in-hand` (retrieved)
2. Asbury, C., Rich, B. (eds.): Learning, Arts and the Brain: The Dana Consortium Report on Arts and Cognition. Dana Press, New York (2008)
3. Hardiman, M., Magsamen, S., McKhann, G., Eilber, J.: Neuroeducation: Learning, Arts and the Brain: Findings and Challenges for Educators and Researchers from the 2009 Johns Hopkins University Summit. The Dana Foundation, New York (2009)
4. Wan, C.Y., Schlaug, G.: Music Making as a Tool for Promoting Brain Plasticity Across the Life Span. The Neuroscientist 16(5), 566–577 (2010)
5. Kraus, N., Chandrasekaran, B.: Music Training for the Development of Auditory Skills. Nature Reviews Neuroscience 11(8), 599–605 (2010)
6. Wilcox, C.: Even a Few Years of Music Training Benefits the Brain. Scientific American (2012), `http://blogs.scientificamerican.com/science-sushi/2012/08/21/even-a-few-years-of-music-training-benefits-the-brain/` (retrieved)

7. Flor, N.V., Holder, B.: Hearing with the Eyes: A Distributed Cognition Perspective on Guitar Song Imitation. In: Proceedings of the 18th Annual Conference of the Cognitive Science Society. LEA, New Jersey (1996)
8. Liarokapis, F.: Augmented Reality Scenarios for Guitar Learning. In: Proceedings of Eurographics UK Theory and Practice of Computer Graphics, pp. 163–170 (2005)
9. Motokawa, Y., Saito, H.: Support System for Guitar Playing using Augmented Reality Display. In: Proceedings of the 5th IEEE and ACM International Symposium on Mixed and Augmented Reality (ISMAR 2006), Santa Barbara, Calif, USA, pp. 243–244 (October 2006)
10. Birchfield, D., Thornburg, H., Megowan-Romanowicz, M.C., Hatton, S., Mechtley, B., Dolgov, I., Burleson, W.: Embodiment, Multimodality, and Composition: Convergent Themes across HCI and Education for Mixed-Reality Learning environments. Advances in Human Computer Interaction, pp. 1–20 (2008)
11. Crayton, J.E.: STEM Art Education Movement: A Creative Approach to Education in Innovation for the 21st century. University of New Mexico, Albuquerque,(2007), http://www.campusclubmilledgeville.org/images/stem_art_education_movement_crayton.pdf (retrieved)
12. Vaughn, K.: Music and Mathematics: Modest Support for the Oft-Claimed Relationship. Journal of Aesthetic Education 34(3/4), 149–166 (2000)
13. Moreno, S., Bialystok, E., Barac, R., Schellenberg, E.G., Cepeda, N.J., Chau, T.: Short-Term Music Training Enhances Verbal Intelligence and Executive Function. Psychological Science 22, 1425–1433 (2011)
14. Skoe, E., Kraus, N.: A Little Goes a Long Way: How the Adult Brain is Shaped by Musical Training in Childhood. The Journal of Neuroscience 32(34), 11507–11510 (2012)
15. Lim, S., Lippman, L.G.: Mental Practice and Memorization of Piano Music. The Journal of General Psychology 118(1), 21–30 (1991)
16. Sherman, W., Craig, A.: Understanding Virtual Reality: Interface Application and Design. Morgan Kaufmann (2003)
17. Klin, A., Jones, W., Schultz, R., Volkmar, F.: The Enactive Mind – From Actions to Cognition: Lessons from Autism. Philosophical Transactions of the Royal Society: Biological Sciences 358, 345–360 (2003)
18. Clark, A.: An Embodied Cognitive Science? Trends in Cognitive Science 3, 345–351 (1999)
19. Alerby, E., Ferm, C.: Learning music: Embodied Experience in the Life-World. Philosophy of Music Education Review 13(2), 177–186 (2005)
20. Krueger, J.W.: Enacting Musical Experience. Journal of Consciousness Studies 16, 2(2), 98–123 (2009)
21. Matyja, J.: Embodied Music Cognition. Philosophy. The University of Edinburgh (2010), http://hdl.handle.net/1842/5325 (retrieved)
22. Tang, A., Owen, C., Biocca, F., Mou, W.: Comparative Effectiveness of Augmented Reality in Object Assembly. In: Proceedings of the SIGCHI Conference on Human Factors in Computing Systems, pp. 73–80 (2003)
23. Raushcer, F.H., Shaw, G., Levine, L.J., Wright, E.L., Dennis, W.R., Newcomb, R.L.: Music Training Causes Long-Term Enhancement of Preschool Children's Spatial-Temporal Reasoning. Neurological Research 19, 2–8 (1997)

Virtual Reality Data Visualization for Team-Based STEAM Education: Tools, Methods, and Lessons Learned

Daniel F. Keefe[1] and David H. Laidlaw[2]

[1] University of Minnesota, Minneapolis MN 55455, USA
keefe@cs.umn.edu
http://www.cs.umn.edu/~keefe

[2] Brown University, Providence RI 02912, USA
dhl@cs.brown.edu
http://www.cs.brown.edu/~dhl

Abstract. We present a discussion of tools, methods, and lessons learned from nearly ten years of work using virtual reality data visualization as a driving problem area for collaborative practice-based STEAM education. This work has spanned multiple universities and design colleges. It has resulted in courses taught to both students majoring in computer science and students majoring in art or design. Within the classroom, an important aspect of our approach is including art and design students directly in real scientific research, often extended beyond the computer science aspects of data visualization to also include the research of collaborators in biology, medicine, and engineering who provide cutting-edge data visualization challenges. The interdisciplinary team-based education efforts have also extended beyond the classroom as art and design students have participated in our labs as research assistants and made major contributions to published scientific research. In some cases, these experiences have impacted career paths for students.

Keywords: STEAM, art, science, computer science, education, virtual reality, visualization.

1 Introduction

Relevance to STEAM Learning. As we look toward the future of education in the 21st century, the importance of a strong STEM (Science, Technology, Engineering, and Mathematics) curriculum is unquestioned. Building upon this foundation, we join a growing chorus of educators and researchers in advocating for strong STEM + Arts or “STEAM”-based education. There are many possible benefits to combining STEM with Arts, and, in our experience, there is a great deal of interest in this combination from both the arts and scientific communities. However, the details of how to most successfully accomplish STEAM integration in the classroom as well as the details of how best to utilize the exciting new technologies that often characterize STEM fields remain underexplored

R. Shumaker (Ed.): VAMR/HCII 2013, Part II, LNCS 8022, pp. 179–187, 2013.

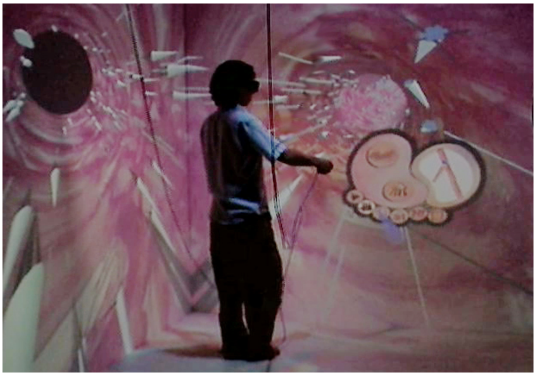


Fig. 1. A CAVE-based VR visualization of simulated blood flow through a branching coronary artery. The visualization and user interface were designed by a small group of art and computer science students as the final project for their course.

and underreported. In this paper, we present a discussion of tools, methods, and lessons learned from nearly ten years of work combining STEM with Arts for education, and in many cases extending beyond the classroom and into research.

We have used virtual reality (VR) technologies as a platform for much of this work. Since VR utilizes advanced technology and computing to create an environment that engages the senses (visual, auditory, haptic) in unique and exciting ways, we have found that it is a particularly engaging platform for students, regardless of whether their background is primarily in a STEM field or in the Arts. Thus, we believe VR may be uniquely positioned as a very powerful tool to facilitate STEAM education.

Bringing VR into the classroom in a way that works for students of all backgrounds also poses considerable challenges, both technical and methodological. We have refined a series of approaches and tools to use VR for team-based STEAM education and are keen to elaborate on a number of the ideas in this paper. Our approach is based in our core area of research, scientific visualization, which can be summarized as using computer graphics and human computer interfaces to understand scientific data, either in an effort to facilitate new discoveries based on the data or to explain data-driven insights to others. We have also found it useful to engage with a number of methods used in traditional art education, such as critique. The following sections provide some additional background on these core concepts.

Visualization. Figure 1 provides a motivating example of scientific visualization in virtual reality. The problem addressed by the visualization is to understand complex patterns in time-varying, three-dimensional fluid flow vector fields; in this case, the data come from a high-end computational simulation of blood flowing through a branching coronary artery [1]. In order for scientists to understand this complex phenomenon they must make sense of patterns across several 3D data fields at once: velocity, vorticity, pressure, shear stress, and so on. Since each of these data fields also changes over time, the amount and spatial complexity of the data make them extremely difficult to interpret, and nearly impossible to interpret without some visual aid. We use VR to create data visualizations for this problem because the increased spatial understanding that is achieved with head-tracked immersive displays can be extremely valuable for analyzing spatial relationships. But, designing effective VR visualizations is extremely challenging, and this is where a STEAM-based approach can be so valuable – not just for the students, but also for end product. For example, Figure 1 is the result of a small group (1-2 computer science students and 1-2 art students) final project for the course *Virtual Reality Design for Science*, taught and cross-registered at Brown University and the Rhode Island School of Design. For a computer scientist looking at this environment, the attention paid to the visual aesthetic of the visualization and user interface is immediately obvious and radically different than the norm. Even the menu (the organic shape to the right of the user's hand) was designed carefully and implemented using hand-painted imagery. The flow data are conveyed using 3D multi-variate glyphs that reshape and reposition themselves interactively in response to the flow data. Again, demonstrating a departure from the norm in computer science, these glyphs were designed with specific reference to forms in nature that respond to flow, e.g., the tentacles of a squid as it propels itself through water. This attention to shape, color, texture, form, narrative, and metaphor is one of the most exciting trends that we see resulting from STEAM efforts and has clear benefit to all who are involved in such collaborations.

Critique. We have found that using virtual environments for STEAM education also requires teaching methodologies that are mostly unfamiliar in computer science but are part of a strong educational tradition in art and design. One example is the use of critique, which we have now adopted in both the classroom and research labs. In the classroom, we have utilized critique in a traditional form. For example, class sessions often begin with students hanging drawings, printouts, or other artifacts from their work since the last meeting on the wall, and we follow with a traditional critique-style discussion of the work. During critique we discuss and teach about the same type of concepts one would encounter in an art class (e.g., shape, color, texture, form, narrative, metaphor) but these are discussed and evaluated with reference to the goal of moving from data to insight. Often, these group critique sessions are extended beyond the traditional classroom space and instead conducted directly in virtual reality environments, such as the CAVE facility pictured in Figure 1. As the problems we discuss often involve interpreting biological or other datasets, scientists from relevant disciplines have often joined these critique sessions as well. We have found this to be invaluable. In fact, the methodology has been so successful in our experience that we have begun to adapt critique to computer science classes, from first-year algorithms to graduate research seminars.

Collaboration. STEAM-based learning is interdisciplinary by nature. We believe one of the great benefits of the approach is that each student who think of

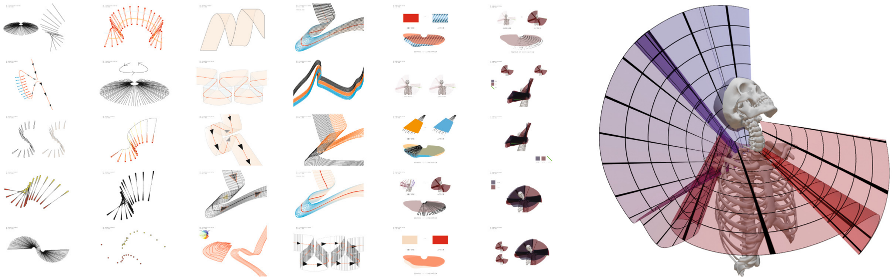


Fig. 2. Left: A selection of representative images from more than 300 sketches created by a graphic design student to explore the use of color, texture, form, narrative, and metaphor in depicting the biomechanics of the spine. Right: One frame from the resulting animated visualization of the data.

him/herself as belonging to a core STEM field becomes a bit more of an artist. Likewise, each artist becomes a bit more of a scientist, technologist, engineer, or mathematician. This isn't to say that these students will necessarily switch disciplines or become an expert in a new discipline, but to facilitate collaboration, students in STEM and Art fields need to learn to speak a bit of each others' language, only then than successful collaborations happen. In our courses and research this type of collaboration takes center stage. We teach students how to work across STEM and Arts disciplines as a starting point, but since the data understanding problems that are our focus involve real, active scientific research, both groups of students also find themselves in the same position of needing to understand how to speak the language of a third group of collaborators, typically from biology, medicine, or some related discipline. This is an excellent skill to develop, and, in the classroom, it helps us to focus on a problem and puts all the students, regardless of discipline, in a similar situation of needing to consider this third party as the user or client who will evaluate the ideas they develop.

2 Methods and Tools

One thing that has enabled our efforts to be successful is the new software tools we developed specifically to support collaboration across disciplines. Indeed, developing these tools has itself become an important area of research. Here, we describe a series of tools and approaches that we have used in the classroom and our research based around the concepts of sketching and prototyping.

2.1 Graphic Design and Traditional Prototyping

Sketching is a staple in arts education and is one of the most powerful design practices that has ever been invented. STEM education can learn from this. In computer science, for example, we find that students tend to be so excited to start programming that they jump right into the mode of implementing a new

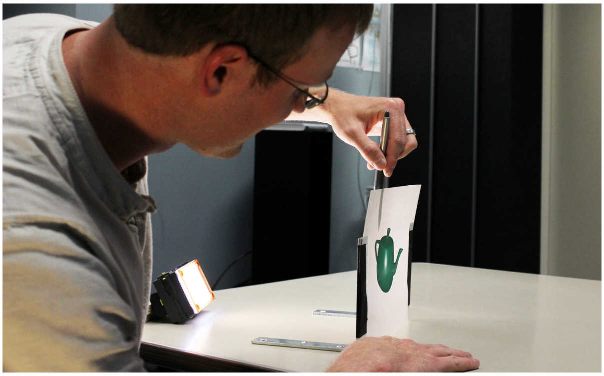


Fig. 3. A graduate student presents an interactive "sketch" of a new user interface idea

idea before it is fully thought out. Unlike sketching, programming is very time consuming; thus, there is a real danger of leaving important design refinements undiscovered when we jump right into programming. This is one concept that we try to teach. Borrowing from the arts tradition, we have found that sketching is an extremely effective method for demonstrating this to students.

To illustrate this, Figure 2 shows an example from our recent research on visualizing human spinal kinematics [2]. The selection of images on the left were created using Adobe Illustrator by a MFA student studying graphic design. At one point, this student generated more than 100 concept sketches in this style in a single week. Her goal was to explore the visual possibilities for how to depict the change over time in the instantaneous axis of rotation for the skull relative to the torso during a head rotation exercise. She refined the use of texture, form, color, and more through these sketches in order to visually emphasize the folds in the 3D surface swept out by these axes over time – our collaborators believe these folds correlate with neck pain, but they have never been analyzed before because it is so challenging to do a 3D analysis of these data. At the end, the goal is not a sketch but instead a data-driven computer graphics visualization. For this example, one frame from the 3D animated visualization result is shown in Figure 2 right. We find that even traditional paper and pencil sketches can be useful for this design and prototyping work; however, we next describe several extensions to the idea of sketching that can be even more useful and are particularly engaging for students.

2.2 Sketching User Experiences

Figure 3 demonstrates how our approach extends beyond traditional sketching to include "sketches" of user experiences. This is critical when the target medium in VR, since the experience of being in an immersive environment cannot really be captured on paper. Our approach draws heavily upon Buxton's notion of sketching user experiences [3]. For example, in Figure 3, a student acts out a

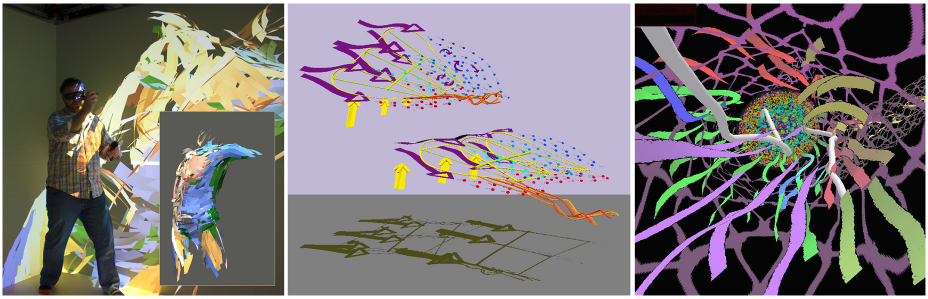


Fig. 4. New user interfaces make it possible to create virtual reality environments simply by sketching in 3D space using a tracked 3D input device

"sketch" of a user experience he has designed as part of a class assignment. He uses a flashlight, a pen, and a picture printed on a piece of paper to demonstrate how it would feel for a user to work with his idea for a new collapsable pen device that provides haptic feedback as it is inserted past the depth of the screen into a virtual world. The shadow of the pen substitutes in this sketch for a virtual pen that he proposes to program using 3D computer graphics. This physical prototyping method enables the class to explore and critique the user experience much more quickly than would be possible if virtual reality programming were required. Students in our courses have used short movies, modeling clay, paperclips, and even helium-filled balloons to sketch user experiences.

2.3 Sketching in Virtual Reality

Figure 4 shows work with the CavePainting virtual reality system [4], which has been used to support artists and designers in both coursework and research [5,6]. This is an engaging way for artists to create in VR without programming (see 3D gesture sketch of a human torso on the left). It can also be used as a 3D prototyping tool. The 3D sketches in the center and right were created by hand in virtual reality by students and depict their ideas for how multivariate fluid flow data (velocity, vorticity, pressure, etc.) could be represented in a virtual reality data visualization. This tool enables artists and designers with no knowledge of programming to design and create directly in virtual reality. In courses, this is critical because it gives artists an entry point to develop custom virtual reality applications, an opportunity that has previously only been available to advanced computer science students.

3 Next Generation Tools and Current Research

The interactive system called "Drawing with the Flow" [7] pictured in Figure 5 begins to address one of the limitations of our prior methodologies. Although we found sketching to be an exceptional entry point for bringing STEM and Arts

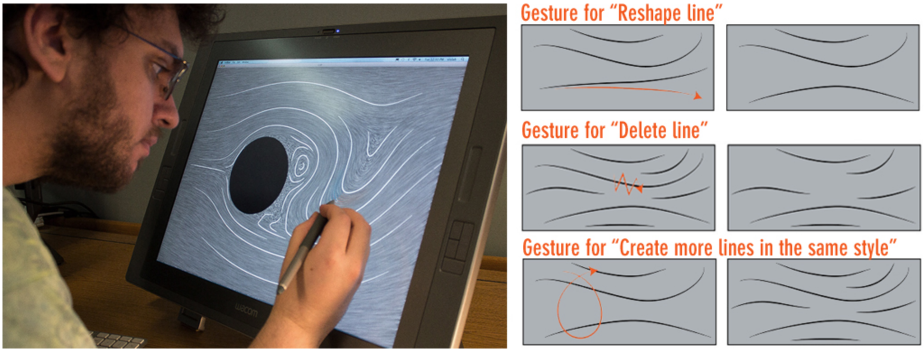


Fig. 5. Extending the productive uses of sketching described earlier, new tools developed through our research make it possible to sketch "on top of" data

together, the sketches discussed up to this point are not accurate with respect to the underlying scientific data. Thus, although they do provide excellent material for discussion and critique, the critique can only go so far until the ideas represented in the sketches have been implemented in an data visualization system. In contrast, the system pictured in Figure 5 is designed such that the user sketches on top of a data canvas. The data in this example describes a 2D flow field. As the user sketches a curve to depict flow lines in this field, the system interprets that sketch and rectifies the line that was drawn so that it remains as close as possible to the original while also staying true to the data. To the user, the result is that after sketching a flow line, he sees the line slowly morph into a line that is valid with respect to the underlying data. In this way, a number a visualization styles can be quickly explored (e.g., visualizations that use flow lines of different lengths, varying density, and so on) but at any point in time, the drawing on the user's screen is accurate with respect to the underlying flow data. The tool is in essence a step toward something like Adobe Illustrator, but for scientific data. We are now extending this concept to work with a variety of other data types and display environments. We believe the resulting tools will open up additional avenues and support for STEAM education.

4 Discussion

We have learned many lessons about how to include virtual reality technologies in new courses and research. Our focus has been on data visualization problems, and we believe this is a rich and productive focus area for STEAM education. Certainly, artists and designers can bring exciting new ideas to these problems, and solutions are often extremely exciting from the standpoint of the STEM concepts involved. The sketch-based tools and approach that we describe above has been a key to the success of our efforts, with one of the main reasons being simply that it makes doing real work (e.g., creating new visual environments) with VR technologies accessible to students trained in the arts. These methods also seem to teach students from STEM backgrounds an aspect of Arts education that can be seen as directly relevant to their work.

Reflections on Interdisciplinary Collaboration. Interdisciplinary collaboration is essential to all of the work reported here; thus, in addition to advocating for STEAM-based education, we hope that this report also serves to make explicit some of the fundamental benefits of this type of collaboration, especially with respect to the doors that it can open for students. As mentioned earlier, data visualization provides a particularly interesting platform for teaching and demonstrating this type of collaboration as it can often require involvement of not just artists and technologists but also "domain scientists" who act as the end users of the tools and techniques developed. Our students seem to benefit greatly from this experience. A graduate of the RISD design program recently reported, "The collaboration between science and art is exactly what my art is about... Much of the abstract work that I have been creating since my last few years at RISD is, in part, inspired by the work that we did in the Lab at Brown."

5 Conclusion

There are many benefits to STEAM-based education, but there are also important challenges relating to bringing together interdisciplinary groups and developing workflows that engage students from both STEM and Arts backgrounds. We believe virtual reality and related technologies can provide an exceptional platform for STEAM education, but our experience suggests that new tools and methodologies are required, especially when the goal is to support collaborative teams in the classroom. We have described several of the most important tools and methods we have adopted to make this work; many of them relate to sketching, which is perhaps not surprising since it is such a fundamental design activity. We encourage educators to think creatively about how sketching can be reinterpreted and integrated into classroom exercises and activities, especially when working with advanced visual technologies, such as VR. We have that doing so can completely change the classroom dynamic for STEAM education; it levels the playing field, empowering artists to create in virtual reality environments, while also teaching science students how to engage with core concepts in arts education. We see a profitable future in continuing to advance and disseminate these tools; when they are combined with methodologies, such as teaching with critique and interdisciplinary collaboration, the result can be a powerful approach to STEAM-based learning.

Acknowledgement. Thanks to the students who have taken part in our courses and to the instructors who have collaborated in developing the courses, especially Fritz Drury, Sharon Swartz, and Peter Richardson. Thanks to David Nuckley and Arin Ellingson who provided data for the spinal kinematics example. This work was supported in part by the National Science Foundation (IIS-1054783, IIS-1218058).

References

1. Sobel, J.S., Forsberg, A.S., Laidlaw, D.H., Zeleznik, R.C., Keefe, D.F., Pivkin, I., Karniadakis, G.E., Richardson, P., Swartz, S.: Particle flurries: Synoptic 3D pulsatile flow visualization. IEEE Computer Graphics and Applications 24(2), 76–85 (2004)
2. Jackson, B., Coffey, D., Thorson, L., Schroeder, D., Ellingson, A., Nuckley, D., Keefe, D.F.: Toward Mixed Method Evaluations of Scientific Visualizations and Design Process as an Evaluation Tool. In: BELIV 2012: Beyond Time and Errors: Novel Evaluation Methods for Visualization, Workshop at IEEE VisWeek 2012 (2012)
3. Buxton, B.: Sketching User Experiences: Getting the Design Right and the Right Design. Morgan Kaufmann, San Francisco CA (2007)
4. Keefe, D.F., Acevedo Feliz, D., Moscovich, T., Laidlaw, D.H., LaViola Jr., J.J.: CavePainting: A fully immersive 3D artistic medium and interactive experience. In: Proceedings of I3D 2001, pp. 85–93 (2001)
5. Keefe, D.F., Karelitz, D.B., Vote, E.L., Laidlaw, D.H.: Artistic collaboration in designing VR visualizations. IEEE Computer Graphics and Applications 25(2), 18–23 (2005)
6. Keefe, D.F., Acevedo, D., Miles, J., Drury, F., Swartz, S.M., Laidlaw, D.H.: Scientific sketching for collaborative VR visualization design. IEEE Transactions on Visualization and Computer Graphics 14(4), 835–847 (2008)
7. Schroeder, D., Coffey, D., Keefe, D.F.: Drawing with the Flow: A Sketch-Based Interface for Illustrative Visualization of 2D Vector Fields. In: Proceedings of ACM SIGGRAPH/Eurographics Symposium on Sketch-Based Interfaces and Modeling 2010, pp. 49–56 (2010)

Architectural Geo-E-Learning

Geolocated Teaching in Urban Environments with Mobile Devices: A Case Study and Work in Progress

Ernest Redondo[1], Albert Sánchez Riera[1], David Fonseca[2], and Alberto Peredo[3]

[1] Universidad Politécnica de Cataluña-Barcelona Tech. Barcelona, Spain
{ernesto.redondo,albert.sanchez.riera}@upc.edu
[2] Architecture School - La Salle, Universitat Ramon Llull. Barcelona, Spain
fonsi@salle.url.edu
[3] CUAAD, Universidad de Guadalajara, Jalisco, México
oderep@gmail.com

Abstract. This work addresses the implementation of a mobile Augmented Reality (AR) browser on educational environments. We seek to analyze new educational tools and methodologies, non-traditional, to improve students' academic performance, commitment and motivation. The basis of our claim lies in the skills improvement that students can achieve thanks to their innate affinity to digital media features of new Smartphones. We worked under the Layar platform for mobile devices to create virtual information channels through a database associated to 3D virtual models and any other type of media content. The teaching experience was carried out with Master Architecture students, and developed in two subjects focused on the use of ICT and Urban Design. We call it Geo-elearning because of the use of new eLearning strategies and methodologies that incorporate geolocation, allowing receiving, sharing, and evaluate own-generated student's proposals, on site.

Keywords: Augmented reality, E-Learning, Geo-Elearning, Urban Planning, Educational research.

1 Introduction

Augmented Reality (AR) technology is based on overlapping virtual information in real space. A framework where technology could be potentially more interesting is the representation and management of the territory, because real scenes could be "completed" with virtual information, which would facilitate a greater awareness and better understanding of the environment. In the field of architecture, for instance, AR allows new buildings proposals visualization, and their impact assessment, on its planned site. To do that, a Geographic Information System (GIS) is needed to provide and manage and filter public queries with different levels of accurate and upgradeable information. In short, we need to link a 3D model to a database which contains all the necessary information associated with it. This process has been the first target of this work.

R. Shumaker (Ed.): VAMR/HCII 2013, Part II, LNCS 8022, pp. 188–197, 2013.

Furthermore, the introduction of new learning methods 3.0 and new collaborative technologies, besides the new ways of information access via 3G phones or tablets, offer new opportunities to provide educational multimedia content. New representation systems and management tools are getting closer and easier to use. As part of the architectural or urban planning, these systems provide new tools for the representation of architectural forms and related content. In adition to this capabilities, cloud computing thechnology which allows sharing applications and services via Internet at any time, generates a workflow where teaching experience becomes a new paradigm of training processes and contextual learning. The evaluation of a methodology focused on the training of architects and planners is the second objective of this investigation.

Our proposal involves methodological changes which include information management through GIS technologies, visualization using RA, and all types of mobile devices interaction. A free application that support all this features is Layar© by SPRXmobile©, initially designed for tourist information. In our case, we used this platform because of its compatibility with all mobile operating systems. Registration of virtual information is based on the use of GPS, which is accurate enough for outdoor environments.

The teaching experience, that was performed to validate the previous premises, arises at Master level subjects. They involve the use of ICTs applied to the analysis and territorial representation, where 3D GIS systems, 3D modeling, and Urban Virtual Reality are combined. The proposed approach is based on the use of smartphones to incorporate virtual models generated by the students in an existing AR platform to view them on site, through their own mobile devices. We tried to promote new learning strategies for sharing, collaborating and transmitting information to other participants. To address the process scientifically, we developed a case focused on large-scale urban projects, in particular on the campus BKC (Barcelona Knowledge Campus), University of Barcelona (UB) and Polytechnic University of Catalonia (UPC).

2 Framework. ICT 3.0. Geo-Elearning

Currently, ICT technologies related to Web 2.0 environments, such as RA or Geolocation, besides mobile devices popularity and their recent advances, open new prospects in Mobile Learning (ML) [1], a specific field of E-Learning (EL). Thanks to this new approach, it is now possible to design teaching activities where student's queries about a particular site are facilitated to share information, experiences, and content, most of the time, own generated. It is known as Web 3.0. These methods can generate extra student's motivation because of the use of their own devices in real environments [2]. Smartphones GPS integration in education [3], and 3D data visualization in outdoor environments [4] has been already tested. In the case of urban planning [5] [6] as well as historical heritage, we can find systems that link graphical information systems with databases, as assitance tools for interpretation and information compilation. Other studies discuss the proper integration of spatial data from different sources

[8], they mostly rely on information mapping and the use of conventional GIS [9]. But, those systems do not usually use AR techniques to geolocate information, and they merely generate a model from photographs or by the use of laser scanner techniques. These projects neither deal with database maintenance and queries filtering, in real time, nor by the implementation of AR in teaching environments.

On the other hand, the emergence of web-based 3D globe viewers with elevations, satellite and aerial images, maps and 3D features, such as Google Earth© or Virtual Earth©, has promoted the exchange and visualization of geo-referenced 3D models in a natural way. Despite its shortcomings, success of these visualization tools is greater than traditional 3D globe viewers based on VRML and X3D [10]. Moreover, the use of an RA urban planning systems to allow consultation through mobile devices, as intended in our trial, has been reported recently [11], [12], [13]. Other authors [14] investigated the use of smartphones as a tool for public participation in urban planning projects. But research addressing these issues is still poorly documented. The introduction of more user-friendly technologies (such as mobile phones, tablet, social networks, etc.), in the learning process, is an educational strategy that removes the traditional and bored lectures. In this case, it helps to address the problem of urban 3d models design, its georeferencing, its consultation and its assessment on site through mobile devices dynamically.

3 Teaching Context

The experiment was carried out by 11 students of Architecture and Planning, in an elective course called "ICT applied to Spatial Analysis", which is tough in the Research *Master in Land Management and Valuation* of UPC, Barcelona-Tech. We worked within the scope of BKC, during the academic year 2011-2012. Total course duration was 60 hours. This method is currently being replicated (during 2012-2013), in the *Master in Processes and Graphical Expression in Architectural Urban Projection*, in the Center of Arts, Architecture and Design (CUAAD, Universidad de Guadalajara, Mexico).

4 Methodology

4.1 Modelling, Geolocation and Virtual Models Visualization

We have started from BKC contents, in particular, documents and planimetric images provided by the authors of the project. Each student should have a mobile device equipped with camera, GPS and 3G connection, and was required to download the free browser *Layar Viewer©*. The process aims to incorporate virtual models generated by students in a mobile application and view them on its planned site. To do this we used a Geolocation-based AR application which uses GPS, compass, and other sensors in the student's mobile phone to provide a "heads-up" display of various geolocated points-of-interest (POI's). In this case, student's Architectural proposals placed on the campus. Students worked with Sketchup© and 3dsMax©, to perform volumetric models and textures design, using real building materials. Secondly, they were divided into different groups A, B, and C. Each group modeled three proposals with the information provided, according to the preset numbers, and set coordinates origin at point 0,0,0 of the modeling program. Then models were exported to *.Obj format, and they were imported from the LayerModelConverter© (LMC, easy installation and free) program, which generates a specific file to be readed by Layar viewer. In addition, UTM coordinates should be recorded to be associated to the model in the database. In order to avoid problems, in this point, students should control the export path, check the units, activate the texture maps and change all YZ coordinates.

Previously an information channel was generated as a developer in the layar platform. The channel was published using BKC basic information and was configured to allow the use of filters. Comprehensive filter settings helped users to find POIs that were interesting easily, and to separate proposals by groups. Database and PHP file was hosted in a public server with PHP, MySQL, Java, support.

Meanwhile, students have installed Layar© RA browser, in their mobile devices. Once installed students were required to locate the particular channel created. In this case, it was located within the category of geo-layers of architecture and buildings, named "Tesis_Albert_app". Students proceed to filter by groups the architectural proposals. In our experiment a group of students evaluated the models of the other groups. The query is sent to the server host, which returns the selected POI. They are shown in the screen superimposed to the real image captured by the camera. As we approach or focus on one in particular, at the bottom of the screen appears a label with the model reference information and distance from the user. Clicking that label, students can acces to questionnaire "iweb" to respond and make comments about appearance, impact and scale of the building.

4.2 On Line and Contextual Questionnaires Design

The next step to complete the "on site" learning process is the contextual questionnaires design, which ought to be answered by students once they had located all proposals generated by different groups. The questionnaire is accessible through a descriptive label in the device display or by pressing on virtual buildings models. We have proposed two different questionnaires: in the case of existing buildings, students can access to questionnaries about the use of technology and system usability assessment, according to a standardized methodology for these experiments; for new buildings, students evaluated and reviewed all information linked to them (Fig 1), as project plans, memory, or project rendered views, which provided aditional insights. In addition, students should choose the best viewpoint to appreciate the integration of the new project with the existing building.Personal responses were sent directly to the teacher who received and analized the information.

Fig. 1. Design and implementation process of on-line contextual questionnaires used in the experience. Source: The authors.

4.3 Case Study

The students have different profiles (they came from 5 different countries, and different disciplines andeducational areas): from Latin America, to Saudi Arabia, most of them were architects from countries where planning and compulsory computer training is not included.

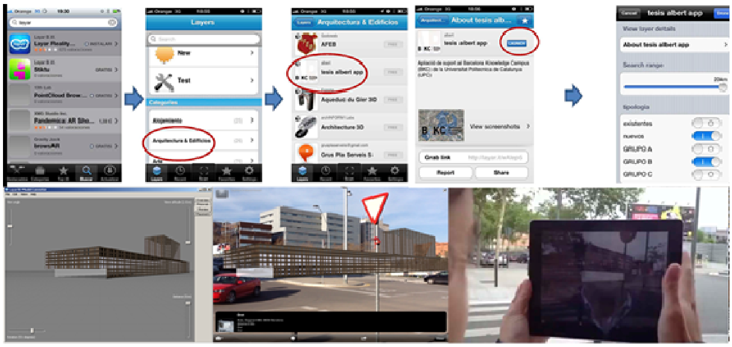


Fig. 2. BKC projects on Layar platform. Feasibility Study of "on site" visualization methodology for 3G mobile devices. BKC projects on Layar platform.

They were divided into three groups: A for existing buildings, and B and C for new buildings, according to planned campus project. They modeled three buildings for each group wich were coded as folows: B1, B2 and B3, C1, C2, and C3. Depending on their computer skills students worked in pairs or individually. They developed the projects assigned in three sessions of four hours each. Once modeled, proposals were geo-referenced, and teachers proceeded to export and upload all generated information to the database that feeds the channel. In the last session, all groups made a guided tour through the campus, where Student's Architectural proposals were located, tested, and evaluated. Since not all students had 3G phones, they were pooled to make the experience, even with the phone of another partner.

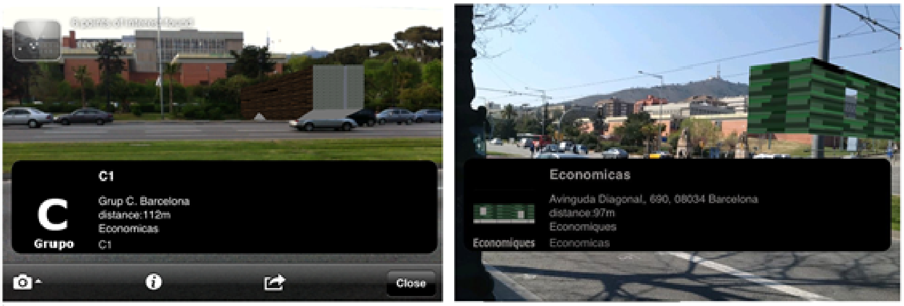


Fig. 3. Sample images of projects from BKC visualization "on site"

5 Results

5.1 Results in the Case Study on BKC

The experience was innovative: 100% students were able to complete the exercise, so they could visualized on site their proposals. Using their mobile devices and geo-locate information, they evaluated all buildings from differents groups. Although GPS accuracy was poor, as expected, both the application of semitransparent textures and the object visualization at a distance, allowed the recreation of a scene true enough to score the proposals and assess their visual impact. Buildings, displayed in their planned site, were assessed using issues such the scale, color, location, heigh, etc…

Questionnaire responses were used as a mechanism to verify student's participation in the exercise. Results showed student's high degree of satisfaction in relation to the activity, and the course contents. The self-assessments about the interest and usefulness of the technolgoy were useful to get an idea of the high level of acceptance achieved.

6 Evaluation

To validate the experiment, we conducted a usability questionnaire as detailed above; in the same way we did in other experiments of this R&D project. The questionnaire is divided in three parts: Personal training and prior knowledge about the technology; Teaching content and course material opinion; and finally, AR technology and software used.

In relation to students personal training and the prior knowledge level of the technology, should be noted that the most often used applications were "Email" and Computer Aided Design (CAD) applications, followed by "Internet browsers" and office applications. The most used operating system was Windows and the less knowledge resulted in LINUX and AR systems. (Scale: 0 = none, 5= advanced). Similar data was obtained from other ICT courses evaluated. Note the high punctuation in CAD applications, possibly because students are on their final phase of their training as architects. Global opinion was rated very positively at the end of the course despite the prior ignorance of AR technology.

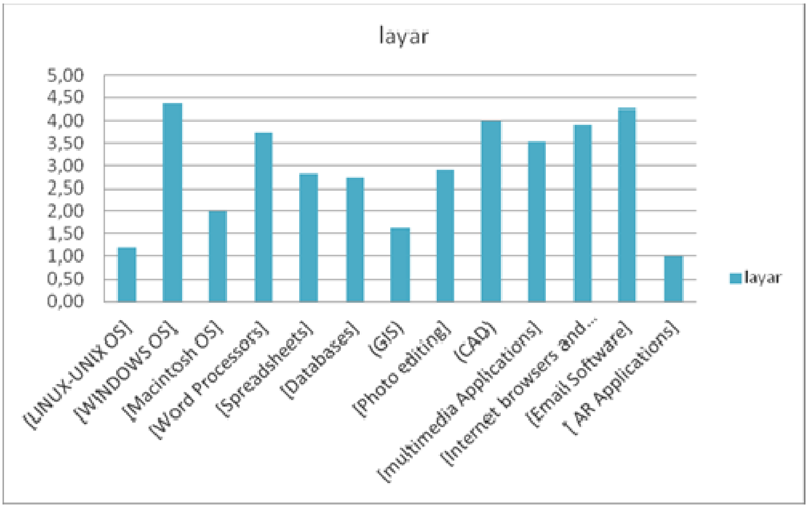


Fig. 4. Average results about personal trainingthe and prior knowledge of AR Technology

Related to the opinion, teaching content, and material of the course, there should be noted that it was very high rated. Material representativeness and the number of exercises in accordance with the objectives were optimal. The final average rating was more than 4.00 out of 5 points (Fig. 5). The worst rated question was referred to the possibility of learning such content independently. Software used was also low rated, probably due to the registration inaccuracy based on GPS.

And related to augmented reality technology and software used, 100% of the students found them useful in the field of architecture and building construction, despite having no prior knowledge of it. Final assessment was 4,27 points out of 5.

In a correlation analysis between the course global opinion and the other variables, a high correlation (0.69) was detected with: the representativeness of the exercises and the quality of the presentation. So these variables are crucial to the success of this teaching experience. Not being so correlated with the fact of being able to solve the exercises independently or with the number of exercises proposed. The strongest correlation (0.86), however, was in the use of appropriate software, and this is, therefore, the most important variable to consider in future work. Variables related to prior knowledge of technology and the use of different software and operating systems were not significantly correlated with the overall opinion of the course.

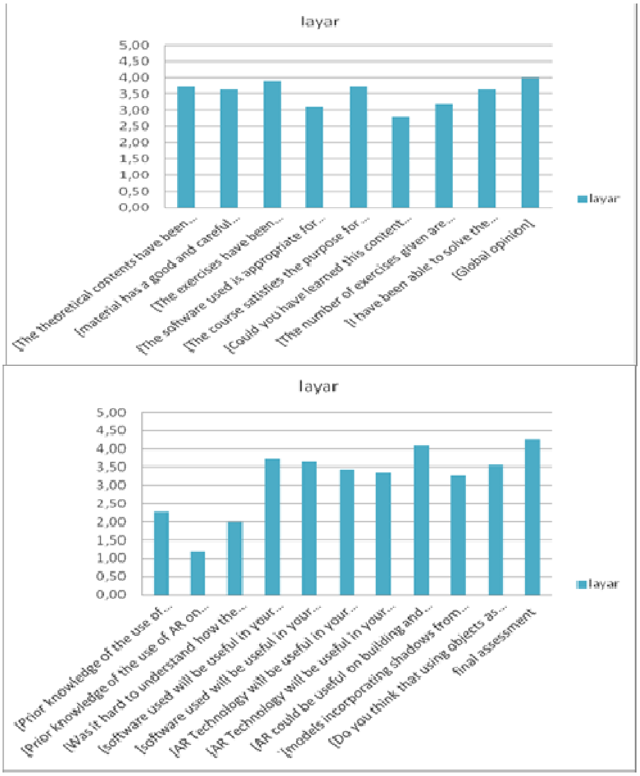


Fig. 5. Top: Rating table about global opinion, content, and material of the course Layar. Botton: Rating table about technology and software used.

7 Conclusions

We've tested the use of new mobile AR technologies in a master's program in architecture as a complement to other experiments. They had been carried out in first and second cycle courses of graduates in the European Higher Education Area (EHEA). In this case, the experience was performed within a larger framework of a R&D project. Thus we've completed an initial assessment in all training levels of architects and planners.

The ability of spatial representation of AR technology has been assessed in models of different complexity and level of realism. The processes to generate and visualize models have been also optimized to suit the capabilities of 3G mobile phones. Moreover we evaluated the ability of spatial location of the Geo-location technology integrated on mobile devices.

Geo-location, RA, and mobile technologies, used in the teaching processes, improve student's academic performance. They allow reducing learning and response times by the student. The teacher, in turn, can extend the learning process anywhere at any time. However, architectural models generation with AR is sometimes complicated and of poor quality in most application. Probably, due to the geometry simplification to reduce the number of triangles to be processed. To counter this effect and to achieve realism in the AR scene is necessary to use lighting simulation techniques to be applied in textures.

This difficulty is lower for higher level students, master's or postgraduate, as they have better computer background of 2D drawings, 3D rendering, and image processing required in these processes. On the contrary if they have no previous experience is better to use high usability programs like SketchUp with special plugins for them. Geo-location of architectural models using 3G phones based on GPS is poor, and only feasible to be seen at a distance of over 25 meters. Therefore we can conclude that these technologies are becoming accessible and easy to use. They increase student satisfaction and interest in the course content as they feel very motivated and are regular users of mobile devices.

In reltaion to the second experiment that is being replicated in the CCU of the UDG, Mexico, we are currently awaiting final presentations of student work to be evaluated. For now, we can say that the number of students who participated in the experiment was 26 and only 41% were able to develop all practices.

Fig. 6. Visualization process on site, of various projects of the CCU Convention Hotel, UDG. Source: The authors and Masters students.

Acknowledgements. Project funded by the VI National Plan for Scientific Research, Development and Technological Innovation, 2008-2011, Government of Spain. No EDU-2012-37247/EDUC.

References

1. Tsvetozar, G., Evgenia, G., Angel, S.: M-Learning - a New Stage of E-Learning. In: Int. Conference on Computer Systems and Technologies. CompSysTech. (2004)
2. Sharples, M.: The Design of Personal Mobile Technologies for Lifelong. Learning Computers and Education 34, 177–193 (2000)
3. Zhang, B., Looi, C.-K., Seow, P., Chia, G., Wong, L.-H., Chen, W., So, H.-J., Soloway, E., Norris, C.: Deconstructing and reconstructing: Transforming primary science learning via a mobilized curriculum. Computers & Education 55, 1504–1523 (2010)
4. Bioret, N., Moreau, G., Servieres, M.: Towards Outdoor Localization from GIS Data and 3D Content Extracted from Videos. In: IEEE International Symposium on Industrial Electronics (ISIE 2010), pp. 3613–3618 (2010)

5. Huang, B., Jiang, B., Li, H.: An integration of GIS, virtual reality and the Internet for visualization, analysis and exploration of spatial data. International Journal of Geographical Information Science 15(5), 439–456 (2001)
6. Kwan, M.-P.: Interactive geovisualization of activity-travel patterns using three-dimensional geographical information systems: a methodological exploration with a large data set. Transportation Research Part C: Emerging Technologies 8(1-6), 185–203 (2000)
7. Droj, G.: Cultural Heritage Conservation by GIS. Nyugat-Magyarországi Egyetem, Geoinformatikai Kar, Székesfehérvár. University of Oradea 1-10 (2010)
8. Belussi, A., Migliorini, S.: A framework for integrating multi-accuracy spatial data in geographical applications. GeoInformatica 16(3), 523–561 (2011)
9. Hosse, K., Schilcher, M.: Temporal Gis For Analysis And Visualisation of Cultural Heritage. In: Proceedings of CIPA XIX International Symposium, Commission V, WG5, Antalya, Munich, pp. 1–6 (consultado January 24, 2013), `http://cipa.icomos.org/index.php?id=61`
10. Breunig, M., Zlatanova, S.: 3D geo-database research: Retrospective and future directions. Computers & Geosciences 37(7), 791–803 (2011)
11. Anagnostou, K., Vlamos, P.: Square AR: Using Augmented Reality for Urban Planning. In: 2011 Third International Conference on Games and Virtual Worlds for Serious Applications, pp. 128–131. IEEE (2011)
12. Lonsing, W.: Architectural models in urban landscapes: Synthesis of markers and virtual structures. In: IEEE International Symposium on Mixed and Augmented Reality - Arts, Media, and Humanities, pp. 109–110. IEEE (2011)
13. Anagnostou, K., Vlamos, P.: Square AR: Using Augmented Reality for Urban Planning. In: 2011 Third International Conference on Games and Virtual Worlds for Serious Applications, pp. 128–131. IEEE (2011)
14. Allen, M., Regenbrecht, H., Abbott, M.: Smart-phone augmented reality for public participation in urban planning. In: Proceedings of the 23rd Australian Computer-Human Interaction Conference on - OzCHI 2011, pp. 11–20. ACM Press, New York (2011)

Part III

Business, Industrial and Military Applications

Mixed Reality Environment for Mission Critical Systems Servicing and Repair

Andrea F. Abate, Fabio Narducci, and Stefano Ricciardi

VRLab, University of Salerno, Italy
{abate,fnarducci,sricciardi}@unisa.it

Abstract. Mixed Reality (MR) technologies may play an important role in assisting on-site operators during maintenance and repair activities. Nevertheless, industrial equipment augmentation requires a high level of precision when co-registering virtual objects to the corresponding real counterparts. In this paper we describe a comprehensive proposal for a mixed reality environment aimed to improve the effectiveness of servicing and repair procedures in mission critical systems, while reducing the time required for the intervention. The tracking of the user's point of view exploits a multi-marker based solution for robust and precise augmentation of the operating field. The architecture also features a diminishing visualization strategy allowing the user to see only the fraction of real equipment that is relevant for the maintenance task. A finger color-based tracking provides powerful interaction capabilities by means of a not-instrumented interface exploiting colored fingertips caps. An evaluation study of the proposed MR environment, performed by technicians with no previous experience of MR systems, highlights the potential of the approach.

Keywords: Mixed reality, diminished reality, finger based interaction, AR based maintenance.

1 Introduction

Mission critical installations (e.g. radar, navigation and communication systems aboard military and civil vehicles or based in airports, ports, etc.) represent examples of high-tech environments featuring complex hardware and software components requiring prompt servicing as, in case of breakdown, the security of a large number of people may be at risk. In this context, Mixed Reality (MR) technologies may play an important role in assisting on-site operators during maintenance and repair activities but, ultimately, issues like tracking accuracy and coverage, augmentation strategies and interaction capabilities determine whether a MR application is useful or not. Henderson and Feiner explored the benefits of AR for maintenance and repair [1]. They also showed how MR can assist military mechanics conducting routine maintenance tasks inside an armored vehicle turret exploting a NaturalPoint OptiTrack tracking system [2]. In 2002, the project ARVIKA [3] fostered the research and the development of AR technologies for production and service in the automotive and aerospace industries, for power/process plants, for machine tools and production

R. Shumaker (Ed.): VAMR/HCII 2013, Part II, LNCS 8022, pp. 201–210, 2013.

gears. Klinker et al. [4] presented an AR system for the inspection of power plants at Framatome ANP, while Shimoda et al. [5] presented an AR solution in order to improve efficiency in nuclear power plants (NPP) maintenance interventions and to reduce the risk of human error. The precision of tracking and visualization strategy adopted represent two key aspects that have to be severely taken into account in industrial components augmentation. Computer vision, in both the marker-based [5-6] and marker-less [7-8] variants, is generally recognized as the only tracking methodology that has the potential to yield non-invasive, accurate and low cost co-registration between virtual and real [9]. Concerning the visualization strategy, how and when to augment real objects and environment may have a great impact on the quality of user assistance provided. For instance, in some situations the scene observed should be possibly simplified rather than augmented. In our proposal, the system is able to remove distracting items from the field of view, leaving as visible only the elements required to perform the desired task.

Another aspect to be considered is represented by the interaction level available and the related interaction paradigm: the user should be able to select what kind of augmenting content to display according to his/her needs by interacting with the MR environment without complicated gear. Wei et al. [10] introduced a MR framework featuring voice commands, but a hand-based interface would rather be more suited to the scope. As a hardware solution (instrumented gloves plus wrists tracking) would provide accurate hands capturing but would also reduce system's acceptability, a more feasible option is to exploit image-based techniques to track hands in real time [11] exploiting the same camera video stream used for the augmentation. The simple and robust approach developed in this paper is based on the recent work by Mistry and Maes [12] and relies on colored caps worn on index and thumb fingers to track their position and gestures, providing effective and natural interaction within the MR environment.

In this paper we describe a comprehensive proposal for a mixed reality environment aimed to improve the effectiveness of servicing and repair procedures in mission critical systems, while reducing the time required for the intervention. As the MR system presented is targeted to rack sized equipment rich of small components and operating in the vicinity of strong electromagnetic fields, it features a computer vision based multi-marker tracking method to deliver high accuracy and robustness and allowing the user to observe the surrounding environment and the computer generated graphics by means of a video see-through HMD. The proposed system also provides powerful interaction capability with the co-registered virtual/real objects by means of a two-hand not instrumented finger based interface exploiting colored fingertips tags.

The rest of this paper is organized as follows. Section 2 describes the overall architecture of the system and each of its main components. Section 3 reports about the experiments conducted in a real industrial case to assess the advantages and the eventual limitations in the proposed approach. Finally, Section 4 draws some conclusions introducing future directions of study.

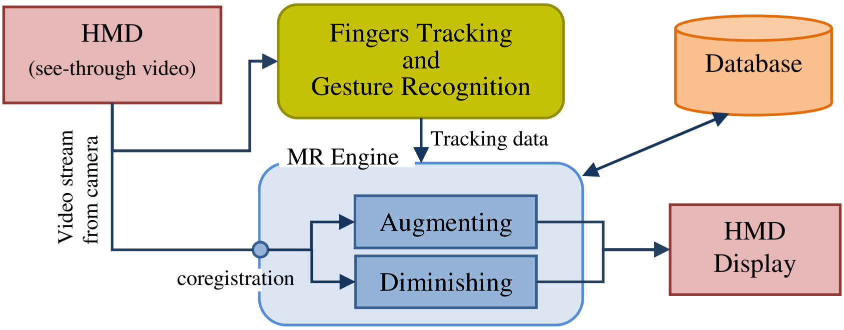


Fig. 1. The overall schematic view of the system

2 System's Description

The schematic view of the proposed system is shown above (see Fig.1). The architecture proposed integrates precise multiple marker based tracking, two modalities for scene augmentation and a not-instrumented finger based interface to provide effective and reliable visual aid during maintenance operations. It is designed around three main components. The Mixed Reality Engine (MRE) is in charge of user's head tracking, scene augmentation/rendering and servicing procedures management. The Finger Tracking module captures fingers position and the Gesture Recognition module recognizes gestures enabling the human-computer interaction, while the Database contains the working environment setup, the virtual contents and the maintenance procedures required for the system to work. To start the assisted servicing procedure, the user has to wear a video-based see-through HMD, a backpack enclosed notebook and a few fingertips caps required for contactless interaction. This architecture is further detailed in the following subsections.

2.1 The Mixed Reality Engine

The tracking system developed exploits the ARToolkit open source AR library for estimating user's perspective. Typical marker based tracking systems operating under controlled conditions (i.e. avoiding or at least reducing strong reflections, extreme shadows and excessive camera noise) are able to track the point of view of the user provided that a single marker is entirely captured in a frame. This simple solution, often adopted for desktop based AR applications, forces the user to continuously aim at the marker, holding it in the center of the visual field to reduce the risk of detection miss. Arranging a marker in the middle of operational environment could simply be unfeasible for many application contexts characterized by uneven surfaces or it could even interfere with the operations. In our proposal, many of these issues are addressed by exploiting multiple markers, thus delivering an inherently more robust and more accurate tracking even using small markers. A useful advantage of the multi marker approach is that it is easily scalable. For instance, by adding other six markers (arranged in two strips of three, placed 60 cm. above and below the basic set) an optimal tracking volume of 180x60x60 cm (adequate for a full-size industrial rack) is seamlessly achieved. Besides an embedded calibration function aimed to measure and correct the lens distortion of the camera, a manual procedure allows the user to fine-tune co-registration between the real camera and its virtual counterpart in charge of rendering the required graphics. Each of the six degrees of freedom, including the focal length of the camera, and the marker thresholds, can be adjusted precisely. This task is performed only once unless physical or environmental changes occur in equipment's configuration. Concerning the augmentation of the scene, the system is designed to support two different approaches. The former is a classic augmentation strategy which consists of displaying different kind of virtual objects (e.g., arrows, labels, 3D models and so on) onto the captured scene. In this way the system literally "augments" the real environment (see Fig.2). However there are situations in which adding extra information to the scene may lead to an even more confusing effect. In particular, we refer to environments characterized by the presence of a large amount of interaction points (e.g., a control board, a rear panel of a complex device etc). In this case, showing further information in addition to those already present may be counterproductive. These are the reasons that led us to develop of the latter augmentation strategy, enabling the system to hide part of the operating environment to let the user focus on the physical elements on which to perform a particular task. This methodology is inspired to the concept of diminished reality [13-15], but it is based on the selective occlusion of unwanted elements rather than on image based object removal (see Fig.3). From a more technical point of view, to the purpose of hiding real objects or part of them, the engine renders an occluding (polygonal) surface on which can be applied either a diffuse map or an opacity map.

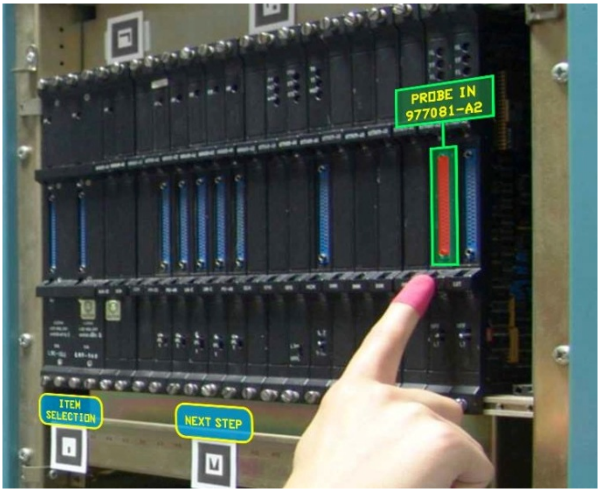


Fig. 2. User interacting with a panel augmented by virtual labels and GUI

By exploiting the device's formal representation stored in the system database (more on this later on), the engine either loads the associated textured polygons or it builds up a procedural texture consisting of a black background featuring white "holes" of various shapes (e.g, circles, squares, polygons and so on) corresponding to the hotspots that should stay visible. All the necessary information to perform this task is available in the working environment database. Once the textured polygons are built, the engine renders it over the real device occluding all the not relevant hotspots. Both the more common augmenting and the diminishing strategies are meant to improve user's operational capabilities. However, there are contexts in which one is more suited than the other. Which of the two visualization methods should be used depends on the total number of hotspots present in the surrounding of the virtual contents visualized at a given step of the intervention. If the density of the hotspots exceeds a threshold then the diminishing modality is preferred. Anyhow, the user may always switch to the other modality in any moment by means of a specific Augmented/Diminished View toggle present in the visual interface. Finally, by combining both augmented and diminished reality a third hybrid visualization approach could be realized, providing a simplified view of the operating field in which only the elements left visible are augmented with additional info. Whatever the strategy adopted, scene augmentation or diminution is made possible thanks to a formal scene representation based on XML, a markup language universally known for being easily extensible and non ambiguous. The XML database consists of a collection of files providing the necessary information to correctly locate each relevant element of the working environment within the 3D space. The MR engine also performs another crucial task: the maintenance procedure management. Each generic maintenance procedure can be represented as a deterministic finite automaton (DFA). According to this approach, a particular state represents a maintenance step and its links define the execution order. DFA result particularly suited to model both simple and complex maintenance procedure in an easy, verifiable and legible way. The DFA representation of a particular procedure is converted in a XML file where a `<step>` tag defines a state. Any possible path through the automaton defines a procedure's file. By this approach a single XML procedure file defines a specific execution order in a maintenance procedure.

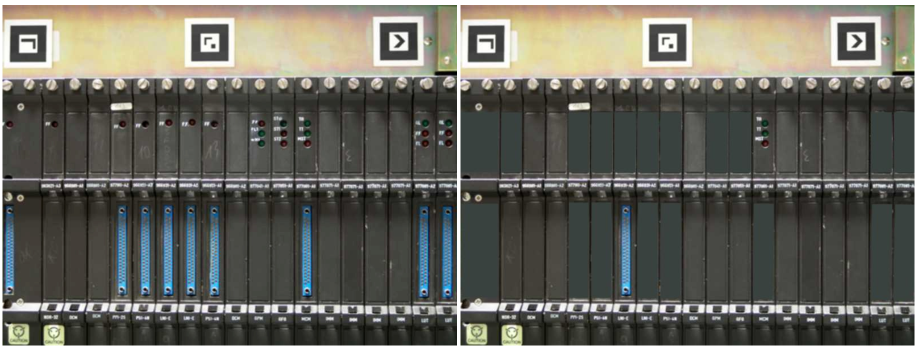
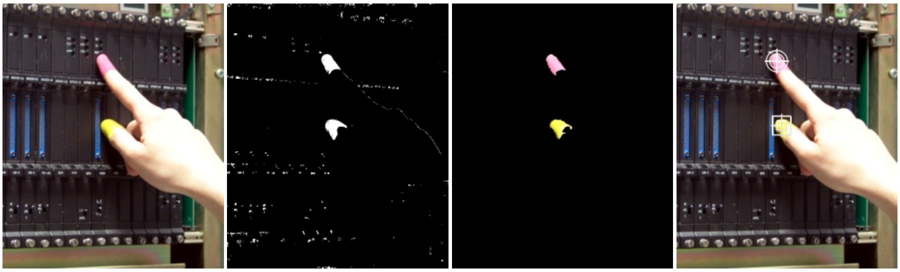


Fig. 3. Original frame (left) and diminished view (right). Only the required elements are left visible, while textured polygons hide potentially confusing items.



(A) (B) (C) (D)

Fig. 4. Fingertips tracking in four steps: (A) original frame grabbed from the camera; (B) candidate pixels highlighted (C); matching regions found; (D) resulting tracking

2.2 Finger Based Contactless Interaction

The contactless interface developed frees the user from the usage of any tangible I/O device to communicate with the system. The user indeed, has only to wear small rubber caps of different colors over his thumb and index fingertips, eventually of both hands. The image-based tracking exploits the same video stream used for marker tracking to achieve fingertips detection and tracking, thus yielding a reduction of computational cost compared to a solution based on a dedicated camera and a simpler hardware configuration. Each finger is associated to a different color according to a simple enrollment procedure repeated for each finger and generally performed only the first time the system startup. The fingertips samples captured are analyzed in the HSL color space to extract the dominant hue and saturation ranges, while the lightness component is used to filter out eventual highlights. At runtime, the image grabbed from the camera is subsampled (by a factor of 4 to 8 times) to both reduce the effect of camera noise and to optimize system's performance. For each pixel in the subsampled image whose HLS levels fall within the ranges defined during enrollment, a recursive search for similar (color wise) neighbors is performed until a region of 20x20 pixels is explored. If at least one half of the pixels inside this region matches with the original pixel, then the engine recognizes that region as one of the colored caps to track (Fig. 4 summarizes the fingers detection steps). This approach resulted both reliable and responsive, granting a sustained frame rate always well above 30 frame per second for an image resolution of 640x480 pixels (typical for most HMD cameras). Finger tracking enables a rich interaction paradigm that can be exploited in many different ways. For instance the user may query the working environment to learn more about it by simply moving the index finger over any hotspot (i.e., a screw, a button, a handle, a led indicator and so on) according to his/her point of view to obtain visual info about a particular component. Moreover, finger tracking enables operating the system by using a graphical user interface (GUI). The main challenge with an intangible GUI is related to the interaction paradigm, which has to manage the lack of physical contact with the interface elements (buttons, slider, toggles etc.). Indeed, when using a conventional (tangible) interface, the kinesthetic feedback provides an important confirm of the operations performed. To address this issue, we exploited a time based interaction paradigm, requiring the user to hold the finger in position for a defined (around one second) amount of time to trigger the associated function. A visual feedback, in the form of a small progress bar drawn over the GUI element selected, inform about the selection state (i.e. hold the finger until the progress is over).

The same paradigm is used during a servicing procedure to move from a step to the next one or previous one as well as to play/pause/rewind an animated virtual tool showing how to perform a specific task. Additionally, as the system is designed to track up to four colored caps, multi-finger gestures can be used to provide more powerful interaction modalities, like object picking, zoom or rotation.

3 Experiments and Results

We conducted two kinds of experiments on the system described above to assess both the performance of the tracking approach and the overall usability of the MR environment applied to a radar system training facility. The hardware used for the experiments includes a notebook, featuring Intel I5 processor and Nvidia GeForce 9 series graphics board and an ARVision-3Dvideo see-through HMD from Trivisio, equipped with two 800x600 LCD display and two 640x480 cameras capturing the surrounding environment at 30 FPS (see Fig.5). Only the left camera has been used for scene capture. During operations the notebook was contained inside a small backpack. In any system evaluation, user testing is of great relevance in confirming the validity and the effectiveness of solutions adopted. To this aim, we prepared a user questionnaire to assess the perceived quality of the interaction after performing a number of tasks significant to the operating context considered. The evaluation sessions involved ten users, selected among specialized technicians with no previous experience of either MR systems or contactless interfaces. The following is a list of the tasks performed by the testers: ***Load*** *a new servicing procedure*; ***Select*** *a particular hotspot*; ***Select*** *a function from the GUI*; ***Toggle*** *between two functions*; ***Perform*** *a servicing procedure*.

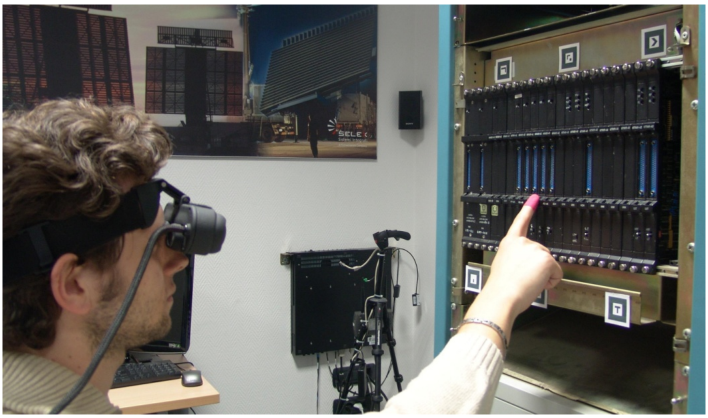


Fig. 5. User wearing HMD and colored fingertip caps

In the final questionnaire, the questions were presented using a five-point Likert scale, where respondents specify their level of agreement to a statement. In order to avoid any bias, some statements were in positive form and others in negative one. This was taken into account in the final assessment of results. The following is the list of the proposed statements:

1. Available finger based functions are easy to perform;
2. Functions are too many to remember them;
3. Interacting by fingers is not intuitive;
4. It is easy to select objects;
5. Visual aids are clear and useful;
6. It is easy to operate the contact-less GUI;
7. The type and number of available functions to interact with objects is not sufficient;
8. Devices worn are not comfortable during operations.

The answers to the questionnaire are summarized in the table below (see Table 1). Most participants reported a good confidence feeling during the usage of the system, and some of them also reported an operational advantage in performing the proposed tasks with respect to their usual operating modality. All the participants to the evaluation sessions have also been interviewed to better understand the motivations behind the answers provided. Most comments showed a general agreement about the finger based interface, although most of them remarked, as we expected, the lack of a physical contact as something strange to which is not easy get used. Both the two visualization modalities were considered useful for improving the confidence and avoiding distraction errors during the operations, while, not surprisingly, the HMD caused a somewhat stressful experience to most users.

Table 1. Scores reported after subjective system evaluation according to five-point Likert scale

Question	I strongly agree	I agree	I do not know	I disagree	I strongly disagree
1	1	6	1	2	0
2	0	2	1	7	0
3	0	1	2	5	2
4	2	6	0	2	0
5	2	7	0	1	0
6	0	6	2	1	1
7	0	1	2	6	1
8	2	4	1	3	0

4 Conclusions and Future Directions

We presented a mixed reality environment for on-site servicing assistance targeted to mission critical systems. According to the testing conducted the main design requirements (within two millimeter co-registration error, context dependent augmentation strategy and natural interaction) have been overall satisfied. The subjective system evaluation, performed by testers in a radar system's training facility, highlights the potential of the proposed approach, though issues related to the hardware used (the reduced HMD's resolution/field-of-view, brief tracking failure under rapid head movement) might sometimes detract from the MR experience. The selective visual removal of possibly confusing components from the actual scene has proved to be an effective option under certain circumstances. According to questionnaire answers, the combination of augmentation and finger-based interface worked well, providing an intuitive interaction paradigm that proved to be suited to the application context. Overall, the MR aided servicing environment produced a valuable improvement in user's confidence during simulated interventions, which could eventually lead to a measurable reduction of time required to tasks completion. For those reasons we are planning and experiencing more formal statistical tests to better evaluate the strength and weak aspects of the proposed system. Tracking accuracy aside, the tests take into account other useful parameters such as the users's effort in terms of concentration, memorization of the maintenance procedures and load on the user's head. We are currently implementing a stereoscopic version of the whole architecture, to improve depth perception of real environment through binocular scene capture and stereo rendering of virtual contents. We are also experimenting infrared-based scene lighting and capture to improve markers extraction from the background and therefore to reduce both tracking errors and markers size. Finally, we are experimenting a predictive approach to head tracking to further improve augmentation accuracy under rapid movements.

References

1. Henderson, S., Feiner, S.: Evaluating the benefits of augmented reality for task localization in maintenance of an armored personnel carrier turret. In: International Symposium on Mixed and Augmented Reality (ISMAR 2009), pp. 135–144 (2009)
2. Henderson, S., Feiner, S.: Exploring the benefits of augmented reality documentation for maintenance and repair. IEEE Transactions on Visualization and Computer Graphics 17(10), 1355–1368 (2011)
3. Friedrich, W., Jahn, D., Schmidt, L.: ARVIKA-augmented reality for development, production and service. In: Proceedings of the IEEE/ACM International Symposium on Mixed and Augmented Reality (ISMAR 2002), pp. 3–4 (2002)
4. Kiyokawa, K., Billinghurst, M., Campbell, B., Woods, E.: An occlusion-capable optical see-through head mount display for supporting co-located collaboration. In: Proceedings of the 2nd IEEE/ACM International Symposium on Mixed and Augmented Reality, p. 133. IEEE (2003)

5. Zhiqiang, B.I.A.N., Ishii, H., Shimoda, H., Yoshikawa, H., Morishita, Y., Kanehira, Y., Izumi, M.: Development of a tracking method for augmented reality applied to NPP maintenance work and its experimental evaluation. IEICE Transactions on Information and Systems 90(6), 963–974 (2007)
6. Wagner, D., Langlotz, T., Schmalstieg, D.: Robust and unobtrusive marker tracking on mobile phones. In: 7th IEEE/ACM International Symposium on Mixed and Augmented Reality (ISMAR 2008), pp. 121–124. IEEE (2008)
7. Maidi, M., Preda, M.: Markerless tracking for mobile augmented reality. In: IEEE International Conference on Signal and Image Processing Applications (ICSIPA 2011), pp. 301–306. IEEE (2011)
8. Comport, A.I., Marchand, E., Pressigout, M., Chaumette, F.: Real-time markerless tracking for augmented reality: the virtual visual servoing framework. IEEE Transactions on Visualization and Computer Graphics 12(4), 615–628 (2006)
9. Fue, P., Lepetit, V.: Vision based 3D tracking and pose estimation for mixed reality, pp. 1–22. Interfaces and Design, Emerging Technologies of Augmented Reality (2005)
10. Wang, W., Qi, Y., Wang, Q.: An Augmented Reality Application Framework for Complex Equipment Collaborative Maintenance. In: Luo, Y. (ed.) CDVE 2011. LNCS, vol. 6874, pp. 154–161. Springer, Heidelberg (2011)
11. Wang, R.Y., Popović, J.: Real-time hand-tracking with a color glove. ACM Transactions on Graphics (TOG) 28(3), 63
12. Mistry, P., Maes, P.: SixthSense: a wearable gestural interface. In: ACM SIGGRAPH ASIA 2009 Sketches, p. 11. ACM (2009)
13. Jarusirisawad, S., Hosokawa, T., Saito, H.: Diminished reality using plane-sweep algorithm with weakly-calibrated cameras. Progress in Informatics 7, 11–20 (2010)
14. Enomoto, A., Saito, H.: Diminished Reality using Multiple Handheld Cameras. In: ACCV 2007 Workshop on Multi-Dimensional and Multi-View Image Processing, Tokyo (November 2007)
15. Herling, J., Broll, W.: Advanced self-contained object removal for realizing real-time Diminished Reality in unconstrained environments. In: 9th IEEE International Symposium on Mixed and Augmented Reality (ISMAR 2010), pp. 207–212. IEEE (2010)

Establishing Workload Manipulations Utilizing a Simulated Environment

Julian Abich IV, Lauren Reinerman-Jones, and Grant Taylor

University of Central Florida (UCF), Institute for Simulation & Training (IST),
Applied Cognition in Virtual Immersive Training Environments Laboratory (ACTIVE Lab)
{jabich,lreinerm,gtaylor}@ist.ucf.edu

Abstract. Research seeking to improve the measurement of workload requires the use of established task load manipulations to impose varying levels of demand on human operators**.** The present study sought to establish task load manipulations for research utilizing realistically complex task environments that elicit distinct levels of workload (i.e. low, medium, and high). A repeated measures design was used to test the effects of various demand manipulations on performance and subjective workload ratings using the NASA-Task Load Index (TLX) and Instantaneous Self-Assessment technique (ISA). This experiment successfully identified task demand manipulations that can be used to investigate operator workload within realistically complex environments. Results revealed that the event rate manipulations had the most consistent impact on performance and subjective workload ratings in both tasks, with each eliciting distinct levels of workload.

Keywords: Workload, simulated environments, complex systems, signal detection, change blindness.

1 Introduction

Despite over 50 years of research related to workload, there is not yet an agreed upon definition that captures the complexity of this construct in its entirety (Knowles, 1963; Taylor, 2012). Early concepts of workload focused on physical load (Meshkati, Hancock, Rahimi, & Dawes, 1995), but later emphasized the cognitive components of the construct. Subsequent efforts focused on defining workload, establishing a workload assessment procedure, and finding the relationship among measurements (Johannsen, 1979), but have yet to establish a single, comprehensive operational definition. A common theme across all proposed workload theories is a dynamic interaction between the task and the operator (Veltman & Gaillard, 1996) with the operator having a limited capacity of resources to allocate to the demands of a task (Wickens & Hollands, 2000). Some theories argue that resources come from a single pool of energy (Kahneman, 1973; Moray, 1967), while others argue that multiple pools or multiple levels exist (Wickens, 2002). These theories have been tested using a range of tasks and tasking environments with workload assessed by various measures. However, establishing distinct levels of workload for tasks have not been

R. Shumaker (Ed.): VAMR/HCII 2013, Part II, LNCS 8022, pp. 211–220, 2013.

identified to directly compare the universality or sensitivity of the various workload measures. To better understand the role of workload within complex task environments, similar demand manipulations that elicit distinct levels of workload have to be determined. The present study sought to develop these manipulations for two theoretically different tasks related to the military intelligence, surveillance, and reconnaissance (ISR) domain within a simulated unmanned vehicle environment for use in future experimentation.

ISR tasks focus on acquiring and processing information on which to base decisions. Two popularly ascribed theories that form the foundation for tasks within the ISR domain are signal detection theory (SDT) and change blindness (CB). SDT is the foundation for threat detection in ISR missions. SDT states that nearly all decisions are made in the presence of uncertainty (Green & Swets, 1966; Heeger, 1997). This uncertainty results from the presence of noise. Noise is either internal (referring to perceptual processing and/or neural activity), external (referring to environmental sources), or a combination of both. Changes in noise increases or decreases workload, which influences decisions made regarding detection of relevant information. Determining the decision-making criteria to reduce unwanted outcomes or experienced workload (See, Howe, Warm, & Dember, 1995) would be highly advantageous in high-risk environments such as impending threat.

Another task required in ISR missions is monitoring changes to operational area and this type of task is founded on theories of CB. CB refers to the observer's failure to notice change in a visual scene (Rensink, 1997), even if it is large and normally easily noticed (Simon & Rensink, 2005). The visual processing involved in first noticing a change requires detection (whether a change occurred), identification (what kind of change occurred) and localization (where the change occurred; Rensink, 2002). Some theories argue that the inability to detect a change might be due to the operator's limited attentional resources (Kahneman, 1973), limited processing of attended information (Rensink, Regan, & Clark, 1997), failure to compare pre- and post-representations of processed visual information (Hollingworth & Henderson, 2002), or that post-representations overwrite pre-representations (Rensink, Regan, & Clark, 1997). Attentional resources are strongly linked to workload level and therefore, varying levels of workload affects performance on CB tasks. This is particularly important to ISR missions because the inability to notice a change might lead to misdiagnosis of activity occurring in the environment, which could result in decreased safety.

Simulated environments are used to experimentally test the types of theoretically - based tasks described above within complex systems. These environments offer a host of advantages in comparison to real-world testing, such as reduced costs associated with developing, running, and maintaining these systems, consistency and control of variable manipulations, logging capabilities for real-time and post-hoc analysis, and increased safety for consequences resulting from operator error. Simulated environments provide the flexibility to test varying levels of task load manipulations that otherwise would not be possible within a real-world system, ideally resulting in recommendations for design of complex systems and improve operator assessment and training.

2 Experiment Overview

Using a simulated unmanned vehicle environment, the aim for the present study was to establish task load manipulations for two theoretically different tasks that induce distinct levels of workload measured by both subjective assessments and performance. The intention was to identify at least one type of manipulation for each task and will determined by the three different types of workload measures. Measures included post-task and online, and subjective and objective assessments. The established manipulations will be implemented in future studies with the goal of observing operator state using physiological measures with the purpose of identifying a universal and comprehensive measure that assesses workload across tasks, domains, and other workload measures.

3 METHODS

3.1 Participants

Fifty-six (34 males, 22 females) volunteers from several universities located in central Florida participated in the experiment with a mean age of 20.6 (*SD* = 3.4).

3.2 Materials

Questionnaires

NASA-Task Load Index (TLX). The TLX is a multi-dimensional questionnaire used to assess perceived workload (Hart & Staveland, 1988, Hart, 2006). It consists of six subscales of workload (mental demand, physical demand, temporal demand, effort, frustration, and perceived performance), each rated on a 100-point scale with five-point increments and 100 being high workload. The average score of the six-subscales provides a separate measure of global workload. The TLX was administered post-task using a customized computer program to automatically activate a visual prompt containing the questionnaire at multiple designated locations (the end of each trial block) throughout each scenario.

Instantaneous Self-Assessment (ISA). The ISA is a single measure used to assess immediate subjective workload during the performance of a task (Tattersall & Foord, 1996). The scale uses a five-point rating scale with 5 being high workload and was administered at multiple designated locations (at 75% of each trial block completion) using a customized computer program to automatically activate an audio prompt containing the phrase "please rate your workload." Participants responded verbally with their rating.

Apparatus

Simulation. The Mixed Initiative eXperimental (MIX) testbed (Reinerman-Jones, Barber, Lackey, & Nicholson, 2010) was utilized to simulate an operator control unit

(OCU) for an unmanned ground vehicle (UGV). The environment the UGV maneuvered through was a generic Middle Eastern town infiltrated by enemy threats. The UGV was fully autonomous and drove itself along a preplanned route while participants identified static enemy targets within the environment. The simulation was presented using a standard desktop computer (3.2GHz, Intel Core i7 processor) with a 22" (16:10 aspect ratio) monitor. Responses were collected using the left mouse button and verbal responses to the ISA measure were collected using a standard external desktop computer microphone.

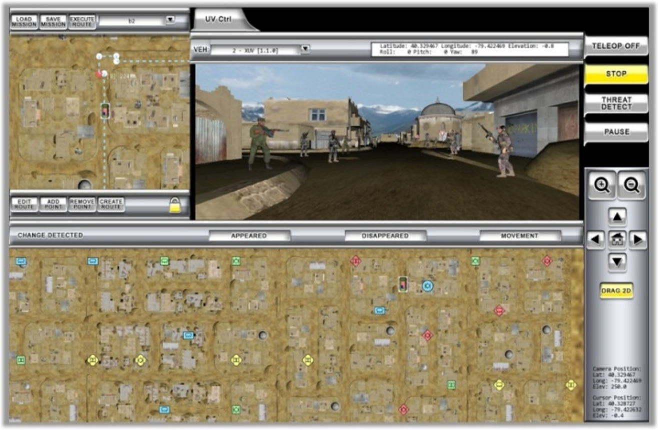


Fig. 1. Screenshot of the MIX testbed

Experimental Tasks. Participants performed two tasks both independently. One task was the threat detection task based on SDT and the other was the change detection task based on CB theory.

Threat Detection (TD) Task. Participants monitored a video feed of the forward perspective of the UGV while it traveled along a pre-planned route and reported any potential threats present in the environment. Four categories of people were present: friendly soldiers, friendly civilians, enemy soldiers, and insurgents (armed civilians). A threat was classified as enemy soldiers and/or insurgents. They were reported by using the computer mouse to left-click a "threat detect" button located along the top right of the OCU and then to left-click on the threat within the UGV video feed. Performance of was rated as percentage of targets correctly identified.

The TD task manipulated both event rate and signal/noise ratio (threat probability). These levels were derived from a meta-analysis of the sensitivity decrement in vigilance (See, Howe, Warm, & Dember, 1995; Table 1). Event rate and threat probability were combined to form five total conditions. Each event rate was presented with a medium threat probability and each threat probability was paired with the medium event rate.

Change Detection (CD) Task. Participants monitored an aerial map positioned at the bottom of the OCU that displayed the location of various entities. Entities were represented by icons borrowed from the DoD, but the associated meanings were not tied to other events. On average 24 icons were present and randomly displayed across the defined area. Icons exhibited three types of changes: appear, disappear and movement. Three change detection buttons labeled after each type of change were located above the aerial map. Icon changes identified were reported using the mouse to left-click on the appropriate change detection button. Performance was rated as percentage of changes correctly detected.

The CD task manipulated both event rate and signal saliency. These levels were derived from previous research (Tollner, 2006; Taylor, Reinerman-Jones, Cosenzo, & Nicholson, 2010; Table 1). Event rate and signal saliency were combined to form five total conditions. Each event rate was presented with a medium signal saliency and each signal saliency was paired with the medium event rate.

Table 1. Levels of manipulations for both tasks

Task Demand	Low	Medium	High
		Threat Detection	
Threat Probability	1/15	2/15	4/15
Event Rate	15/min	30/min	60/min
		Change Detection	
Signal Saliency	4 icons	2 icons	1 icon
Event Rate	6/min	12/min	24/min

3.3 Procedure

Participants were trained on the tasks followed by a brief practice session. Each participant completed two scenarios. The two scenarios consisted of the: CD task with the five conditions of workload and TD task with the five conditions of workload (Table 2). Scenario and workload condition order were counterbalanced for each participant. Each block was six minutes, totaling 30 minutes per scenario, with the entire experiment lasting roughly two hours.

Table 2. Example of full experiment run

Tasks	Change detection	Threat Detection
Manipulations	*Signal saliency/Event Rate ratios:* 2:6, 2:12, 2:24, 1:12, 4:12	*Threat ratios:* 2:30, 4:30, 8:30, 2:15, 8:60

4 Results

Evaluation of each manipulation on subjective and performance data was conducted using a series of repeated measure ANOVAs. Bonferroni and Greenhouse-Geisser corrections and pairwise deletions were applied where appropriate. Due to a logging error, the sample size for threat detection performance was reduced to 29.

4.1 Threat Detection

Results revealed that the event rate manipulation had a significant main effect on all dependent variables, Mental Demand, $F(1.562, 82.781) = 19.944$, $p < .001$, $\eta^2 = .273$, Physical Demand, $F(2, 106) = 6.354$, $p = .002$, $\eta^2 = .107$, Temporal Demand, $F(1.436, 76.118) = 18.484$, $p < .001$, $\eta^2 = .259$, Effort, $F(1.572, 83.305) = 16.147$, $p < .001$, $\eta^2 = .234$, Performance, $F(1.798, 95.307) = 13.512$, $p < .001$, $\eta^2 = .203$, Global, $F(1.639, 78.654) = 17.065$, $p < .001$, $\eta^2 = .262$, ISA, $F(2, 98) = 48.926$, $p < .001$, $\eta^2 = .500$, and percent correct detected, $F(1.549, 43.371) = 34.305$, $p < .001$, $\eta^2 = .551$, with the exception of Frustration, $F(2, 106) = 4.631$, $p = .012$, $\eta^2 = .080$ (Table 3). Means and standard deviations (in parentheses) for tables 3-6 are provided for each level of demand. Means designated with subscripts are significantly different from means with equivalent subscripts ($p < .0167$ in each case).

Table 3. Results for event rate manipulation on threat detection task

	Level of Demand		
Variables	Low	Medium	High
Mental Demand	18.056_a (20.50)	23.148_a (21.57)	30.648_a (24.32)
Physical Demand	8.80_a (14.37)	10.28 (15.06)	13.24_a (18.54)
Temporal Demand	15.00_a (17.88)	20.65_a (20.31)	27.59_a (25.31)
Effort	19.81_a (21.50)	24.63_b (23.59)	31.20_{ab} (25.23)
Performance	7.41_a (9.85)	9.26_b (9.24)	13.70_{ab} (13.07)
Global	27.13_a (11.91)	30.32_a (13.39)	34.27_a (15.34)
ISA	1.42_a (.54)	1.82_a (.60)	2.26_a (.75)
Percent Detect	58.27_a (16.73)	68.30_a (7.81)	80.94_a (6.63)

The threat probability manipulation had a significant main effect on Temporal Demand, $F(1.478, 78.34) = 6.83$, $p = .005$, $\eta^2 = .114$, Global, $F(1.651, 79.256) = 6.216$, $p = .005$, $\eta^2 = .115$, ISA, $F(1.567, 76.798) = 22.424$, $p < .001$, $\eta^2 = .314$, and percent correct detected, $F(1.660, 46.485) = 10.342$, $p < .001$, $\eta^2 = .270$, with the exception of Mental Demand, $F(1.704, 90.302) = 5.384$, $p = .009$, $\eta^2 = .092$, Effort, $F(2, 106) = 3.343$, $p = .039$, $\eta^2 = .059$, Physical Demand, $F(2, 106) = 2.152$, $p = .121$, $\eta^2 = .039$, Frustration $F(1.518, 80.45) = 1.103$, $p = .323$, $\eta^2 = .020$, and Performance, $F(2, 106) = 3.074$, $p = .05$, $\eta^2 = .055$ (Table 4).

Table 4. Results for threat probability manipulation on threat detection task

	Level of Demand		
Variables	Low	Medium	High
Temporal Demand	17.22_a (19.00)	20.65 (20.31)	24.44_a (23.77)
Global	28.15_a (12.66)	30.32 (13.39)	32.30_a (14.21)
ISA	1.52_a (.61)	1.82_a (.60)	2.12_a (.77)
Percent Detect	68.10_a (17.35)	68.30_b (7.81)	78.74_{ab} (8.32)

4.2 Change Detection

Results revealed that the event rate manipulation had a significant main effect on Mental Demand, $F(2, 98) = 18.460$, $p < .001$, $\eta^2 = .274$, Physical Demand, $F(2, 98) = 6.775$, $p = .002$, $\eta^2 = .121$, Temporal Demand, $F(1.754, 85.502) = 24.958$, $p < .001$, $\eta^2 = .337$, Effort, $F(1.627, 79.723) = 15.111$, $p < .001$, $\eta^2 = .236$, Frustration, $F(2, 98) = 10.227$, $p < .001$, $\eta^2 = .173$, Performance, $F(2, 98) = 8.849$, $p < .001$, $\eta^2 = .153$, Global, $F(1.732, 84.862) = 20.007$, $p < .001$, $\eta^2 = .290$, ISA, $F(2, 96) = 27.466$, $p < .001$, $\eta^2 = .364$, and percent correct detected, $F(2, 94) = 116.856$, $p < .001$, $\eta^2 = .713$ (Table 5).

The signal saliency manipulation only had a significant main effect on percent correct detected, $F(2, 94) = 86.472$, $p < .001$, $\eta^2 = .648$. Signal saliency did not have a significant main effect on Mental Demand, $F(1.476, 72.325) = 1.779$, $p = .184$, $\eta^2 = .035$, Physical Demand, $F(2, 98) = .826$, $p = .441$, $\eta^2 = .017$, Temporal Demand, $F(1.462, 71.638) = 2.189$, $p = .133$, $\eta^2 = .043$, Effort, $F(1.622, 79.472) = 2.114$, $p = .137$, $\eta^2 = .041$, Frustration, $F(1.536, 75.271) = .061$, $p = .899$, $\eta^2 = .001$, Performance, $F(2, 98) = .783$, $p = .460$, $\eta^2 = .016$, Global, $F(1.366, 66.910) = 1.602$, $p = .213$, $\eta^2 = .032$, and ISA, $F(1.532, 73.528) = 1.582$, $p = .215$, $\eta^2 = .032$ (Table 6).

Table 5. Results for event rate manipulation on change detection task

Variables	Level of Demand		
	Low	Medium	High
Mental Demand	47.90_a (26.75)	56.70_a (27.53)	63.40_a (25.30)
Physical Demand	19.00_a (17.87)	22.70 (21.93)	27.20_a (25.86)
Temporal Demand	41.00_a (27.31)	47.70_a (28.56)	58.60_a (26.97)
Effort	49.40_a (25.85)	53.60_b (26.03)	62.50_{ab} (25.10)
Frustration	31.20_a (25.73)	33.00_b (25.50)	43.30_{ab} (28.24)
Performance	34.50_a (20.08)	35.90_b (22.08)	43.20_{ab} (22.20)
Global	42.33_a (15.04)	46.30_a (16.11)	51.97_a (15.99)
ISA	2.47_a (.89)	2.77_a (.94)	3.39_a (.86)
Percent Detect	61.16_a (16.82)	59.41_b (15.74)	38.93_{ab} (11.13)

Table 6. Results for signal saliency manipulation on change detection task

Variables	Level of Demand		
	Low	Medium	High
Percent Detect	70.06_a (15.87)	59.41_a (15.74)	48.55_a (11.20)

5 Discussion

The goal for the present study was achieved. Event rate for both threat detection and change detection elicited distinct levels of workload as shown by the TLX, ISA, and performance. These conditions will be used to investigate physiological responses to distinct levels of workload for two different theoretically driven tasks. That will enable improved adaptive trainers for complex systems, enable direct human-robot implicit communication, and better objective workload assessments.

On the surface, threat probability revealed some distinction between low, medium, and high workload levels. Global TLX showed a trend for these distinct level differences, but not all were significant as indicated by post-hoc comparisons. The same was the case for performance. Furthermore, the TLX sub-scale of temporal demand indicated differences in workload for the low and high levels of signal probability. Thus, the driving factor of this manipulation appears to be time, which is likely due to the amount of time participants felt they had to click on the increased threats in the environment. ISA results did indicate distinct levels of low, medium, and high workload. The inconsistency between the measures indicates that further investigation into signal probability is needed. A future study should look at different probabilities of the signal to noise ratio with perhaps greater increments in the probabilities. Additionally, the type of event should be further investigated. The present study used enemy soldiers and insurgents as signals. It is possible that this manipulation would provide distinct levels of workload if objects like Improvised Explosive Devices (IEDs) were used.

Signal saliency for CD yielded inconsistent results across the three measures of workload. Thus, the manipulation of saliency is not as clearly understood as that of event rate and might not be a manipulation of workload at all, but perhaps drawing on different cognitive capacities. Alternatively, the popularly used subjective measures of workload might not be sensitive to saliency changes that actually do increase or decrease operator workload. Further research should investigate additional subjective workload measures for signal saliency.

The present study illustrates the importance of systematically investigating manipulation choices for experiments and careful consideration of the generalizability of a type of manipulation to various tasks driven by different theories and cognitive processing requirements. It is also important to understand the strengths and limitations of measures for a given phenomenon and not just blindly accept the most popularly used measures.

Acknowledgements. This work was supported by the US Army RDECOM (W91CRB08D00150068). The views and conclusions contained in this document are those of the authors and should not be interpreted as representing the official policies, either expressed or implied, of RDECOM or the US Government. The US Government is authorized to reproduce and distribute reprints for Government purposes notwithstanding any copyright notation hereon.

References

1. Green, D.M., Swets, J.A.: Signal detection theory and psychophysics. Wiley, New York (1966)
2. Hart, S.G.: Nasa-Task Load Index (NASA-TLX); 20 Years Later. Proceedings of the Human Factors and Ergonomics Society Annual Meeting 50(9), 904–908 (2006)
3. Hart, S.G., Staveland, L.E.: Development of NASA-TLX (Task Load Index): Results of empirical and theoretical research. In: Hancock, P.A., Meshkati, N. (eds.) Human Mental Workload, pp. 139–184. Elsevier Science Publishers, Holland (1988)
4. Heeger, D.: Signal Detection Theory. Standford University, Department of Psychology (1997)
5. Hollingsworth, A., Henderson, J.M.: Accurate visual memory for previously attended objects in natural scenes. Journal of Experimental Psychology: Human Perception and Performance, 113–136 (2002)
6. Johannsen, G.: Workload and workload measurement. In: Moray, N. (ed.) Mental Workload, pp. 3–12. Plenum Press, New York (1979)
7. Kahneman, D.: Attention and effort. Prentice Hall, Englewood Cliffs (1973)

8. Knowles, W.B.: Operator loading tasks. Human Factors 5(2), 155–161 (1963)
9. Meshkati, N., Hancock, P., Rahimi, M., Dawes, S.M.: Techniques in mental workload assessment. In: Wilson, J.R., Corlett, E.N. (eds.) Evaluation of Human Work: A Practical Ergonomics Methodology, 2nd edn., pp. 749–782. Taylor & Francis, Bristol (1995)
10. Moray, N.: Where is capacity limited? A Survey and A Model Acta Psychologica 27, 84–92 (1967)
11. Reinerman-Jones, L., Barber, D., Lackey, S., Nicholson, D.: Developing methods for utilizing physiological measures. In: Applied Human Factors and Ergonomics, Miami, FL (2010)
12. Rensink, R.A.: How much of a scene is seen? The role of attention in scene perception. In: Investigative Ophthalmology and Visual Science, vol. 38 (1997)
13. Rensink, R.A.: Change detection. Annual Review of Psychology 53, 245–277 (2002)
14. See, J.E., Howe, S.R., Warm, J.S., Dember, W.N.: Meta-analysis of the sensitivity decrement in vigilance. Psychological Bulletin 117(2), 230–249 (1995)
15. Simons, D.J., Rensink, R.A.: Change blindness: Past, present, and future. Trends in Cognitive Sciences 9(1), 16–20 (2005)
16. Rensink, R.A., Regan, J.K.O., Clark, J.J.: To see or not to see: The need for attention to perceive changes in scenes. *Psychological Science 8(5), 368–373 (1997)*
17. Taylor, G.S.: Comparing types of adaptive automation within a multi-tasking environment. University of Central Florida, Orlando (2012) (Dissertation)
18. Tattersall, A.J., Foord, P.S.: An experimental evaluation of instantaneous self-assessment as a measure of workload. Ergonomics 39(5), 740–748 (1996)
19. Tollner, A.M.: Individual and Team Susceptibility to Change Blindness. University of Cincinnati, Cincinnati (2006) (Dissertation)
20. Veltman, J.A., Gaillard, A.W.K.: Physiological indices of workload in a simulated flight task. Biological Psychology 42, 323–342 (1996)
21. Wickens, C.D., Hollands, J.G.: Engineering psy-chology and human performance, 3rd edn. Prentice Hall, Upper Saddle River (2000)
22. Wickens, C.: Multiple resources and performance prediction. Theoretical Issues in Ergonomic Science 3(2), 159–177 (2002)

Interactive Virtual Reality Shopping and the Impact in Luxury Brands

Samar Altarteer, Vassilis Charissis, David Harrison, and Warren Chan

Glasgow Caledonian University
School of Engineering and Computing,
Division of Computing and Creative Technologies,
Glasgow, UK
samaraltarteer@gmail.com

Abstract. This paper investigates the impact of human-computer interaction in virtual reality online shopping interface on the consumer experience. In particular, it measures the effectiveness of visualising a three dimensional photorealistic item, the real-time interactivity with the product and the real-time, fully interactive product customization service. The proposed VR system employs a sophisticated approach of interaction with primary objective to simplify and improve the user experience during online shopping. The proposed interface was evaluated through a preliminary questionnaire designed to simulate the typical decision making process prior to a luxury object purchase. The paper presents the outcomes of this usability trial on a group of ten luxury brands customers, the challenges involved in the HCI design are discussed, the visual components of the interface are presented in addition to an analysis of the system evaluation. Adhering to the derived feedback, our future plan of work entails additional development of the interactive tools with a view to further enhance the system usability and user experience. Furthermore we aim to introduce more object choices and customisation covering a larger group of luxury brands.

Keywords: Virtual Reality, HCI, 3D Visualization, Luxury Marketing, Luxury Brands.

1 Introduction

In the twenty-first century, technology has become an essential part in people's daily routine contributing to the quality of life and well-being of modern societies. E-commerce, which is a relatively new form of commerce, utilises the online obtainable technology and offers the opportunity for customers to have a round the clock access to selected services. The capability of being available everywhere and anytime for customers is achieving impressive results in increasing sales profits. However, for luxury brands, a particular challenge for e-shopping stations is preserving their

R. Shumaker (Ed.): VAMR/HCII 2013, Part II, LNCS 8022, pp. 221–230, 2013.

prestige amid their faithful customers, whilst simultaneously providing competitive services. Accomplishing individual desires is a major interest of luxury houses, therefore they provide their clients with made to order service to allow integrating sides of the customer personality within the end product for ultimate satisfaction. Only a small number of luxury companies has already launched online product customisation, however, the presentation and extent of available options for visualisation and customisation are not quite there yet. Along with the aforementioned facts, the visual display units have improved greatly in recent years. The three-dimensional visualization feature elevated the display monitors up to an advanced stage. The technology used in these monitors is constantly improving and the accompanied image-translating tool is progressing remarkably.

To this end, the paper introduces a novel HCI e-purchasing interface for luxury brands with embedded 3D Virtual Reality (VR) real-time interactive product visualisation and customisation service. Furthermore, the system is designed to be presented in 3D enabled monitors to utilise the huge capacity in drawing advance technological experience for the customer. In order to primarily assist the usability of the trial version of the system, 10 users whom represent the expected customers have evaluated the interface. Qualitative and quantitative feedback has been acquired, in the form of questionnaires, video recordings and eye-tracking (and interaction detector) methods. The analysis of the data offered promising results and insight to potential improvements required.

Overall, this paper discusses the implementation of a VR (HCI) interactive interface for luxury brands e-commerce for 3D monitor presentation. The paper investigates the current luxury online trading trends and the expected future orientation. The following section describes the system rational and outlines the issues. Section three describes the system design, the hardware and the software requirements. The fourth section evaluates the system and analyses the result of the trial.

Finally, the evolution of the system design as a result of ongoing evaluation and user trials is discussed with suggestions for a tentative plan for future research work.

2 Luxury Marketing Trends

Purchasing luxury products intentions and motivations have various social, economic and psychological destinations. Brand knowledge [7], purchasing convenience, products availability, purchasing timing and place [16], sales interaction by sale person and point of purchase stimuli [17] are some of the factors that affect decision making in buying products. However, beside these facts, although luxury brands' consumers share similar motivations, they seek to differentiate themselves from others and like to be treated uniquely. These are exactly the concerns of high-end brands marketing, to be always accessible but exclusive [9]. The demand of the online distance shopping has become essential for those people who look for convenience, time saving, limited social interaction, and an extended time to compare and contrast products [8]. In 2010, e Marketer predicted that the digital retailing in the US will reach 223.90 billion Dollars revenue in 2014, which is about 0.7 times higher than that in year 2010 [3]. In international level, a study carried by Bain and Company in 2011 compared the e-shopping average to the total retail in six countries for the period between 2005 and 2011. The percentage have been increasing proportionally, while the United kingdom has witnessed dramatic increase to reach almost 9% of the retail percentage, which is 3 times higher than most of the other countries for the same period [2].

Even closely the luxury brand online market has also witnessed spectacular expansion and the targeted audience in particular have presented heavily in online environment. The luxury items internet shopping registered 9.5 million in 2008 and the expected increase was 27+ per cent growth annually. With 6.6 billion Internet users, 0.6 were luxury purchasers [11]. Relatively, figures show that up to 90% of affluent consumers frequently buy products online [9], and they spent 9% more in 2012 than they did in the year before [2].

This adoption also benefits the brands themselves since it opens a wider gate for loyal customers around the world to purchase with considerably less cost and effort. Luxury brands create strong relationship with customers through commitment and trust, however in electronic shopping by nature; these two factors are less easily maintained and might lead to uncertainty [13].

3 The Impact of HCI in Online Luxury Trading

A website has a non-physical presence, however a functional and uncluttered web application that provides useable and clear information is deemed essential if used to represent a company that would traditionally base its customer retention on excellent customer service and interpersonal relationships [13]. Furthermore user-friendly interface design, easy navigation methods and guidance as well as attractiveness are important website features that enhance online shopping experience and affect consumer shopping motivations [14]. Sejin and Stoel [15] identified four main factors affecting the perception of the e-shopping site apparel quality: website content, functionality and atmospheric/experiential value, which have the greater stimulus on shopping satisfaction and contribute in electronic shopping intention, in addition to the security, privacy and customer service, which affect the intentions only. Okonkwo outlined some of the main factors that had impact on the online luxury brand customer in particular. The Internet availability and the easiness of access, digital technology and interactive media as well as the social web influence facilitates much closer relationship with brands and the brands' fans [12]. A succinct description of the aforementioned factors is presented in figure 1.

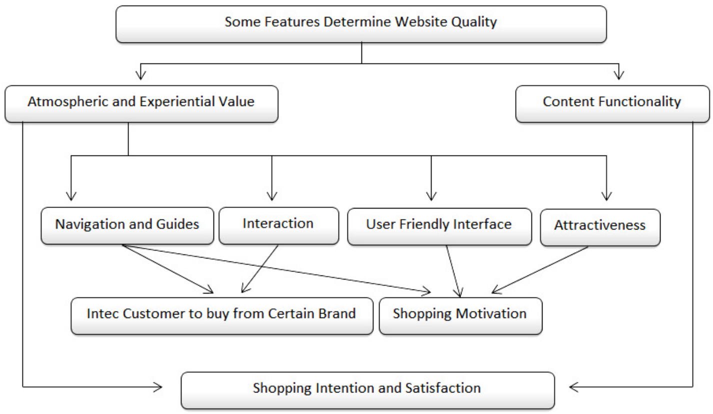


Fig. 1. Main Factors Affect Shopping Decisions in Virtual Shops

4 System Rationale

To this end, the online presence and services were readily attainable for such companies; however, the main challenge is serving the luxury concept and providing special services. As such it became essential and imperative for high-end brands to adopt new approaches using up to date technologies to engage the consumers and increase the attractiveness of their e-commerce sites aiming to ultimately create a large loyal group of customers. It is evident that firms should deliver exceptional experience and services in addition to unique products as the customers expect the virtual experience to be as immersive and simulating as a physical one [12]. Proceeding from the purchasing decision motivations and products value estimation considerations, the customers need to saturate their demands and desires to get the best of their journey in the web shopping. Using 3D visualization is estimated to be an effective way that enhances the characteristics and the features of products visualization in the consumers' conception. Previous studies proved that high level of fidelity 3D virtual model improves audience reaction towards e-purchasing [10]. Furthermore it helps in the imagination of the products as well as decreases the expected risk [11].

Algharabat and Dennis [1] classified the effect of 3D products authenticity on value opinion; on utilitarian value and on hedonic value. Enhancing the consumer awareness towards products' features, attributes and characteristics [4,10,18] and improving the customer involvement in seeking and gathering extra information about the product [4] all enhance the utilitarian value of 3D models.

On the other hand, it is claimed [5,11] that the gained hedonic value in 3D product visualization is superior to the capability of utilitarian value production. The higher interactivity the consumer gets with the web graphics, the higher the expected hedonic value perceived [4]. The main attractions for the customer towards the 3D products are entertainment, enjoinment and fun [10]. In comparison with multiple-pictures and video-clips interfaces, 3D VR interface produce greater Telepresence for consumers [18]. High level of interaction, possible dialog and communication, collaborative features such as customization and co-creation are key things to engrave the company name on the most favorite list [12]. Okonkwo suggested that offering services like personalization or bespoke features would enhance the client self-expression. In addition to the spoiling treatment the consumer gets from the luxury brands, they seek to have made-to-measure and customised items specially for the maturity consumptive clients [12]. 34% of the luxury customers believe that products and services should be customized to their needs and desires [6].

5 System Design

In view of the aforementioned facts and observations, the paper proposed HCI for online luxury brand through fully immersive VR customer service points that have been designed for 3D monitor representation. The intention behind the system is to enable the potential user to obtain more comprehensible and clear information through real-time 3D investigation of the photo-realistic virtual models. The 3D content has been generated in Autodesk 3Ds Max software. Adobe Photoshop has been used for editing and refining HD photos of actual products material to be used in shaders and normal maps to obtain high detailed realistic textures. Consequently, the scene has been exported to Unity 3D to create the real time virtual environment as illustrated in Figure 4. The real-time interface functions were built with C# programming language.

Evidently customers' adoption of shopping luxury brands online is mostly acceptable as these names have a global well-known reputation and trustworthiness which is an influential factor on online-purchasing intention. The Louis Vuitton brand has been chosen for our prototype interface design as they are highly associated with luxury, have a great historical presence within the fashion world and in addition are well known for their innovative techniques used to advertise their products. The precise and photorealistic visualisation of a specific group of objects (i.e. travel bags) from only one brand offered a more controllable environment for context development. Obviously the system uses the specific brand as a sample case. The main objective of this virtual reality interface is to identify the efficiency of such system and highlight the positive or negative impact it might have in the decision making process of the potential luxury brands customers.

The customization menu is placed in the right hand side of the screen, where the user can choose between 3 types of material and 3 vertical hierarchy choices to customise the front compartment of a travelling bag as depicted in figure 2 below. The interaction with the product can be made either by the control panel on the left bottom of the screen or directly with item by click and drag the mouse to rotate and zoom in and out of the product as presented in figure 3.

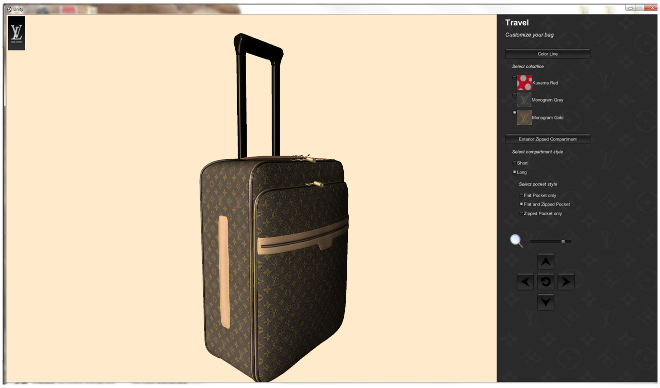


Fig. 2. Screenshot of the prototype VR interface

The system was evaluated in the Virtual Reality and Simulation laboratory (VRS lab) of Glasgow Caledonian University (Glasgow, UK). The particular testing environment offers state of the art equipment such as HD/3D projector, 3DTV, and surround audio, all supported by custom-built workstations which entail NVidia Quadro 4000 Fermi graphics cards. A wireless Bluetooth mouse was used in this case as the mean of interaction with the Virtual Reality environment as it was deemed the simplest and most cost effective way.

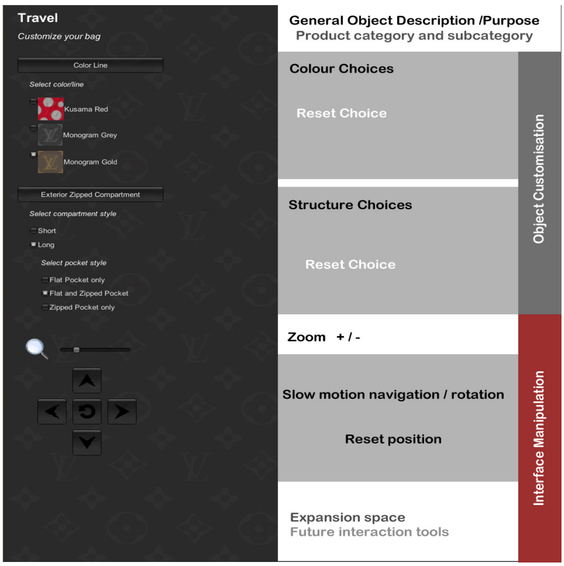


Fig. 3. Analysis of proposed VR interface components

6 Evaluation

In order to evaluate the plot system, 10 users who represent the current customer of the online luxury shopping have tested the system and answered a pre and post-test questionnaire. The aim of the questionnaire was three fold. Firstly, to identify the current actual shopping tendencies of the local population with regards to luxury brand products; Secondly to measure the efficiency of the proposed interface through customer satisfaction; Lastly, to receive further feedback regarding the interface design and functionality. As such the questionnaire had three distinctive set of questions. In this paper we will focus mainly on the second set of questions related to the interface performance and usability. The participants were asked to build a bag of their preference in all three colour schemes available. As such the users had the opportunity to investigate the VR environment through a task, imitating a typical customisation online process of their favourite products.

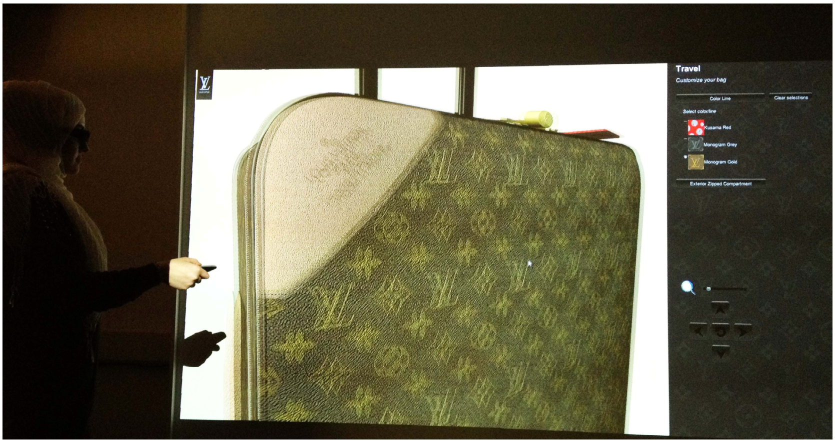


Fig. 4. Screenshot of the user during the trial session

7 Results and Discussion

The results of the second group of questions of the adapted QUIS (Questionnaire for User Interface) presented a positive initial appraisal of the system. The questions were marked in scale of 1 to 5, with 1 representing "*Not satisfied*" and 5 "*Extremely satisfied*". The main set of questions and their scores are presented below in figure 5.

Notably the system scored particularly well in all seven interface related questions with the lowest average of 78% for Q2 and highest average of 96% for Q6. The latter is in contrast to the typical performance of the majority of online luxury shopping experience which is usually hindered by overloaded graphics' context and elaborate animated introductions. As such the users in this trial appreciated the speed and the simplicity of the information provided. The lowest score was mainly due to the conduit of interaction which was a Bluetooth wireless mouse. The specific device although offers freedom of hand motion it is very difficult to point out and activate menus due to the motion sensitivity. As such we currently investigate alternative methods for navigating freely in the virtual environment, with the use of Kinect tracking device.

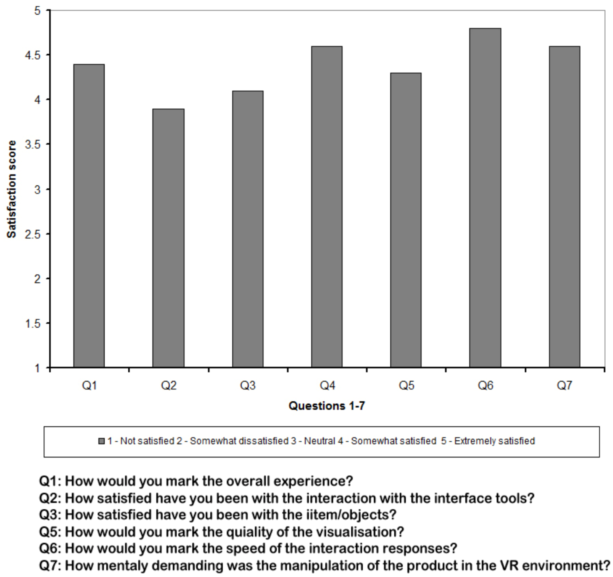


Fig. 5. Graph of User Experience satisfaction score with the use of VR shopping environment

Further thematic analysis of the open questions elucidated a number of intriguing themes useful for further development. In particular, 60% of the users responded that they would prefer to do online shopping for luxury brands. The main reasons described were split equally between better prices and more options available.

Almost 80% of the participants responded that the would like to see this system applied in other products even non-luxury ones.

The users also responded unanimously that they would prefer to do online shopping in a website that offers this 3D real-time visualisation and customisation experience as it offers significantly better quality of visual information regarding the product. The same positive response applied to four questions investigating the proposed system acceptability and user future preferences. As such the users were in favour of the system over traditional 2D image depictions, animations and videos. Furthermore they stated that such application would increase their shopping motivation of high-end brands as the real-time 3D visualisation and interaction elucidates clearly all the aesthetic and functional aspects of the product that they would like to buy. This is not possible currently with the use of the traditional online illustrations and for this reason the majority prefer to identify the product online, visit the store to experience the real-product and return to online shopping to complete the transaction which is usually of better price. Our proposed system circumvents this intermediate level and offers a photorealistic experience to the user.

Notably the system lacks any tactile interaction or appraisal capacity of object's weight. This drawback is inevitably the only weak point which cannot be simulated in a cost efficient manner. As we opt for off-the-shelf 3DTV hardware in order to improve the cost-efficiency of the system, haptic devices (i.e. Phantom arm, or cyber-gloves) are prohibitive. In contrary, the rapid market distribution of large scale 3D TVs is providing an excellent opportunity for such online applications which could visualise photo-realistically and in scale 1:1 the vast majority of the luxury brand objects (i.e. bags, shoes, clothing, watches and other accessories).

8 Conclusions

Overall this paper presented the design considerations and development process of a novel VR interactive environment designed particularly for luxury brands shopping. This novel application aims to improve the speed and volume of online shopping through clear, photorealistic and fully interactive object 3D models which will enrich significantly the user experience. The clarity of visual information aims furthermore to enhance the trust towards the online information which are typically scrutinised by the online shoppers during their purchases of high cost luxury brand products. By improving the above, our hypothesis is, that we can further improve the decision making process of the user which will conclude with a successful purchase. As such we evaluated the proposed VR system and received feedback from ten users.

A concise but enlightening appraisal of the proposed VR shopping environment and advanced user-interface was provided throughout this trial. The paper presents the qualitative information, explores the empirical evidence regarding 3D visualisation and analysed the user suggestions and future tendencies regarding such type of interaction. The encouraging feedback and the suggestions derived will further be transformed to interface solutions in order to accommodate a more flexible and functional interface. Our future plans entail a number of additional yet simplified interaction tools, that will increase significantly user's purchasing confidence and will enable the user to investigate extensively the potential product prior to purchase.

References

1. Algharabat, R., Dennis, C.: 3D Product Authenticity Model for Online Retail: An Invariance Analysis. Journal of Customer Behaviour 9(2), 97–155 (2010)
2. Bain and Company (2012). Online Retail Sales Growth, by Country and by Category, Front Of Store, (December 27, 2011), `http://frontofstore.org/2011/12/27/online-retail-sales-growth-by-country-and-by-category/`
3. Marketer, E.: Retail E-Commerce Resumes Double-Digit Growth (April 1, 2010), `http://www.emarketer.com/Article.aspx?R=1007605`

4. Fiore, A.M., Jihyun, K., Hyun-Hwa, L.: Effect of Image Interactivity Technology on Consumer responses toward the online retailer. Journal of Interactive Marketing 19(3), 38–53 (2005a)
5. Fiore, A.M., Jin, H.-J., Kim, J.: For Fun and Profit: Hedonic Value from Image Interactivity and responses toward an online store. Psychology & Marketing, zurseasons (2005b)
6. Fourseasons: The Luxury Consumer in the New Digital World: Then and Now (2012), http://www.fourseasons.com/content/dam/fourseasons/web/pdfs/landing_page_pdfs/2012_TRD_Report_final.pdf
7. Guizani, H., Triguero, H., Valette-Florence, P.: Development of French consumer brand equity scale. Advances Cons. Res. 2, 198–199 (2008); Guizani, H., Merunka, D., Valette, F.: The impact of brand personality and sales promotions on brand equity. Journal of Business Research 64, 24–28 (2011)
8. Hassouneh, D., Brengman, M.: Shopping in Virtual Worlds: Perceptions, Motivations, and Behaviour. Journal of Electronic Commerce Research 12(4) (2011)
9. Ikeler, M., Jackson, P.: Digital Strategies for Luxury Brands: Insights, Examples, and Implications (March 01, 2010), http://www.slideshare.net/marciikeler/digital-strategies-for-luxury-brands
10. Kim, J., Forsythe, S.: Hedonic Usage of Product Virtualization Technology in Online Apparel Shopping. International Journal of Retail and Distribution Management 35(6), 502–514 (2007)
11. Kim, J., Forsythe, S.: Hedonic Usage of Product Virtualization Technology in Online Apparel Shopping. International Journal of Retail and Distribution Management 35(6), 502–514 (2007)
12. Okonkwo, U.: Luxury Online: Styles, Systems, Strategies. Palgrave Macmillan, London (2010)
13. Park, C., Kim, Y.: Identifying Key Factors Affecting Consumer Purchase Behaviour in an Online Shopping Context. International Journal of Retail & Distribution Management 31(1), 16–29 (2003)
14. Roy, M., Dewit, O., Aubert, B.: The Impact of Interface Usability on Trust in Web Retailers. Internet Research 11(5), 388–398 (2001)
15. Sejin, H., Stoel, L.: Online Apparel Retailing: Roles of e-Shopping Quality and Experiential e-Shopping Motives. Journal of Service Management 23(2), 197–215 (2012)
16. Singhavolakul, N., Lee, T.: Identifying types of stores in commercial zones and their marketing strategies. Int. J. Enterprise Network Management 4(3) (2011)
17. Solomon, M.R.: Consumer Behaviour: Buying, Having, and Being, pp. 291–322. Pearson Education Inc. (2004); Singhavolakul, N., Lee, T.: Identifying types of stores in commercial zones and their marketing strategies. Int. J. Enterprise Network Management 4(3) (2011)
18. Suh, K.-S., Chang, S.: User Interfaces and Consumer Perceptions of Online Stores: the Role of Telepresence. Behav. Inform. Tech. 25(2), 99–113 (2006)

Multiple Remotely Piloted Aircraft Control: Visualization and Control of Future Path

Gloria Calhoun[1], Heath Ruff[2], Chad Breeden[1], Joshua Hamell[3], Mark Draper[1], and Christopher Miller[3]

[1] Air Force Research Laboratory (AFRL), Wright-Patterson Air Force Base, Ohio
[2] Ball Aerospace and Technologies Corporation, Fairborn, Ohio
[3] Smart Information Flow Technologies, Minneapolis, MN
{gloria.calhoun,heath.ruff,chad.breeden, mark.draper}@wpafb.af.mil, {jhamell,cmiller}@sift.net

Abstract. Advances in automation technology are leading to development of operational concepts in which a single pilot is responsible for multiple remotely piloted aircraft (RPAs). This requires design and evaluation of pilot-RPA interfaces that support these new supervisory control requirements. This paper focuses on a method by which an RPA's near-term future flight path can be visualized and commanded using the stick and throttle. The design decisions driving its symbology and implementation are described as well as preliminary quantitative data and pilot feedback to date.

Keywords: remotely piloted aircraft, unmanned air systems, flight path, display symbology, RPA, UAS, flexible automation.

1 Introduction

Advances in automation technology are leading to development of operational concepts in which a single pilot is responsible for multiple remotely piloted aircraft (RPAs). With RPA flight highly automated in this vision, multi-RPA systems will necessarily involve supervisory control with requirements for the pilot to frequently shift attention between RPAs. Displays that facilitate rapid retrieval of each RPA's state and associated tasking are required. Moreover, new control methods will be necessary to enable the pilot, when time and attention are available, to quickly and precisely redirect any RPA's path and tailor the supporting automation.

To tackle this design challenge, the Air Force Research Laboratory "Flexible Levels of Execution – Interface Technologies" (FLEX-IT) effort developed a demonstration simulation illustrating four different RPA control modes: manual (pilot controls the RPA's flight with stick and throttle), noodle (enabling the pilot to visualize and command the RPA's future path), micro-plays (quick maneuvers initiated by verbal command), and plays (complex tasks initiated by the pilot, for

R. Shumaker (Ed.): VAMR/HCII 2013, Part II, LNCS 8022, pp. 231–240, 2013.

example, command to monitor a specified target). The simulation (Fig. 1) also illustrates the symbology and controls to support seamless transition between any and all of the four control methods. Early and refined versions of the simulation are described elsewhere [1, 2, 3].

The FLEX-IT multi-RPA simulation was used to support operator-centered efforts to acquire feedback from operators, many of them USAF pilots with experience flying RPAs. The methodology and RPA operator feedback have been reported [3, 4]. Pilots indicated that a FLEX-IT approach is indeed promising in terms of providing intuitive multi-level control methods to interact with automation. A key to the utility of FLEX-IT is the symbology that provides feedback on the automation's processing, proposed plans, and state of execution.

Fig. 1. Flexible Levels of Execution – Interface Technologies (FLEX-IT) Simulation

Of the four control modes integrated, one was a mode enabling rapid command of a vehicle's immediate future path using the same flight controls (stick and throttle) as for normal flight. It also provided visualization that made use of the vehicle's flight model to show the result of these actions on the future path of the vehicle in the form of a flexible 'noodle' that extended from the front of the aircraft symbol and reacted, bending and shifting, as if the pilot were currently flying the commanded path. The noodle was particularly novel and viewed favorably by the pilots. In one evaluation, all six pilot participants rated this feature as a likely aid for future control of multiple RPAs. Specifically, they selected one of the top two favorable ratings on a 5-point scale [3]. Additionally, the pilots' comments were aligned with the ratings. Some described the noodle concept as "very intuitive" and providing a good quick visual representation of the vehicle's future flight path [3]. They indicated that the ability to input the future path of the vehicle with automated tools via the stick and throttle manipulations would be very useful, at least for the existing pilot community. For example, the tool provides a quick means of rerouting an RPA to avoid an area that suddenly became restricted. With the noodle, a complex maneuver can be quickly defined and commanded to the RPA (as constrained by the vehicle's real flight capabilities) ahead of time, allowing the pilot to turn attention to another vehicle/task while the first RPA flies its noodle directed path.

The remainder of this paper will focus on this novel control mode. First the approach will be described in more detail. This will include examples of alternative designs entertained in the process of developing this automated tool. Next, the results of a study designed to collect data on how long it takes to employ this control method to designate a vehicle's future path will be presented. This quantitative data supplements the qualitative data reported earlier [3, 4].

2 RPA Candidate "Noodle" Control Mode

This control mode is termed "noodle" as its symbology includes a flexible line segment resembling a variable length bendable noodle emerging from the nose of an RPA symbol on the map display. Two noodle symbology sets were implemented and demonstrated. With the *Predicted Noodle,* a line (colored to match the respective RPA's symbol) is drawn to show the forecasted flight path the RPA will fly, given the current state of control inputs is maintained. By employing faster-than-real-time software simulations, the RPA's position is predicted as a series of points in the future based on current autopilot or manual joystick control. These points are then connected into a line that shows the expected flight path for the specified duration. Several alphanumeric readouts are presented along the noodle. The Predicted Noodle example shown in Figure 2 illustrates that at the end of the purple line, the RPA will have traveled for 119 seconds and be at an airspeed of 86 KIAS, an altitude of 14,000 feet MSL, and be oriented at a heading of 275 degrees. This prediction is perfect in the simulation, since the same model is used to predict aircraft path as is used to actually 'fly' it. In a real world deployment, winds and other factors will introduce error and will need to be compensated for by estimates and approximations. In the majority of control modes, the presentation of the Predicted Noodle is under pilot control in the simulation: either displayed or turned off.

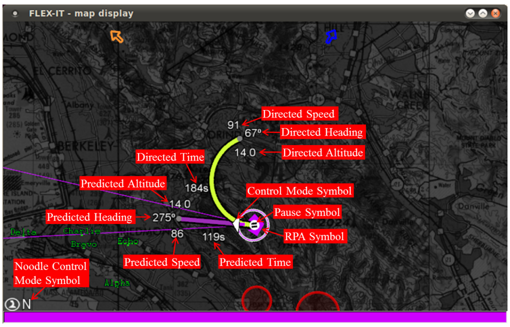


Fig. 2. Illustration of the Noodle Symbology in the FLEX-IT Simulation. (Purple line is Predicted Noodle, yellow curve is Directed Noodle. Note that red flags are annotations for this graphic and are not presented during use.)

In contrast to the Predicted Noodle that solely displays future path given current state, the *Directed Noodle* mode serves as both a display and control, enabling the pilot to specify the future path of a vehicle. For this mode, the functionality of the stick and throttle are remapped. Inputs on the right console control stick specify the RPA's destination change in heading and altitude at the end of the noodle. One throttle button controls noodle duration, while throttle movement controls desired airspeed. Exact noodle length is derived from noodle duration and airspeed setting. Once the pilot approves the setting (via a switch on the throttle), the RPA follows the noodle path.

When invoked, the Directed Noodle is visualized with a yellow “directed” path shown in addition to the Predicted Noodle (see Fig. 2). The yellow Directed Noodle is manipulated by the stick and throttle as a possible future path of the aircraft, but it is not carried-out unless the operator pushes a switch on the throttle to confirm the path. At this point, the Predicted Noodle reads that the control commands for the RPA have changed, and re-calculates accordingly, effectively taking the place of the previous Directed Noodle segment. This provides a very natural method, using the familiar flight controls, to input a future flight path that (because it is computed and shown using the same flight control algorithms used to control the aircraft itself) obeys all known constraints and capabilities of the RPA in its current state.

Figure 2 shows the effect of switching to noodle control after an automated play was enacted to monitor a target. As soon as the operator executes the noodle control mode, a pause symbol (two parallel lines) appears on the RPA icon to let the operator know that the play has been interrupted. The letter “N” next to the joystick symbol (lower left) shows that the joystick is in noodle control mode and the arrow icon on the RPA symbol shows that the aircraft is in a semi-autonomous state of control (in between auto-pilot/waypoint following mode and manual control). The purple-colored Predicted Noodle indicates the vehicle’s current flight path. In Figure 2, the Directed Noodle is being maneuvered to reach an altitude of 14,000 feet, a targeted maneuver speed of 91 KIAS, and a heading of 67 degrees after 184 seconds of flight time.

A new Directed Noodle can be added at the end of the accepted noodle, and manipulated for additional flight path specification. Thus, multiple Directed Noodles legs can be “chained” together to command complex near-term flight maneuvers. Future segments can also be overridden by pulling back on a throttle switch. Each backward switch employment moves the currently manipulated Directed Noodle back one segment in the path. This allows for the pilot to redefine undesired noodle segments without discarding the entire sequence.

The intuitive use of manual inputs by the pilot, along with supporting automation, provides a means of quickly specifying a vehicle’s future path with more precise control/granularity without employing detailed menus/procedures associated with complex route planning software. More importantly, this new method of flight path control may enable a pilot to devote more dedicated attention to other supervised RPAs.

3 Design of the RPA Noodle Control Mode

The noodle control mode described above represents the results of a series of design and evaluation cycles, each considering an alternative approach for presenting required information and supporting control inputs. Each cycle involved obtaining the expert opinions of RPA and manned aircraft pilots, while the methodology employed varied (ranging from PowerPoint illustrations to actual instantiations in simulation). Some design questions included: What should the color, style, and width of the line be to ensure visibility and yet minimize clutter on the map? What other information is needed to describe the future flight path? How should this information be presented (e.g., labels on the noodle) to minimize clutter and be visible in relation to the map? How should the pilot initiate the noodle control -- what input should control the functionality of the stick and throttle? How best can the operator be informed of the current mode of these controllers and the state of the path under noodle construction?

One question that received considerable attention was how best to indicate altitude variations that reflect the pilot's inputs during the construction of the Directed Noodle. One limiting factor in providing this feedback is that the map display is two-dimensional. Several methods of presenting altitude information directly on the map were considered (see Fig. 3). The first option, adding plus and minus readouts at the end of the noodle was determined to be inadequate in terms of saliency (Fig. 3a). However, symbology was needed to depict climb/descent for each noodle segment/chain, so "+" or "-" symbols were added in-line to reflect the individual segment altitude change (Fig. 3b). Symbology currently used by USAF pilots to denote climbing and descending turns was also considered (Fig. 3c) – both along the line (Fig. 3d), or in-line with the noodle (Fig. 3e). However, it would be difficult to depict this symbology for tight turns, so one design idea was to widen the noodle so as to completely surround the climb/descent symbology (Fig. 3f). Several coding methods involving the noodle line itself were considered, such as depicting the line segment in different shades (Fig. 3g) or colors (Fig. 3h). These methods were fairly visible on the map without adding clutter and helped portray whether the altitude was increasing or decreasing across the corresponding segment. However, color is a key coding method to differentiate symbology associated with one RPA versus another. All the aforementioned approaches were similar in that they only indicate whether the altitude increased or decreased in respect to some threshold. This binary information fails to provide an indication of the magnitude of the change.

A method commonly used to depict altitude changes on maps was next considered. Contour maps use circles to depict specific altitudes (e.g., every 100 ft of altitude change), with the proximity of each circle to other circles an indication of how rapidly altitude is changing. This convention can be applied to symbology overlaid on the noodle line, with the proximity of the symbols indicative of the severity or degree of altitude change. However, increases in altitude still need to be distinguished from decreases. Thus, it was decided to use two symbol elements to denote this difference. A caret symbol (^; similar to the notation suggested by pilots; Fig. 3c) is used in the current simulation to denote a vehicle's climb in altitude (Fig. 3i). A straight line perpendicularly intersecting the noodle represents a descent (Fig. 3i). These two symbols are also easier to code and draw on the noodle, although the width of the noodle had to be increased slightly to increase the visibility of the symbology. Additionally, explicit altitude readings are displayed at each segment barrier.

Figure 3i illustrates how both the direction and magnitude of altitude change are concurrently presented. The effect of chaining multiple Directed Noodles into an extended flight path is also shown. The RPA is approaching the start of two chained established (i.e., Predicted) noodle segments (blue-colored route) in which the predicted altitude will start at 15,000 feet. The RPA will travel over two noodle segments in 163 seconds of flight time. The first segment has the RPA descending, which is denoted by the hash-marks (| symbols) added in-line to the noodle symbology for every 50 feet of descent, to reach an altitude of 14,800 feet at the start of the next segment. This next noodle segment has the RPA climbing, which is denoted by the symbol (^) added in-line to the noodle at every 50 feet of climb, directing the aircraft to climb to ‘15.2’ thousand feet by the end of the noodle segment. At this point, the chained noodle segments have ended and the Directed Noodle is made available for further manipulation in order to continue chaining noodle segments into a flight path.

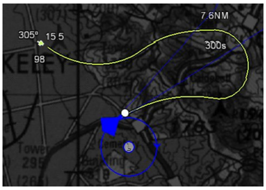


Figure 3a.

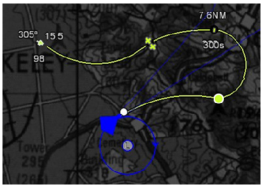


Figure 3b.

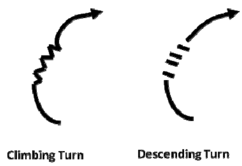


Figure 3c.

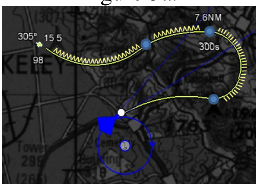


Figure 3d.

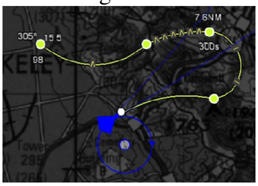


Figure 3e.

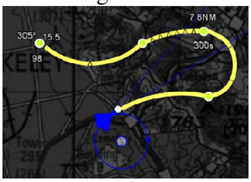


Figure 3f.

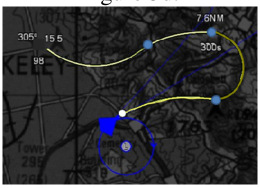


Figure 3g.

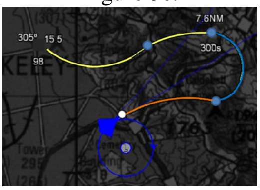


Figure 3h.

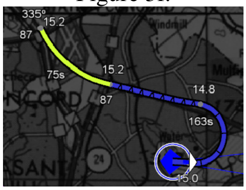


Figure 3i.

Fig. 3. Illustration of Different Methods Considered for Conveying Segment Altitude: 3a. “+” symbol. 3b. In-line “+” and “-” symbols. 3c. Pilot drawing symbology. 3d. Noodle augmented with pilot symbology. 3e. Pilot symbology in-line with noodle. 3f. Pilot symbology in-line with wider noodle enclosing pilot symbology. 3g. Shading. 3h. Color. 3i. Current implementation.

4 Evaluation of Noodle Construction

In evaluations conducted to date, feedback has been very favorable about the utility of the noodle tool, both from pilots [3, 4] and video gamer [4] participants. However, these data on the tool’s efficiency have only been qualitative. Thus, a pilot study was conducted to collect some preliminary quantitative data to characterize the effort required in employing this control method. In this study, only the noodle control was exercised; other FLEX-IT features were not employed. Specifically, data were collected on the time and accuracy in constructing noodles with varying complexities.

4.1 Noodle Construction Task

Six tracks (approximately 16.7 NM; 10.5 in centered on 22 in display) comprised of seven segments were designed and overlaid on the map of the simulator. Each track provided cues to the participants on what heading and altitude inputs were required during the construction of the noodle. Figure 4 illustrates a track that required both multiple heading and elevation changes. The change in the path's direction cued participants that a heading change would be required during noodle construction. Changes to the altitude were cued by the color of the path segment's borders (red: decrease altitude, white: maintain altitude, blue: increase altitude; all altitude changes were 500 ft). Changes in the heading and altitude were accomplished using the joystick. A throttle switch was used to change the length of the noodle and associated throttle switches were used to confirm or erase constructed noodle segments. Each track segment was designed to be approximately twice the distance of the default noodle length. Participants were not allowed to change the vehicle's airspeed setting.

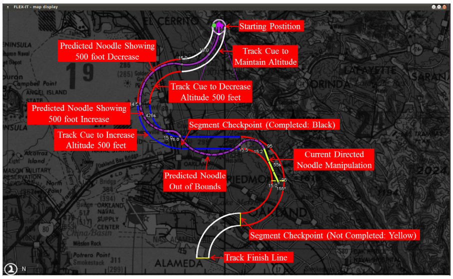


Fig. 4. Sample Segmented Track used during Experimental Trials

4.2 Experimental Design

Six different tracks were utilized varying in complexity in terms of the number of required changes in heading and/or altitude (see Table 1). Tracks that were mirror images of the six initially generated tracks were also used, to increase the variety of heading changes. Also, each of these 12 variations was presented both with the vehicle starting from the south and from the north (bottom and top of the display, respectively). These variations resulted in 24 track configurations (6 complexity levels X 2 mirror images X 2 directions) and each track configuration was considered an individual trial. The order of the 24 trials (i.e., 24 track configurations) was independently randomized for each participant in a within-subject design with the constraint that any track was not presented more than twice in a row and any other factors (e.g., direction flown) did not occur more than three times in a row.

Five volunteers (4 male, 1 female) with an average age of 34.8 years (SD=11.37) served as participants. All reported vision correctable to 20/20. None of the participants were rated pilots; none played more than 4 hours of video games per week.

Table 1. Number of heading and altitude changes commanded for each of the six track complexity levels as well as total absolute change in altitude across tracks

	TRACK COMPLEXITY					
	1	**2**	**3**	**4**	**5**	**6**
# Heading Changes	2	4	6	2	4	6
# Altitude Changes	0	0	0	3	4	5
# Altitude Increase	0	0	0	2	2	2
# Altitude Decrease	0	0	0	1	2	3
Total Absolute Altitude Change (ft)				1500	2000	2500

4.3 Procedures

Experimental sessions lasted 50-60 minutes with participants tested individually. After receiving training on the noodle construct, features of each controller, associated symbology, and test procedures, participants practiced constructing noodle paths for several vehicles using a "training track" with a variety of heading and altitude changes. This training took approximately 20 minutes to complete. Then, each subsequent trial was initiated by the experimenter's input that caused a new track to be presented with the RPA symbol at the beginning of the track where it remained suspended for the duration of the trial. Recording of noodle construction time began when the participant activated a joystick switch that moved an indicator on the display to the "N" option. Participants were required to immediately start constructing a noodle that followed the track and overlaid each checkpoint (yellow lines dissecting the track) at the end of segments. (Once the accepted noodle passed a checkpoint, the yellow line turned black.) When a Directed Noodle segment was accepted the line became purple, denoting a Predicted Noodle. Upon completion of the new flight path (construction and acceptance past all checkpoints), the word "FINISHED" appeared in large text.

Participants were instructed that they could take as long as they wanted to construct the noodle, as well as employ any strategy they deemed most efficient (e.g., employ multiple short noodles versus fewer longer ones). However, they were also told that the time it takes for them to designate the future path of the RPA through the track using the noodle tool would be measured, as well as the extent to which the resulting path crossed outside the track's borders. While the Directed Noodle (yellow line) could exit the track during construction, participants were instructed that the Predicted Noodle (purple; appears when Directed Noodle segment is committed with throttle switch selection) should be within the track's borders. A questionnaire with 14 rating scales and 5 open-ended questions was administered after all trials were completed.

4.4 Results

Quantitative. Participants were very accurate in path construction: path was within borders on 97.5% of the trials. Mean time to construct the noodle was 62.84 s (range:

22.59-192.31). Results of a within-subject repeated measures 6 (complexity) X 2 (mirror images) X 2 (directions) analysis of variance (ANOVA) indicated a significant effect: mean noodle construction time significantly differed across complexity levels (Fig. 5, $F(5,20)=71.834$, $p=0$). Post-hoc tests indicated Track 6 with 6 heading and 5 altitude changes took significantly longer than others. The percentage of time savings realized by use of the noodle tool was estimated on each trial by subtracting construction time from the time it would take an RPA to hypothetically fly the designated route at the fixed airspeed (dividing that difference by the total flight time). An ANOVA of these data indicated that the time savings significantly decreased as the complexity of the track increased (Fig. 5; $F(5,20)=71.832$, $p=0$). Participants employed between 5 and 18 segments when constructing the noodle (mean=10.43, SD=3.69) and rarely used the erase feature (mean=0.267, SD=0.63).

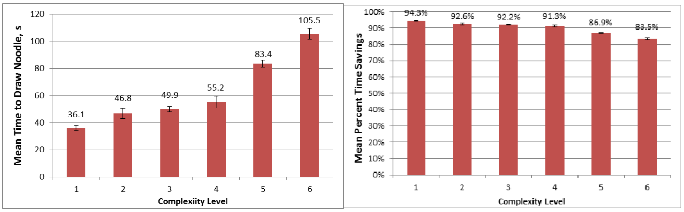


Fig. 5. Mean noodle construction time (left) and percentage of time savings for each track complexity level (right). Error bars are standard error of the mean.

Qualitative. The first questionnaire item asked for an overall rating on the participants' experience in constructing noodles (scale: very difficult, moderately difficult, moderately easy, very easy, or no opinion). Four of the five participants rated noodle construction as moderately or very easy. One participant selected 'moderately difficult' for this as well as 5 of 12 specific steps in creating the noodle, including changing heading, changing altitude, and heading south (as opposed to north). One other participant also indicated heading south was difficult, commenting that the southerly direction made it more likely to make erroneous left/right joystick inputs to change heading. Otherwise, the participants generally rated specific construction steps as either 'moderately easy' or 'very easy.' Four of the five participants indicated very little hand fatigue, while one selected 'somewhat.' Tracks that required a change in both altitude and heading for a segment were slightly more time consuming. There were also specific comments on some of the controls (e.g., too easy to activate the switch that erased a noodle segment). Participants' strategy in terms of the length of segments used in noodle construction varied.

5 Conclusions/Future Directions

The quantitative data collected in this effort compliments the qualitative data collected earlier indicating that this control mode for visualizing and commanding near-term future path is a definite candidate for future multi-RPA control. For example, the data demonstrates how the future flight path can be constructed very quickly (typically less than a minute). Comparing these data with an estimate of time to hypothetically fly the paths suggests a 90% attentional time savings for typical operations (albeit, the operator would still periodically check status of all vehicles even with the noodle method and would complete cross checks during some manual flight phases). Moreover, it is anticipated that pilots would construct noodles more quickly and realize additional time savings, compared to this study's non-pilot participants. Nevertheless, the ability to easily set the complex flight path of one vehicle should allow more time to be focused on other vehicles, enhancing supervisory control of multiple RPAs.

To apply this noodle concept operationally, however, there are several challenges still to be addressed. The predicted flight path is not likely to precisely match the actual path in the real world due to limitations in prediction models, effects of external disturbances such as significant winds, and inaccurate, delayed, or interrupted telemetry data. Near term applications will probably be more successful with high performance flight systems that are less impacted by external disturbances and have more sophisticated flight models and auto pilot systems. Another design issue to consider is how to represent the path's uncertainty to the operator. If the magnitudes of inaccuracies are known, then the predicted flight path could include augmented symbology that depicts the nature and source of the uncertainty, with an indication of how the degree of uncertainty increases with distance and time in the future.

Acknowledgments. Research was supported by AF/A2Q and managed by AFRL. Support from Ball Aerospace & Technologies Corp. and Smart Information Flow Technologies personnel was provided under Air Force Contract FA8650-08-D-6801.

References

1. Miller, C., Hamell, J., Barry, R., Ruff, H., Draper, M., Calhoun, G.: Adaptable Operator-Automation Interface for Future Unmanned Aerial Systems Control: Development of a Highly Flexible Delegation Concept Demonstration. In: Infotech@Aerospace 2012 Conference AIAA-2012, pp. 25–29 (2012)
2. Draper, M.H., Miller, C.A., Calhoun, G.L., Ruff, H., Hamell, J., Benton, J., Barry, T.: Multi-unmanned Aerial Vehicle Systems Control via Flexible Levels of Interaction: An Adaptable Operator-automation Interface Concept Demonstration. In: Infotech@Aerospace 2013 Conference (2013)
3. Calhoun, G.L., Draper, M., Ruff, H., Miller, C., Hamell, J.: Future Unmanned Aerial Systems Control: Feedback on Highly Flexible Operator-automation Delegation Interface Concept. In: Infotech@Aerospace 2012 Conference, AIAA-2012, pp. 25–49 (2012)
4. Calhoun, G., Draper, M., Miller, C., Ruff, H., Breeden, C., Hamell, J.: Adaptable Operator-automation Interface for Multi-unmanned Aerial Systems Control: Preliminary Usability Evaluation. In: Proc. Human Factors & Ergonomics Society (September/October 2013)

The Virtual Dressing Room: A Perspective on Recent Developments

Michael B. Holte

Department of Architecture, Design and Media Technology
Aalborg University, Esbjerg, Denmark
mbh@create.aau.dk

Abstract. This paper presents a review of recent developments and future perspectives, addressing the problem of creating a virtual dressing room. First, we review the current state-of-the-art of exiting solutions and discuss their applicability and limitations. We categorize the exiting solutions into three kinds: (1) virtual real-time 2D image/video techniques, where the consumer gets to superimpose the clothes on their real-time video to visualize themselves wearing the clothes. (2) 2D and 3D mannequins, where a web-application uses the body measurements provided by the customer, to superimpose the standard sizes to fit a customized 2D or 3D mannequin before buying. (3) 3D camera and laser technologies which acquire 3D information of the costumer, enabling estimation of the body shape and measurements. Additionally, we conduct user studies to investigate the user behavior when buying clothes and their demands to a virtual dressing room.

Keywords: Human-computer interaction, virtual reality, augmented reality, user interface design, computer graphics, interaction design, review, survey, clothing industry, 3D imaging, and 3D scanning.

1 Introduction

Today people are increasingly shopping online; most of them are satisfied buying certain types of goods online like books, electronics, tickets etc. However, when it comes to buying clothes online, they are not entirely satisfied [1]. Many people choose not to shop online due to privacy (e.g., sharing body measurements) and security issues [2]. Online shopping has undergone many changes in the past decade; the sellers are always pushing the boundaries to provide the ultimate shopping experience for the consumers. From online cataloguing of products (amazon.com) to live online shopping assistants (3LiveShop, Sweden), the industry has made many innovations in the delivery models. Products like books and electronics, given the nature of the values that they provide the consumers, serves appropriate for the see-click-buy online shopping protocol. However, shopping for clothes online raises some challenges.

A report shows that there is approximately a 25% return rate of the ordered goods in the online clothing industry in Denmark. The reason for the returns at this moment can be speculated as, the clothes do not fit the customers properly or the customers simply

R. Shumaker (Ed.): VAMR/HCII 2013, Part II, LNCS 8022, pp. 241–250, 2013.

dislike the cloth when they actually wear it. As a result, there is an increase in the costs for the online retailers and dissatisfaction among the consumers. Apparently, the consumers are looking for more reliable solutions for buying clothes online. The industry is beginning to recognize that new technologies like virtual-reality and 3D camera-based systems have great potential to solve this problem. Hence, the virtual dressing room (see Fig. 1) addresses this problem by enabling the consumers to, e.g., try on the virtual version of the clothes on their virtual 3D avatar/profile before buying the real clothes, from the convenience of their home computers, TV or hand-held devices.

The advantages of using a virtual dressing room from your home/office are several: reducing the proportion of returned items; competitive advantages since the customer can try the clothes by using the application from the web shop before buying and then reduce risk in the buying process; increased opportunities for customization - the customer are able to create their own customized clothes; customers can easily record and upload a video showing them with different clothes on social medias to have their friends opinions; the possibility to make different visualizations/augmented reality, e.g., if you try a bikini, the background can be changed to a sunny beach, so you are able to see the clothes in the situations you actually use it. In addition, installing a virtual dressing room physically in a retail store can have the following advantages: faster trying of different clothes, reducing queues to the dressing rooms; reducing inventory (the need for having a lot of garments in different colors and designs in the store); after you have found a piece of cloth using the virtual dressing room, you are able physically to try the clothes in the store and get a feeling of the garments; generates fun and attention.

In this paper we give a review of recent developments and future perspectives, addressing the problem of creating a virtual dressing room, and discuss and compare the performance of existing solutions. First, we review the current state-of-the-art and discuss the applicability and limitations. Based on this review and discussion we explore the integration of virtual-reality technologies in a user-friendly fashion for the online clothes shopping industry, dealing with both design and technological research aspects. Next, we conduct user studies to investigate the user behavior and their demands. To this end, we discuss existing solutions and their usability based on the environment and settings: retail store, home/office web-shopping, mobile platform or for social networking.

Fig. 1. A concept sketch of a virtual dressing room

2 Existing Solutions

The industry is attempting to improve the aforementioned scenario by developing a more holistic virtual dressing room with an acceptable quality of use. The current state-of-the-art solutions can be categorized into three kinds: (1) virtual real-time 2D image/video techniques, where the consumer gets to superimpose the clothes on their real-time video to visualize themselves wearing the clothes. (2) 2D and 3D mannequins, where a web-application uses the body measurements provided by the customer, to superimpose the standard sizes to fit a customized 2D or 3D mannequin before buying. (3) 3D camera and laser technologies; these types of solutions are new to the industry, yet to prove their full benefits. In the following we look further into exiting solutions in each of the three categories.

2.1 Virtual Real-Time 2D Image/Video Techniques

This set of techniques superimposes 2D graphical models of clothes on top of input image/video of the user in real-time. These 2D solutions only give a simulated impression of how the clothes might look on the customers, and are often inaccurate, failing to take into account the consumer's 3D body proportions. Furthermore, the models of the cloth are rigid and need to be adjusted in size and translated to the correct position manually. In Fig. 2 some examples images of these solutions are shown.

Metaio Inc. Business. A product where clothes are applied onto a webcam image of yourself. The clothes consist of images on the center of the screen which you can scale up and down in size to get a close as possible fit to the size (distance to the webcam) of your body. Funny, but does not show clothes in a natural and photorealistic way. Uses non-flexible garment models and is not a virtual mirror [3].

Facecake's Swivel. A virtual dressing room, where clothes are applied to a real-time image of yourself. The clothes are images shot from different angles which change to the rotation of your body. While the clothes do not fit perfectly, it does allow you to see how it looks in a more appealing way. Only uses non-flexible 2D models of clothes, and does not follow the person's curves and movements [4].

ImmediaC. A digital dressing room, where clothes are applied to a real-time image of yourself. Again, the clothes are images shot from different angles which change to the rotation of your body. While the clothes do not fit perfectly, it does allow you to see how it looks in an attractive way. Uses 3D model of clothes, however they are non-flexible, meaning that garments does not follow your curves and movements [5].

EyeMagnet. A virtual dressing room, where clothes are applied to a real-time image of yourself. The clothes are images shot from different angles which change to the rotation of your body. Again, the clothes do not fit perfectly but it does allow you to see how the clothes look in a more appealing way. Uses non-flexible 3D clothes [6].

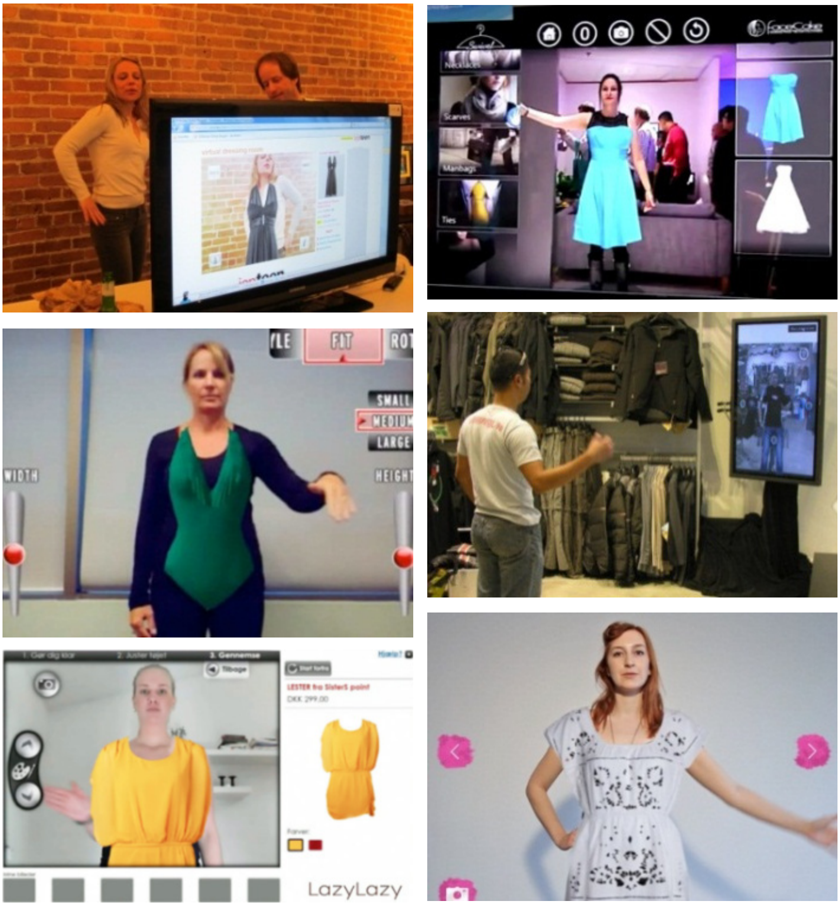


Fig. 2. Example images of the virtual real-time 2D image/video techniques, where the consumer gets to superimpose the clothes on their real-time video to visualize themselves wearing the clothes. Read row-wise: Metaio Inc. Business [3], Facecake's Swivel [4], ImmediaC [5], EyeMagnet [6], LazyLazy [7] and Indiska's Jade Jagger [8].

LazyLazy. A virtual dressing room, where clothes are applied to a real-time image of yourself. The clothes consist of images in the center of the screen which you can scale up and down in size to get a close as possible fit to the size (distance to the camera) of your body. Funny, but does not show clothes in a natural and photorealistic way [7].

Indiska's Jade Jagger. Also a virtual dressing room, where clothes are applied to a real-time image of yourself. This solution uses a place holder object (a piece of paper with a certain pattern) to place the cloth on top of the costumer. Hence, it is up to the user to navigate the augmented cloth correctly on top of the input video. The need of a place holder object and manual placement is obviously not desirable [8].

2.2 2D and 3D Mannequins

This is a set of solutions where a web-application uses the body measurements provided by the customer, to superimpose the standard sizes to fit a customized 3D mannequin before buying. These solutions simply lack real-time simulation, are not virtual mirrors, and do not provide the kind of user interaction as the virtual real-time 2D image/video techniques. Fig. 3 shows some examples images of these solutions.

Fits.Me. A solution where manufactures send their clothes to the Fits.me company in Estonia. They then apply the clothes to a shape shifting mannequin and take pictures of each shape. All these photos are then stored in a database and used in their online application, and the closest fit is shown to the user when they input their measurements [9].

My Virtual Model. A gaming/social network solution where 3D clothes are generated by the user base and applied to a generic female 3D mannequin, which can be altered in shape and look by the user. The model can be rotated so it is in theory similar to how you see a mannequin in a store. Funny, but does not show clothes in a natural and photorealistic way [10].

Upcloud. A webcam-based solution for body measurement. This is not really an interactive virtual dressing room, but serves solely for measuring the proportions of the costumer. The body measures are store and used for determining the size of the clothes when shopping online [11].

HM Virtual Dress Room. A gadget solution with a dull that you can dress up with a few different garments that look like some of the HM models. This is solely a 2D photo-based dressing room with no inputs related to the look of the costumer [12].

2.3 3D Camera and Laser Technologies

These solutions use 3D/range cameras or lasers to acquire 3D information of the costumer, enabling estimation of the body shape and measurements (see Fig. 3).

Bodymetrics. One such solution is Bodymetrics, which uses a rig setup of multiple Microsoft Kinect sensors to measure the user's body proportions, followed by an expert in the shop floor guiding the buyer with the fitting. This is a costly and time consuming process. Hence, this is not a real-time interactive virtual dressing room, but serves solely for measuring the proportions of the costumer [13].

Fitnect. A solution similar to the virtual real-time 2D image/video techniques, however, instead of standard image/video input, the Microsoft Kinect sensor is used to acquire input video and depth/3D information of the user. The additional 3D information of the body facilitates automatic fitting of the cloth to the body, which follows your movements. The clothes are flexible but do not fit the body shape and follow your movements accurately. The garments look alright but not photorealistic [14].

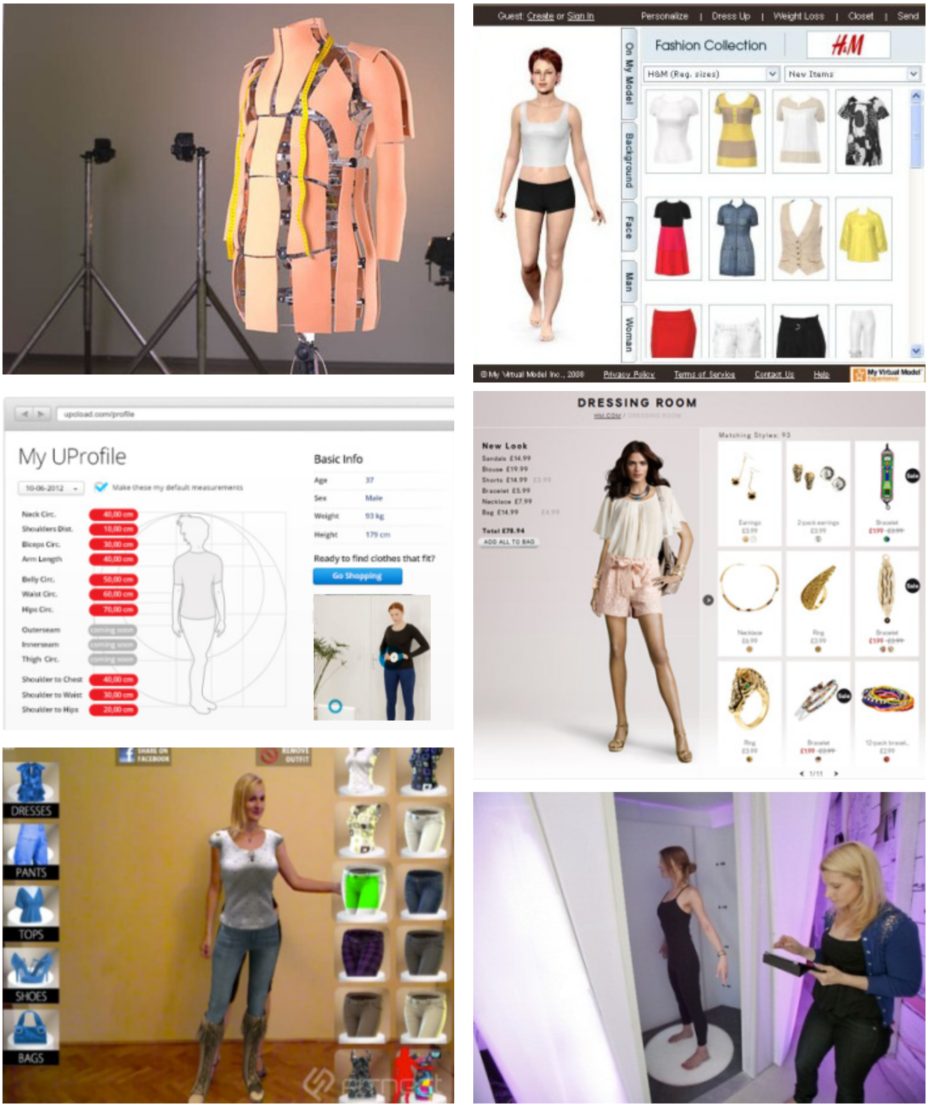


Fig. 3. Example images of 2D and 3D mannequins, and 3D camera and laser technologies. These solutions let the costumer dress up 2D or 3D mannequins for better visualization of the clothes and how it matches. In addition some acquire 3D information of the costumer, enabling estimation of the body shape and measurements. Read row-wise: Fits.Me [9], My Virtual Model [10], Upcloud [11], HM Virtual Dress Room [12], Fitnect [14] and Bodymetrics [13].

In Table 1 some significant characteristics, limitations and usability of the exiting virtual dressing room solutions are listed for an easy comparison. Note that the usability of the different systems varies, hence, some systems are applicable for retail stores, home/office, mobile platforms and/or social networking. Commonly for the virtual real-time 2D image/video techniques, is that they can be applied both in retail stores and at home/office, while some of the 3D mannequins and 3D camera/laser technologies are mostly applicable in retail store or similar settings. Several of the systems can be adapted to mobile platforms but only the simple web application-based solutions are so far suited for mobile usage. Additionally, some systems have utilities allowing the user to share resulting images of them wearing the virtual clothes through a social network media, e.g. Facebook.

Table 1. Significant characteristics, limitations and usability of the exiting virtual dressing room solutions.

Product	**Input / Interface**	**Dim.**	**Real-time**	**Limitations**	**Usability**
Metaio Inc. Business [3]	Camera Monitor	2D	Yes	Rigid models	Retail store Home/office
Facecake's Swivel [4]	Camera Monitor	2D	Yes	Rigid models	Retail store Home/office
ImmediaC [5]	Camera Monitor	2D	Yes	Rigid models	Retail store Home/office
EyeMagnet [6]	Camera Monitor	2D	Yes	Rigid models	Retail store Home/office
LazyLazy [7]	Camera Monitor	2D	Yes	Rigid models	Retail store Home/office
Indiska's Jade Jagger [8]	Camera Monitor	2D	Yes	Rigid models Requires place holder object	Retail store Home/office
Fits.Me [9]	Mechanical mannequin	3D	No	Not a virtual mirror	Retail store
My Virtual Model [10]	Web application	2D	No	Not a virtual mirror	Home/office
Upcloud [11]	Camera Monitor	2D/ 3D	No	Only for body measurements	Retail store Home/office Mobile
HM Virtual Dress Room [12]	Web application	2D	No	Not a virtual mirror	Home/office Mobile Social network
Bodymetrics [13]	Laser	3D	No	Only for body measurements	Retail store
Fitnect [14]	Kinect Monitor	3D	Yes	Inaccurate fit to body movements	Retail store Home/office Social network

3 User Studies

In this section we present user studies, which we have conducted to investigate the user behavior when buying clothes and the users' demands of an acceptable virtual dressing room solution.

3.1 Cloth Shopping Behavior

To study the shopping behavior of costumers, we asked 39 (14 males and 25 females) subjects within the age range of 16 - 62 years with an average age of approx. 30 years. The subjects were mainly working full-time or students with part-time jobs, and with an average income. Nearly all subjects buy clothes monthly or more often, while only a few buy clothes once or a few times a year. 95% of the subjects buy clothes in retail store, 54% from online web shops, 21% in second hand stores and 8% on flea markets (multiple selection were possible, hence, the percentages do not add up to 100%). The online shoppers prefer online web shops, since it is cheaper, faster, easier and more convenient, and have a larger selection than the retail stores. On the other hand the costumers buying clothes in retail stores prefer to see and feel the fabric and texture. Additionally, the size can vary especially for pants, dresses and suits, and they want to check the fit before buying it. Actually, 18% of the subjects check the quality in retail stores before buying it online.

The need of returning cloth bought online varies; from never to a couple of times in total, while some return cloth for nearly every time they make an order. Mostly because it does not fit, match or the appearance differs from the online version. This seems to correspond well with the fact that there is approximately 25% return rate in the online clothing industry in Denmark. In conclusion, over the half of the asked subjects buy clothes online, due to several advantages. However, there is relative high return rate of the bought cloth, due incorrect size or the cloth does not live up to the expectations. A virtual dressing room could address some of these issues by estimating the body measures of the individual and let costumer virtually try the cloth, e.g. to see if it fits or matches other pieces of clothes.

3.2 User Tests and Demands

Next, we conducted an experimental study to get feedback on how satisfied the users are with the exiting virtual dressing rooms. For this purpose we tested LazyLazy's solution, which represents the current standard of the virtual dressing room well. In this experiment 30 subjects (within the age range of 13 - 47 years with an average age of approx. 21 years) tested the virtual dressing room, and we received the following feedback. 7% though the system was very good, 43% good, 27% neutral, 17% bad and 0% very bad. Hence, the overall impression was fairly good. The system was fun, interesting and for most fairly easy to use. However, the test subjects had a lot of comments and complaints during the test. Some of the most important ones were: slow reaction/lag; the cloth does not adapt to the body shape; the graphic of the clothes is poor; not the same feeling of trying real physical cloth; you cannot feel the fabric and how it fits your body; you cannot turn and see how the cloth looks form behind or different viewpoints; it is difficult to reach and press the correct buttons; takes some time to get used to the navigation; would be nice if the cloth moved on top of your body automatically; it is easier to move yourself than the cloth; the cloth should follow and adapt to body movements; the cloth should adjust to the correct size automatically, so you can determine you size.

Most of the subjects were not more likely to buy clothes online using this system. Mainly because the system did not help them much in the end, since they only got a poor idea of how they look wearing the cloth, due to poor graphics and rigid models of clothes. Furthermore, the navigation and reaction time were not satisfactory. In conclusion the tested virtual dressing room has to be improved significantly to become a useful tool for shopping cloth online than just a fun gadget.

4 Conclusion and Future Perspectives

In this paper we have given a review of exiting virtual dressing room solutions, and conducted user studies to reveal users' cloth shopping behaviors and their demands to a virtual dressing room. Based on this review and user studies, one can conclude that none of the exiting solutions have, so far, managed to develop a holistic solution with an acceptable quality of use. However, the current solutions are more fun to have gadgets than a need to have tool.

Concretely, a turnkey-solution needs to meet all the demands of (a) user satisfied interfacing, (b) accurate fitting of the virtual clothes, (c) photorealistic virtual clothing, (d) reliable and low cost, while (e) meeting real-time performance. The recent 3D camera and laser technologies, e.g. Microsoft Kinect, make it feasible to fuse the advantages of a real-time virtual video-based mirror and the estimation of 3D body measurements similar to the 3D mannequins. Combined with an improved user interface, more photorealistic graphical models of garments, and more accurate 3D fitting of virtual clothes, a turnkey virtual dressing room solution meeting the users' demands might not be far away.

Furthermore, a user-friendly buying process is much preferred among the online shoppers [15], it is also one of the biggest challenges this new interaction domain poses to the designers and developers. The online shoppers also like comparing the prices of the clothes they want to buy [16], especially due to the flexibility that the web medium provides. In order to encourage and enhance the consumers' online shopping experience, one needs innovative ideas to make the shopping a fun and more social experience [16], that can be compared to outlet shopping, which offers social interaction with shopping partners, people enjoy this.

In this work we have been focusing on the customer interface of the virtual dressing room. Another part of designing and developing a virtual dressing room is the company interface, where a major concern is how to design and produce digital clothes for the virtual dressing room. Creating digital clothes using specialized programs, e.g. Marvelous Designer [17], is a time consuming process, and it might be difficult to convince the manufactures of an investment in digital cloth.

Acknowledgements. The research leading to these results has received funding from The Danish National Advanced Technology Foundation under the research project "Virtual Dressing Room". The author would like to thank Yi Gao, Aalborg University Esbjerg, for conducting user studies.

References

1. Schaupp, L.C., Belanger, F.: A Conjoint Analysis of Online Consumer Satisfaction. Journal of Electronic Commerce Research 6(2), 95–111 (2005)
2. Hoffman, D.L., Novak, T.P., Peralta, M.: Building consumer trust online. Communications of the ACM 42(4), 80–85 (1999)
3. Metaio Inc Business, 1000 Sansome Street, San Francisco, CA 94111-1323, http://www.seventeen.com/fashion/virtual-dressing-room
4. Facecake Marketing, 4603 Deseret Drive, Woodland Hills, CA 91364-3719, http://www.electrictv.com/?p=2752
5. ImmediaC: upper water street, Halifax, NS Canada (1969), http://www.youtube.com/watch?v=XjiQ7MXe1-8
6. EyeMagnet: Level 3 B, 1 Nelson St, Auckland, New Zealand, http://www.youtube.com/watch?v=XU1XhzCHj4I
7. LazyLazy.com., http://lazylazy.com/en-DK/content/shoppingfeatures
8. Indiska, Malmvägen 3 Värtahamnen, Stockholm, http://vimeo.com/24165498
9. Fits.Me, Veerenni 24C, 10135 Tallinn, Estonia, http://www.fits.me
10. My Virtual Model, 80 Queen Street, Montréal, QC H3C 2N5, Canada, http://www.myvirtualmodel.com
11. UPcloud, Ziegelstr. 13c, 10117 Berlin, Germany, http://www.upcload.com
12. Hennes, H.M., Mauritz, L.P.: 215 Park Ave. South, 15th floor, New York, N.Y. 10003 USA., http://www.hm.com/us/dressingroom/LADIES
13. BodyMetrics, Tottenham Court Road, London, Great Britain, http://www. bodymetrics.com
14. Fitnect Interactive Kft, Budaörs, Hungary (2040), http://www.fitnect.hu
15. Lim, K.: Security and Motivational Factors of E-Shopping Web Site Usage. In: Decision Sciences Institute 2002 Annual Meeting Proceedings, pp. 611–616 (2002)
16. Limayem, M., Khalifa, M., Frini, A.: What Makes Consumers Buy from Internet? A Longitudinal Study of Online Shopping 30(4), 421–432 (2000)
17. Marvelous Designer. Product, http://www.marvelousdesigner.com/marvelous/Default.aspx

Making Sense of Large Datasets in the Context of Complex Situation Understanding

Marielle Mokhtari[1], Eric Boivin[1], and Denis Laurendeau[2]

[1] Defence Research and Development Canada, Quebec City, Canada
[2] Computer Vision and Systems Laboratory, Université Laval, Quebec City, Canada
{Marielle.Mokhtari,Eric.Boivin}@drdc-rddc.gc.ca,
Denis.Laurendeau@gel.ulaval.ca

Abstract. This paper presents exploration prototype tools (combining visualization and human-computer interaction aspects) developed for immersive displays in the context of the IMAGE project. IMAGE supports collaboration of users (i.e. experts, specialists, decision-makers…) for common understanding of complex situations by using a human guided feedback loop involving cutting-edge techniques for knowledge representation, scenario scripting, simulation and exploration of large datasets.

1 Introduction

Informed decision-making usually results from a proper comprehension of a situation/phenomena. Nowadays, in the military domain, like in many other domains, achieving a good understanding of a *complex situation* (CS) is a challenge when rapid changes, both operational and technological, occur. This challenge, combined with the objective of increasing the agility of Canadian Forces in dealing with such situations, is investigated in the IMAGE project (Lizotte, Bernier, Mokhtari & Boivin, 2012). IMAGE aims to support collaboration between users (i.e. experts, specialists, decision-makers…) to develop a common CS understanding, which can be shared at different levels. The IMAGE concept is supported by four modules: (1) KNOWLEDGE REPRESENTATION, which builds and shares a vocabulary and conceptual graphs making the CS understanding explicit; (2) SCENARIO SCRIPTING, which transforms conceptual graphs into an executable simulation model; (3) SIMULATION (SIM), which exercises the space of scenario variables and browses through simulations; and (4) EXPLORATION (EXP), which exploits visualization and interaction metaphors for investigating large datasets dynamically so data become meaningful to the user. The work presented in this paper focuses on this fourth module, which puts emphasis on the proper use of technologies (and the tools developed for supporting them) in response to the user's exploration needs. Finally, the paper concentrates on non-desktop toolsets developed to help users in making sense of data.

In a wide range of scientific and technological domains (e.g. economics, climate and seismic modeling, melting modeling, astronomy, neuroscience and archaeology), experts need to make sense of and extract useful knowledge from large datasets

R. Shumaker (Ed.): VAMR/HCII 2013, Part II, LNCS 8022, pp. 251–260, 2013.

composed of various types of data. To achieve these goals, experts need tools combining visualization and rich user interactions to guide them to incrementally and interactively explore (large) datasets, to organize data, to process information, and to experience and understand these datasets. The EXP module consists in making datasets (generated by the SIM module) explicit in (static and/or dynamic) tailored views, which can be used by users to augment their individual or collective CS comprehension models.

2 The EXP Concept

Fig. 1 illustrates the EXP concept, which aims to produce individual or collective explanatory and tailored (static and/or dynamic) views in ways that words cannot communicate clearly. These views, in order to bring a meaning to (raw and/or processed) data, have to generate, stimulate, increase or accelerate the CS understanding. The EXP concept is supported by tools, which implement visualization and interaction approaches by taking advantage of the potential of different platforms in terms of technology and in terms of exploitation of human sensory information (mainly *perception* and *action*). The EXP concept has been developed through two main prototypes: the EXP Version 1 (EXP V1) was part of the IMAGE V1 cognitive experiment focusing on individual understanding (Lizotte et al., 2012, sections 7.2 and 8.3) while the EXP Version 2 (EXP V2) was part of the IMAGE V2 feasibility effort and focused on immersive virtual exploration and collaboration aspects (Lizotte et al., 2012, section 7.3).

As represented in Fig. 1, hardware components composing the platforms supporting the EXP tools span from a traditional hardware setup composed of LCD screen / keyboard / mouse to a more high-tech platform exploiting immersive technologies leading to user immersion in a virtual environment (VE) explicating the simulation world initiated by the SIM module. In fact, the EXP module puts emphasis on the proper use of technologies in order to fulfill the user's exploration needs: from (1) conventional non-immersive systems (no immersion, limited interaction), to (2) hybrid systems combining the adequate use of non-immersive systems and immersive interactive systems (in terms of head-centered interaction, access to traditional interaction tools augmented by virtual gadgets) up to (3) high-end fully immersive systems (in terms of large displays, large field of view, head- / hands-centered interaction, more natural interactions). Immersion can be achieved to varying degrees by using advanced human machine interfaces (involving sensory modalities – visual, auditory, tactile, olfactory – and body tracking, etc), and by exploiting more natural interaction with the immersive VE (IVE).

Virtual reality (VR) technology immerses one or more users in a VE representing a real or fictive world that allows users to become agents capable of changing the properties of the environment and interact with its entities (Burdea & Coiffet, 2003). IVEs, such as CAVEs (Creagh, 2003; Cruz-Neira, Sandin, DeFanti, Kenyon, & Hart, 1992) support, among other things, the ability to analyze and interpret a large amount of data faster and to explore data in a more intuitive way through interactions with visually rendered data. The CAVE used in IMAGE is a multi-screen projection-based VR system with three walls and a floor that are arranged in a cube for *total* immersion.

One of the major advantages of the CAVE is that it makes (collaborative) exploration of large-scale content possible through manipulation and navigation within an IVE. The user navigates the virtual content by naturally moving around inside the cube, while his/her field of view is filled with 3D representations of components of the environment. To achieve realistic interactions with the CAVE, the tracking system implements fast refresh rates, low latency and smooth tracking. Consequently, it provides smooth and precise estimates of the position and orientation (6 degrees of freedom) of the user's head and hands while not interfering with his/her immersive experience. The tracking system offers ergonomically designed devices: (1) a head tracker attached to stereo glasses; (2) trackers attached to *Cybergloves*, which are composed of 22 flexible sensors that can accurately measure the position and movement of the fingers and wrist; and (3) a tracked *wand*, which includes buttons and a joystick for interacting with the virtual content. Stereoscopic rendering based on active glasses provides the user with depth perception of the virtual content.

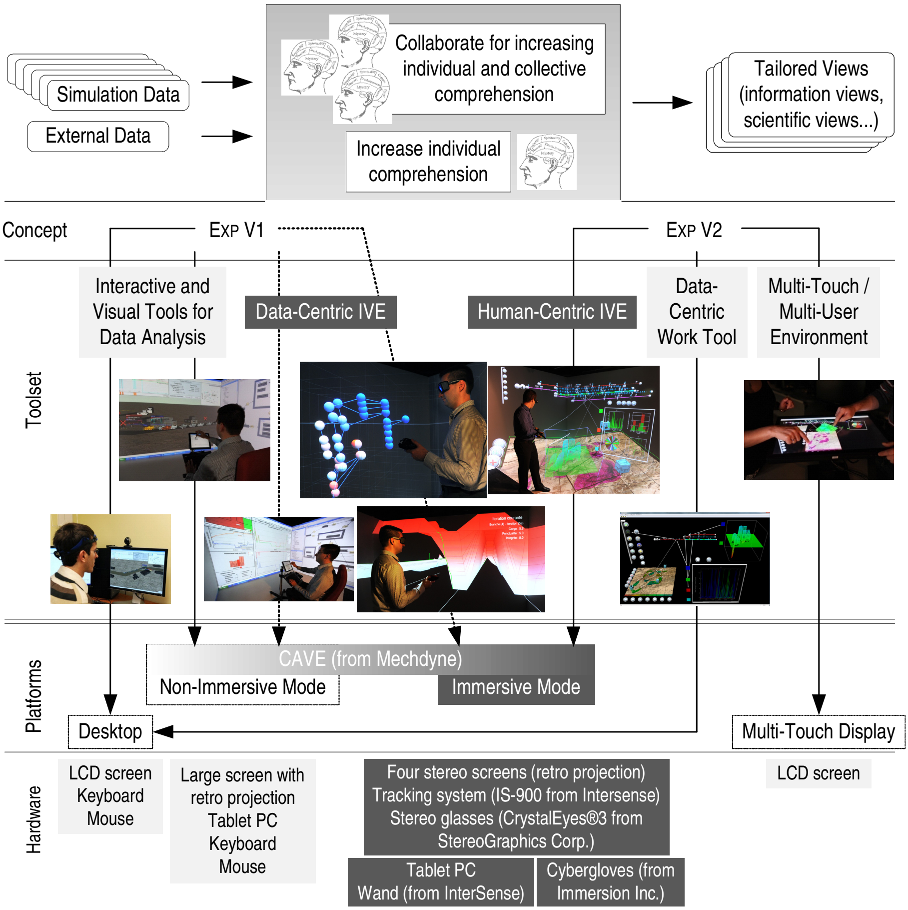


Fig. 1. The EXP module at a glance

3 EXP Toolsets in General

The two versions, EXP V1 and EXP V2, implement interactive and informative simulation analysis tools for exploring different aspects of CSs. To address as much as possible exploration needs with respect to diversity and richness, the toolset proposed in the EXP V1 is a combination of turn-key and commercial off-the-shelf (COTS) tools while the EXP V2 toolset is mainly a custom turn-key tool. The toolsets developed mainly offer the user rapid prototyping of views as well as tailored views and focus on the *best view* notion rather than the *perfect view* notion. In the spirit of the *Visualization* domain, the main task of which is to allow *information* to be derived from *data*, the EXP toolsets have been thought to assist users understand and interpret large and rich datasets, by exploiting information rendered graphically (e.g. graphs, diagrams, charts, etc.), which are visual artifacts representing data, illustrating relationships and patterns among data, presenting *information* (appropriate conversion of *data*), and also allowing analysis and comparison between graphical representations of datasets. In addition, the EXP toolsets address the dynamic display of trends in the data and human interaction with the displayed data. Consequently the toolsets have been designed to: (1) explore and compare large datasets; (2) extract (hidden) information; (3) discover unexpected trends in the data or find patterns; and, finally, (4) understand specific non-intuitive aspects of CSs. To achieve these goals, the EXP module uses different techniques such as (1) filtering (for focusing on relevant subsets of data) and data brushing (support visual linking of various data and address the visual fragmentation problem of multivariate data representations) (Becker & Cleveland, 1987), (2) datasets comparison, (3) multi-level datasets exploration, and (4) real-time and stereoscopic rendering.

4 EXP V1 Toolset - From Desktop Tools to CAVE Tools

The EXP V1 toolset has been designed to be used on a traditional desktop and on a large screen as well as, at least for a part, on a display wall and in a CAVE, both providing active stereoscopic visualization. Although traditional techniques of data visualization remain conceptually interesting and viable when the user faces a (very) large amount of data to be processed, some techniques become more difficult to implement on a traditional desktop screen. VR allows for the ability to simultaneously analyze a large amount of data / information as well as for a more intuitive exploration in IVEs. The VR gives the impression to the user that he/she interacts and manipulates data directly and not just its graphical representation (Bryson, 1996). Furthermore, multi-dimensional visualization in an immersive context allows users to better understand all the relations between data (Arms, Cook, & Cruz-Neira, 1999; Raja, Bowman, Lucas, & North, 2004; Van Dam, Forsberg, Laidlaw, LaViola Jr, & Simpson, 2000). IVEs provide many benefits for understanding complex systems (Knight & Munro, 1999), namely the visualization of immersive 3D graphics and different immersion paradigms offering intuitive navigation to explore and analyze data / information. Besides, IVEs facilitate the learning process (Dalgarno & Lee,

2010) by, among other factors, promoting the spatial representation of knowledge and augmenting both user's motivation and commitment.

Desktop tools (and parts of) have then been adapted with this in mind so that the user can visualize and explore the data as being "in-the-box i.e. inside looking out" (CAVE version) rather than as being "out-of-the-box i.e. outside looking in" (desktop version). A tablet PC is used to switch between immersive and non-immersive modes, and also to select parameters and data to configure and display. The user can navigate inside the 3D virtual content at will. He/she uses the wand to manipulate (translate, rotate, scale or reset to default values) the 3D graphics. Data selection is not needed because any manipulation involves the whole view. Fig. 2 shows a user manipulating datasets with the wand and visualizing them with the stereo glasses. The data-centric IVEs developed are supported by a custom open architecture, mainly based on *openGL* and the 3D graphics toolkit *OpenSceneGraph* (OSG), a 3D graphics rendering engine, used to develop applications in scientific visualization, VR and modeling.

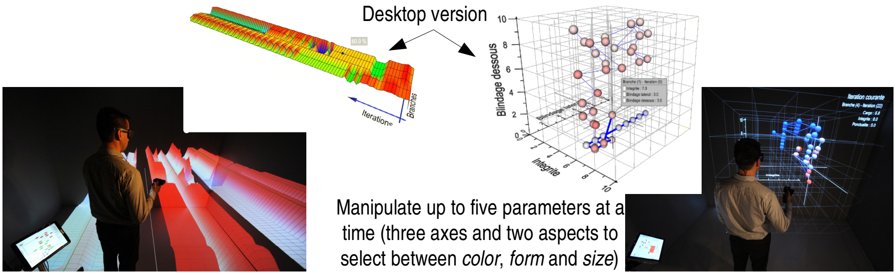


Fig. 2. Exp V1 toolset – From desktop tools to CAVE tools

5 EXP V2 Toolset

While the development of the EXP V1 toolset has been constrained by cognitive experiments, the EXP V2 is free of all cognitive experiment constraints, hence to the possibility to exploit VR concepts to their full potential and to immerse the user in a dynamic working IVE he/she can build, organize and manage at will by exploiting a toolset for simulation, visualization and analysis. In this second version, the SIM and EXP modules (at least the GUI) are combined within the same environment. A first version of the EXP V2 toolset was presented in (Mokhtari, Boivin, Laurendeau, Comtois, Ouellet, Lévesque & Ouellet, 2011).

5.1 Architecture

The EXP V2 engine is coded in C++ and is based on OSG. The EXP V2 environment can be used on a flat display wall or in a CAVE, both providing active stereoscopic visualization. The architecture also extends to desktop and multi-touch platforms, and supports collaborative work between users. The architecture has been designed to be independent of graphics and physics rendering engines, and is built as a tree, in which each branch is an entity, some entities having no correspondence with graphics and physics rendering trees. Three categories of entities have been defined:

1. Manipulator: represents a hardware device (e.g. data glove, wand, keyboard / mouse…) used by human to interact with the objects in the environment. Several Manipulators of the same type can coexist in the EXP V2 environment;
2. User: corresponds to a real user in the environment. It establishes the correspondence between the user and his/her Manipulator(s) and also manages object manipulation to avoid conflict between users;
3. Object: represents an element that is not a Manipulator and with which User can interact. 3D Objects are special instances of Object: their pose (position, rotation, pivoting point and scale factor) can be modified and manipulations (move, rotate, zoom) can be applied on them.

Interaction is considered as a communication mechanism (i.e. flow of messages) allowing Manipulators to send interactions to Objects. Interactions are not directly coupled with Manipulators so the Manipulator design is independent of the Object design. Drag-and-drop is a mechanism that is implemented at the entity level (1) to transfer information from one Object to another (via a ghost concept) and (2) to instantiate associations between two Objects. Another flexible and powerful feature of the architecture is the metadata concept, which encapsulates all data types provided to the EXP module by the SIM module.

5.2 Fully Immersive Exploration Prototype

The EXP V2 environment (see Fig. 3) is entirely built, managed and controlled by the user during his/her work session (including object creation, manipulation, destruction, association of objects, etc.). The various objects available, which can be considered as a set of work tools, are presented succinctly in the following subsections (one instance – or several – of the four first work tools is – are – visible on Fig. 3).

Simulation Tree (one scenario, one parameter space). The Simulation Tree is the cornerstone of the EXP V2 environment. It is the direct link between the SIM and EXP modules. The Simulation Tree corresponds to a 3D implementation of the multichronic tree proposed by (Rioux, Bernier & Laurendeau, 2008), which is the visual representation of the simulation conceptual framework used by IMAGE V1. The Simulation Tree encapsulates both (1) simulations logs in a dynamic layout and (2) simulation data. It can be related to the concept of hierarchy visualization (a branch of *Information Visualization*) whose visualization and layout techniques have been identified and listed in (Schulz, 2011). Within this classification, the Simulation Tree belongs to the category of 3D dimensionality, an explicit edge (node-link) representation and as exhibiting principal axis-parallel node alignment. Several Simulation Trees can coexist in the EXP V2 environment, each one representing one simulation space created by the user. To mark a point of interest on the Simulation Tree, the user can create sliders, objects that stick to the Simulation Tree and slide along its branches. Combining concepts of data and/or information visualization, the Simulation Tree provides users with (1) a visually explicit history of their experiments; (2) a self-explanatory representation of the simulations; and (3) interaction metaphors with running simulations for which parameters and resulting data, rendered in a proper way, can assist users to better understand the simulated CS.

Geospatial View (a means of expliciting a simulation). The Geospatial View allows the simulation status to be visualized as a snapshot at each simulation step (i.e. defined by a slider in the Simulation Tree). The Geospatial View corresponds to a terrain representation on which the scenario elements are displayed symbolically. Information layers and gauges / indicators (being updated at each simulation step) can be added to inform the user of the status of simulation parameters. This tool also permits a simulation to be *played* (like a video). This object is active if and only if it is associated with a given simulation (in the Simulation Tree) by a drag-and-drop operation.

Scientific Views (comparing simulations in different ways). At any time, to visualize, analyze and explore simulation data or compare simulations, the user can configure Scientific Views (already associated or not to one or several simulations) according to different plot functions: (1) 2D plots such as line (plot), bar (histogram, stacked), and area (pie) graphs. The user can create up to four different graphics within a figure window, each representing different data combinations; and (2) 3D plots that are similar to the 2D plots but in 3D. To associate a simulation to one Scientific View, users can drag and drop this simulation onto a Scientific View and, by a set of menus, select data they wish to visualize and/or analyze (and compare) as well as the type of plot (2D or 3D) to be used. Scientific Views are based on the PLplot library.

Association (communication channel between objects). In order to create an environment in which objects can interact with each other, the concept of *association* has been implemented at the entity level. This mechanism allows two objects to communicate and exchange data in an asynchronous way using a common protocol. In the environment, an association is represented graphically as a physical link between objects (a bidirectional arrow). Objects' associations can be deleted by the user when they are no longer needed.

Control menus. Two types of control menus are available to users:

1. Main menu (see Fig. 3 and Fig. 4, on the left): is a traditional linear menu composed of a set of icons that represent actions that can be either selected or dragged onto the target object; is always visible; allows to connect the EXP module to the simulator, to create objects (some by changing the value of parameters), and to associate information with elements of the environment. *Note*: object deletion is achieved by throwing the object above the head of the user (the asymmetric bimanual gestural interface for achieving this operation – and others – is succinctly explained in section 5.3).
2. Circular menu (see Fig. 3 and Fig. 4, on the right): is a hierarchical menu that can be attached to specific objects for additional configuration capabilities (Huot &

Lecolinet, 2007); is visible on-demand, and allows selecting which and how data should be visualized and displayed. A circular menu is an interesting technique for displaying many items in a confined space. Besides, it speeds up and optimizes the display options without causing information overload. The item selection time is constant due to the circular organization around the activation point unlike linear menu where the item selection time varies depending on the item position in the menu (Bailly, 2009). Circular menu management is achieved with the asymmetric bimanual gestural interface.

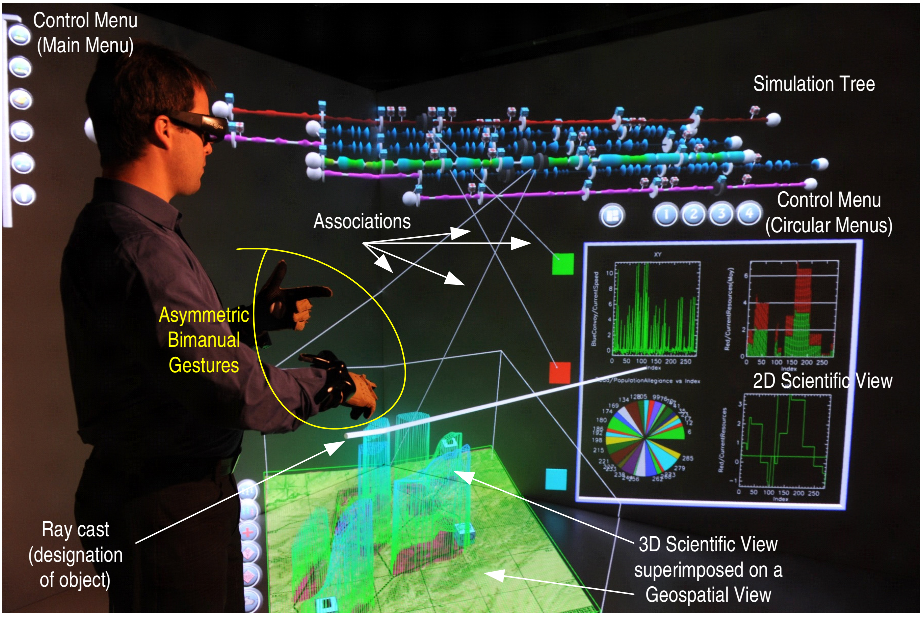


Fig. 3. An instance of the EXP V2 CAVE environment

5.3 Asymmetric Bimanual Gestural Interface

To enable the interaction between the user and the IVE, a 3D bimanual gestural interface using Cybergloves has been developed (Lévesque, Laurendeau & Mokhtari, 2011, 2013). It is built upon past contributions about gestural interfaces and bimanual interactions to create an efficient and intuitive gestural interface tailored to IMAGE needs. Based on real world bimanual interactions, the interface uses the hands in an asymmetric style, with the left hand providing the mode of interaction and the right hand acting at a finer level of detail. The user's actions in the environment have been separated into four categories of gestures: (1) designation (based on ray-casting) and selection of objects; (2) generic manipulations, which group all interactions related to moving / positioning, resizing and rotating objects; (3) specific manipulations, which are interactions tuned for specific objects (e.g. the playback capabilities associated to the Geospatial View – allowing for control of the playback using the right hand's gestures and or pose while the left hand maintains the mode); and (4) system control, which represents all actions that are related to (main and circular) menus and modifying the way the environment behaves. Symbolic input of numerical values is also implemented through a behavior similar to that of circular menus, i.e., by rotating the right hand and accepting or cancelling with the left hand. Action examples are presented in Fig. 5. There is no need for travelling interactions since the user is located at the *center* of the IVE and has access to objects without *really* navigating in the IVE.

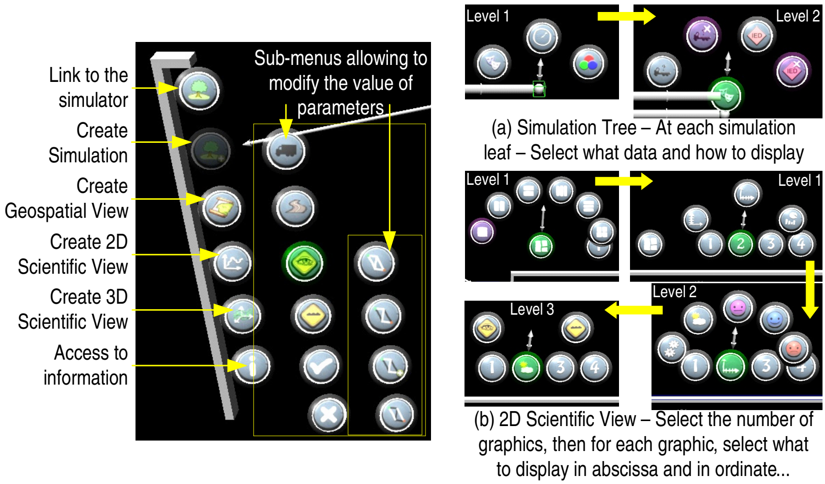


Fig. 4. Control menus of the EXP V2 CAVE environment: main menu + submenus (left) and examples of circular menus + part of their hierarchy (right)

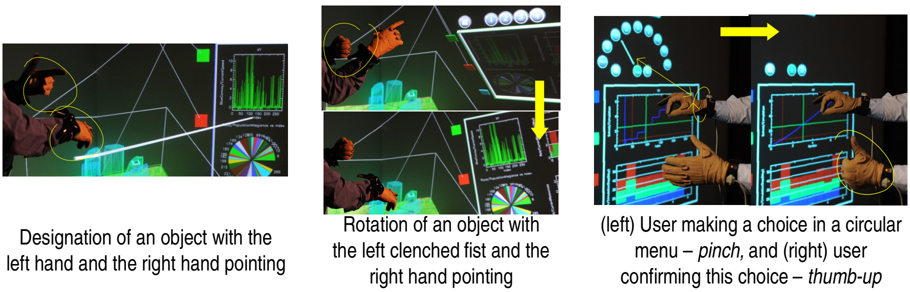


Fig. 5. Actions: selection (left), generic manipulation (middle), and system control (right)

6 Conclusion

In this paper, we introduced different tools that were developed for immersive displays to help human to deal with large datasets composed of different types of data. The EXP V1 tools are complementary tools to turn-key and COTS desktop tools while the EXP V2 toolset is a custom turn-key toolset developed by our team. We are currently exploring ways of connecting several CAVEs in order to implement collaborative work in IVEs in the context of IMAGE.

References

1. Arms, L., Cook, D., Cruz-Neira, C.: The Benefits of Statistical Visualization in an Immersive Environment. In: Proc. of Virtual Reality Conference (1999)
2. Bailly, G.: Techniques de menus: Caractérisation, conception et évaluation. Ph. D. thesis (2009) (in French)
3. Becker, R.A., Cleveland, W.S.: Brushing Scatterplots. Technometrics 29(2), 127–142 (1987)
4. Burdea, G.C., Coiffet, P.: Virtual Reality Technology, 2nd edn. Wiley-IEEE Press (2003)
5. Bryson, S.: Virtual Reality in Scientific Visualization. Commun. ACM 39(5), 62–71 (1996)
6. Creagh, H.: CAVE Automatic Virtual Environment. In: Proc. of Electrical Insulation Conference and Electrical Manufacturing Coil Winding Technology Conference (2003)
7. Cruz-Neira, C., Sandin, D.J., DeFanti, T.A., Kenyon, R.V., Hart, J.C.: The CAVE: Audio Visual Experience Automatic Virtual Environment. Commun. ACM 35(6), 64–72 (1992)
8. Dalgarno, B., Lee, M.J.W.: What are the Learning Affordances of 3-D Virtual Environments? British Journal of Educational Technology 41(1), 10–32 (2010)
9. Huot, S., Lecolinet, E.: ArchMenu et ThumbMenu: contrôler son dispositif mobile ”sur le pouce”. In: Proc. of the 19th International Conference of the Association Francophone d'Interaction Homme-Machine (2007) (in French)
10. Knight, C., Munro, M.: Comprehension with(in) Virtual Environment Visualizations. In: Proc. of the 7th International Workshop on Program Comprehension (1999)
11. Lévesque, J.-C., Laurendeau, D., Mokhtari, M.: Bimanual Gestural Interface for Virtual Environments. In: Proc. of Virtual Reality Conference (2011)
12. Lévesque, J.-C., Laurendeau, D., Mokhtari, M.: An Asymmetric Bimanual Gestural Interface for Immersive Virtual Environments. In: Proc. of the 15th International Conference on Human-Computer Interaction (2013)
13. Lizotte, M., Bernier, F., Mokhtari, M., Boivin, E.: IMAGE Final Report. Technical report 2012-397, DRDC Valcartier (2012)
14. Mokhtari, M., Boivin, E., Laurendeau, D., Comtois, S., Ouellet, D., Lévesque, J.-C., Ouellet, E.: IMAGE – Complex Situation Understanding: An Immersive Concept Development. In: Proc. of Virtual Reality Conference (2011)
15. Raja, D., Bowman, D., Lucas, J., North, C.: Exploring the Benefits of Immersion in Abstract Information Visualization. In: Proc. of the Immersive Projection Technology Workshop (2004)
16. Rioux, F., Bernier, F., Laurendeau, D.: Multichronia – A Generic Visual Interactive Simulation Exploration Framework. In: Proc. of Interservice/Industry Training, Simulation and Education Conference (2008)
17. Schulz, H.-J.: Treevis.net: A Tree Visualization Reference. IEEE Computer Graphics and Applications 31(6), 11–15 (2011)
18. Van Dam, A., Forsberg, A.S., Laidlaw, D.H., LaViola Jr, J.J., Simpson, R.M.: Immersive VR for Scientific Visualization: A Progress Report. IEEE Computer Graphics and Applications 20(6), 26–52 (2000)

Evaluating Distraction and Disengagement of Attention from the Road

Valentine Nwakacha[1], Andy Crabtree[1], and Gary Burnett[2]

[1] Mixed Reality Laboratory, University of Nottingham, Nottingham, UK
{vqn,axc}@cs.nott.ac.uk
[2] Human Factors Research Group, University of Nottingham, Nottingham, UK
Gary.Burnett@nottingham.ac.uk

Abstract. Drivers use sat nav for navigation assistance but research links sat nav with risk of distraction [10]. Visual and cognitive workload can be increased as drivers divert their attention from the road [1, 8]. Mitigating such risks is vital and head-up displays (HUDs) can be beneficial [9]. HUDs present images on the windshield to reduce diversion of drivers' attention from the road. This paper presents a driving simulator experiment which examined how 30 participants behaved with three navigation interfaces; novel virtual car HUD, arrow HUD and sat nav to outline potential benefits of the virtual car HUD over the arrow HUD and sat nav. Distraction-related data (speed, headway, lane position and peripheral detection) were gathered. The findings showed participants were better at navigation performance and peripheral detection with the virtual car HUD. Subjective data showed participants rated the virtual car HUD easiest to use, least distracting and most preferred interface.

Keywords: Driver distraction, head-up display, user interface design.

1 Introduction

The market for In-Vehicle navigation systems has risen significantly since the first commercial satellite navigation system (sat nav) for vehicles was arguably designed by Steven Lobbezzo [5]. Sat navs are useful as they can track the location of vehicles on the route and provide turn-by-turn navigation instructions using audio and visual mechanisms [4]. The typical sat nav is mounted on the dashboard as shown in Fig. 1a but alternative designs can be placed on the windshield as shown in Fig. 1b.

(a)  (b) 

Fig. 1. a) Dashboard-mounted sat nav b) Windshield-mounted sat nav

R. Shumaker (Ed.): VAMR/HCII 2013, Part II, LNCS 8022, pp. 261–270, 2013.

According to a 2007 Gallup survey for the European Union (EU), 35% of the EU citizens (which accounts for approximately 159 million people) currently use or intend to purchase a sat nav [6] (even though these figures will undoubtedly have risen since then). Despite their benefits, there are several pitfalls which have been associated with sat nav use while driving. Research highlights that they are a potential source of distraction [10] which can cause vehicle drivers to disengage their attention visually (eyes-off-the-road) and/or cognitively (mind-off-the-road) [2]. Studies in [7] have shown that drivers can fail to detect the changes on the road when glancing at head-down displays (i.e. sat nav) which can increase the risk of crash. Significant attention has been directed towards the design of head-up displays which research has outlined can deal with the issues involving drivers disengaging their attention from the road. Head-up displays present virtual images on the windshield so that drivers can reduce the diversion of their attention from the road [3] when perceiving instructions. In essence, head-up displays can reduce the shift in the locus of work for obtaining the required instructions from the road. The possible outcomes are reduction in the driver's visual and cognitive workload, increase in visual awareness of events on the road and reduction in the response times to any change.

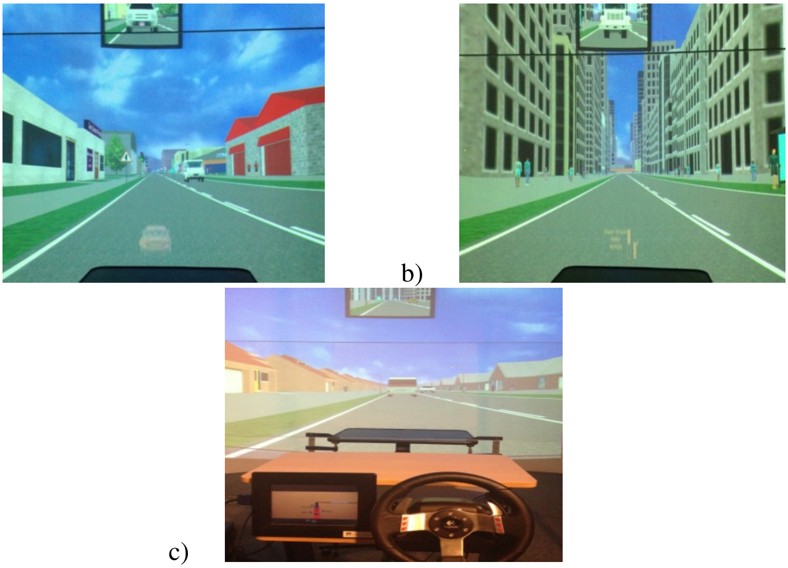
In this paper, a novel virtual car head-up display concept is proposed as an alternative to current navigation systems (sat nav and arrow head-up display) for presenting the required turn-by-turn navigation instructions during navigation to vehicle drivers. The virtual car head-up display is a novel multimodal interface that can be projected on the windshield to present drivers with visual and sound practices which are employed in regular driving (e.g. following vehicles, turning and indicating direction of turn with sound). The virtual car image appears embedded on the road in front of drivers as a lead vehicle to reduce the shift of the driver's visual attention from the road. The integrated indicating sound prompts drivers to know when a turn is about to be made. This is beneficial for enhanced turn anticipation, preparation and execution.

Furthermore, the virtual car head-up display uses its two states (the active and inactive states) so that drivers can safely control their vehicle movement without the need to be continually engaged with it. The active state is the state where the virtual car image provides drivers with the navigation instructions at the turning points e.g. indicating and turning at junctions. The inactive state is the state where the virtual car image remains in a forward, idle position which indicates to drivers that no turn actions need to be taken. It is proposed that the inactive state can be very useful for reducing the driver's visual workload (glancing away from the road to the navigation interface) and cognitive workload (translating instructions from the virtual car to the road) with the interface. The predicted outcome is the ability for drivers to quickly detect and respond to changes which occur on the road.

The virtual car head-up display was evaluated along with the arrow head-up display and sat nav by 30 participants in a desktop driving simulator to identify the extent to which each of the interfaces could cause participants to disengage their attention from the road. More so, the potential benefits of the virtual car head-up display over the arrow head-up display and sat nav were sought. The arrow interface was projected on the windshield and used only visual symbols (the arrow symbol and written information e.g. street name, direction of the next turn and estimated distance to the turn) to present participants with instructions for better turn anticipation, preparation and execution. The aerial map view of the dashboard-mounted sat nav was complimented by spoken commands which participants had to listen to, process in bits and execute sequentially in order to follow the route. The sat nav also presented participants with other information (current street name, next turn direction and estimated distance to the turn) on the visual interface for better turn anticipation, preparation and execution.

2 The Experiment Overview

The three navigation interfaces evaluated in the experiment are shown in Fig. 2.



a) b) c)

Fig. 2. a) Virtual car HUD b) Arrow HUD c) Sat nav

The hardware setup comprised of an interconnected game steering wheel and pedals system for controlling the movement of the vehicle on the simulated road. The STISIM software was used to design the simulated environment for the drives. The head-up display interfaces (Fig. 2a and b) were projected on an improvised windshield (perspex glass) from a monitor that was located about three meters in front of the participants. The sat nav interface (Fig. 2c) was located one meter in front of the participants on the dashboard. Two video recorders were placed around the participants to capture data on their driving behavior. The first recorder was placed at a 45° angle in front of the participants at a distance of about three meters away to capture their eyes and head movements during the tasks. The second recorder was placed at the rear of the simulator room and recorded the visual behavior of the participants.

Thirty participants (twenty males and ten females, average age: 27.8 years) who were residents in Nottingham took part in the experiment. They each had a valid UK driver's license with driving experience of at least one full year. The participants were divided into three groups of ten and a counter-balanced format for each group of participants with the navigation interfaces was adopted as shown in Table 1.

Table 1. Format for participant groups and interface use in the drives

Group no.	1st drive interface	2nd drive interface	3rd drive interface
1	Virtual car HUD	Arrow HUD	Sat nav
2	Arrow HUD	Sat nav	Virtual car HUD
3	Sat nav	Virtual car HUD	Arrow HUD

The participants carried out three tasks in each of the drives; driving, navigation and peripheral detection. The driving task involved safely controlling the vehicle movement on the road using the steering wheel and pedals. The navigation task involved following the correct turns to reach the destination by using one navigation interface per route. The peripheral detection task involved detecting the appearance of an attention symbol (an arrow that randomly appeared on the left/right side of the road scene shown in (Fig. 3a and b)) on five different occasions; three of which occurred when navigation instructions were provided.

(a)  (b) 

Fig. 3. a) Attention symbol on the left b) Attention symbol on the right

The peripheral detection task was directly linked with hazard awareness which was considered a vital aspect of knowing the potential perceptual tunneling effect with each interface [7]. The participants were required to press a button on the steering wheel when they detected the appearance of the attention symbol. The participants were initially required to take a test drive to familiarize with the driving simulator controls. There was no data collected or navigation interface used. During the main drives, a within-subject design was employed with the experimental conditions counter-balanced. Several distraction-related data were collected, for instance, in the driving task, speed, lateral lane position and headway to lead vehicles were recorded. In the navigation task, the number of correct turns taken was recorded for navigation performance. Also, glance frequency and durations away from the road were recorded. In the peripheral detection task, reaction times and success rates for detecting the attention symbol were recorded for the potential visual tunneling behavior with each navigation interface. After each drive, the participants filled out a NASA-TLX questionnaire, providing responses based on their experience with the navigation interface used. The information provided in the questionnaire included the physical and mental demand, driving performance, ease of use, level of distraction and overall preference.

3 The Key Findings

A repeated measures ANOVA (Analysis of Variance) with Sphericity assumed for variables measured showed a statistical difference for the speed values with the navigation interfaces ($F(2,58) = 130.39$, $p<0.05$). Bonferroni post hoc tests revealed differences in the mean speed and variation comparing the virtual car head-up display and sat nav (29.5 ± 0.9 vs. 27.5 ± 1.0 mph) ($p = .00$). A higher mean and variation was obtained with the arrow head-up display (32.3 ± 1.2 mph) ($p = .00$) as shown in Fig. 4. The lowest speed values recorded with the sat nav suggested the possibility that the participants may have found the tasks more difficult to carry out with the sat nav. For the lateral lane position, the test showed that there was no significant difference in the mean values for the navigation interfaces ($F(2,58) = 0.8$, $p>.00$). Bonferroni post hoc tests revealed no significant difference in the lateral lane position with the sat nav (8.5 ± 0.22 feet) when compared with the virtual car head-up display (8.4 ± 0.17 feet) ($p = .485$) and arrow head-up display (8.5 ± 0.18 feet) ($p = 1.00$) as shown in Fig. 5. It was concluded that change of navigation interface had no significant impact on participants' lane keeping behavior.

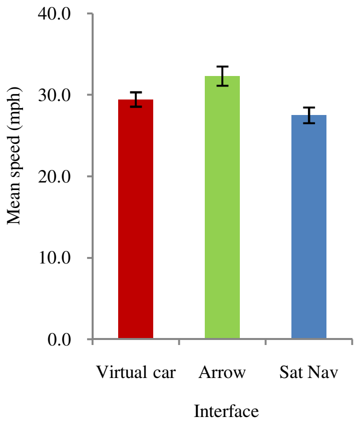


Fig. 4. Mean speeds

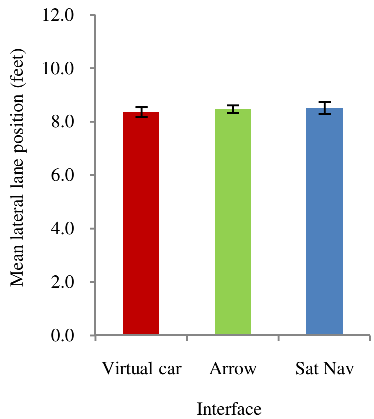


Fig. 5. Mean lateral lane positions

For the headway, there was a statistical difference between the navigation interfaces ($F(2,58) = 41.37$, $p<0.05$). The Bonferroni post hoc tests revealed a significant difference in the mean and variation values comparing headway to the lead vehicles with the head-up displays (186.4 ± 67.0 vs. 279.1 ± 63.44 feet) ($p = .00$).

Also the mean headway and variation comparing the virtual car head-up display and sat nav (186.4 ± 67.0 vs. 167.1 ± 125.9 feet) (p = .00) differed. The arrow head-up display and sat nav headways and variations also significantly differed (279.1 ± 63.4 vs. 167.1 ± 125.9 feet) (p = .00) as shown in Fig. 6. The higher headway value for the arrow head-up display was attributed to participants presumably requiring visual acuity to read information on the windshield thus leaving bigger gaps to the lead vehicles. The sat nav had the least impact on headway as no visual interface was on the windshield to engage with. More so, participants could choose to only listen to the audio instructions if desired. The arrow head-up display was thus associated with a higher risk of increasing the headway allocated to vehicles in front when compared with the virtual car head-up display and sat nav.

For the navigation performance, the participants took all the correct turns with the virtual car and arrow head-up displays which indicated an average success rate of 100%. With the sat nav, the participants missed one turn on average which indicated an average success rate of 80% as shown in Fig. 7. Since the head-up displays information were projected on the windshield, it was assumed that this helped to reduce the visual scanning process to obtain the navigation information needed to take the correct turns. The angular displacement of the sat nav from the driver's visual field meant that participants often glanced away from the road to obtain the visual navigation instructions which impacted on their ability to correctly take the turns on the route. It was concluded that the head-up displays were able to support better navigation performance than the sat nav. Also, the participants did not glance away from the road scenery while driving with the head-up displays. However, with the sat nav, an average of 42 glances (min 17, max 75) was recorded for drives with a mean time of 7 minutes. The mean glance duration for the participants was 1.5s (min 0.5s, max 2s).

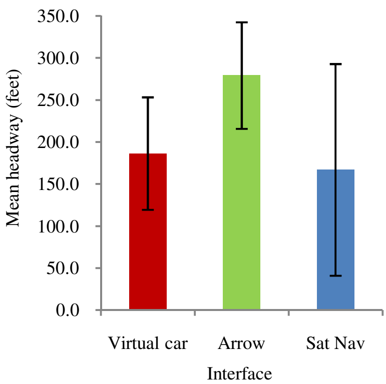


Fig. 6. Mean headways

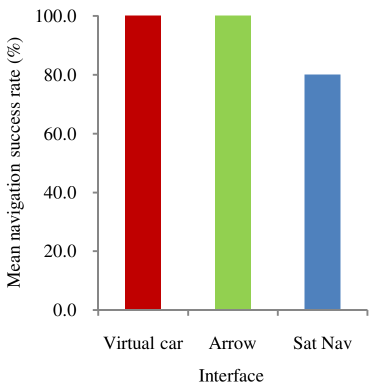


Fig. 7. Mean navigation success rates

The mean button pressing times after detecting the attention symbol in the peripheral detection task was 1.14s for the virtual car head-up display, 1.23s for the arrow head-up display and 1.3s for the sat nav as shown in Fig. 8. The faster reaction times to pressing the button after detecting the attention symbol recorded with the virtual car head-up display was attributed to the fact that participants did not have to continuously engage with the interface which allowed them to divert their attention towards attending to other tasks. It was concluded that the virtual car head-up display supports faster detection of critical events on the road than the arrow head-up display and sat nav. The average success rates in the peripheral detection task were 98% for the virtual car head-up display, 96% for the arrow head-up display and 94% for the sat nav as shown in Fig. 9. The head-up displays allowed participants to have a good visual awareness of the road scenery and were associated with higher rates for detecting the attention symbol when compared with the sat nav. Also, it was identified that participants were less occupied with attending to the virtual car head-up display due to the inactive state thus allowing for better detection of the attention symbol when compared with the arrow head-up display and sat nav. It was therefore concluded that the virtual car head-up display has the ability to support better detection of hazardous situations than the arrow head-up display and sat nav.

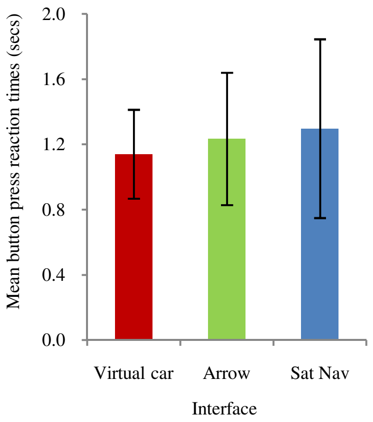


Fig. 8. Mean button press reaction times

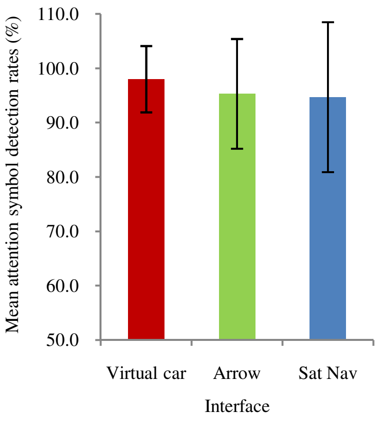


Fig. 9. Mean attention symbol detection rates

From the questionnaire feedback, the virtual car head-up display was associated with the least physical and mental demand and level of distraction followed by the arrow head-up display and sat nav. The virtual car head-up display was associated with the best performance followed by the arrow head-up display and sat nav. The virtual car head-up display was rated as the easiest to use and most preferred interface followed by the arrow head-up display and sat nav. The findings led to the conclusion that the virtual car head-up display supported better driving and navigation behavior when compared with the arrow head-up display and sat nav because it was able to reduce the workload and distraction for the participants.

4 Discussion

The virtual car head-up display concept was explored as a potential means of reducing driver workload whilst improving behavior and performance in relation to existing user interfaces. The fundamental benefits predicted with the virtual car head-up display are its ability to reduce the shift of attention away from the road and the need to constantly engage with the display while driving. The virtual car head-up display reduces the shift in the locus of work from the road by ensuring that the driver's attention is fixed on events in their field of view while driving. The benefit is that drivers are less exposed to stress and fatigue which can lead to inattention. This is because any increase in the driver's workload e.g. visual scanning for information that is not in the driver's field of view or the need to process complex information that takes time to complete is avoided. The reduced need to constantly engage with the interface ensures that the drivers are able to focus their attention instead on the critical activities that are needed to safely control the vehicle on the road. Furthermore, there is the potential for better hazard awareness where the driver is able to see the road scenery in a short period of time. This can allow for quicker reaction times to avoid any unwanted occurrences on the road.

The novelty of the virtual car head-up display concept implies that the navigation instructions are presented to drivers using new techniques. For example, the virtual car used in the head-up display presents navigation instructions using regular practices which are employed in real-world driving. The competent knowledge of drivers is exploited through a set of practices displayed which are based on their familiarity with how vehicles behave on the road. Also, the inactive state of the virtual car head-up display is a novel way of reducing the need for drivers to be continuously engaged with the navigation interface. When the virtual car image remains in the forward, idle position, drivers can perceive that the virtual car image is not presenting any instructions that are needed to turn and know that they are required to keep driving straight based on their familiarity with what happens when following real-world vehicles.

Another useful aspect of the virtual car head-up display is that it replaces any abstractions that are used by current navigation interfaces to provide instructions (e.g. arrow, speech, written information etc) with visual driving actions (indicating and turning) that are potentially easy to understand and require little or no time to process. This eliminates the need for drivers to map the abstractions to specific executable actions. The drivers basically carry out the instructions that are presented by the virtual car image and in essence, mimic its behavior. Thus, the virtual car head-up display helps reduce the mental workload of the navigation task when compared to that which might be experienced with sat nav. For example, the driver may receive the following instruction from a sat nav; "after 200 yards turn left". This instruction is a vague abstraction which the driver has to mentally process and map to the real world to follow the route. The driver has to estimate 200 yards from his/her current location, project that distance down the road and identify the exact location of the turn. Carrying out these tasks can increase the driver's visual and cognitive workload which can impact on how they allocate their attention.

Furthermore, mapping the instructions to specific actions and being able to correctly carry out the sequence of activities that are needed to take the turn at the same time may pose risk of work increase for the drivers. The virtual car prompts the driver to prepare to turn by indicating (with the indicating light and sound) at a certain distance away from the turn and uses the vehicle turn movement to signal the arrival at the turning point. The indicating sound used by the virtual car head-up display does not require as much time to process as spoken words thus reduces the processing time for the instruction. This ensures that drivers are able to allocate more time to the driving task than when spoken audio commands are issued.

Despite the benefits of the head-up displays over traditional in-vehicle sat nav, their impact on headway allocation due to the presence of information on the windshield is an area of interest. This was particularly evident when participants drove using the arrow head-up display where they were presumably reading the information on the windshield and visual acuity was required. The result was that they left bigger gaps between their vehicle and the vehicle in front which was considered to be a potential distraction behavior associated with possible tunneling effect of the interface. With the virtual car head-up display, an area of interest was that because the virtual car image was a virtual object that appeared on the road in the same way that the simulated vehicles did and the virtual car was responsible for providing instructions to the participants, it may have been possible that the participants perceived the virtual car image to be another vehicle on the road rather than an image on their windshield. It is proposed that road trials should be done as part of future work to investigate any possible effects which the virtual car head-up display can have on the allocation of headway to vehicles where there is a clear distinction between the real and virtual car on the road.

5 Conclusion

The virtual car head-up display has been identified to have the potential to reduce the driver's workload when compared with an arrow head-up display and sat nav. Automobile designers can benefit from the potentials which the virtual car head-up display offers and use them to inform future designs of In-Vehicle Navigation Systems. The safety implications of the virtual car head-up display are consistent with the philosophy of key documents in this area including the AAM (Alliance of Automobile Manufacturers), European State of Principles and JAMA (Japan Automobile Manufacturers Association). The virtual car head-up display is in its development stage and additional work is needed before the interface can be considered for integration into real-world vehicles. Road trials should be carried out to provide more validity to the findings from the driving simulator.

Acknowledgement. We acknowledge all the participants who contributed to this research.

References

1. Bach, K.M., Jaeger, M., Skov, M.B., Thomassen, N.G.: Interacting with invehicle systems: Understanding, measuring and evaluating attention. In: Proc. of the 23rd British HCI Group Annual Conference on People and Computers: Celebrating People and Technology, Cambridge, United Kingdom, pp. 453–462 (2009)
2. Brown, I.: Driver Fatigue in Human Factors. Human Factors and Ergonomics Society 36(2), 298–314 (1994)
3. Chao, C.W., Huang, C.H., Doong, J.L.: See-Through Head-up Display of In-Vehicle Information Systems for Vision Enhancement. In: 3 rd International Association of Societies of Design Research, Seoul, South Korea, pp. 877–886 (2009)
4. Farber, E., Foley, J., Scott, S.: Visual attention design limits for ITS in-vehicle systems: The Society of Automotive Engineers standard for limiting visual distraction while driving. In: Transportation Research Board Annual General Meeting, Washington DC, USA (2000)
5. `http://en.wikipedia.org/wiki/Automotive_navigation_system`
6. Keith, O., Burnett, G.: Learning-oriented Vehicle Navigation Systems: A Preliminary Investigation in a Driving Simulator. In: Proc. of the 10th International Conference on Human computer Interaction with Mobile Devices and Services, Amsterdam, Netherlands, pp. 119–126 (2008)
7. Mancero, G., Wong, W., Amaldi, P.: Looking but not seeing: Implications for HCI. In: Proc. of the 14th European Conference on Cognitive Ergonomics, London, UK, pp. 167–174 (2007)
8. Pettit, M., Burnett, G., Stevens, A.: Defining driver distraction. In: 12th World Congress on Intelligent transport systems, San Francisco, USA, pp. 1–12 (2005)
9. Wu, W., Blaicher, F., Yang, J., Seder, T., Cui, D.: A prototype of landmark-based car navigation using a full-windshield head-up display system. In: Proc. of the 2009 Workshop on Ambient Media Computing, Beijing, China, pp. 21–28 (2009)
10. Young, K., Regan, M., Hammer, M.: Driver distraction: A review of the literature. Monash University Accident Research Centre - Report #206 (2003)

DCS 3D Operators in Industrial Environments: New HCI Paradigm for the Industry

Manuel Pérez Cota and Miguel Ramón González Castro

Dpto. de Informática, Universidad de Vigo, Vigo, Spain
mpcota@uvigo.es, miguelrgc@terra.es

Abstract. The Distributed Control Systems (DCS) are electronic control devices used in continuous process industry, in which the operator becomes an essential part, and he/she should take decisions of operation that can lead to dangerous situations and/or with heavy losses. This paper shows the work done in the design, implementation and tests of a DCS console operator, which used 2.5D or 3D systems to facilitate the intuitive understanding of the state that it was in the industrial process. Also explains how different input devices were used to facilitate navigation and selection of components in the graphic display, and how different graphical concepts (geometries, colors, animations) were integrated in order to do the industrial process more understandable.

Keywords: DCS, HCI, 3D, 2.5D, Java, OPC, Jinput, RMI.

1 Introduction

Global losses in the industry with continuous process have an amount to $20 billion, equivalent to 5% of its annual budget [1]. 40% of these losses were attributable to mistakes or failures of operation, which implies that the efficiency of the operators becomes a critical element of this industry.

This industry is characterized by a critical productive process, in which a failure or a shutdown can generate situations very dangerous, both for the environment and people; outages for several days due to complex breakdowns because of those unplanned failures or shutdowns; or startups that can delay for several days until the process can be stabilized. This thing requires the use of specific control systems for this type of industry, which are referred to as Distributed Control Systems (DCS). The DCS managed thousands of analog and digital signals that are controlled through control loops, PID or systems of Artificial Intelligence (Neural Networks or Multivariate Analysis), with some sampling rates ranging between 100 ms and 2 second [2-3]. Configuration is done through logical blocks, templates or programming libraries designed to the these type of industrial processes that must be controlled, making tasks much easier for the teamwork composed by software engineers and productive process engineers. Equally, the DCS has Object Graphics Libraries that facilitate the tasks design of the operation displays that should be supervised by the operator.

R. Shumaker (Ed.): VAMR/HCII 2013, Part II, LNCS 8022, pp. 271–280, 2013.

The DCS, like those ones that they are appreciated in "Fig. 1", they have got a distributed architecture, because from the engineering station the configuration is loaded in the centralized and redundant hardware ("control and configuration module/s"), while the code is executed in systems ("process modules") independent and distributed. It allows that in case of failure of a module, the other modules of the DCS can continue working; what increases the reliability, availability and security of these control devices considerably.

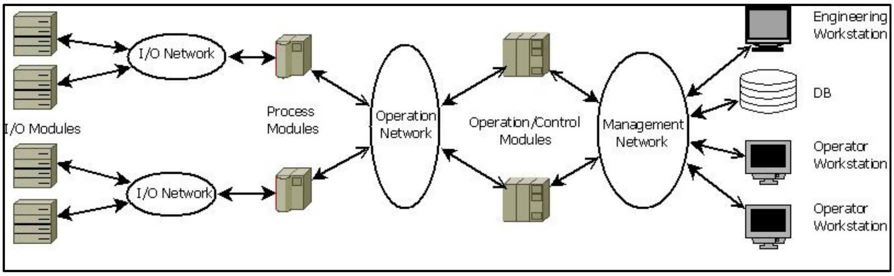


Fig. 1. DCS architecture

However, the intensive interaction with the operator is the main characteristic of these systems, because the operator is who should take key decisions on the operation of the process that is being controlled. The operator uses those set-points of the control loops, to fix the reference values (levels, pressures, intensities, etc.) of the physical variables that the DCS is controlling, and on to his/her way to modify the state of the process. Besides, they should decide if the control loops work in automatic, remote or manual mode. In automatic mode, the output values are determined by the operative of the control loop; in the manual mode, the operator forces the output values; and in the remote mode, besides the automatic operation, the set-point it is fixed by another control loop. The operator also starts the execution of operation sequences, or he/she operates valves and/or motors that are in manual mode. Equally, the DCS informs him/her about the state of the process, pointing out him/her the existence of warning or critical conditions in the process, as well as shortcomings in the operation of the several actuators or instruments of the process.

In the previous paragraph, the importance of the operator was analyzed, as fundamental element in the good operation of those industrial processes that are controlled by a DCS. It suggests that it should improve the quality of the information that receives the operator, to make it more intuitive, friendly and comprehensible [4]. So it decided to develop a "Advanced Operator Interface of DCS" that shows the industrial process in a 3D / 2.5D display and it allows different views with animations or the use of advanced navigation devices.

2 Actual Operator Interfaces

The operator interfaces of DCS [4] that they were marketed until the moment of writing of this paper, they had monitors 2D and, the displays contained piping or process diagrams in format 2D. It implied it that, if he/she wanted to have different views of the process, it was necessary to create new displays. The environmental noise and the existence of multiple consoles impeded the appropriate use of voice alarms; since it was complicated to understand their message and it increased the environmental noise, generating stress and fatigue. The use of directional speakers, although it reduced the environmental noise lightly, he/she forced to that the operator was always near to its operating console; because he/she could not hear the alarms or messages that were emitted. The alarms of the DCS were shown in displays, changing color of the device and they were available in an orderly chart for priorities and/or areas, and accessible from different options of the displays. All the operating displays showed the state of the components by means of a color code, and some allowed the access to programming manuals, device manuals (valve, motor, etc.), programs of logic blocks, wiring diagrams, piping an instrumentation diagrams or direct interconnection to other management applications (Maximo or SAP). Likewise, all the components allowed that the operator selected the operation mode (automatic, remote or manual); as well as to force their input and/or output values. The typical input interface of the DCS was the keyboard and the mouse, although, in some cases, a touch screen was used as advanced device. Equally, it admitted default configurations by the user, that it facilitated the access and navigation through the operation display.

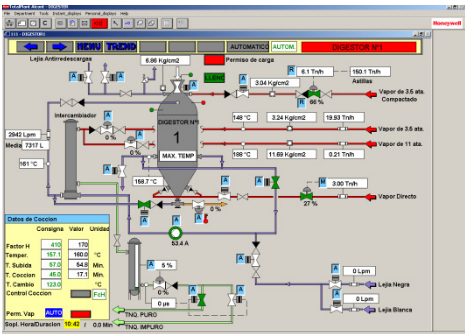


Fig. 2. Example of an DCS operator display DCS – Honeywell

Also, the console operators used to work until half hundreds of displays with several thousands of I/O signals, with what it could have displays or I/O signals that were not visualized by them during several working shifts. In occasions it happened that these displays or devices were not watched or viewed until an alarm was activated, what could cause serious breakdown in the process. So, it suggested that it was necessary to create a tool that forced to the operator to see, at least once each lapse of time, all the displays or devices of the industrial process.

3 Application Design

We tried to design the application "Advanced Interface of DCS", being always in mind the objectives that we want to reach and limited by the technological that they were imposed by software and hardware. The main objective of the application was to create software that it showed to the DCS operator a friendly graphic interface in 2.5D/3D environment, with data refreshment in real time.

The hardware limitations were imposed by the 3D devices, that they were available in the market, for example: memory and speed of the graphic cards, 3D monitors, etc. The process speed of the graphic card conditioned the complexity of the graphic objects and the type of animation, which were showed in the display [5].

The software limitation was mainly conditioned by the features of the programming language, which will create the 2.5D / 3D applications; because of it will determine the complexity of programming and design tasks of the application. This first and main decision in the realization of the interface was to test the features of each language. Finally, Java [6] was selected to develop the application and their election was based on the arguments that subsequently are detailed:

- Graphic Libraries: Java had a high-level API for the development of graphics 2.5D called “Java 3D” that it allowed the creation of displays and graphic objects.
- Remote data access: Java permitted the use of remote data, by means of the use of RMI (Remote Method Invocation), JDBC (Java Database Connectivity), socket…
- Portability: Java was designed to operate in multiple platforms that they had different Operating Systems (Windows, Linux, AIX, Solaris, etc.).
- Web - Internet: Java was designed to be executed in an Applet, what facilitated the immediate integration in a web page for any navigator.
- Nonstandard devices: Jinput, JOAL, JMF or JOGL are API's of Java that facilitated the haptics administration[7], the use of interfaces for non standard input/ouput, the integration of multimedia objects and/or the control of 3D sound effects.
- Free Software Reuse: There is a great quantity of libraries and free software available in internet to use directly or to facilitate the elaboration of new libraries.
- Industrial Devices: Java had specific libraries that they were adapted to main Industrial Communications Protocol. (OPC, ModBus, etc)
- Other: JNI (Java Native Interface) it allowed that Java code link with native code written in other languages (C++, assembler, etc.).

4 Application Architecture

After selecting Java as the programming language, the following stage was the design of hardware and software architecture to carry out the tasks of retrieving data from the industrial process, administration of the industrial information and graphic presentation to the operator. It was decided to use an architecture based on the pattern Client-Server "Fig. 3", where the server retrieved data of the industrial devices and the client had to show data in the operating display.

4.1 Server Process

The process "server" had two sub processes, where a sub-process read or wrote data in the DCS via OPC and the other sub-process put this information to disposition of all clients, through a RMI-Server. The direct connection to the DCS to read or to write data was non-viable, because of the DCS had a proprietary protocol that allowed them to communicate with the operator interface, located in user's computers. It obliged to use OPC (OLE (Object Linking and Embedding) for Process Control) [8-9], to communicate the operator interface with the DCS; because OPC was a standard open to industrial data communications based on DCOM (Distributed Component Object Model) technology of Microsoft, that it was supported by the immense majority of makers of DCS. The communication among the OPC server and the DCS were carried out by means of a proprietary protocol, but the data of the OPC server were accessible to any OPC client that could communicate by means of DCOM. So, it was decided to install the process server in the computer where it was installed the OPC server, because it avoided to use OPC tunnel or a DCOM connection among computers. The sub-process that they connected with OPC server, they used the "JEasyOpc" [10-11] library to create a OPC client that was developed in technology Java. This sub-process sent and received the data from the DCS, via OPC, to store them in a temporary memory, that it was consulted by a sub-process of RMI server.

We could have decided to install the OPC communication service inside of "Operator Interface" of application and to avoid the use of RMI. This option was discarded, due to it obliges to install Operating System "Windows" in the operator computer, because it would be necessary to install DCOM or tunnel-OPC software to connect with the OPC server. However, this was very dangerous, because OPC didn't have any type of security and anyone could force the state of any logical-block in the DCS and to cause serious mishaps or incidences. The installation of free OPC-client in the OPC-server or in the OPC-tunnel, it would allow the access without restrictions to any logical block of the DCS, with unsuspected consequences for the security of people and of the industrial facilities. Also, the use of DCOM would oblige to that the client and server computers had to belong to the same Windows-Domain, what would complicate the administration of computer networks. If it was selected a communication with tunnel-OPC software, they would be necessary to buy two use licenses and it would have an associate over cost, because this software didn't have free license.

The option of using a server that included a OPC client and a RMI server, it was considered the best solution; due to it shielded the data access of the DCS, because the connections weren't permitted to the OPC server from the operator computer. The computer of the "applications server" only permitted net connections forward to the RMI server port and later, it sent these queries to the OPC server. So, this design didn't permit connections from the operator computer to the OPC server and it protected the DCS data of any undue and/or malicious access. Also, the security in the connections to the RMI server [12] it was negotiated by the API of Java, since RMI had the security integrated in core as in the RMI client as in the RMI server. RMI was a technology created for and to Java, witch integrated the security as essential part of its design. Also, RMI provided more flexibility and power in the use of Java objects, because it allowed the transmission of complex objects between the client and the server. Everything favored the election of architecture with RMI server, since it impeded the illegitimate accesses to the DCS from malicious users and also, it facilitated the installation of any operating system in the client computer.

4.2 Client Process

The client process had three sub-processes called "RMI client", "Process-main" and "Graphic Unit". The process "RMI client" exchanged (sent and received) information from the server process and, later, it stored the data received in an intermediate memory. Subsequently, the information of the intermediate memory was used in the "Process-main". The refresh speed of this memory in "RMI client" was controlled by a thread-timer; nevertheless, the data sending to the DCS was done instantly, to avoid that a sequence of operations would lose the order and temporary distance, if they were grouped in a data block with the same time stamp.

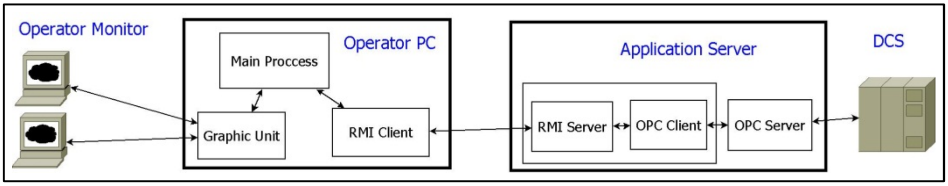


Fig. 3. Application architecture

The "Main Process" managed the received data of the DCS and to indicate how this numeric information should be became in visual information, that it was more friendly and intuitive to operator. The visual information was grouped in a related data blocks where the color and/or tonality, speed of the animations or dimensions of a geometric figures are controlled. This information was sent to the "Graphic Unit", so that would generate the new graphic scene associated with the instantaneous state of the variables of the DCS. The fluid pressure in a pipe or the voltage in a line was showed in function of the tone intensity of the assigned color, avoiding the extreme tones (light or dark) that could cause some confusion, in spite of belonging to different colors. The fluid flow or the electric current intensity were indicated by means of an animation of the solid color in the corresponding line or pipeline. The levels or volumes were showed with changes of the size of the interior geometric element that represented them. Equally, the "Main Process" had a group of temporized threads, whose mission were to check the state of the non standard peripherals (for example the 3D-joystick), as well as to create the visual illusion of the animations. The programming of the 3D-joystick was carried out using JInput [13-14], that was an API "Open Source" in Java, developed to control non standard devices, by means of the technique of "polling". The information obtained on the 3D-joystick state was sent to the process "Graphic Unit", so that it would locate in the graphic scene, the new position of the 3D-pointer.

The alarms or warnings generated by the DCS were shown in the graphic scene, by means of the change of the device color (valve, motor, etc.). Each color indicated the state in which was the device (motor in run or stopped, critical alarm, warning, thermal relay alarm, safety switch alarm, etc.). This information was analyzed in the "Main Process", and subsequently, it was sent to "Graphic Unit" the color that the device should show in the graphic scene.

The use of the whole colors range in the representation of the 2.5D or 3D figures, which were visualized in the graphic scene, should not be considered strictly incompatible, with the application of the usability rule in the computer systems. It was owing to the fact that "risks evaluation" for workplace of a console operator indicated that these tasks could not be carried out by daltonic people or with other visual or physical deficiencies. This evaluation would be also applicable to electricians, since an electrician with daltonism would not be qualified to carry out his/her work correctly, because he/she would not differentiate all colors wires of a cable bundle. The "risks evaluation" for workplace was carried out by the Health and Safety Department, that analyzed all tasks assigned to this workplace as well as their associate risks.

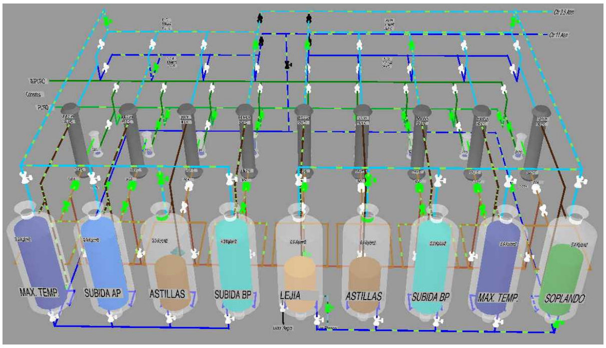


Fig. 4. Frontal view of the industrial process

The mission of the process "Graphic Unit" was the creation of a display in 2.5D or 3D that would have the graphic representation of the industrial process that he/she should supervise the console operator of DCS. This process developed in "Java 3D" [15-18], it allowed the visualization of different views of industrial process in the display, for example it admitted front views, lateral views, oblique views, etc. The navigation along the graphic scene was done by means of a 3D-joystick, while the selection of a graphic component was carried out with the mouse (only 2.5D graphic scene) and/or 3D-joystick. Also, independently of all visual helps to the operator, the graphic scene also showed in textual format all physical magnitudes of the process, as well as the state of all machines and industrial devices (valves, motors, switches, etc.).

This workspace also reported of the appearance of an alarm or warning, by the emission of a vocal message with emotion. The emission of the mentioned vocal messages with emotion facilitated the immediate understanding to the operator of the upsets that were happening; because the emotional intensity indicated the importance of the alarm, while the message indicated exactly, which was the device that had generated it. This module was developed using JLayer [21] that was a Java library with license LGPL that permitted the reproduction of files of audio, that would have the format of compression "mp3" (MPEG-1 or MPEG-2 Audio Layer III).

5 Graphic View

The graphic scene represented a part of the industrial process of obtaining of cellulose (pulp wood) from wood Eucalyptus. It was carried out by a group of 9 production lines that it operated in parallel cooking wood. Each production line was composed by a digester that cooked wood, together with her heat-exchanger, steam pipelines, bleach pipelines, pulp pipelines and a foul condensed tank. The "fig. 3" shows the typical display of the DCS Honeywell-TotalPlant that was used to control one of the nine lines of the production process. The new console operator was the result of merging 13 old displays, in an unique operation display in 2.5D / 3D , "fig. 4 and 5"; since it grouped the nine displays of each production line, a display of steam distribution, a display of liquor distribution and two general views of the productive process.

The new operator display, "fig. 4 and 5" was based on a flow diagram of the process that showed the devices in a multilayer structure that facilitated their understanding. However, it was planned that the future upgrades of this application would include other different views. These new views would be the following ones:

- Real Image: This view will show images or real pictures of the industrial process.
- Outline of Real Vision: This view will show a plane to scale, in which symbolic or conceptual elements will be showed in their real position in the factory.

Because it there were 9 identical lines, it facilitated that geometric objects could be shared for as minimum 9 devices (one for line), avoiding the excessive use of the system resources. The digester was represented by a geometric figure with semi-transparent appearance, that contained in his interior a cylinder of variable height that indicated the load level of the digester; while, his color indicated the cooking state. The foul condensed tank had a similar graphic design as the digester, but the interior cylinder already indicated the foul condensed level.

The graphic representation of the pipeline had an animation to indicate the liquid flow that it transported. The animation design imposed the creation of a geometrical figures sequence that should be alternated in a synchronized way, to pretend a movement appearance. It obligate to compose a pipeline by joining of 4 geometrical figures, where each one of them was composed by a group of cylindrical rings, same-spaced, at a distance equal to 4 times the length of each ring. That was the reason why, between two rings of oneself geometrical figure, there was 3 ring, which belonged to each one of the other 3 geometrical figures that composed the pipeline. The movement simulation was achieved assigning the representative color from the pipe to 3 geometrical figures and assigning an attractive color to the fourth geometrical figure. Finally, thanks to a temporized thread in the "main process", a sequence alternated the attractive color among all the geometrical figures and it created the illusion of the animation movement.

The presentation of 2.5D and 3D objects in a 2D screen, it imply that the axes of coordinates converged toward the point $(0, 0, -\infty)$, which is to say, it converged towards a point centered in the bottom of the display; while, the operator had the optic illusion that the axes of coordinates followed a parallel trajectory to the real world in which he was. So, the action of to select an object was complicated, because one could have the illusion of being pointing to an object and truly it was pointing out to the empty space. To eliminate this spatial discrepancy in the objects selection of a 2.5D or 3D graphic display, it was decided to use a 3D-joystick and the mouse (only 2.5D graphic scene), like double system to facilitate the objects selection for the console operator. The mouse selected objects by means of "picking", while the 3D-joystick selected objects by the collision of the 3D-pointer with the object. This double method was very useful in the rotation, displacement or shifting of the graphic scene, because they lost the spatial references, as well as the exact position of the mouse and/or 3D-joystick (3D-pointer) [19-20].

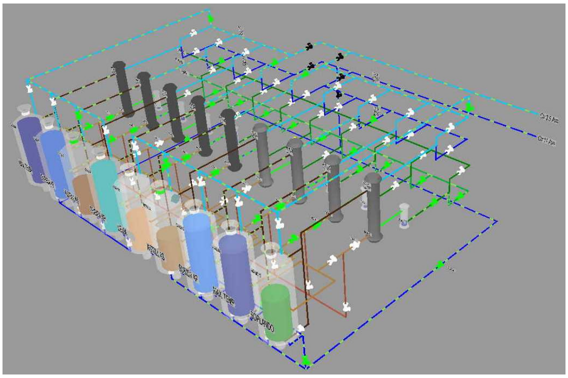


Fig. 5. Oblique view of the industrial process

The process "Graphic Unit" had functions that made displacement, rotation and zoom of the graphic scene, to facilitate the visualization of the industrial process from different points of view (frontal, upper, lateral, rear, oblique, etc.). These functions modified the disposition of the scene and the location of the 3D-pointer, to facilitate their spatial location by the operator.

Also, the illumination of the scene was designed so that it were visualized the borders in the geometric ways sharply, improving this way the perfect identification of the contour of the objects. Java had several types of illumination [19-20]: ambient light, directional light, point light (omni-directional light) and spot light (concentrated light). The ambient light didn't enhance the edges and contours of the geometrical forms, so it was discarded. The spot light (concentrated light) created scenes with very illuminated areas and dimness areas, so also it was discarded. The directional light was discarded, because it neither attenuated with the distance nor varied her direction; so the edges were not showed with the enough clarity. Finally, the point light (omni-directional light) was selected to illuminate the scene, because it raised a lot of clarity of edges in the geometrical figures; so, her intensity attenuated with the distance and her direction varied for the points that doesn't belong to the same light beam. Two points light, in left lateral and right lateral, were selected to illuminate the scene, and to increase the contrast in the geometrical figures from any observation point. Equally the grey color was chosen as background color to intensify the difference between the l0 graphic objects and the empty space.

6 Conclusions

It was created a new concept of Operator Interface of DCS for real industrial console that it differed of the habitual pattern that was used to show the information.

This Interface used a 2.5D/3D graphic scene to show the industrial process, because of visual information was more intuitive for the operator, and then it improved the capacity of understanding of the real situation of the industrial process. Equally, this graphic information allowed to observe the process from different views (frontal, lateral, etc.) and to show the interior of any industrial devices (digester, foul condensed tank, etc.). The operator console joined the visual information with the use of input interfaces like the mouse (only 2.5D graphic scene) or the 3D-joystick that facilitated the navigation and picking through the 2.5D / 3D scene.

References

1. Breibvold, H.P., Olausson, M., Timsjo, S., Larsson, M., Noble, R.T.B., Sneddon, N.: El Operario Eficaz, Revista ABB 4|10, pp. 6–11 (October 2010)
2. Nelson, B., Stauffer, T.: PLCs and DCSs Convergence, Control Engineering (March 2008), `http://ecmweb.com/mag/electric_plc_dcs_better/#`
3. Shaw, T.: PLC Or DCS: Which Is Better For You? EC&M (March 1, 1999), `http://ecmweb.com/mag/electric_plc_dcs_better/#`
4. Pérez Cota, M., González Castro, M.R.: Interfaz Avanzada de Operador de DCS. In: 6ª Confêrencia Ibérica de Sistemas e Tecnonologias de Informacao, Chaves (Portugal), pp. 37–41 (Junio 2011)
5. Nvidia 3D Vision (2011), `http://www.nvidia.es/object/3d-vision-pro-requirements-es.html`
6. Oracle. "Java" (July 2011), `http://www.oracle.com/technetwork/java/index.html`
7. Worst, C.B., Voiz, R.: Evaluation of Haptic Mixed Reality System for Interactions with a Virtual Control Pantl, Presence, vol. 14, pp. 677–696. MIT (2005)
8. OPC Foundation (July 2011), `http://www.opcfoundation.org/`
9. Iwaniitz, F., Lange, J.: OPC Fundamnetals, Imentatiton and Application, 3rd rev. edn. Hutting GmbH (2006)
10. Opc Programers Conection (July 2011), `http://www.opcconnect.com/java.php`
11. JEasyOPC Client (July 2011), `http://sourceforge.net/projects/jeasyopc/`
12. Oracle. Java™ Remote Method Invocation API (Java RMI) (July 2011), `http://download.oracle.com/javase/7/docs/technotes/guides/rmi/index.html`
13. Java.Net “Jinput" (July 2011), `http://java.net/projects/jinput/`
14. Davinson, A.: Pro Java 6 3D Game Development. Apress (2007)
15. Java.Net API Java 3D (July 2011), `http://java3d.java.net/`
16. Java.Net. Java 3D Project (July 2011), `http://java.net/projects/java3d`
17. The Java 3D Community Site (July 2011), `http://www.j3d.org/`
18. Java 3D TM Graphics (July 2011), `http://www.java3d.org/`
19. Oracle Sun Developer Network. Java 3D API Tutorial (July 2011), `http://java.sun.com/developer/onlineTraining/java3d/`
20. Selman, D.: Java 3D Programming Ed. Manning Publication Co. (2002)
21. JLayer - MP3 library (September 2011), `http://www.javazoom.net/javalayer/sources.html`

Natural Feature Tracking Augmented Reality for On-Site Assembly Assistance Systems

Rafael Radkowski and James Oliver

Iowa State University, Virtual Reality Applications Center,
1620 Howe Hall, Ames, IA, 50011, USA
{rafael,oliver}@iastate.edu

Abstract. We introduce a natural feature tracking approach that facilitates the tracking of rigid objects for an on-site assembly assistance system. The tracking system must track multiple circuit boards without added fiducial markers, and they are manipulated by the user. We use a common SIFT feature matching detector enhanced with a probability search. This search estimates how likely a set of query descriptors belongs to a particular object. The method was realized and tested. The results show that the probability search enhanced the identification of different circuit boards.

Keywords: Augmented Reality, Natural Feature Tracking, Assembly Assistance.

1 Introduction

An assembly assistance system is a computer terminal, which provides assembly work instructions such as the assembly sequence, the components needed for a product, the handling of tools, etc. They are located at assembly stations on a factory floor and are commonly used in a variety of industries. These systems are critical for novice assemblers who typically refer to them regularly. However, even experienced assemblers are required to use assembly assistance systems because product variants are difficult to memorize and these systems are also used to track production efficiency. Most assembly assistance systems are comprised of simple alphanumeric lists of instructions with perhaps links to associated 2D schematic drawings. To enhance the effectiveness of such systems, we developed an Augmented Reality (AR) assembly assistance system for a major manufacturer of electrical components that superimposes a live video image of a manual assembly station with 3D models, 2D texts, and annotations. It tracks the parts to assemble, shows the assembly sequence, and provides information about the assembly method.

The assembly assistance system must track multiple rigid objects; in our case, planar printed circuit boards that are manipulated by the user (Figure 1). Therefore, we have developed a natural feature tracking (NFT) system. It relies on the so-called SIFT feature tracker [1]: feature maps are created that represent the objects to track. To identify and track an object, features need to be identified in a video stream and

R. Shumaker (Ed.): VAMR/HCII 2013, Part II, LNCS 8022, pp. 281–290, 2013.

compared with the stored feature database. This and similar techniques are well known for image tracking. For example, today they are employed in magazines and newspapers.

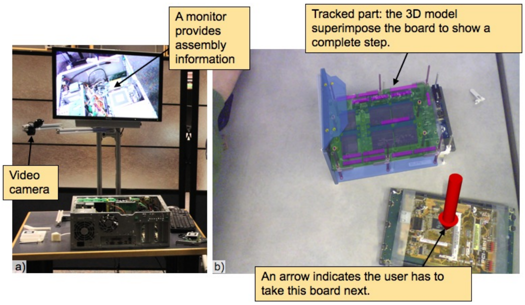


Fig. 1. a) A prototype of the AR on-site assistance system. b) The application tracks multiple circuit boards and adds assembly instructions.

Despite its ubiquity, the SIFT method presents several challenges when tracking multiple circuit boards. In general, circuit boards are difficult to track: they look similar to on another, provide only a limited number of good feature points, and they are 3D objects, even though they are flat.

Our contribution is a method to distinguish multiple circuit boards and to identify the related feature map in real time as well as an optimized feature map structure for rigid object feature maps. We employ a probability search that estimates how likely a set of given query features belongs to a particular circuit board. The method utilizes statistical similarity of features and clustering methods to calculate a probability value. We use a tree data structure to compare all relevant feature maps on a different level of complexity.

The next section reviews the relevant related work that drives our method. In section 3, we present our realization of the tracking system and explain the probability search. We present an application example in section 4 and close the paper with a summary and an outlook.

2 Related Work

In general, natural feature tracking (NFT) is a vision-based tracking approach. According to Zhou [2], vision-based tracking can be classified in model-based tracking techniques and feature-based techniques. This classification considers the amount of previous knowledge the tracking system needs to have about the scene. NFT belongs to the feature-based techniques, which relies on natural features. A large share of research is devoted to NFT for AR applications since AR relies on tracking and NFT facilitates the usage of physical objects in the environment. Thus, the review will only highlight some research that fosters our approach.

Lepetit et al. [3] introduced a keypoint-based tracking method that automatically builds different view sets of a training image in order to improve performance and robustness. Multiple keypoints are extracted from these images and stored as a classification database. They use a randomized *kd*-tree to classify the feature points of a sample image. The method works robustly, it facilitates tracking of a wide range of images, and also copes with cluttered and distorted objects. Nevertheless, it is trained for only one object.

Klein et al. [4] has developed a method that simultaneously estimates the pose of a camera and creates a feature map. The idea is to split tracking and mapping into two different threads. This enables the use of computationally expensive optimization methods in order to build an optimized feature map. The approach is robust and works with a large set of keypoints. Nevertheless, it is intended for navigation purposes and cannot identify particular objects.

Chen et al. [5] have developed a keypoint tracking system that copes with different lighting conditions. The authors employ a FAST algorithm to extract keypoint features and descriptors. The descriptors are organized in a *kd*-tree for fast keypoint retrieval. To improve the robustness, a Kanade-Lucas-Tomasi (KLT) tracker [6] has been added that delivers additional information for pose estimation. This enhances the probability of obtaining good features to track. The method utilizes an additional matching algorithm to improve the robustness. Nevertheless, their method does not distinguish different objects.

Cagalaban et al. [7] introduce a tracking method that allows tracking of multiple 3D objects in unprepared environments. The method incorporates a KLT tracking and color tracking to detect multiple moving objects. However, the authors' test objects were relatively simple (cars), object segmentation relies on background separation and the tracked objects cannot be identified.

Uchiyama et al. [8] present a tracking method that relies on a method called locally likely arrangement hashing. The authors intend to track 2D maps, which are difficult to track because the arrangement of a map looks similar from different viewpoints. Their tracking approach utilizes the intersections on maps to retrieve a robust feature map. In addition, the authors use online learning to be able to cover a large map.

The Fours Eyes Lab conducted research in keypoint optimization and keypoint selection in order to optimize the keypoint database in such a way that only the best, most robust, features maintain tracking (i.e., [9], [10]). For instance, they explore the effect of different texture characteristics on tracking. They also evaluate the influence of different tracking parameters using a large database of 2D images that show different light conditions and geometric changes. Their research is aimed at developing a robust tracking system. Nevertheless, the research does not address circuit boards, particularly, nor in particular, the problems associated with tracking physical tracking targets with keypoints.

3 Natural Feature Tracking for Circuit Boards

The objective of the tracking system is to determine the pose of a video camera to enable spatial registration for AR. Our tracking system relies, like many others, on matching keypoints from training images with keypoints obtained from a run-time video stream. Usually, all training keypoints are stored in one database, and the query set is matched against this database. The challenge, when tracking physical objects like circuit boards, is to receive a sufficient number of correctly matched feature points to enable pose estimation and tracking. Since circuit boards roughly look similar, the number of false matches increases with the overall number of features. In this section, we first explain the suggested tracking method. Afterwards, we introduce our feature map optimization strategy.

3.1 Feature Point Tracking and Matching

Figure 2 presents an overview of the natural feature tracking method. The method implements the functions keypoint extractions, descriptor computation, descriptor matching, and pose estimation. Our implementation is based on OpenCV [11], an open source computer vision library that provides the required core image processing functionality. In addition, a rendering function generates the output image. However, since the rendering function is not part of part of the tracking system, it is not explained.

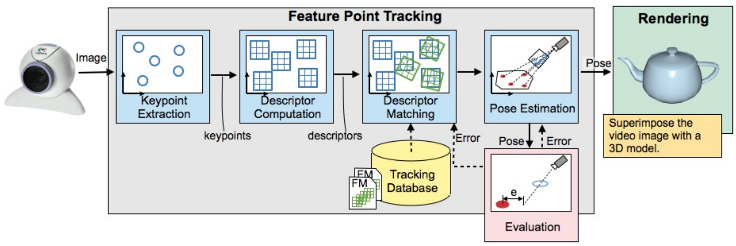


Fig. 2. Overview about the entire natural feature tracking system

Consider a feature map $F_i = \{k_1, ..., k_N | d_0...d_N\}$, with N keypoints $k \in K_i$ and N associated descriptors $d \in D_i$. Each set F_i enables tracking of a physical object O_i, with i, the index of the object, also referred to as tracking targets. We use SIFT feature points and descriptors [1]. A keypoint describes the location of a feature. The descriptor is a vector with 128 values that represent the magnitude and the orientation of gradient vectors that surrounding the keypoint. A training database DB_{ref} with $DB_{ref} = \{F_0, ..., F_N\}$ stores all feature keypoints K_N and descriptors D_N of all physical objects to track.

For object tracking, let $I(x, y)$ be the input image fetched from a video camera. First, we identify a set of keypoints K^* and extract keypoint descriptors D^* in I. Initially, the keypoints and descriptors are unidentified; they may belong to one or to multiple tracking targets. To identify them, descriptors D^* are matched against the reference descriptors $D_N \in DB_{ref}$. A k-nearest-neighbors method (KNN) is a common way to match data. The descriptors in DB_{ref} are organized as a randomized *kd*-tree [12]. A *kd*-tree is a space-partitioning data structure [13]. It splits *k*-dimensional data into half-spaces considering the variance of each dimension. Randomized *kd*-trees use a limited number of dimensions to split the state-space of data. The *kd*-tree is trained in advance. The nearest neighbors are retrieved by traversing the *kd*-tree.

We added a probability search in order to facilitate the tracking of circuit boards. During our work on the assembly assistance station, with increasing number of descriptors in the database we encountered too many false-matches. The probability value is added in order to eliminate a set of descriptors. Thus, the number of descriptors that need to be considered is reduced. Therefore, we use a limited number of descriptors with a limited complexity. Figure 3depicts an abstraction of the tree structure spaces utilized. The nodes carry the descriptors $d_{i,j}$, with i, the tracking target identifier, and j, the descriptor number. Each node splits the search space into half-spaces. The upper nodes of the tree use the probability search. A probability value estimates how likely a descriptor belongs to a particular feature map.

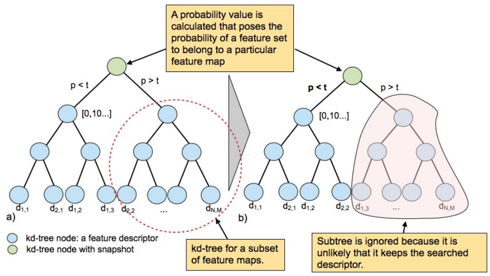


Fig. 3. kd-trees are used for a fast k-nearest neighbor search. An additional probability search allows determination of a subset of descriptors that likely belong to the searched feature map.

Let d_t^* be the query descriptor that needs to be matched. First, a probability value is calculated. We compare the vectors of each descriptor with a feature map signature $S = \{r_0, r_1, ..., r_N\}$ and the signature centroid Sc_0 and Sc_1. The signature is a composed descriptor vector that contains the most robust gradient vector direction and magnitude ranges $r = \{r_{min}, r_{max}\}$ of a feature map. The centroids are the center of the majority of descriptor directions. The probability value is calculated using:

$$\forall d_t^* \in D^* : \Delta s = min(\frac{|Sc(d_t^*)_i^* - Sc_0|}{var(S)}, \frac{|Sc(d_t^*)_i^* - Sc_1|}{var(S)}) \tag{1}$$

with Δs, the distance value, and Sc_i, the centroid of the descriptor d_t^*, and $var(S)$, the variance of the signature. The calculation is carried out for all feature maps in the database. Using this approach, one descriptor can be associated to more than one feature map. Thus, we consider all query descriptors to calculate the probability value by counting the possible assignment of a descriptor to a particular feature map:

$$P_i = \frac{1}{N} \sum_{j=0}^{N} k| \begin{cases} k = 1 & if \Delta S_i \in FM_i \\ k = 0 & if \Delta S_i \notin FM_i \end{cases} \tag{2}$$

With *P* the probability value, *N*, the number of descriptors. This method facilitates the determination of outliers. All descriptors of the training database that belong to a feature map with a low probability value are not considered in the following step.

The probability search is carried out only when the ratio between matched and unmatched descriptors of a scene is below a threshold *T*. We empirically determined a $T = 0.4$ considering three circuit boards at once in the scene. If a major set of descriptors cannot be matched, this ratio indicates that a new circuit board is in the scene. As long as the tracking quality is acceptable we assume that only the identified subset is needed. The *kd*-tree was adapted to consider the probability value; we remove the nodes and retrain the sub-tree.

To match the feature d_t^*, we employ a *k*-nearest-neighbor (KNN) matching to find the best match in the training database [14]. The KNN method is a non-parametric method. It calculates *k*-distances for each vector of the input data to the reference data, where *k* represents the number of neighbors the method returns when calculating the output distance. We calculate the k=2 distances, thus, for each query descriptor we find the two best matches in the reference dataset D_{ref}:

$$dist_i(d_i, d_{ref}) = \sqrt{(d_{i,1} - d_{ref,1})^2 + \cdots + (d_{i,N} - d_{ref,N})^2} \tag{3}$$

$$y = \frac{dist_1 + dist_2}{2} \tag{4}$$

with $dist_i$, the distance between each input descriptor and reference descriptor, and *y*, the final output calculated from the two best matches $dist_1$ and $dist_2$. The KNN method returns the two nearest neighbors for each query descriptor; the matching set is denoted as *M*. We utilize the OpenCV implementation of the Fast Library for Approximate Nearest Neighbors for KNN matching [15].

Next, a ratio test is employed to find the best matches of the entire output set *M*. The ratio test checks whether the matches found violate a threshold [6]:

$$r < r_t \tag{5}$$

$$r = \frac{dist_1}{dist_2} \tag{6}$$

With *r*, the ratio, and r_t, the threshold. Usually, the ratio ranges from 0.4 to 0.6. Only high quality matches pass this test. All other matches are deleted.

Finally, we employ an epipolar search to find feature points that meet the fundamental epipolar constraint of a 3D projection: all keypoints of the query set must lie on the epipolar line of the reference keypoint set. We use an 8-point RANSAC algorithm to calculate them. The RANSAC algorithm is an iterative technique to estimate the parameters of a model from a set of given input points [16]. The inliers of the RANSAC test comply with this constraint and are used to estimate the pose.

To estimate the pose, we calculate the extrinsic camera parameters using a textbook technique, namely, Direct Linear Transformation (DLT), solved with Singular Value Decomposition [15]. The DLT uses corresponding keypoints of the tracking targets and the image to solve a camera model equation. Figure 4 presents a result. The left image shows the training image and its keypoints. The right image depicts the video stream and the matching keypoints. The lines indicate the corresponding keypoints.

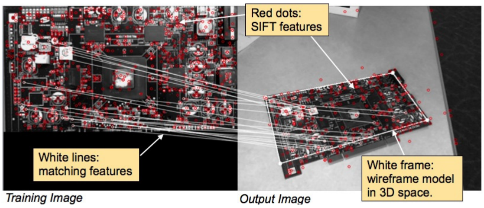


Fig. 4. Tracking of a circuit board with SIFT feature tracking

3.2 Feature Map Signature

The feature map signature is a vector with the most robust descriptors. It facilitates fast recognition and classification of feature descriptors, which are determined by counting. A reference image is used and its feature points and descriptors are extracted. To get query images, multiple affine transformations are applied. Also noise is added to several images and the resolution is changed. The feature descriptors of these images are extracted and matched against the reference image descriptors. Each match in the reference image is counted. After N samples (N=400 for our tests) a matching count for each reference descriptor available. For the further processing, we select 20% of the most matching features when they occur in 50% of all frames.

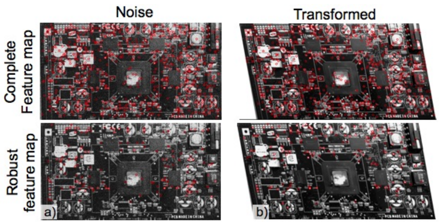


Fig. 5. Feature maps after optimization

The idea of the probability search is to determine a probability value for each query descriptor considering the majority of all robust descriptors of a feature map and their variance. Therefore, we calculate a signature and the feature map centroids. Both are determined using a k-means clustering method. Let $D = \{D_0, D_1, D_2, ..., D_N\}$ be the set of all robust descriptors with $D_i = \{d_0, d_1, d_2, ..., d_{127}\}$. First we cluster the descriptors with the k-mean clustering with k=2:

$$argmin \sum_{i=1}^{k} \sum_{d_j \in D} |d_j - \mu_i|^2 \tag{8}$$

with μ, the mean of the descriptor values. The clustering results in a set of descriptors, which are most similar, and the cluster centroids. Figure 6 shows a result generated from three circuit boards[1].

[1] The circuit boards used for testing are computer graphics cards, network adapters, etc. The application has been developed for circuit boards of controllers. The board layout cannot be published due to a confidential declaration.

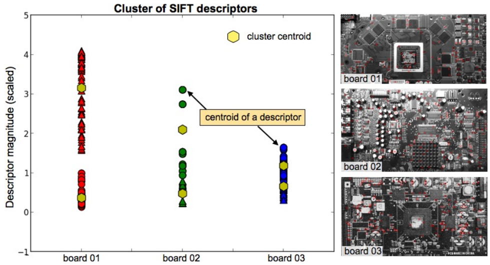


Fig. 6. Cluster of centroids of descriptor magnitudes

This approach relies on different cluster centroids and different descriptor distributions that result in different variances. Thus, we can use a statistical approach to anticipate a best match. As shown in Figure 6, the clusters are distinguishable. Each of the clusters poses one signature *S*. The centroids are used as mean values μ for each feature map. As presented before, Equation 1 and 2 are used to calculate the probability of new descriptors to belong to a particular feature map.

The signature S of a feature map is a 2xN vector that keeps the range of the majority of all descriptor values per descriptor element:

$$S = \{s_0, s_1, s_2, \ldots, s_N\},\ with\ s_i = (d_{max}, d_{min}) \tag{9}$$

with d_{max} and d_{min}, the min and max values for every descriptor element d_i.

4 Results

We compared the tracking performance with a regular natural feature tracking that works without a probability search. The same SIFT feature tracker was applied and the *k*-nearest neighbor method was used for feature matching. For repeatability, the tests have been carried out with testing software and images of circuit boards and not with a live video and manual operated camera. We used eight computer-related circuit boards (graphics card, network adapter etc.) as reference data. The query dataset used eight images of the same boards plus eight additional colored advertisements from journals. All query images have been automatically transformed (size, rotation, perspective, shear) in order to get five query images for each sample. All sample images have been matched against the training database. The tests have been carried out on a Dell Precision T3500 computer, with an Intel Xeon 3.47 GHz processor, NVidia Quadro 5000 GPU, and 6GB RAM.

Table 1 shows the results. The first column shows the number of tracking targets in the database, the second the overall number of descriptors in the database. The next three columns, the percentage of matched features and the percentage of false matched features for the standard KNN algorithm. The ratio is the quotient of false to all matches. The last three rows show the same for suggested method.

Table 1. Results from feature matching experiments

No. DB	Overall descriptors	KNN	ratio	KNN false	KNN + P	KNN + p false	ratio
1	145	95%	0%	0%	94%	0%	0%
2	538	94%	5%	5%	90%	4%	4%
4	2211	75%	59%	44%	80%	13%	16%
6	3689	72%	76%	55%	79%	10%	13%
8	4946	82%	60%	49%	82%	7%	9%

The results show that the optimized descriptor matching algorithm yields a higher matching ratio when the number of tracking targets, i.e., the number of descriptors in the database, is increased. The usual matching algorithm provides too many false matches, matches that belong to different tracking targets. The subsequent applied data tests cannot remove all of these false matches. Thus, the pose estimation does not work correctly. The probability method helps to remove data from the database before the KNN matching is applied. Thus, the number of wrong matches were reduced.

However, we also used both methods to track regular 2D color images. In this case that the suggested algorithm does not yield significant advantage. Regular color images produce many more feature descriptors than circuit boards do. The feature descriptors are also more distributed; a major descriptor direction often does not exist. Thus, the *k*-means clustering cannot identify descriptor clusters because the variance of descriptors is similar. In general, circuit boards provide fewer good feature descriptors than color images do. Thus we assume that cluster centroids and a majority of descriptor directions and magnitudes can be identified.

5 Summary and Outlook

This paper introduces an augmented reality on-site assembly assistance system that helps assemblers during the assembly process. The system identifies and tracks the objects to assemble and provides assistance if necessary. For our purpose, the assembly assistance system needs to track and to identify physical planar circuit boards. We encountered problems with regular approaches when more than two circuit boards must be tracked. Circuit boards do not provide a sufficient number of distinguishable features points. Thus, we enhanced the identification probability by introducing a probability value. It relies on the sparse descriptor set of circuit boards: it is possible to identify clusters and to distinguish boards using the cluster centroids. It can be used to enhance the probability of identifying the correct feature keypoints and, finally, the correct tracking target.

However, the method is limited to tracking targets with a sparse descriptor set. We conducted a limited set of experiments. We know, the approach works for our system and the product we need to track. In the future, we seek to analyze the overall capability of the approach presented. Therefore, we will work on a larger dataset for training and query.

References

1. Lowe, D.G.: Distinctive image features from scale-invariant keypoints. International Journal of Computer Vision 60(2), 91–110 (2004)
2. Zhou, F., Duh, H., Billinghurst, M.: Trends in Augmented Reality Tracking Interaction and Display: A Review of Ten Years of ISMAR. In: International Symposium on Mixed and Augmented Reality (ISMAR 2008), Cambridge, UK, vol. 2008, p. 10 (July 2008)
3. Lepetit, V., Fua, P.: Keypoint Recognition using Randomized Trees. Transactions on Pattern Analysis and Machine Intelligence 28(9), 1465–1479 (2006)
4. Klein, G., Murray, D.: Parallel Tracking and Mapping for Small AR Workspaces. In: Proc. International Symposium on Mixed and Augmented Reality (ISMAR 2007, Nara) (2007)
5. Chen, Z., Li, X.: Tracking based on Natural Feature for Augmented Reality. In: International Coriference on Educational and Information Technology, ICEIT 2010 (2010)
6. Laganière, R.: OpenCV Computer Vision. Packt-Publishing, Birmingham (2011)
7. Cagalaban, G., Kim, S.: Multiple Object Tracking in Unprepared Environments Using Combined Feature for Augmented Reality Applications. In: Kim, T.-h., Chang, A.C.-C., Li, M., Rong, C., Patrikakis, C.Z., Ślęzak, D. (eds.) FGCN 2010. Communications in Computer and Information Science, vol. 119, pp. 1–9. Springer, Heidelberg (2010)
8. Uchiyama, H., Saito, H., Servières, M., Moreau, G.: Camera tracking by online learning of keypoint arrangements using LLAH in augmented reality applications. In: Virtual Reality, vol. 15(2-3), pp. 109–117 (2011)
9. Gruber, L., Zollmann, S., Wagner, D., Schmalstieg, D., Höllerer, T.: Optimization of Target Objects for Natural Feature Tracking. In: Proc. IAPR/IEEE ICPR (20th International Conference on Pattern Recognition), 2010, Istanbul, Turkey, August 23–26, pp. 3607–3610 (2010)
10. Gauglitz, S., Höllerer, T., Turk, M.: Evaluation of Interest Point Detectors and Feature Descriptors for Visual Tracking. International Journal of Computer Vision (IJCV) 94(3), 335–360 (2011)
11. Webpage of the OpenCV Wiki: last seen: (February 22, 2013), `http://opencv.willowgarage.com/wiki/`
12. Silpa-Anan, C., Hartley, R.: Localization using an imagemap. In: Australasian Conference on Robotics and Automation, Australia, December 6 - 8 (2004)
13. Freidman, J.H., Bentley, J.L., Finkel, R.A.: An algorithm for finding best matches in logarithmic expected time. ACM Trans. Math. Softw. 3, 209–226 (1977)
14. Jacob, E.G., O'Rourke, J., Indyk, P.(ed.): Nearest neighbours in high-dimensional spaces. In: Handbook of Discrete and Computational Geometry, 2nd edn. CRC Press (2004)
15. Muja, M., Lowe, D.G.: Fast Approximate Nearest Neighbors with Automatic Algorithm Configuration. In: International Conference on Computer Vision Theory and Applications, VISAPP 2009 (2009)
16. Fischler, M.A., Bolles, R.C.: Random Sample Consensus: A Paradigm for Model Fitting with Applications to Image Analysis and Automated Cartography. Comm. of the ACM 24(6), 381–395 (1981)

Augmented Reality Interactive System to Support Space Planning Activities

Guido Maria Re, Giandomenico Caruso, and Monica Bordegoni

Dipartimento di Meccanica, Politecnico di Milano, Via La Masa, 1
20156 Milano, Italy
guido.re@mail.polimi.it,
{giandomenico.caruso,monica.bordegoni}@polimi.it

Abstract. The Space Planning (SP) is a process that allows making an environment more ergonomic, functional and aesthetically pleasing. The introduction of Computer Aided tools for this kind of practice led to an increase of the quality of the final result thanks to some versatile support used for the generation of different options to consider for the evaluation. In particular, those based on Augmented Reality (AR) technologies allow evaluating several options directly in a real room. In this paper, an AR system, developed with the aim of supporting Space Planning activities, is proposed. The system has been developed in order to overcome some problems related to the tracking in wide environments and to be usable in different typologies of Space Planning environments. The paper also presents a qualitative evaluation of the AR system in three different scenarios. The positive results obtained through these evaluation tests show the effectiveness and the suitability of the system in different Space Planning contexts.

Keywords: Augmented Reality, Space Planning design, HCI.

1 Introduction

The Space Planning (SP) is a process that leads to arrange buildings, rooms, factories or other generic environments in a practical manner. Very often, this process involves more than a single category of specialists (interior designers, architects, ergonomists, engineers, etc.), who have to collaborate together to define an optimal layout of furniture and equipment, so as to correctly arrange an environment. The definition of an optimal layout implies taking into account several aspects, which depends on the environment that has to be arranged. In a working space, for instance, safety and ergonomics are issues that have to be assessed to guarantee liveability to the employees. Usually, these aspects are regulated by specific guidelines that allow correctly defining the right furniture and equipment positioning [9]. More in general, skilled professionals follow several other guidelines to optimize the furniture layout. These guidelines have been elaborated, over the years, starting from the personal experience and often are regulated by the common sense [11, 18].

R. Shumaker (Ed.): VAMR/HCII 2013, Part II, LNCS 8022, pp. 291–300, 2013.

Nowadays, this process is supported by Computer-Aided tools that allow the easy rearranging of virtual furniture and equipment in a virtual environment. Simplest tools, which are based on bi-dimensional drawings, can be useful to assess some of the above-mentioned criteria. Actually, it is possible to simply evaluate the encumbrance, the position and the orientation of each object with respect to the dimensions of the room. However, this simplified representation does not allow correctly evaluating other criteria, such as the aesthetic impact of the final layout.

3D modelling and rendering techniques, instead, give the possibility of improving the virtual representation both of the objects and of the surrounding environment, where the objects have to be arranged. The result of this virtual simulation is very effective and the visualization of the different layouts provides a high level of realism. However, making this kind of virtual simulation is a time-consuming process and requires specific skills. Furthermore, while the 3D models of objects can be imported from pre-existent databases, often, the surrounding environment has to be modelled from scratch. Consequently, the time needed for modelling the surrounding environment is a task that requires more time than that required for the arrangement of the environment [4].

By using Augmented Reality (AR) technology is possible to overcome this issue, since it allows the user to see virtual objects overlaying the real scene [1]. Therefore, AR has been already used to develop some applications for SP, in order to enable the direct rearrangement of the virtual objects in the real environment. [15] is one of the first examples and it grounds on the augmentation of a housing space by static pictures previously taken. This simple approach of working with pictures has been also used in industrial contexts to plan the disposition of machineries in a factory [5, 13].

The use of the real environment increases the comprehension of the final result as well. In fact, if the virtual objects arrangement is displayed in the real world, no abstraction activities are required to imagine the final results. In addition, the better understanding of the final result allows the user to directly interact with the environment to be planned and to be fully involved in the arrangement of the spaces. In this way, every kind of user can participate in the design process even if he/she has no skills in SP and in 3D modelling. This is the case of customers and final users for whom some solutions based on Mobile AR are under investigation [10, 16].

However, performing SP activities in an AR environment implies using systems that allow the tracking of the whole working area. In this way, it is possible to correctly place virtual objects according to the real environment. High-accuracy systems, such as the optical ones, are able to cover an entire room but require the installations of several and expensive devices. Cheaper tracking systems, instead, use Computer Vision (CV) algorithms to detect the camera. Some of these algorithms estimate the camera pose by detecting natural features in the environment [3] but they cannot work in case of completely empty spaces. Another more reliable solution is by adding known objects, such as fiducial markers, within the real scene. However, also in this case, the tracked environment requires to be structured by positioning and calibrating several markers in the working area [7, 2]. This activity is time-consuming and not versatile.

This research proposes the use of an AR system integrating marker-based tracking with the tracking ability of a commercial mobile robot. The robot can be seen as a mobile point of reference, which is automatically controlled by means of the device used to visualize the AR scene. A fiducial marker has been placed on the robot to maintain the marker always traceable. Consequently, starting from an absolute reference point, the AR system is always able to track its position in the space, by moving also the robot within the environment. Several authors have investigated the use of mobile robots in AR environment for different purposes [19, 6, 8]. However, in none of these works the tracking ability of the robot is used to extend the working area of the tracking techniques based on fiducial marker.

In order to exploit the characteristics of this AR system, the authors have developed a software application to support the SP activities. This software application enables the user to work with virtual objects located in the real scene. By means of a Graphic User Interface (GUI), the user can see the real environment and manage in real time the virtual objects that are placed within the real environment.

The paper presents also the description of some evaluation tests in order to validate the versatility and the usability of the AR system in different SP scenarios. During these evaluation tests, the users had to arrange different objects and in different environments. In particular, the scenarios, which have been analysed, concern the field of interior design, the setting up of an exhibition and for configuration of the machineries to be placed in a factory.

2 Description of the AR System

The design of the whole AR system grounds on a versatile and cost-effective architecture. The system has been developed by using not encumbrance devices, which are available on the mass market. Consequently, the AR system is quite inexpensive and can be easily transported. The versatility of the AR system allows using it in different kinds of SP contexts and its strong points relate to the ease of installation in different working environments and the ease of managing different virtual objects. The AR system mainly consists of two parts: an AR Interface and a mobile robot, as showed in Fig. 1.

The AR Interface provides all the interactive AR tools to support SP activities. It consists of a laptop and an external USB camera settled on a trolley. In this way it can be easily moved within the augmented environment, thus reducing the workload of the user.

The mobile robot, instead, has been equipped with a fiducial marker, which is used for tracking. The robot manages the position of this marker, placed on its top, with the purpose to extend the AR working area. The mobile robot used in this research is an iRobot Roomba 560[1], which is a commercial mobile robot. The choice of this mobile robot was mainly due to its robustness, its availability on the market and the remote-control easiness. The system uses a tracking approach able to estimate the pose of the

[1] iRobot Roomba 560 – store.irobot.com/product/index.jsp?productId=3881236.

camera in the environment by merging data coming from the mobile robot and the ones coming from the marker-based tracking. The communication between the robot and the laptop has been obtained by using two XBee[2] devices that allow a wireless bidirectional transmission of the data coming from the serial port mounted on the top of the robot to the USB port of the laptop. Thus, the robot position can be remotely controlled by the user either in a manual or in automatic way.

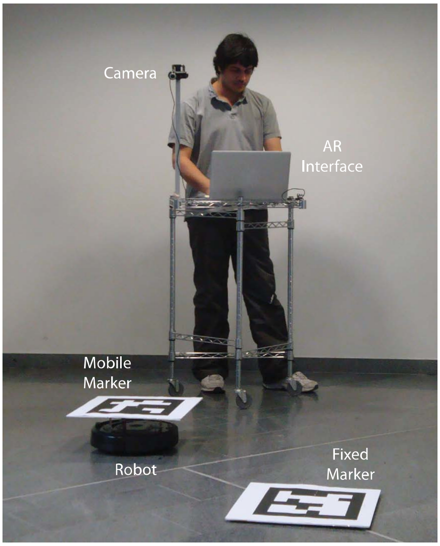


Fig. 1. The components of the AR system

2.1 Tracking Technique

The tracking technique used by the AR system relies on a particular approach that can combine two different kinds of data to estimate the camera pose [14]. These two data refer to the camera pose estimation performed by means of ARToolkit Plus [17] with two squared planar markers of 320mm size and to the position of the robot obtained by its own odometric system. One marker is fixedly placed on the floor of the room (fixed marker) and defines the position of the absolute reference system. In this way, the AR system estimates the camera pose according to the position of this marker. Moreover, the fixed marker is also used to set the initial position of the mobile robot, so as to have the robot position always coherent with the defined reference system. Every time the fixed marker is framed by the AR system, only the data obtained by tracking it by means of a standard marker-based approach are used to estimate the camera pose.

The other marker is placed on the top of the robot (mobile marker) and it can be moved, in an automatic or manual mode, every time the user moves the AR Interface to frame another part of the scene. In this way, this marker can be always visible to

[2] Xbee – Wireless RF Modules, `www.digi.com/xbee/`

the camera. Every time only the mobile marker is in the framed scene, the AR system estimates the camera pose by exploiting also data coming from the robot. Actually, the marker tracking is performed on a mobile support, whose location is known by means of the odometric data, which the robot continuously sends to the AR Interface, to convey its position. In this way, the camera pose is calculated as a linear combination of data from the encoders and the mobile marker tracking.

In order to perform tracking, the AR system includes an initialization phase that consists in calibrating the initial position of the robot according to the reference fixed marker. Since the environment is not structured, the initialization is very fast and takes less than a minute. In order to perform this initialization, the only request is to frame both the markers by the camera. Afterwards, once the reference marker is arranged in the defined position of the working space and the robot is placed on the floor, the AR system automatically performs an auto-calibration and it can be used.

2.2 AR Visualization and Interaction Tools

The AR Interface has been designed to provide the users with two main functions, which enable supporting SP activities. These two functions are the AR visualization and the management of the virtual models directly in the AR environment. The user can see the augmented environment through the main window of the AR Interface. The visualization exploits the above-described tracking approach in order to have spatial and temporal coherence between the real environment and the virtual objects. The user can autonomously manage virtual objects in the real environment by means of a dedicated GUI, which is integrated in the AR Interface. The GUI shows all the available virtual objects, which are stored in a database, by means of a preview window. Once the user has selected and added the desired virtual object through this GUI into the real space, it is automatically placed in the scene and visualised in front of the camera point of view, to a distance of 1.80 metres. Afterwards, the user can change the object position in order to correctly place it in the desired location or in the most consistent one with the defined specifications. This operation is performed by means of six buttons, two for each axis, which enable the user to modify the position and the orientation of the virtual objects according to a step-by-step value.

The system also provides the user with two further functionalities. The first one regards the designed plans of the environments, which are useful to provide a further reference for locating the virtual objects in their specific position. These plans are previously prepared by the user and, once loaded, they are visualised on the ground of the working space. The second functionality regards the virtual light configuration. A correct illumination enhances the coherency between the real and the digital worlds and the level of immersiveness of the augmented environment. In this way, the user's perception of the designed space is improved and the AR interface helps evaluating the aesthetic impact of the virtual object more deeply. The lights settings are externally designed according to the different real illumination of the working environment and the user can load them within the AR scene to assess the visual impact on the virtual objects.

Finally, the AR Interface allows saving the current solution in a file during or at the end of the Space Planning process. In this way, the user can store different configurations and quickly switch from one to another, in order to quickly show the results to other people (e.g., customers) or keep on working on a previous space plan.

3 Evaluation

The main feature of the AR system is the possibility of being used in different SP contexts and, hence, some evaluation sessions have been organized in order to evaluate and validate the usability of the AR system in various scenarios. In particular, the AR system has been tested for the interior design of an apartment, the setting up of an exhibition in a museum and the planning of machineries layout in a factory. Each of these scenarios has allowed the authors to test the AR system with different typologies of virtual objects and in different spaces.

During the evaluation session, 5 users have been involved with the purpose to qualitatively assess the usability of the system. These ones are experts in the context of interior design, exhibit organization and engineering. The evaluation has been performed through specific heuristics, which have been developed starting from the ones proposed by Nielsen [12]. Therefore, it has been asked to the user to fill in a questionnaire after using the AR system, wherein they have to express their opinions with a scale of points from 0 (bad) to 5 (excellent).

3.1 Interior Design Scenario

The interior design scenario has concerned the definition of a furnishing for an apartment about 40 square metres wide. Two users, expert in interior design, have tested the AR system in this scenario. Their tasks concerned the placing of virtual furniture in the apartment, according to a plan. The system provided the visualization of the designed plan on the floor so that the user could have reference points for the disposition of the objects. Fig. 2 shows: one of the two users during the initialization procedure (on the left), the virtual plan, in which the furniture layout is defined (centre), furnishing of a room of the apartment (on the right).

Fig. 2. Example of using the AR system for interior design purposes

3.2 Exhibition Planning Scenario

In the scenario related to the exhibition planning, the test was conducted in an empty area of a museum about 30 square meters wide and the tasks concerned the setting up of an exhibition. The users were two professional exhibits organizers, who have been involved in arranging some statues in the area of the museum. The AR system allows the user to have a preview of the final layout without moving the real statues. Moreover, it is possible to use the system to check if the pieces of art can fit in the space without any troubles or risks to be damaged. The statues are indeed valuable and, often, fragile objects that have to be moved with extremely care. In this way, all the exhibit supervisors can participate in the arrangement of the statues and change also the final configuration, without the risk of damaging the real one. Fig. 3 shows some images taken during the arranging of the statues.

Fig. 3. Example of using the system as support for exhibition planning

3.3 Factory Scenario

The factory scenario consists in defining a functional layout of industrial machineries. In this scenario, an engineer, who is expert in factory planning, has been involved with the aim of testing the system in the evaluation of the machinery layout before buying all the industrial equipment. In particular, the aim of the assessment relates to the encumbrances of the machineries within a factory area about 15 square meters wide.

In this scenario, primitive 3D shapes have been used rather than virtual models of machineries, since the encumbrance evaluation does not need a complete definition of the shape. In addition, it is important to provide the user with the working volume of the machinery that usually does not coincide with the size of the machine at rest (as in case of a robotic arm). For these reasons, during the test, the whole functional encumbrance has been represented by using simple parallelepipeds with different colours. Fig. 4 shows the user during the execution of the definition of the machinery layout.

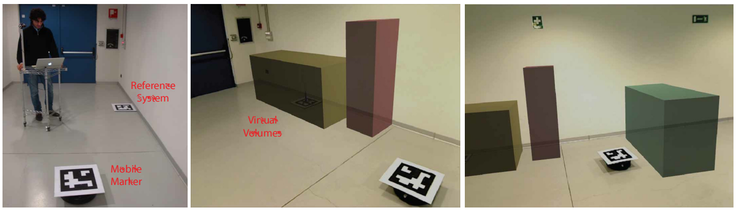


Fig. 4. Example of using the system for encumbrance evaluation in a factory

4 Discussion

The proposed evaluation scenarios have demonstrated that the AR system can be effectively used in different fields of SP. In particular, the system enables the user to overcome some issues regarding the tracking. The tracking solution adopted turned out to be advantageous in terms of ease of installation and usability. This solution has allowed the user to avoid the typical time-consuming phase of initialization and setup of the traditional tracking methods. In addition, the precision of the tracking solution adopted is reasonably acceptable for SP planning purposes and it represents a good compromise between precision and costs.

Moreover, the AR system resulted to be an effective interactive device for SP in the different scenarios analysed without any distinction. Actually all its functionalities worked well in all of the three environments analysed without any problem concerning the typologies of objects involved and their visualization.

The heuristics elaborated to assess the usability of the AR system show positive results, as shown in Tab.1. The AR Interface designed to visualize and manage the virtual objects turned out to be effective for the designers and engineers that tested the application (Q2, Q7 and Q8). Concerning the visualization, the AR Interface provides a good level of integration of the virtual objects with the real environment (Q1 and Q3). Moreover, it comes up with an easy method to manage the objects, which is also easy to learn (Q5). As a matter of fact, each user was able to autonomously use the AR system after few minutes (Q4 and Q6).

However, some users noticed the lack of some common tools that are present in the traditional VR systems for modelling. The first one is the possibility of selecting and moving different objects together. This is particularly evident in the case of interior design, where the designer has to deal with objects that are grouped for function (as for instance, the chairs and the table in a living room or the night table and the bed in a bedroom). The second one is related to the possibility of having a snap tool to help the user placing the virtual object in the space. In this way, it would be easier to attach to objects together, as in the case of the modules of which a kitchen is made up.

Table 1. Result of the questionnaires for the three scenarios

Questions	Interior Design	Exhibition	Factory
Q1. Level of integration of the virtual objects	4	3.5	4
Q2. Effectiveness for layout evaluation	3.5	4	5
Q3. Effectiveness of the whole AR visualization	4	3.5	4
Q4. Easiness and intuitiveness	3.5	4	4
Q5. Effectiveness for managing of the virtual objects	4	4.5	4
Q6. Learnability of the positioning methodology	4.5	4	5
Q7. Overall comfort	4	4.5	4
Q8. Satisfaction in using the AR system	4	4.5	4

5 Conclusion

AR can be a useful tool for Space Planning, since it allows the user to visualize in the real environment the final layout before buying or installing the objects inside. This kind of approach provides an easier comprehension of how the space will appear by using a traditional designed plan. The user, in fact, has to perform a minor abstraction effort to understand the result. In addition, an AR system provides a better interaction solution compared with the VR systems since the objects are located in real time in the real world.

The AR system proposed in this research provides a user-friendly tracking solution for the end user, who is able to cover a wide working area by using a mobile robot, avoiding the time-consuming procedure to structure the environment. Moreover, the system provides interactive software to easily manage virtual objects to effectively arrange them in the space. In this way, the designed system turned out to be versatile since it is able to provide an AR support, which is easy to install in various Space Planning scenarios and it can deal with different typologies of working environments and objects.

Therefore, some evaluation sessions with expert users have been carried out in order to validate the system. In particular, the AR system has been validated in three different scenarios: the interior design of an apartment, the configuration of the layout of machineries in a factory and the design of the arrangement of statues for an exhibition. The system was easy to install and it worked well in all the three scenarios. Thanks to the feedbacks provided by the users through a questionnaire, it is possible to state that the AR system provides an easy interactive approach with different typologies of virtual object in an effective manner.

The authors intend to continue investigating this field of research and, in particular, they will study how to improve the interactivity of the system, in order to make it more suitable for users in the space design sector and common users as well. For this reason, new techniques to manage the virtual objects will be developed, in particular by proposing new interactive metaphors. Finally, some quantitative tests with users will be carried out in order to assess the usability and interactive performances of the proposed system in different Space Planning scenarios.

References

1. Azuma, R.T.: A survey of augmented reality. Presence 6(4), 355–385 (1997)
2. Baratoff, G., Neubeck, A., Regenbrecht, H.: Interactive multi-marker calibration for augmented reality applications. In: International Symposium on Mixed and Augmented Reality, pp. 107–116 (2002)
3. Castle, R., Klein, G., Murray, D.W.: Wide-area augmented reality using camera tracking and mapping in multiple regions. Computer Vision and Image Understanding 115(6), 854–867 (2011)
4. Clifford, R., Clark, A., Rogozin, M.: Using augmented reality for rapid prototyping and collaborative design to model 3D buildings. In: Proceedings of the 12th Annual Conference of the New Zealand Chapter of the ACM Special Interest Group on Computer-Human Interaction, pp. 7–10 (2011)

5. Doil, F., Schreiber, W., Alt, T., Patron, C.: Augmented reality for manufacturing planning. In: Proceedings of the Workshop on Virtual Environments, pp. 71–76 (2003)
6. Fiala, M.: A Robot Control and Augmented Reality Interface for Multiple Robots. In: Canadian Conference on Computer and Robot Vision, pp. 31–36 (2009)
7. Galantay, R., Torpus, J., Engeli, M.: 'living-room': interactive, space-oriented augmented reality. In: Proceedings of the 12th Annual ACM International Conference on Multimedia, pp. 64–71 (2004)
8. Hashimoto, S., Ishida, A., Inami, M., Igarash, T.: TouchMe: An Augmented Reality Based Remote Robot Manipulation. In: The 21st International Conference on Artificial Reality and Telexistence (ICAT), pp. 1–6 (2011)
9. ISO-9241-11. Ergonomic requirements for office work with visual display terminals (VTDs) – Part 11: Guidance on usability. International Organizations for Standardization (1998)
10. Kim, H., Reitmayr, G., Woo, W.: IMAF: in situ indoor modeling and annotation framework on mobile phones. Personal and Ubiquitous Computing, 1–12 (2012)
11. Lyons, G.H.: Ten Common Home Decorating Mistakes and How to Avoid Them (2008)
12. Nielsen, J., Molich, R.: Heuristic evaluation of user interfaces. In: Proceedings of the SIGCHI Conference on Human Factors in Computing Systems, pp. 249–256 (1990)
13. Pentenrieder, K., Bade, C., Doil, F., Meier, P.: Augmented Reality-based factory planning - an application tailored to industrial needs. In: IEEE and ACM International Symposium on Mixed and Augmented Reality, pp. 1–9 (2007)
14. Re, G.M., Caruso, G., Belluco, P., Bordegoni, M.: Hybrid technique to support the tracking in unstructured augmented reality environments. In: Proceedings of the ASME International Design Engineering Technical Conferences & Computers and Information in Engineering Conference IDETC/CIE, pp. 1–10 (2012)
15. Siltanen, S., Woodward, C.: Augmented interiors with digital camera images. ACM International Conference Proceeding Series 169, 33–36 (2006)
16. Sukan, M., Feiner, S., Tversky, B., Energin, S.: Quick Viewpoint Switching for Manipulating Virtual Objects in Hand-Held Augmented Reality using Stored Snapshots. In: IEEE International Symposium on Mixed and Augmented Reality, pp. 217–226 (2012)
17. Wagner, D., Reitmayr, G., Mulloni, A., Drummond, T., Schmalstieg, D.: Real-time detection and tracking for augmented reality on mobile phones. IEEE Transactions on Visualization and Computer Graphics 16(3), 355–368 (2010)
18. Ward, L.: Use What You Have Decorating. Perigee Trade 240 (1999)
19. Young, J.E., Xin, M., Sharlin, E.: Robot expressionism through cartooning. In: Proceedings of the ACM/IEEE International Conference on Human-Robot Interaction, pp. 309–316 (2007)

Empirical Investigation of Transferring Cockpit Interactions from Virtual to Real-Life Environments

Diana Reich and Elisabeth Dittrich

Research Training Group prometei, Berlin, Germany
{diana.reich,elisabeth.dittrich}@zmms.tu-berlin.de

Abstract. Human-cockpit interaction is an innovative and promising field of automotive research. Indeed, automakers need to ensure safety and user satisfaction for their cockpit development concepts, if driving and interacting occurs simultaneously. One suggested approach is to evaluate simple cockpit prototypes within virtual test environments. Hybrid prototyping allows a more realistic experience with the prototype in early stages of development. With our research study we focused on important basic parameters within hybrid test environments (e.g. shutter glasses and virtual projected car model) and evaluated their potential of influence. There are no hints to assume that shutter glasses influence user behaviour. Interestingly, we found significant faster task completion times within a virtual projected car model, which indicate that immersive environments increase user performances. In summary, we can suggest hybrid prototyping within immersive test environments for evaluating human-cockpit interactions.

Keywords: human-cockpit interaction, virtual environment, hybrid prototyping.

1 Introduction

There are different research and development concepts in the area of car cockpits (Ablaßmeier, 2009). The „Cockpit of the Future“ is a hot topic and on every automakers lips. Regarding to different automotive literature research, there are three big and influential development streams in the field of innovative human-cockpit systems, beside voice entries: touch (e.g. MyFord Touch), spin control (e.g. BMWs iDrive) and gestures (e.g. Daimlers Dice). Accordingly, different cockpit concepts and interfaces require specific ways of human-cockpit interaction (Broy, 2007). Because of different interaction techniques, divergent cognitive processes and parameters get relevant, e.g. distribution of attention and situation awareness (Jahn et al., 2005). But how can industry, developers and researchers define the gold standard for their systems regarding human abilities?

Not all of these innovative and for developers seemingly usable and intuitive cockpit concepts are in fact user-friendly: while driving and interacting with the cockpit

R. Shumaker (Ed.): VAMR/HCII 2013, Part II, LNCS 8022, pp. 301–309, 2013.

simultaneously difficulties can arise easily. Different studies reported negative ratings and user experiences concerning usability, user satisfaction and distribution of attention while driving and interacting with touch and spin controlled cockpits (e.g. ORF, 2012).

One way to ensure usability for cockpit interfaces and interactions is to ascertain specific user needs prospectively during the conceptual design phase of the cockpit development (Mayhew, 1999). With the usability engineering approach for cockpit development unusable cockpit innovations could be avoided. Therefore virtual test environments (e.g. CAVE) are a good alternative to traditional validation processes during the development process (Krause, 1999). Human-cockpit interactions should be evaluated virtually during different levels of the development process to ensure safety, usability and user-satisfaction. If misuses or dissatisfaction occur while interacting, rightsizing of the virtual car cockpit concept during an early stage of development process will be more easily, compared to advanced development stages with physical car cockpit prototypes (Stark, 2009). Thereby automakers will be more flexible and avoid high costs. However, prospective evaluations with physical prototypes are still underrepresented, because adaptive prototype evaluation is still time-consuming, inflexible and expensive (Wang, 2002). A logical consequence and an innovative solution would be to evaluate simple cockpit prototypes within virtual test environments. This hybrid prototyping approach allows a more realistic experience with the prototype (Stark, 2009).

Now, we are interested in the differences and explanatory variables between user performances and impressions within different test environments (modalities).

2 Research Study

2.1 Task Description and Experimental Design

In order to realize evaluations of human-cockpit interactions within virtual test environments presenting a physical interface prototype, a human-cockpit interaction set-up was simulated by using a tablet with an implemented navigation system interface, which was operated by a spin controller, within a virtual test environment (CAVE). A fully immersive five-sided 2 x 2 m CAVE system was equipped with shutter glasses to give the users stereoscopic views of the car and cockpit. Head-tracking devices mounted on the users' shutter glasses controlled the CAVE viewpoint according to the users' body and head movements. The stereoscopic application displayed a complete model of a car, including e.g. car body, seats, steering wheel, and cockpit (see figure 1).

The user tasks were to enter six predefined addresses (country/city/street) into the navigation system by interacting and manipulating the spin controller. The addresses were announced by the study manager. The navigation system was physically implemented in the virtual car cockpit, when the stereoscopic view was switched on (see figure 1). In the condition without virtual car model the navigation interface was right in front of the users chair within a white 2 x 2 m CAVE (see figure 2).

We realized a completely crossed within-design with 20 participants. The experimental design of the study had two independent variables: shutter glasses (yes/no) and modality (virtual car model vs. no virtual car model) in a within-factorial design with dependent measures of errors, task completion times and subjective data (AttrakDiff).

Twenty participants took part in the study (10 female and 10 male). The average age of the subject group was 26 years, 17 were right-handed and three were left-handed, with good stereopsis, and all with valid driver`s license. Each subject performed the task within the three test environment modalities: (A) no virtual car model; without shutter glasses, (B) no virtual car model; with shutter glasses and (C) virtual car model; with shutter glasses. We didn`t realize a complete test design, because the hypothetic modality (D) virtual car; without shutter glasses, wouldn`t make any sense.

The dependent variables were used to analyze the differences with regard to performance and impressions between the different modalities (A, B, C).

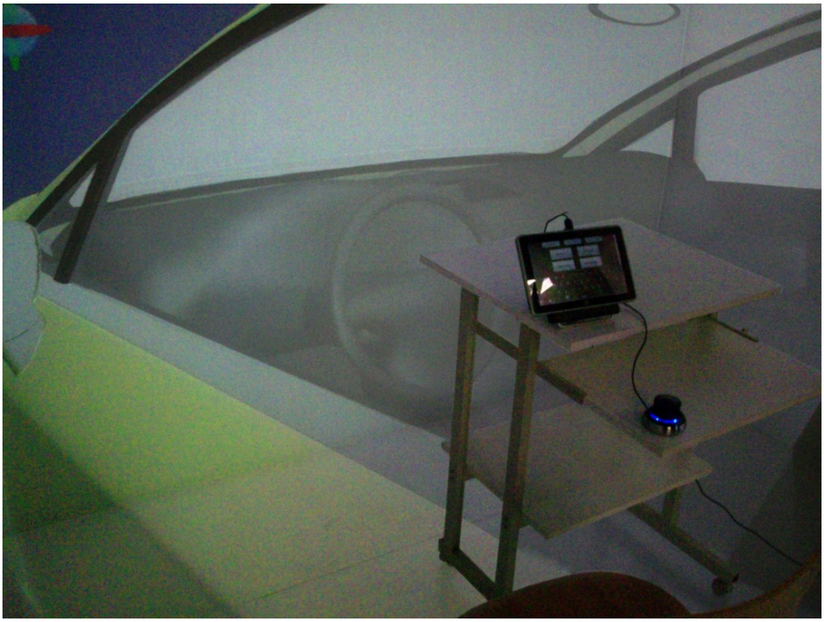


Fig. 1. Virtual car model with physical navigation interface and spin controller

Fig. 2. Participant with shutter glasses within a non-immersive test environment

2.2 Research Questions

The objective of this empirical investigation was to develop a hybrid test scenario, involving a virtual projected car model (car body and car cockpit) within a CAVE and a user interaction task with a physical interface combined with a spin controller, to answer the following research questions:

- Do shutter glasses influence users' behavior per se? Because they are slightly darkened and shutter alternating with 60Hz on each eye. The following sub-questions were asked to analyze the potential influence:
  - Do users need more time for data entries with shutter glasses?
  - Do users make more mistakes with shutter glasses?
  - Do users get different impressions with shutter glasses, compared to no shutter glasses?
- Does the virtual simulated car model influence users' behaviour? Because there will be a more realistic (immersive) test scenario within the virtual projected car, compared to a white box containing one single chair, little white table, tablet interface and spin controller. Although, users need to cognitively process more parameters of perception, majorly visual and motion parameters. The following sub-questions were asked to analyze the potential influence of the virtual car model:
  - Do users need more or less time within a virtual car model?
  - Do users make more or less mistakes within a virtual car model?
  - Do users get different impressions within a virtual car model, compared to no virtual influences?

An additional research question was prospected beside the main study intention:

- How do users rate the usability of the implemented navigation interface?

2.3 Procedure

First, the participants were required to fill out a demographic questionnaire and had to pass a stereopsis eye-test on a table outside the CAVE. After that, they were placed in the CAVE with a comfortable distance to the interface and the spin controller. Afterwards the participants had the chance to try out the navigation system by operating the spin controller.

Then the main experiment started with the first session (modality A or B). Half of the participants (n = 10) started their first trial without shutter glasses (modality A) and the other half (n = 10) with shutter glasses (modality B) within a non-immersive test environment. After completing the first session the participants were required to fill out the AttrakDiff questionnaire. Hereafter the next session within the immersive test environment (modality C) started. All participants got shutter glasses and the virtual car model was projected. They had to tackle the same task and got the AttrakDiff questionnaire once again. In the third and last trial (modality A or B) half of the participants worked within the modality they didn`t perform within the first session and reverse for the other half. The third trial was also realized within a non-immersive test environment. The main experiment was completed after two short questions (impressions about proximity to a real car).

The last ten minutes were used to fill out the usability questionnaire ISONORM about the implemented navigation system on a separate table outside the CAVE.

3 Results

3.1 Descriptive Analyses

For a descriptive and visual analysis of the results, the completion times and errors over all participants are visualized in the following table.

Table 1. Mean task completion times in seconds and average number of errors

	Modality A Without shutter glasses and no virtual car model	**Modality B** With shutter glasses and no virtual car model	**Modality C** With shutter glasses and virtual car model
Time	505 s*	499 s	451 s*
Errors	2,7	2,8	2,3

*Significant differences

The first two columns contain the times and the number of errors for the interaction task with and without shutter glasses (both without virtual car model). There is a marginal difference of 6 seconds, whereby the time is surprisingly shorter with shutter glasses. In the last column the time for the modality with a projected car model show lowest time and error rates of all three modalities.

3.2 Task Times

Task times were analyzed using a repeated measure ANOVA, with task environment (modality) as independent variable. The little difference of 6 seconds between the modalities A and B (non-immersive test environments) is not significant. It indicates that the shutter glasses do not have an influence on user performances.

With regard to the second research question, whether the virtual cockpit (immersive test environment) has an influence on user performance the repeated measure ANOVA shows a significant difference between modality B and C ($F(1,19) = 4.81$, $p < 0.005$, $n^2part = 0.202$). Hence, task completion times were faster within immersive environments, compared to non-immersive environments (see table 1), which indicates that immersive environments increase user performances. Indeed, we found undoubtful training and repetition effects. All participants were slowest in the first session and fastest in the last one.

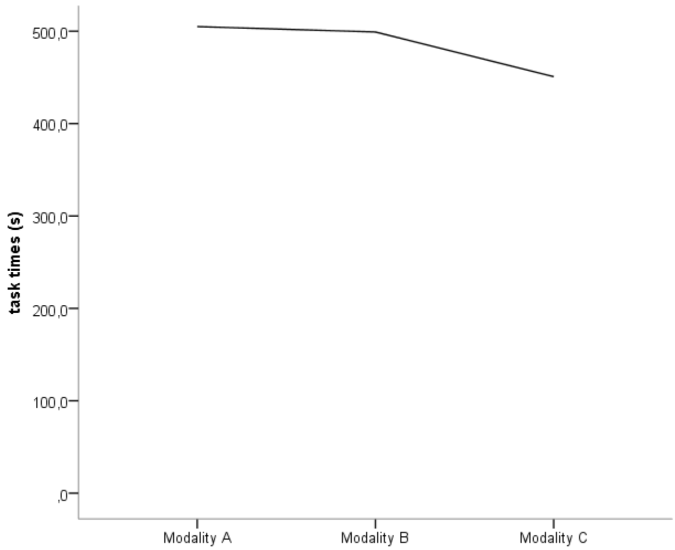


Fig. 3. Task completion times within different modalities

3.3 Task Errors

Task errors were also analyzed using a repeated measure ANOVA, with task environment (modality) as the independent variable. There were no significant differences found. But absolute mean task errors were lowest in the immersive modality C, compared to the two non-immersive conditions A and B (see table 1). This indicates a performance benefit within virtual test environment and no influence of shutter glasses, too.

3.4 Subjective Data

The AttrakDiff questionnaire (Hassenzahl, Burmester & Koller, 2003) was used to analyze subjective impressions within different test environment modalities. The AttrakDiff consists of 21 pairs of adjectives and is suitable for repeated assessments during a test session. Adjective pairs are semantic differentials rated on a seven point Likert scale. The items are grouped into three dimensions: pragmatic quality (e.g. impractical – practical), hedonic quality (e.g. tacky – stylish) and overall attractiveness (e.g. bad – good). The used AttrakDiff shows no significant differences for one of the three dimensions regarding different task environments.

3.5 Usability

To analyze the usability of the implemented navigation system the ISONORM questionnaire ISONORM 9241/10 (Prümper & Anft, 1997) was used. According to the authors, the ISONORM represents an operationalization of the seven dialog principles in ISO 9241, part 10 (ISO, 1996): suitability for the task, self-descriptiveness, controllability, conformity with user expectations, error tolerance, suitability for individualisation and suitability for learning. The ISONORM comprises five bipolar items (on a scale from -3 to +3) for each of the seven principles of ISO 9241 Part 10. In our survey we had only six dimensions (see figure 4). We eliminated the dialog principle suitability for individualization, because we decided in advance that individualization is not necessary and desired for our cockpit-interaction experiments.

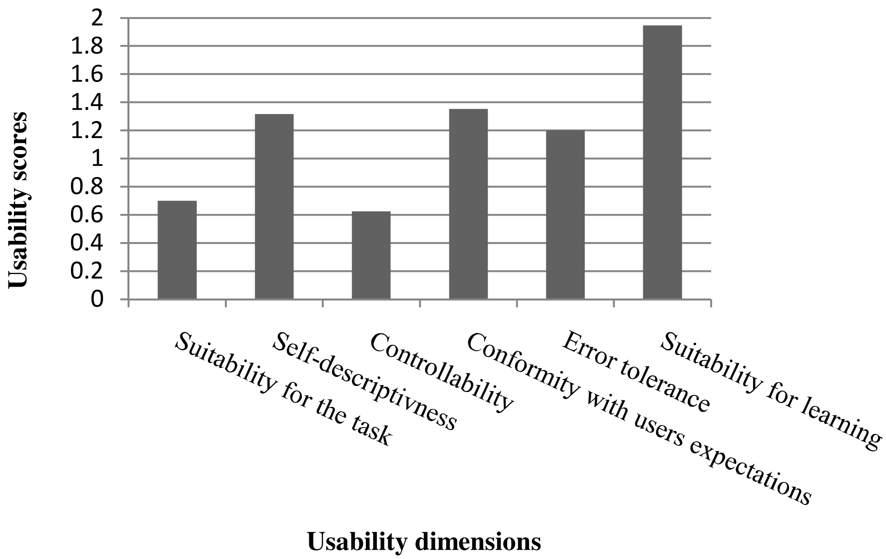


Fig. 4. Mean usability scores for all participants (ISONORM)

The intention was to detect usability issues (e.g. confusing menu structure) and/or buggy programming (e.g. non-functional menu buttons), because this navigation interface will be used for several upcoming comparative studies. The cut-off score was 0 (arithmetic mean). Figure 1 shows that all dimensions were above 0 within a positive value range. The most important dimensions for our field of application were self-descriptiveness (M = 1.3; SD = 1.2) and conformity with user expectations (M = 1.4; SD = 1.1), to avoid time-delays due to usability issues instead of modality influences.

4 Discussion and Outlook

There seems to be an explicit need for more systematic user involvement in the early stages of development processes, especially in virtual test environments. With our research study we focused on important basic parameters and evaluated their potential of influence.

Three different test environments were presented and evaluated regarding different task performances and subjective ratings: Two non-immersive modalities (with or without shutter glasses) and one immersive modality with a virtual projected car model. The aim was to analyze differences (task performance and subjective ratings) and explanatory variables (shutter glasses and virtual projected car model). To do so, different research questions were operationalized.

The first difference was analyzed regarding an important explanatory variable: the shutter glasses. The question was, if shutter glasses influence users' behavior per se, because of a slightly darkened sight and alternating shutter effects with 60Hz on each eye. There were no significant effects found for any differences between with or without shutter glasses regarding to task performance (times and errors) or subjective ratings (AttrakDiff). It is quite reliable that shutter glasses do not influence users' behavior or impressions per se. Participants do not need more time for data entries or make more mistakes with shutter glasses. Participants also do not rate different impressions with shutter glasses, compared to no shutter glasses. In summary, there are no hints to assume that limiting influences through shutter glasses will occur.

The second difference was analyzed with regard to the next important explanatory variable: the virtual projected car model. The question was, if the virtual simulated car model influences user behavior (time and errors) and subjective impressions (AttrakDiff). There are some hints that virtual simulated car models influence users' behavior positively.

Participants need significant less time within a virtual car model, compared to non-immersive conditions. Faster task completion times within immersive environments indicate that immersive environments increase user performances. Furthermore, there seems to be a tendency that participants make fewer mistakes within a virtual car model, but that effect didn`t become significant yet. Surprisingly, participants didn`t get different impressions within a virtual car model, compared to non-immersive test environments. We assume that effects will occur, when driving within hybrid test environments will be added.

After the main experiment users were asked about the usability of the navigation interface. Participants rated all of the usability dimensions from the ISONORM questionnaire within a positive range, which indicates an adequate usability value. Special focus was lying on the dimensions self-descriptiveness and conformity with user expectations, to avoid time-delays due to usability issues instead of modality influences. Both were within a good usability range. In summary, the navigation interface will be useful for several upcoming comparative studies in that field, too.

The findings presented in this paper leave much room for further research. For example, it would be interesting to see if the presented findings are applicable to different cockpit interactions (e.g. touch and gestures) or if the subjective perception differs more within a virtual driving environment compared to a conventional 2D driving environment.

References

1. Ablaßmeyer, M.: Multimodales, kontextadaptives Informationsmanagement im Automobil (Dissertation). Technische Universität München (2009)
2. Broy, V.: Benutzerzentrierte, graphische Interaktionsmetaphern für Fahrerinformationssysteme (Dissertation). Technische Universität München (2007)
3. Hassenzahl, M., Burmester, M., Koller, F.: AttrakDiff: Ein Fragebogen zur Messung wahrgenommener hedonischer und pragmatischer Qualitäten. In: Szwillus, G., Ziegler, J. (eds.) Mensch & Computer 2003: Interaktion in Bewegung, pp. 187–196. B.G. Teubner, Stuttgart (2003)
4. Jahn, G., Oehme, A., Krems, J.F., Gelau, C.: Peripheral detection as a workload measure in driving: Effects of traffic complexity and route guidance system use in a driving study. Transportation Research Part F-Traffic Psychology and Behaviour 8(3), 255–275 (2005)
5. Mayhew, D.J.: The Usability Engineering Lifecycle. Morgan Kaufmann, San Francisco (1999)
6. Krause, F.-L., Franke, H.-J., Gausemeier, J.: Innovationspotenziale der Produktentwicklung. Carl Hanser Verlag, München (2007)
7. ORF Online und Teletext GmbH & Co KG, `http://salzburg.orf.at/m/news/stories/2517908/`
8. Prümper, J., Anft, M.: ISONORM 9241/10, Beurteilungsbogen auf Grundlage der Internationalen Ergonomie-Norm ISO 9241/10. Büro für Arbeits- und Organisationspsychologie, Berlin (1997)
9. Stark, R., Beckmann-Dobrev, B., Schulze, E.E., Adenauer, J., Israel, J.H.: Smart Hybrid Prototyping zur multimodalen Erlebbarkeit virtueller Prototypen innerhalb der Produktentstehung. In: Proc., vol. 8, pp. 437–443 (2009)
10. Wang, G.G.: Definition and review of virtual prototyping. Journal of Computing and Information Science in Engineering (Transactions of the ASME) 2(3), 232–236 (2002)

Mixed and Augmented Reality for Marine Corps Training

Richard Schaffer, Sean Cullen, Phe Meas, and Kevin Dill

Lockheed Martin Mission Systems & Training, Burlington, MA, USA
{richard.l.schaffer,sean.cullen,phe.meas,kevin.dill}@lmco.com

Abstract. The United States Marine Corps faces numerous challenges in preparing Marines for current operations; among them are the cost of specialized training environments and the difficulty of realistically representing the deployed environment. This paper reports on two Office of Naval Research efforts to address these challenges. The first employs Mixed Reality, which combines real-world and virtual elements to create a Hollywood-set-like representation of an Afghan village where Marines can train prior to deployment. The second explores the use of Augmented Reality to train USMC observers. Observers are responsible for directing artillery and mortar fires and aircraft attacks in the proximity of friendly forces. While the live environment has numerous advantages, the costs of supporting troops, ammunition, and equipment are considerable. Augmented Reality can replace live supporting forces, resulting in lower cost use of training areas during down time and enabling almost any area to become an augmented training area.

Keywords: Mixed Reality, Augmented Reality.

1 Introduction

Training our military forces is challenging. Effective training requires the construction of specialized environments, expenditure of ammunition, use of high-cost equipment such as aircraft or artillery, and supporting personnel (including military personnel, civilian role-players, and staff). In addition, it can be difficult to accurately portray conditions in the deployed environment, and even more difficult to rapidly convert training environments from one remote location to another.

In this paper we report on two Office of Naval Research (ONR) efforts that explore techniques for combining live training with virtual elements in order to create more realistic, more immersive training environments at lower cost and with decreased requirements in terms of materiel and manpower.

1.1 Mixed Reality at the Infantry Immersion Trainer

Since its inception in 2007 at US Marine Corps Camp Pendleton, California, the goal of the Infantry Immersion Trainer (IIT) has been to recreate as accurately as possible

R. Shumaker (Ed.): VAMR/HCII 2013, Part II, LNCS 8022, pp. 310–319, 2013.

a small village similar to what the Marines could expect to face in deployed environments, first in Iraq and more recently in Afghanistan. This includes not only the physical features – the sights, the sounds, the smells – but also the socio-cultural factors – the language, the culture, and the behaviors of the locals [1]. The latter is typically achieved through the use of civilian role-players who are first-generation Afghan immigrants, and who play roles such as the village elder, the fertilizer salesman, the insurgents, and so on. Immersive training has become increasingly popular in recent years due to its effectiveness in preparing Marines for contemporary combat operations. Since 2007, ONR has continued to enhance the system and IIT facilities have been constructed at two additional USMC sites, although the Camp Pendleton IIT remains the site for ONR experimentation.

The creation of an immersive training experience requires careful construction of the physical environment and selection of appropriate role-players, but also includes a variety of Mixed Reality (MR) technologies, where an MR technology is anything that combines real-world and virtual elements into a single system. Among the MR elements at the IIT are scent generators, speakers that provide ambient sounds, Hollywood-quality pyrotechnics for simulating IEDs and rocket-launched grenades, and digital avatars. In this paper we will focus on the digital avatars, which are virtual characters who are projected onto the wall using a standard video projector.

1.2 Augmented Immersive Team Training

Augmented Reality (AR) uses a head-mounted display (HMD) to overlay virtual characters onto the physical world. For example, the HMD might show the physical world but insert an aircraft, or an insurgent, or an explosion. The Augmented Immersive Team Training (AITT) program is a five-year program, currently in its third year, which is examining the viability of using AR for training USMC observers. Observers are responsible for directing artillery and mortar fires, and for controlling aircraft attacks, in the proximity of friendly forces. While the live environment has numerous advantages, the costs of supporting troops, ammunition, artillery, and aircraft are considerable, and training areas are not always available. We anticipate that the use of AR-generated forces will allow the Marine Corps to decrease costs and to use almost any area as an augmented training area when the Marines are deployed or physical training areas are unavailable.

2 Digital Avatars at the IIT

While the role-players do an excellent job of portraying culturally authentic characters, they require expenditures of thousands of dollars each day. In addition, there are certain roles that can't easily be portrayed with role-players. Appropriate personnel may simply not be available (as is often the case with the elderly), or there may be legal and moral prohibitions against using them (as is the case with children or badly wounded characters). Thus the digital avatars are not intended to replace the role-players, but rather to supplement them with additional characters in specific roles [2].

Fig. 1. Digital avatars of a father and son (left) and insurgent and hostage (right)

Figure 1 shows two examples of digital avatars. The image on the left is a screen-shot of the virtual characters, while that on the right shows characters in the context of the surrounding room.

Originally, the digital avatars at the IIT were used for highly scripted "shoot-or-no-shoot scenarios" in which the Marines would see the avatars upon entering a room and immediately decide whether or not to open fire. These scenarios typically lasted only a few seconds, so there was limited need for complex decision making; with the exception of detecting hits from hostile fire, high-level decisions could be handled by an instructor/operator who was controlling the character from behind the scenes. Our more recent work strives to make the avatars capable of a broader, more expressive set of behaviors. Consequently we have implemented a new Artificial Intelligence (AI) system. The characters are still semi-autonomous, capable of responding to both their own AI decisions and operator control, but their degree of autonomy has significantly increased.

The digital avatars present us with three distinct HCI challenges. First, if the avatars are to respond appropriately to the situation in the physical world then they must be able to sense the actions of the trainees. Second, even the best machine sensing and AI algorithms are not as good as those of a human. Consequently, while we want our avatars to be as autonomous as possible, they should still be responsive to direction from the instructor/operators. Finally, we need to enable the instructor/operators to modify the characters' performances and to create new scenarios. This is particularly challenging because, while the instructor operators are quite capable and are experts in their domain, they are not deeply knowledgeable with respect to computer programming or artificial intelligence (AI).

2.1 Sensing the Environment

The first step of nearly every approach to AI is to sense the current state of the environment, so that appropriate actions can be selected. The AI that we see in video games and training simulations has improved significantly over the last decade or two, but these AIs have the advantage of perfect knowledge of everything that happens in the game or simulation. In a mixed reality scenario, on the other hand, the trainees take their actions in the physical world. As a result, we need to be able to sense the trainees' physical actions in order to allow our characters to respond.

One form of sensing technology, included at all IITs, is a shot-tracking system that was implemented by NAVAIR Orlando. This system uses an infrared (IR) camera to detect hotspots on the screen. These hotspots correspond to the impact of a DITS laser (which is part of the training system used by the Marines to register hits when weapons are fired with blank rounds), so we can use them to determine not only that the Marines have shot at the screen, but also whether or not they hit one of our avatars.

More recently we have integrated the Microsoft Kinect, which is a motion sensing input device originally developed for use with the Xbox 360. The Kinect includes a visible light camera, an active IR sensor, and a multi-array microphone. It provides:

- The position and orientation of each Marine in range
- Skeletal animation data for each Marine in range
- Speech recognition, as well as the direction of any sound

Using this data, we are able to provide the AI with the following sensory information:

- Whether there are any Marines in the room.
- If Marines are in range of the Kinect, we can determine their location and the direction that they are facing. If we have multiple Kinects in the room, we can use their microphone arrays triangulate estimated positions of any Marines that are speaking even when they are not within the Kinect's usual detection range.
- We can watch for specific poses, such as whether a Marine's weapon is raised or lowered. We can couple this with the orientation of the Marine's torso in order to get a rough idea of the direction that they are aiming their weapon (so avatars can play appropriate responses when the Marines aim at them), although obviously this is inexact. Figure 2 shows a user with their weapon raised, the skeleton detected by the Kinect, and the avatar's response (inset in the bottom right).
- We can use the Kinect's speech recognition capabilities to allow the avatars to respond to specific utterances (such as "surrender!" or "get down!"). In addition, even if we don't have a response to a particular utterance, we can still recognize the fact that the Marines are speaking and play a generic response (perhaps something like "No speak English! Don't shoot! Don't shoot!").

Both sensing systems (the NAVAIR shot tracker and the Kinect) are imperfect. They may fail to detect valid events or falsely detect events that didn't occur. In particular, the Kinect can be very sensitive to placement (it has a fairly limited sensing area), and when there are more than two humans in the room it will only detect skeletons for two of them (though it will provide position and orientation for the rest). Frustratingly, we don't get to select which users have their skeletons tracked. In addition, because the Kinect uses an active IR sensor, if two Kinects are placed in close proximity or if the Kinect is placed too close to the shot tracker, the systems may conflict. With that said, these systems are good enough to allow the avatars to display realistic responses to many trainee actions – and the operators can take control when they don't.

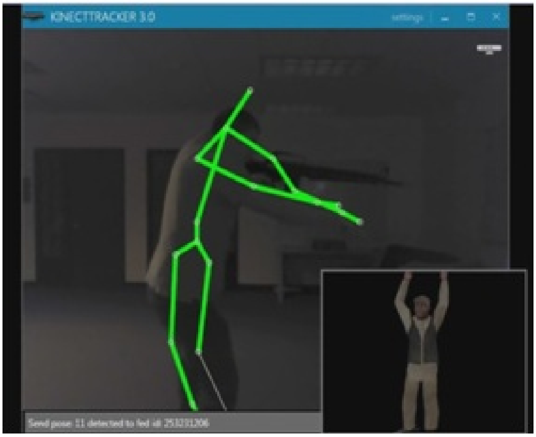


Fig. 2. Pose detection with the avatar's response

2.2 Semi-autonomous Control

Although the Kinect gives us a far greater ability to sense the Marines' actions than we had previously, it is still limited. Furthermore, even if we had perfect sensing the best current AI techniques are not at the level of human intelligence. Consequently, it makes sense to retain the ability for instructor/operators to take control.

Unfortunately, humans are also imperfect. The IIT instructor/operators have two interfaces that they can use to directly control the avatars, each of which has limitations. One is via the instructor/operator station (IOS), shown in Figure 2, which is located in a remote command center where the instructors observe and react to situations viewed in video and audio feeds available throughout the facility (this is far more difficult than one might think). The other is via a standard TV remote control programmed to a limited set of commands and controlled by an instructor who is moving with the trainee unit. Either way, their attention is often divided, and their reaction time can never be as fast as that of a machine. While we do want to provide the instructors with high-level directorial control over the avatars, we certainly don't want them to have to micro-manage every action that the characters take. Thus the ideal approach is to have an autonomous AI which can respond to most likely actions on the part of the trainee and maintain moment-to-moment believability, while allowing the operator to provide high-level commands that change the overall direction of the performance.

With all of that in mind, each scenario specifies both scenario-wide commands and entity-specific commands. The scenario-wide commands change the behavior of all entities in the scenario, while the entity-specific commands are sent to a specific individual. Examples of common commands include things like "fire at the Marines," "surrender," or "wave your weapon and rant." The instructor/operator still retains ultimate override capability using the IOS and remote control interfaces, but the AI augments the instructor's ability to manage a realistic, immersive scenario.

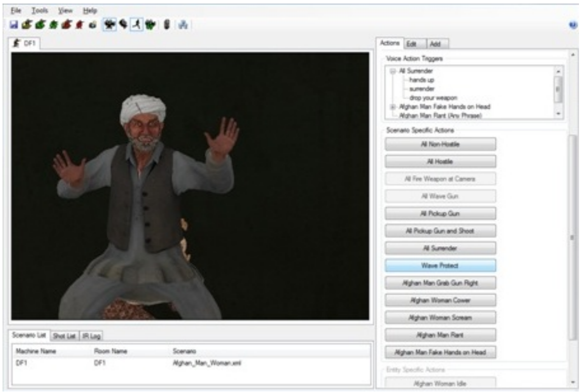


Fig. 3. The Instructor Operator Station. The actions of the digital avatars are shown on the left. The voice triggers appear on the top right, and buttons for action selection appear below them.

A full description of the AI architecture is beyond the scope of this paper, but the general approach is similar to that taken by the Angry Grandmother character created as part of ONR's Future Immersive Training Environments (FITE) Joint Capabilities Technology Demonstration (JCTD) [2], and is built on top of our Game AI Architecture (GAIA) [3]. That work has been extended with the notion of *roles*. Roles are assigned to characters when the scenario is created, and define the purpose that each character serves in the scenario. Currently, most scenarios assign the characters to "hostile" and "non-hostile" roles. The roles limit the actions that will be available to each character, preventing the instructor/operators from unintentionally selecting a nonsensical action. They also define the avatars' actions when under AI control or when they receive a scenario-wide command.

2.3 Scenario Creation

In addition to having overarching control of the digital avatars during scenario execution, the instructor/operators also want to be able to create their own training scenarios. They may do this on their own or in conjunction with the unit commanders for the Marines being trained. We can and do support them in this endeavor, but they shouldn't have to turn to us every time they need a new scenario constructed.

GAIA's modular architecture was designed in large part to make it easier to build the AI by "plugging together" large, conceptually meaningful pieces [4]. Nevertheless, GAIA components still require significant configuration, making the architecture most appropriate for use by somebody with a strong technical background. With that in mind, we have worked to find standard configurations for the GAIA modules, so that they can be dropped into a character's AI without further tuning [5]. This is made simpler because of the relative lack of complexity in our inputs. Thus a GAIA configuration for one of our digital avatars contains a dual utility reasoner. The reasoner's options each contain one or more considerations that are used to determine which option to select. The considerations that we support are:

- **Operator Command**: Activates when the appropriate command is received from the instructor/operator.
- **Track Detected**: Activates if Marines have been detected in the avatar's room.
- **Shot Detected**: Activates if shots are detected by the shot tracking system.
- **Utterance Detected**: Activates if a specific phrase is uttered by the Marines. Alternately, "any phrase" can be selected, in which case this consideration will activate any time a Marine speaks.
- **Pose Detected**: Activates if a particular pose is detected (e.g. if the Marine raises his weapon).
- **Cooldown**: Prevents the option from running twice in close succession.

In the terminology we previously proposed [5], all of these are "opt-in" considerations except for the Cooldown consideration, which is "opt out." A full discussion of dual utility reasoners and GAIA configuration can be found in our previous papers [2-5].

3 Augmented Reality

"Augmented reality" (AR) is a term that has been used in a variety of ways. Often, it simply refers to the overlay of information (such as intelligence data) on top of the physical world. Under the FITE JCTD, and more recently as part of AITT, we have been working with our teammates at SRI International to develop technology that, rather than overlaying information, can overlay virtual characters and events. So, for example, you might be looking at a physical landscape but we would insert a virtual plane, or a virtual tank, or a virtual explosion. This requires:

- A Head-Mounted Display (HMD) or augmented prop
- A 6-axis head/prop tracker
- A 3D model of the static physical environment
- A man-wearable computer
- A Remote-Enhanced Instructor Station (R-EIS)

In brief, the data from the head tracker is used to determine the position and orientation of the user's head in the physical environment. From this, we can determine where on the HMD to draw the image of our virtual objects. The model of the physical environment is used to determine which portions of the virtual objects would be visible to the user, and occlude them as appropriate. All computation and rendering occurs on the man-wearable computer, and the R-EIS controls the training scenario. A more detailed description of each subsystem appears below.

The Head-Mounted Display. Two approaches exist to combining a view of the physical world with virtual objects on the HMD. The optical see-through approach uses a translucent HMD. In this case the user can see the physical world directly, and we simply augment their view by displaying the virtual objects in the appropriate positions, appropriately occluded. This approach would generally be preferable, but the technology is not as advanced. The video see-through approach uses a head-mounted camera to display the physical world, along with augmentation, on an opaque HMD.

The ideal HMD would support the resolution and field of view of normal human vision. In addition, it would support near and far vision, allow us to attach the head-tracking system directly to the HMD, and would be compatible with arbitrary external optical devices (such as binoculars). Unfortunately, currently available HMDs fall far short of this. For example, where human vision is roughly 160x135 degrees, typical commercial HMDs only support a 60 degree diagonal display. The resolution is lower as well – typically about 2 arc minutes per pixel, as opposed to 50 arc seconds for 20/20 vision or 1.5 arc minutes for 20/30 vision [6].

Augmented Props. Observing the enemy, determining, and marking his location is often accomplished through the use of optical devices. Consequently our training needs to support these as well. The ideal solution would be to use HMDs with the standard optical devices, but this is impractical. Instead we have created props that mimic the form, fit, and function of the real devices but replace the optical paths with cameras and miniaturized displays. Currently we support two common devices used by the Marine Corps: the Vector 21 (sophisticated binoculars with orientation and laser range-finding capabilities in both 7x and 10x magnifications) and the Portable Lightweight Designator Rangefinder (PLDR), which is used to mark a target for attack. The Vector 21 prop is fully functional and can be connected to a Defense Advanced GPS Receiver (DAGR) for calculating the location of the virtual target in the exact same manner that the trainee would use with a live target.

The Head Tracker. Our head-tracking technology is provided by SRI International. It incorporates numerous sensors, including cameras for performing both visual feature landmark tracking and visual odometry, an inertial measurement unit, a GPS, a magnetometer, and a barometric pressure sensor. This system combines these inputs and then communicates the position and orientation of the user's head, along with the image from the forward looking camera [7].

The 3D Model of the Physical Environment. The rendering system needs to ensure that it only displays objects that are observable from the trainee's current position. To accomplish this, a 3D terrain skin is developed using the best data available, and this is supplemented with real-time models of dynamic objects (such as vehicles, structures, furniture, and other humans). The better the data, the better the system will be at properly occluding objects in the scene (such as a helicopter flying over a hill, or an insurgent peering around a corner). Currently occlusion works poorly for nearby targets or when high precision is required, but reasonably well on distant targets.

The Man-Wearable Computer. This computer runs both the tracking system software and the rendering software. We currently use a high end laptop with a modern graphics card, but the goal for the program is to transition to a single board solution utilizing embedded mobile graphics chips. This would allow the computer to be carried in a lightweight backpack or embedded directly into the props. Because the graphics system is only displaying a limited number of entities and weapons effects (and not a detailed environment with complex terrain and numerous characters) the typical rendering load should be manageable on these less powerful systems.

The Remote-Enhanced Instructor Station. The R-EIS provides the ability to inject friendly, enemy, and neutral entities into the scenario, create Combat Air Support aircraft, and perform Call for Fire missions with simulated artillery units. The interface is built on an iPad (Figure 4), allowing the instructor to move around with the trainees while controlling both friendly and enemy forces.

Fig. 4. The Remote Enhanced Instructor Station

3.1 Augmented Immersive Team Training

In theory, we could alter reality in whatever ways we want. For example, we could build a simple concrete training environment and then "paint the walls" with whatever materials we want (wood, stucco, wallpaper, etc.) to change the appearance to match the target environment. We could create virtual opponents who move in the physical world and are indistinguishable from physical opponents. This level of fidelity has the potential to truly transform the ways in which we do training. Unfortunately, the accuracy of the head tracker and the 3D model of the physical environment are not yet good enough to make this possible – although they have been rapidly improving.

In contrast, Forward Observer/Forward Air Controller (FO/FAC) training is easily conducted with this system. A typical scenario requires only a Vector 21 prop and an R-EIS. The trainee would use the Vector 21 to observe the enemy units placed by the instructor, and then connect his DAGR to the Vector 21 to calculate the real-world location of the virtual enemy (or use a paper map to determine the location by hand). Next, a fire mission is developed and communicated to the instructor. The instructor uses the R-EIS to insert virtual artillery detonations at the appropriate time and location. The student uses the Vector 21 to observe the incoming rounds and provide corrections back to the instructor. This entire scenario is possible with minimal staff and resources, and yet is still conducted outside in a realistic working environment, where the trainee has to deal with real-world weather situations.

Forward Air Controller training is likewise possible with the system. The FAC can track/locate the enemy in a similar fashion, and then use the HMD to observe and clear attack aircraft in bombing missions. By adding the PLDR laser designator and flight simulator, a much more complex scenario can be employed that has the trainee designating the target and directing the pilot who is flying on the simulator.

4 Conclusion

Mixed reality and augmented reality are new paradigms for Human-Computer Interaction with the potential to transform the ways in which we use computers. The supporting technologies for them are available now, and are improving by leaps and bounds every year, so now is the time to start thinking about how we can make use of them. In this paper we described two efforts that are currently using these technologies in the context of training US Marines.

References

1. Fry, D.J., Meas, P., Cerritelli, L., Schaffer, R.L., Pair, J.: Mixed-Reality Urban Combat Training with the Infantry Immersive Trainer Program. In: 2008 SISO Spring Interoperability Workshop (2008)
2. Dill, K.: A Game AI Approach to Autonomous Control of Virtual Characters. In: 2011 Intraservice/Industry Training, Simulation and Education Conference (2011)
3. Dill, K.: Introducing GAIA: A Reusable, Extensible Architecture for AI Behavior. In: 2012 SISO Spring Interoperability Workshop (2012)
4. Dill, K., Pursel, E.R., Garrity, P., Fragomeni, G.: Achieving Modular AI through Conceptual Abstractions. In: 2012 Intraservice/Industry Training, Simulation and Education Conference (2012)
5. Dill, K., Pursel, E.R., Garrity, P., Fragomeni, G.: Design Patterns for the Configuration of Utility-Based AI. In: 2012 Intraservice/Industry Training, Simulation and Education Conference (2012)
6. Australian National University, College of Engineering and Computer Science, `http://escience.anu.edu.au/lecture/ivr/sight/eyeMainCharacteristics2.en.html`.
7. Kumar, R., Samarasekera, S., Chaudry, A., Zhu, Z., Chiu, H.P., Oskiper, T., Villamil, R., Branzoi, V., Hadsell, R., Pursel, E.R., Dean, F., Garrity, P.: Implementation of an Augmented Reality System for Training Dismounted Warfighters. In: 2012 Intraservice/Industry Training, Simulation and Education Conference (2012)

Proactive Supervisory Decision Support from Trend-Based Monitoring of Autonomous and Automated Systems: A Tale of Two Domains

Harvey S. Smallman and Maia B. Cook

Pacific Science & Engineering, San Diego, CA, USA
{smallman,maiacook}@pacific-science.com

Abstract. The digital technology revolution continues to roil work domains. An influx of automation and autonomous systems is transforming the role of humans from *operators* into *supervisors*. For some domains, such as **process control**, supervisory control is already the norm. For other domains, such as **military command and control,** the transformation to autonomous supervision is just beginning. In both domains, legacy operation-centric, real-time data displays and tools provide inadequate task support, leading to unproductive user work-arounds. They give rise to a *reactive* monitoring stance, and will not scale to meet the new, different task needs. We review advanced display design projects in each domain that, in contrast, provide *proactive* supervisory decision support. We identified key perceptual and cognitive challenges in supervision, and applied cognitive science concepts to the design of novel trend-based interfaces. We drew lessons from process control to combat the challenges likely to arise in military command and control.

Keywords: Visualization, automation, supervisory control, decision support, proactive monitoring, cognitive science, human factors, user-centered design.

1 Introduction

With *A tale of two cities*, his famous novel set during the French revolution [1], Charles Dickens created a story with a message. Dickens used the work to draw pointed lessons for his affluent readers in London of the perils of persistent social inequity and oppression that had recently led Paris to bloody insurrection. Those two cities, London and Paris, here, are metaphors for two work domains shaken by the modern digital technology revolution. Like Dickens's, our tale is cautionary. We draw lessons for a domain bracing for the influx of sophisticated automated and autonomous technologies - military command and control (C2) - from another domain that has experience weathering a similar influx of it, industrial process control.

The digital technology revolution in the latter half of the twentieth century created dramatic changes in the enablers of worker productivity [2]. Now, there are lofty ambitions for sophisticated new automation and autonomous systems to radically improve productivity and transform the roles of its human users. For example, in the

R. Shumaker (Ed.): VAMR/HCII 2013, Part II, LNCS 8022, pp. 320–329, 2013.

military domain, it currently takes two or more humans to *operate* one unmanned aerial vehicle (UAV) performing one mission [3-4]. The future vision is for one human to *supervise* multiple autonomous systems performing multiple missions [5].

There are many questions about how to achieve this role shift. Here, we focus on how to support the information display requirements of the future multi-system, multi-mission supervisor. Will current, operation-centric display formats scale to meet the new task demands? Or is a fundamental shift to *supervisory decision support* tools and visualizations required? If so, what new display metaphors are needed and what science can be brought to bear to constrain their design?

To begin to answer these questions, we examine another domain for insights and lessons learned. Industrial process control and military C2 share a marked semblance. They are similar in overall task structure, organizational roles, and even control room layouts. In both domains, supervisors are faced with the task of monitoring unfolding situations and deciding whether, when, and how to intervene in processes or missions.

What *differs* across the two domains is the current state of automation adoption. Industrial processes have run on supervisory control [6] for decades. Process control experienced a revolution in the 1970s through the introduction of distributed control systems (DCS). DCS changed the nature of human work from active physical operation and control of elements across the plant, to passive configuration and oversight conducted from remote, central control centers via visual displays and automation [7].

We have taken advantage of the earlier adoption of automation in process control to apply lessons learned and to anticipate problems likely to arise as more automation is introduced into the military C2 domain. Additionally, we have capitalized on the similarities in task structure and shortfalls of current supervisory displays in both domains to develop solutions that apply *across* domains.

2 User Task Requirements: Proactive vs. Reactive Monitoring

To determine the information display requirements for supervisory control, we performed cognitive task analysis [8] in both application domains. This analysis also revealed the limitations and gaps in support for current practice. The task analysis was an integral part of our user-centered design (UCD) process. Our UCD approach stresses analysis of the cognitive and perceptual challenges of user task performance, and the application of relevant science to those challenges. The aim being to center tool design around users' work domain and task requirements, tailored to human cognitive and perceptual capabilities.

The goal of supervisory control is to stay abreast of processes and situations and intervene to keep things on track. We focused on monitoring, as it is the most time-consuming and challenging aspect of supervisory control [6]. We created descriptive task flows of current monitoring practice and prescriptive flows of ideal practice. The descriptive flows were created from site visits and interviews with subject matter experts (SMEs) from process control and military C2 domains. The process control SMEs were interviewed in site visits to a large oil refinery and a chemical processing facility in Texas in 2011. 27 military unmanned system SMEs were interviewed in 4 site visits conducted across the US in 2012; detailed results reported separately [4].

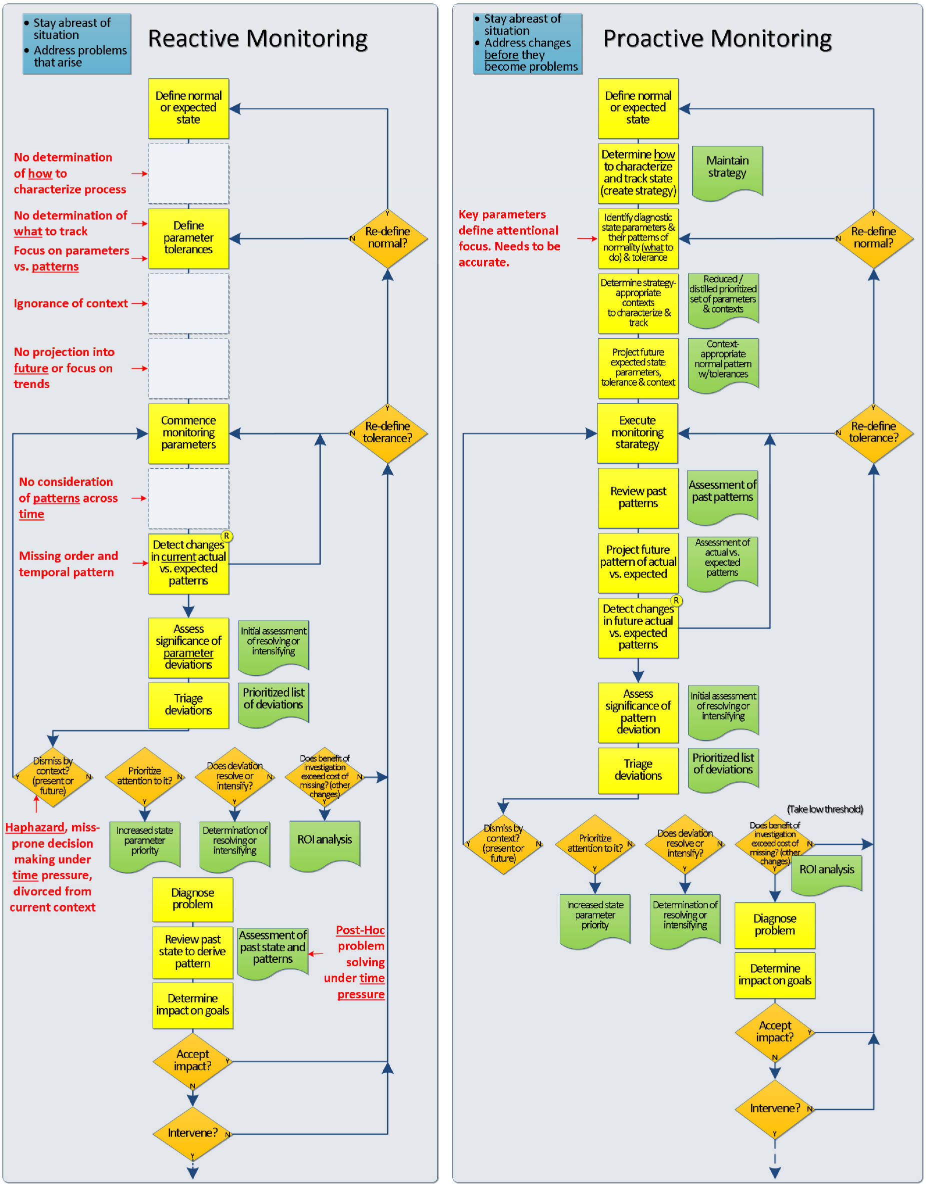


Fig. 1. Task flows for reactive (descriptive, left) and proactive (prescriptive, right) supervisory monitoring. Goals (blue), tasks (yellow), products (green), decisions (orange), and cons (red).

Essentially the same main supervisory monitoring control loop was found across domains, so only one is reported. The descriptive task flow is shown on Fig 1, left. It is centered on responding to system alarms and alerts. It is labeled ‘reactive monitoring’ because controllers diagnose and act on problems only after they’ve manifested.

The prescriptive flow of ideal monitoring practice reflects input from the SMEs [4], previous task analyses [9], literature [10] and doctrine [5]. It is labeled '*proactive monitoring*' in Fig 1, right, after terminology in widespread usage [9-11]. Burns [10] previously listed three phases of proactive monitoring as (1) deviation detection, (2) problem prediction, and (3) compensatory action. We expanded these phases into a detailed task and decision flow. The flow is *proactive* in stressing recognizing deviations from normality before they become critical problems, to give time for appropriate diagnosis. Further, proactive monitoring is characterized by defining a monitoring strategy before engaging in monitoring. This strategy is geared around hypothesis testing (e.g., "the process is stable"). Proactive monitoring has also been referred to as "cognitive," or "knowledge-driven" monitoring, to reflect its top-down nature [9].

The juxtaposition of the task flows in Fig 1 reveals their relative strengths and weaknesses (shown in red text). Reactive monitoring has many disadvantages [9-10, 12-13]. The absence of an explicit monitoring strategy can result in delayed or missed detection of deviations. Because diagnosis tasks are only performed after problems manifest, they must be performed under time pressure, often accompanied by a further stream of distracting and uninformative alarms. In addition, a reactive posture can result in an overall sub-optimal system operation, in continual alarm state, instead of in more optimal state conditions [12].

The only downside of proactive monitoring listed in Fig 1 is its reliance on an accurate definition of diagnostic, key performance indicators to track. The restricted attentional set, when accurate, streamlines proactive monitoring, but when inaccurate, may inadvertently delay problem detection.

SMEs from both domains are aware of the overall advantages of proactive monitoring [9]. To the extent they are able to be proactive, though, it is despite their displays and tools, not because of them. We turn to this issue next.

3 Legacy Display Metaphors: Naïve Realism and Reactivity

In a long-running research project, we studied the basis of military situation display design [14-19]. The project was motivated by a U.S. Navy interest in moving from conventional top-down geospatial displays to *spatially realistic* three-dimensional (3D) perspective view displays. In a series of studies, we found a consistent preferences and positive intuitions for spatially realistic displays that performed poorly. For example, both naïve and expert participants expressed a preference for realistic 3D icons (Fig 3, left). However, in controlled testing, participants identified track attributes more slowly and less accurately with the 3D icons than with less realistic symbols [14].

Conventional situation displays are also *temporally realistic* - they show the current state of the situation. In controlled studies, we revealed the parallel limits of temporal realism. Real-time situation displays performed poorly for monitoring for tactically-relevant changes in naval airspace and reporting those changes after an interruption [15]. Further, a temporally realistic instant replay tool actually degraded performance for these tasks below baseline [16]. Alternative, less realistic approaches that presented automation-extracted changes were shown to be far superior, though their utility was underestimated by participants [15].

We developed a theory called *Naïve Realism* to explain the paradoxical desire for spatially and temporally realistic information display, theorizing that it was based in folk fallacies about the workings and output of perception [17-18]. Naïve Realism has led to display metaphors that unhelpfully mimic spatio-temporal aspects of physical task domains that then predispose their users to monitor reactively.

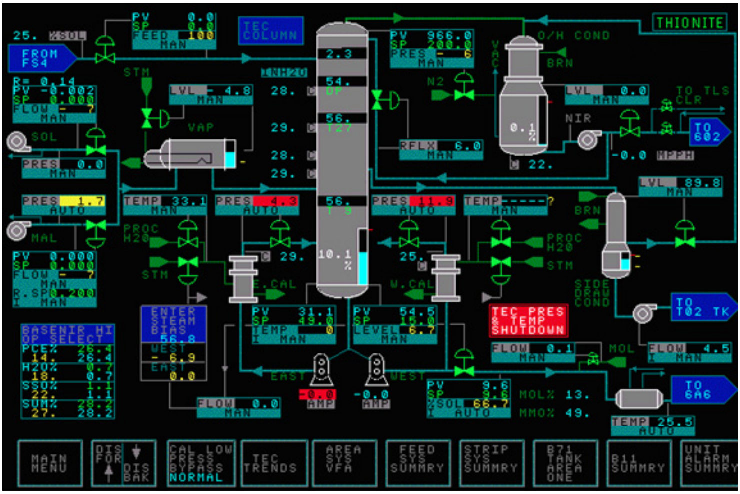


Fig. 2. Example of the traditional piping & instrumentation diagram (P&ID) mimic display format of process control (Source: [20] *InTech* magazine, Nov/Dec 2012. © 2012 ISA. Reprinted by permission. All rights reserved).

In process control, for example, the predominant display metaphor is the piping and instrumentation diagram, or “P&ID,” which is a mimic of the plant engineering diagram. The P&ID is often densely superimposed with real-time process parameters (see Fig 2 for an example). Like our experiment participants who struggled to detect changes in real-time Navy situation displays [15], users of P&IDs must extract temporal context and trends from real-time data with their eyes, and stitch it together over time in memory to infer potential deviations. Process control experts may engage in workarounds to combat this difficulty. For example, in the nuclear industry, operators have been observed to print out copies of control screens and subsequently write down observed values to externally reconstruct temporal context missing from their displays [9]. Expert users may also overlay miniature trend graphs next to process elements on P&IDs, or create separate custom displays of arrays of trend graphs [11]. Unfortunately, these work-arounds are of limited utility as they lead to more cluttered and inconsistent P&IDs to monitor, and more displays to track and mentally relate.

Naively realistic users pack P&IDs with too many undiagnostic parameters partly because they overestimate their ability to extract information from cluttered displays and underestimate the effect of that clutter on their search performance [19]. This clutter makes scanning for deviations a slow and error-prone process [20]. This exacerbates the reactivity problem, putting operators in the mode of playing constant catch-up when monitoring, and hindering their ability to get ahead of problems.

Finally, the alerting schemes for P&IDs invariably use data-level, context-insensitive thresholding, issuing alerts simply when a parameter exceeds some threshold value. During crisis episodes, this can result in unhelpful "alarm flooding" with little or no context available to help triage and rationalize alerts and resolve problems [12-13].

4 Trend-Based Monitoring for Process Control

In a UCD project conducted for a large process control manufacturer, we addressed the limits of the legacy P&ID format, applying lessons from the Naïve Realism research to the design of display concepts for proactive, trend-based monitoring. In the military domain, we had previously combatted the shortfalls of spatially realistic 3D icons through the design of novel hybrid symbol-icons (***Symbicons***) that graphically encapsulate, caricature and emphasize key task-relevant attributes of military tracks [14], see Fig 3. Instead of the complex and indiscriminable 3D shaded icon, the Symbicon simply, quickly, and unambiguously conveys a friendly fighter aircraft and its heading. Symbicon advantages were confirmed by significantly faster and more accurate identification performance in controlled studies [14]. In the process control domain project, we combatted temporal realism in an analogous way by designing and prototyping ***Trendicons*** (patent pending) to graphically encapsulate, caricature, and emphasize key task-relevant attributes of parameter trends. The Trendicon in Fig 3 simply and quickly conveys a process parameter that is above normal (black bar thermometer), trending upward (triangle) and getting worse (bold outline) but not yet reached alarm state (grey fill color, instead of yellow or red).

Trendicons transform monitoring. Instead of monitoring real-time values, or even *trends* of those values [11, 21], users monitor the *task-relevant aspects* of trends that the Trendicons emphasize. Instead of extracting features like actual vs. target value and rate of change (Fig 3, right) themselves, this burdensome work is offloaded to the display, freeing users to devote their efforts to what is most critical: monitoring proactively. Unlike real-time parameters, trends buried in cluttered P&IDs, and banks of trend graphs, the visual format of Trendicons is tailored to human attention and perceptual capabilities. Users can quickly and accurately identify, compare, and prioritize key attributes across several parameters. Further, Trendicons provide a natural task entry point for navigating to more detailed parameter information, through a semantic zoom into underlying trends. This approach allows users to naturally progress into deeper levels of detail when needed, building context useful for diagnosis, *before* problems manifest in alarms and alerts. Other approaches, have targeted the shortfalls of P&IDs through solutions grounded in sophisticated work domain analysis [e.g., 22] but less focused on design solutions matched to users' perceptual and cognitive limitations and their flawed metacognition.

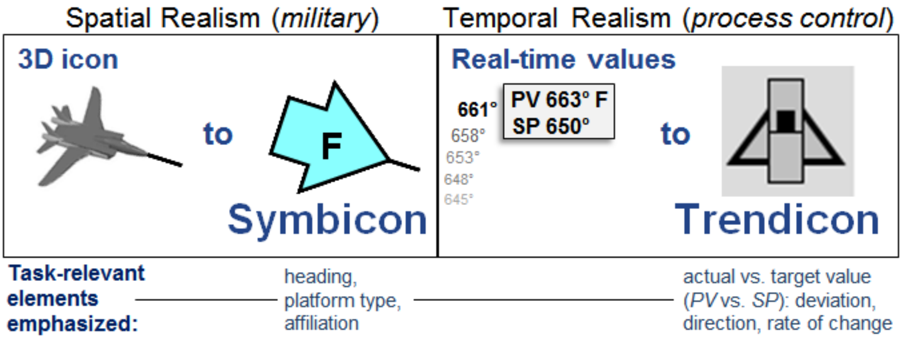


Fig. 3. Combatting Naïve Realism in the military (left, ***Symbicons*** [14]) and process control domains (right, ***Trendicons***, patent pending)

5 Proactive Decision Support for Autonomous Military C2

We are currently folding the lessons learned from process control into advanced concept development for proactive supervision in military C2. Many of the issues and problems are similar across domains. As in process control, today's military unmanned vehicle control displays are characterized by dense, real-time depiction of vehicle and sensor parameters. As in process control, alarms and alerts are generally uninformative, simplistic data-level thresholds not prioritized or aligned to task, and lacking the context necessary to process them effectively. As in process control, SMEs report engaging in work-arounds to facilitate monitoring UAV status from dense displays of real-time vehicle parameters [4]. In interviews, military SMEs reported circling key parameters on their screens to help guide attention, and annotating starting values as references for comparison against current values. The common theme is that legacy, operation-centric display metaphors are not providing good task support. Instead, users are burdened to develop, train, and implement work-arounds external to the systems that should be *intrinsically* supporting their task needs. With the conservatism of large military system development, absent clear guidance and demonstration of these issues, these metaphors are likely to persist.

Motivated by the failures of the current displays and tools, we scoped a UCD effort to define new display metaphors to support the desired shift to multi-system, multi-mission supervision. As in the process control work, the effort was centered around the requirements of the proactive monitoring task flow of Fig 1. However, unlike the process control work that supported *current* activities, the military UCD work defined a *future* vision of how to support multi-system, multi-mission supervision.

For adaption to the future military autonomous C2 domain, our display concepts for current process control were tailored and evolved in several key ways. The future timeline and different scope of the military domain application enabled us to invoke several additional capabilities. These included (1) supervision of multiple instead of single processes and missions, (2) monitoring performance indicators instead of individual operating parameters, (3) parameterizing and integrating context into trended indicators vs. context not parameterized and not integrated, and (4) automated projection of future state vs. no projection automation.

The scope and span of supervisory oversight determines what type of indicators to show. For example, for today's process control, we focused on proactive monitoring of process *parameters*. For today's military C2, we will separately apply these concepts to help unmanned vehicle operators proactively monitor vehicle and sensor performance indicators. However, for future military C2, the focus is on proactive monitoring of multiple missions. This entails aggregating vehicle and sensor indicators to mission level *performance indicators* so that supervisors can better answer questions such as "how are my missions doing? are they meeting their objectives?"

The current state of the prototype is an initial "*Contextualized Trend Board"* user interface concept, shown in Fig 4. The aim is to support proactive monitoring of multiple missions through an interface that summarizes trends in key status icon indicators, and that helps direct attention to those with the most pressing, developing issues.

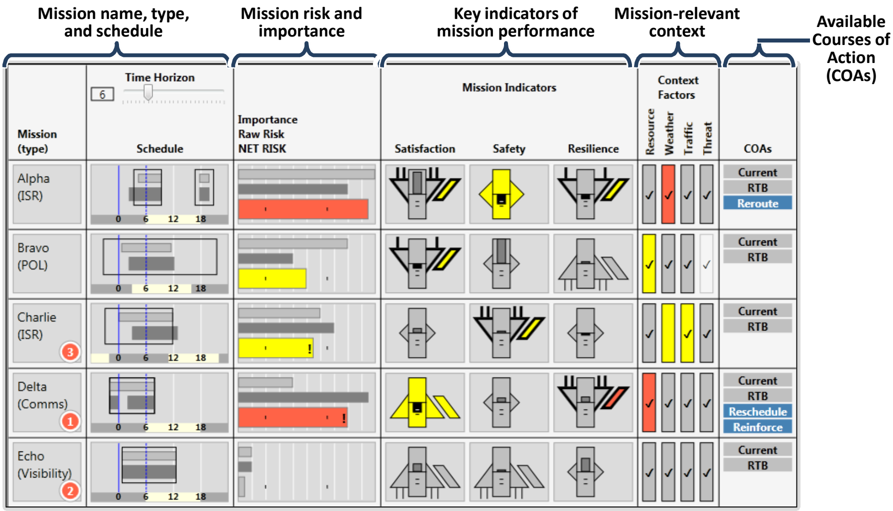


Fig. 4. Contextualized *Trend Board* concept prototype. A top-level overview in a hierarchy of displays for proactive monitoring of multi-mission autonomous military C2.

Triage and diagnosis tasks are supported by a key innovation of directly factoring context into the trended values. In the fourth row of Fig 4, the large red bar for mission risk directs a supervisor's attention to a potential problem in an ongoing communication mission. The resilience of the mission is degrading and projected to become critical, as indicated by the red flange to the right of the status icon. But rather than have no context to understand the reasons for this degradation, the interface shows that mission-relevant weather context is available, impactful, and factored into the indicators (penultimate column) providing the supervisor with an awareness that the resilience problem may well be weather-related, facilitating triage of the problem.

All elements of the proactive monitoring task flow have been woven together in the interface concept and are currently being refined in a rapid prototyping spiral with military SMEs. Fig 4 shows only the top overview layer of a more extensive tool and display set composed of progressive layers that will enable navigation into other layers for detailed diagnosis and intervention tasks related to each mission.

The majority of research and development efforts aimed at reversing the many-operators to single-vehicle ratio have been focused on the technology capabilities, which are necessary but not sufficient to achieve the vision. The future multi-vehicle, multi-mission supervisor will need displays and tools to oversee the autonomy and missions. The work shown here, though initial, provides new metaphors and concepts to define and guide development on a path to the vision.

6 Conclusions

We are entering the age of autonomy with high hopes for advanced technologies, on the one hand, yet only questionable legacy interfaces for them, on the other. These legacy displays often mimic superficial aspects of the work domain, either by showing information realistically in space or time, or through recapitulating paper artifacts from an operation-centric past. These displays provide inadequate support for current tasking, as evidenced by the poor performance they support in controlled testing and the similarity of workarounds they force users to engage in across application domains. Complicating matters, users may believe they are able to extract more information from these displays than they actually can. Legacy display metaphors do not provide a promising basis on which to build needed supervisory decision support for managing the influx of automation.

The autonomy revolution is nearly upon us. All is not bleak, but now is the time to clean house of outmoded display formats.

Acknowledgements. We gratefully acknowledge the technical assistance of Dan Manes and Heather Kobus of PSE. Sponsored by the Office of Naval Research of the US Department of Defense, program officers Dr. Julie L. Marble (current) and Dr. Jeffrey G. Morrison (previous), under contract N00014-12-C-0244. Also sponsored by Emerson Process Management under the technical oversight of Mark Nixon.

References

1. Dickens, C.: A tale of two cities. Chapman and Hall, London (1859)
2. Rasmussen, J.: Human factors in a dynamic information society: where are we headed? Erg. 43, 869–879 (2000)
3. Gugerty, L.: Using cognitive task analysis to design multiple synthetic tasks for uninhabited aerial vehicle operation. In: Schiflett, S., Elliott, L., Salas, E., Coovert, M. (eds.) Scaled Worlds: Development, Validation, and Application, pp. 240–262. Ashgate Publishers, London (2004)

4. Cook, M.B., Smallman, H.S.: Human-centered command and control of future autonomous systems. In: 18th International Command & Control Research Technology Symposium (Paper presented at Symposium), Arlington, VA, June 19-21 (2013)
5. Office of the Secretary of Defense: Unmanned Systems Integrated Roadmap 2009-2034. Office of the Secretary of Defense technical memo, Department of Defense (2009)
6. Sheridan, T.: Supervisory Control. In: Salvendy, G. (ed.) Handbook of human factors and ergonomics, 3rd edn., pp. 1025–1052. John Wiley, New York (2006)
7. Jamieson, G.A.: Human Factors: to compete or cooperate? Mech. Eng., `http://www.memagazine.org/contents/current/webonly/wexa0208.html`
8. Diaper, D., Stanton, N.: The handbook of task analysis for human-computer interaction. Lawrence Erlbaum Associates, Mahwah (2004)
9. Mumaw, R.J., Roth, E.M., Vicente, K.J., Burns, C.M.: There is more to monitoring a nuclear power plan than meets the eye. Hum. Factors. 42, 36–55 (2000)
10. Burns, C.M.: Towards proactive monitoring in the petrochemical industry. Safety Sci. 44 (2006)
11. Yin, S., Tan, A., Helander, M.: Proactive Process Control Monitoring using Trends. In: Proceedings of the Human Factors and Ergonomics Society 52nd Annual Meeting, pp. 2003–2007. HFES, Santa Monica (2008)
12. Nimmo, I.: Adequately address abnormal situation management. Chem. Eng. Prog. 91, 36–45 (1995)
13. Bransby, M.: Design of alarm systems. In: Noyes, J., Bransby, M. (eds.) People in control: Human factors in control room design, pp. 207–222. IEEE, London (2001)
14. Smallman, H.S., St. John, M.S., Oonk, H.M., Cowen, M.B.: 'SYMBICONS': A hybrid symbology that combines the best elements of symbols and icons. In: Proceedings of the Human Factors and Ergonomics Society 45th Annual Meeting, pp. 110–114. HFES, Santa Monica (2001)
15. Smallman, H.S., St. John, M.S.: CHEX (Change History EXplicit): New HCI concepts for change awareness. In: Proceedings of the Human Factors and Ergonomics Society 47th Annual Meeting, pp. 528–532. HFES, Santa Monica (2003)
16. St. John, M., Smallman, H.S., Manes, D.I.: Recovery from Interruptions to a Dynamic Monitoring Task: The Beguiling Utility of Instant Replay. In: Proceedings of the Human Factors and Ergonomics Society 49th Annual Meeting, pp. 473–477. HFES, Santa Monica (2005)
17. Smallman, H.S., St. John, M.S.: Naive realism: Misplaced faith in realistic displays. Erg. Design 13, 6–13 (2005)
18. Smallman, H.S., Cook, M.B.: Naïve Realism: Folk fallacies in the design and use of visual displays. Topics in Cog. Sci. 3, 579–608 (2011)
19. Hegarty, M., Smallman, H.S., Stull, A.T.: Choosing and using geospatial displays: effects of design on performance and metacognition. J. Exp. Psych. Appl. 18, 1–17 (2012)
20. Hollifield, B.: The high performance HMI: process graphics to maximize operator effectiveness. Intech (November/ December 2012), `http://www.isa.org/intech/20121204`
21. Hajdukiewicz, J., Wu, P.: Beyond trends: A framework for mapping time-based requirements and display formats for process operations. In: Proceedings of the Human Factors and Ergonomics Society 48th Annual Meeting, pp. 1885–1889. HFES, Santa Monica (2004)
22. Jamieson, G.A., Miller, C.A., Ho, W.H., Vicente, K.J.: Integrating task- and work domain-based work analyses in ecological interface design: A process control case study. IEEE Trans. Sys., Man, and Cyber. - Part A: Sys And Hum 27, 887–905 (2007)

The ART of CSI: An Augmented Reality Tool (ART) to Annotate Crime Scenes in Forensic Investigation

Jan Willem Streefkerk, Mark Houben, Pjotr van Amerongen, Frank ter Haar, and Judith Dijk

TNO, Kampweg 5, 3769 DE Soesterberg, The Netherlands
{j.w.streefkerk,mark.houben,pjotr.vanamerongen,frank.terhaar, judith.dijk}@tno.nl

Abstract. Forensic professionals have to collect evidence at crime scenes quickly and without contamination. A handheld Augmented Reality (AR) annotation tool allows these users to virtually tag evidence traces at crime scenes and to review, share and export evidence lists. In an user walkthrough with this tool, eight end-users annotated a virtual crime scene while thinking aloud. Qualitative results show that annotation could improve orientation on the crime scene, speed up the collection process and diminish administrative pressure. While the current prototype suffered from technical limitations due to slow feature tracking, AR annotation was found to be a promising, usable and valuable tool in crime scene investigation.

Keywords: Augmented Reality, Indoor Positioning, User Walkthrough, Forensics.

1 Introduction

The main aim of crime scene investigation (CSI) is to find, examine and secure traces left by perpetrators. Forensic professionals involved in CSI have to collect evidence in a secure, traceable manner for juridical purposes. In recording the crime scene, it is very important to not erase, alter or contaminate evidence. Most often – due to costs and clearance of the area – it is also important to work quickly. Augmented Reality (AR) annotation enables adding virtual information to surfaces or objects in real life. In this way, evidence traces can be tagged and visualized right at the crime scene. Applying AR annotation to forensics may accelerate investigation while avoiding interference with evidence by virtually highlighting relevant items and locations. Furthermore, it may assist collaboration and coordination, as well as transfer of findings between forensic investigators. At a glance it is known which trace has been found where and which areas still need to be searched, which helps the spatial perception of the evidence at a crime scene and minimize miscommunication.

R. Shumaker (Ed.): VAMR/HCII 2013, Part II, LNCS 8022, pp. 330–339, 2013.

This paper presents the design and evaluation of an AR annotation tool for forensic professionals. The question guiding this research is: ‘To what extent will the concept and features of an augmented reality application provide added value for forensic professionals in their work process for securing traces?’ First, related work in AR annotation, also for forensic professionals, is presented. Then, the work processes involved in CSI and the subsequent design of the AR annotation tool is presented. The method and results of a qualitative evaluation with twelve forensic end-users are detailed, in which the added value of this tool for the forensic process is assessed. Finally, the main findings, benefits and drawbacks of the tool are discussed.

2 Related Work

Augmented Reality (AR) refers to adding virtual information to real life scenes. AR has become fashionable since the 1990’s, and has been applied in education, medicine, military, entertainment, and industry [1]. The most common use of AR is attaching annotations or virtual “tags” to real world objects, which can be seen through a display device (e.g. an iPad with camera). AR annotation is contrasted with 2D map annotations in that it displays tags in a 3D real world view. Until recently, this has been implemented mainly for entertainment purposes, such as museum, games or city-tours [2], and relying on content that is generated off-line by others. Not many applications allow users to create new or modify annotations on-the-go (online). To our knowledge, online annotation has not yet been implemented to benefit professional domains such as law enforcement or military organizations.

2.1 Enabling Technology for AR Annotation

Enabling technology for AR annotation has made significant progress. However, two important technical issues need to be addressed in order for online, indoor AR annotation to work: perceptual issues and positioning issues. Perceptual issues in working with AR systems include: inaccurate depth perception, inaccurate depth ordering, object matching, and low visibility due to displays [3]. [3] concluded in their survey that: “Users are regularly unable to correctly match the overlaid information to the real world,…”. This is a serious risk for user acceptance when AR techniques are applied to professional domains. These perceptual issues are further influenced by whether a handheld or head-mounted device (HMD) is used. While the head-mounted devices leave users’ hands free for other tasks, it has been shown that using HMD’s over longer periods of time can cause visual discomfort or even motion sickness [4]. Also inputting information using a HMD is more difficult than using a handheld device, but here gesture recognition could be a solution (see e.g. [5]).

Augmented Reality annotation systems rely on position and orientation information to accurately display tags. Outdoor Augmented Reality systems can use GPS tracking to obtain location information (see e.g. [6]), but indoor positioning remains a challenge. To solve the issue of indoor navigation and positioning, [7] propose a vision-based location positioning system based on a large, preconstructed image set.

Their solution recognized key features from the real world and matched these to the image set to determine the location of the user. Using this setup, they achieved almost 90% reliable recognition of location. This approach could be feasible in predictable, small-scale and known environments, but not for forensic professionals. Crime scenes are by definition unknown and unpredictable, where creating an image set for the whole scene would take too much time and effort. In this paper, we propose a similar approach based on feature tracking, but processing is done in real-time. This has its own challenges, as will be elaborated in the discussion.

2.2 AR Annotation in Forensics

As application domain, we are primarily concerned how crime scene investigation would be aided by AR annotation. A related effort to implement AR for forensic professionals used GPS and indoor localization with Ultra Wide Band beacons to create a virtual incident map [8]. Interestingly, this allowed users to tag annotations to objects, take pictures and send these to a central processing bureau. This seems a useful approach, although no users were involved in the design or assessment of the application. Also, relying on UWB beacons requires bringing more equipment to a crime scene that has to be calibrated and can potentially contaminate evidence. A recent related effort in the forensics domain employed a HMD display to support crime scene annotation and collaboration (e.g. with an expert or coordinator) across distances [9]. Users wear a tailor-made, non-see-through head mounted display. An expert at a distance could "look over the shoulder" for collaboration. However, the tags that a user could virtually "leave" at a crime scene primarily indicate the no-go areas and not traces of evidence that we focus on. Consequently, their analysis focused more on the added benefit of AR for collaboration than for the primary forensic process (securing traces). As said before, securing traces and inputting the necessary information in a system will not be very comfortable using a HMD over longer periods of time.

2.3 Involving End-Users in the Design

To optimize the potential of an Augmented Reality tool for forensic professionals, these end-users should be included in the design and evaluation stages of such a tool [10]. Perceived usefulness is a high determinant of whether or not Augmented Reality applications will be accepted and used [11]. This stresses the need to incorporate end-users early in the design approach of AR tools to maximize the perceived usefulness for a specific target group. When innovative solutions, such as AR tools, are applied to new domains, evaluation methods should be selected carefully and tuned to domain specific criteria [12]. In the current paper, end-users are included in the design stage (to acquire operational demands and user requirements for the AR tool) as well as in the prototype evaluation (to acquire qualitative feedback on added value of the AR tool for forensic work processes). This approach is similar to that of [6] in that it uses an AR annotation prototype to collect feedback on different measures from end-users.

3 AR Tool Design

Contrary to earlier efforts, this paper presents an entire design and evaluation iteration of a prototype handheld AR annotation tool for CSI. Starting out, interviews with forensic experts explicated the operational demands of forensic investigation. High-level requirements were outlined and detailed into interface features, which were implemented in the prototype to gather end-user feedback in an early stage.

3.1 Operational Demands

To understand the operational demands in crime scene investigation and to guide our design effort, we held three interviews with two forensic professionals about their primary work process (securing traces at a crime scene), the preconditions for their work and the challenges they encountered. Currently, their work process encompasses the collection of physical trace evidence ("traces") at a crime scene. For each trace, the physical location, type, method of collection and remarks are recorded on a paper-based evidence list. Traces are uniquely numbered, and based on these numbers, photographs of the traces are collected and added to a database at a later stage. Forensic professionals create a 2D map of the crime scene, showing locations of all traces. In addition to this paperwork, the traces have to be carefully collected, packaged, labeled and transported to the CSI lab. At the lab, the traces and evidence lists need to be reviewed for completion and correctly filed for the judicial process.

From the interviews, the following preconditions for this process were identified:

Crime scenes should not be contaminated. No evidence at the scene should be altered or destroyed, requiring that the process of securing traces proceeds fast and sterile. All equipment brought on the scene is carefully decontaminated and personnel at the scene wear special clothing (mask, gloves, suit). The first officer arriving at the scene is responsible for keeping it as untouched as possible.

Crime scenes should not be visited longer than necessary. This lowers the risk of contamination, but any number of reasons can increase the time pressure on the forensic team. For example, victims wanting to return to their home, mortal remains that have to be cleared, or decaying biological evidence. Forensic professionals work as fast as possible while at the scene, as there is one chance of visiting a crime scene.

Crime scenes should not be visited by more people than necessary. Forensic professionals need to coordinate and collaborate with others while collecting evidence. With the supervisor or trace coordinator, task allocation is coordinated (who collects which traces at which parts of the scene?), but also collaboration with trace experts is required (e.g. radiological experts). The need for coordination and collaboration increases with the complexity of the crime scene.

Crime scenes should be documented as completely and accurately as possible. This includes all traces found at the scene, who visited the scene, methods of investigation, collection and transportation used. This is required to aid the judicial process and increase the chance that the evidence can be used for a conviction in a court of law.

Major challenges that forensic professionals face in their work process include inefficient documentation and administration (e.g. copying paper-based evidence lists in multiple databases, digitizing a paper-based map sketched at the scene), time pressure while having to work accurately (due to the preconditions outlined above), the inability to visit the crime scene again, and limited means of communication.

3.2 Design Requirements

Due to these challenging operational demands, the design requirements for the AR tool focused on making interaction with the tool as intuitive and effortless as possible. Adding tags should proceed via a point-and-shoot interaction mechanism, using a touchscreen and a camera. Trace information, photographs and details should be automatically coupled, to avoid double work. Storing, retrieving and exporting this information to relevant databases should be easy. Importantly, previously annotated crime scenes should be accessible even when not at the scene. This is expected to improve briefings, reduce collection time and improve documentation. Furthermore, users should be able to work with the tool wearing protective clothing such as gloves, and the equipment should be easily cleanable. Its functioning should not rely on bringing additional equipment (e.g. indoor positioning beacons) to the scene [8, 9].

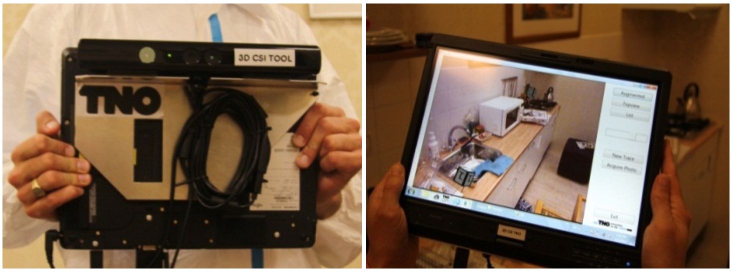


Fig. 1. The implemented AR tool on the tablet in operation at the simulated crime scene

3.3 Hardware and Sensor Technology

The most challenging part to display the virtual tags was generating accurate position and orientation information without relying on indoor positioning systems. The AR tool was implemented on a Toshiba tablet PC with touchscreen (Fig. 1) and programmed in Windows Presentation Foundation. To the tablet PC, a Microsoft Kinect camera was connected (through USB), powered by a portable battery. The 3D position and orientation is estimated using feature tracking on color images combined with depth images from the Kinect. Dedicated feature tracking software (written in C++) for a handheld 3D scan device ran in the background to provide position and orientation information. This enabled the tool to display AR annotations (tags) with sufficient accuracy and redisplay the tags when the tool pointed in the same direction again. While this setup allowed relatively accurate 3D positioning, it required that the tablet started a recording from a fixed starting point and that the tablet did not move too fast to allow the feature tracking to keep up.

3.4 AR Tool Interface

The AR tool interface consisted of a camera view that filled up most of the screen, and a button toolbar at the right of the screen. Users could select between three views: an Augmented Reality view (Fig. 2), a map view (Fig. 3), and a trace-list view (Fig. 4). The toolbar contained buttons to switch between views and two buttons to add a new tag or a photo. The slider on the toolbar adjusted the size of the tags (numbered gray cubes) in the AR view and map view. The exit button below closed the tool.

When adding a new tag, a green cross-hairs in the AR view could be used to point to the location of a trace. Pressing the button "New Trace" added a virtual tag to the indicated surface (e.g. a table), shown in AR view as a numbered cube (Fig. 2). Clicking on this tag opened a trace property window in which trace information could be added or changed (number, date, location, trace type and subtype). Clicking on "Acquire Photo" switched to the AR view to add photos to the current tag. In the trace-list view the complete list of traces could be reviewed, saved and exported. Also, a previously annotated crime scene could be loaded.

Fig. 2. The AR tool interface showing the AR view with a tag near the sink

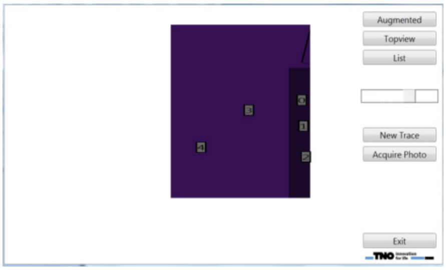


Fig. 3. The AR tool interface showing the 2D map view

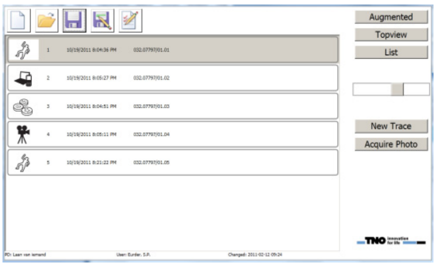


Fig. 4. The AR tool interface showing the trace-list view

As said, proper display of tags required accurate position information. If no position information was available, due to loss of feature tracking, the tool showed cross-hairs in red, to indicate that no tag could be added at that moment. In addition, a small still image appeared in the top left corner, showing the last image the feature tracking software had recognized. When users lined up the AR view with this image, feature tracking resumed as normal and the cross-hairs turned green again.

4 Evaluation Method

To ascertain the added value of the AR tool, forensic professionals conducted a walk-through of the tool at a simulated crime scene followed by a plenary discussion.

4.1 Participants

Twelve male participants, aged between 30 and 55 years, attended the evaluation. All were experienced forensic professionals from either the Dutch police force, the Dutch Police Academy or the Netherlands Forensic Institute. Their specializations included technical detectives, forensic detectives and forensic advisers. All participants attended the plenary sessions, while eight of them performed the user walkthrough.

Fig. 5. A forensic professional working with the AR tool during the user walkthrough (*left*) and an overview of the simulated crime scene (*right*)

4.2 Procedure and Data Collection

In total, the evaluation lasted six hours. Three walkthrough sessions of 40 minutes each were held in small groups (2-3 participants) with eight participants in total, guided by two test leaders. First, participants were briefed on the functionalities of the AR tool. Next, they were introduced to the tool at the simulated crime scene (Fig. 5). In turn, each participant moved from the starting point around the room with the AR tool to conduct a series of standard forensic tasks. First, they looked to see which traces already had been secured. Then, each participant secured a number of interesting traces themselves (e.g. knife in the kitchen sink, cigarettes in the ashtray). They added these traces to the trace list, included a photo and detailed information. Upon completion, they reviewed the trace list and the map view of the crime scene. During the walkthrough, the test leaders gave extra instruction to participants when necessary. They posed questions and stimulated participants to discuss specific features aloud while working with the tool. After the three groups completed the walkthrough, a plenary discussion was held with all twelve participants. All participants in the plenary discussion had either completed the walkthrough or watched the walkthrough via remote video connection. Video and audio recordings were made for analysis purposes. In the plenary discussion, specific statements were posed to the group and based on the subsequent discussion, remarks from the group were recorded. All qualitative feedback from the participants (statements) were grouped into five categories: added value of the concept, added value of the features, usability of the tool, impact on the forensic operational process, and necessary improvements and additions to the tool.

4.3 Results

Added Value of Concept and Features. Most participants were positive about the added value of the concept of adding virtual tags. They expect that it will improve their awareness of a crime scene and reduce the time needed for a first orientation on the scene. They found the usability of adding and reviewing tags very intuitive and easy. The screen size and accuracy of displaying tags in the AR view was sufficient for this application as well as for coordination tasks away from the crime scene. Most criticisms dealt with the current prototype implementation using online feature tracking. This required a fixed starting point and severely limited the speed of movement due to the computationally intensive feature tracking. Participants felt they had to move unnaturally slow with the prototype tool. The tablet PC, Kinect and battery made the current tool too heavy for prolonged use.

Usability. Concerning the visualization of tags, the gray cubes could obscure an object in the AR view. Scaling the size of the tag with the slider did not mitigate this completely. Instead, participants mentioned that tags should be moveable and displayed next to an object with a connector line. Another option is to display a circle around an object as a tag. Minor usability drawbacks were that the tool requires a certain level of lighting (which is absent in some crime scenes) and that numbering of tags should be improved (showing trace numbers in photographs). Also, the “exit” button should be moved to a proper place. Participants thought this button saved the trace-list to file, while in reality clicking on it quit the application altogether.

Impact on Forensic Processes. Participants were very positive about the expected impact of the tool on the forensic processes. They think it will be very useful for speeding up the whole trace collection process, specifically because traces are captured digitally and trace numbers are directly coupled to a crime case. They expect that digital collection of trace lists and tags on a 2D map will reduce time needed for documentation. Furthermore, having a quick overview of traces improves coordination, specifically in deciding on collection methods, and collaboration with trace experts at the bureau. The tool is applicable for most types of crime scenes, except for very complex ones for which a higher accuracy is needed. It will especially help forensic professionals in routine cases (high-volume crimes) that involve a lot of paperwork (e.g. burglary). The tool allows for remote viewing of a digital video stream from the crime scene (not implemented). Participants thought this would improve collaboration with external team members, such as the trace coordinator or subject matter experts.

An interesting objection to implementing an AR tool for forensics was raised by one participant. He was concerned that users might focus attention too much on the tool, thereby overlooking or even destroying evidence. Two others objected to 'yet another tool'. Should this AR tool be implemented, it should substitute paper-based trace lists and be used during the whole operational and juridical process.

Further Improvements. Participants required more freedom in manipulating the tags. For example, they like to use their fingers on the touchscreen to pinpoint the exact location of a new tag, instead of having to point cross-hairs. This allows adding multiple tags in a single camera view, possibly speeding up the first orientation on the scene even further. Suggestions for added functionality include a chronological timeline of added tags to evidence traces, instead of only a numbered list, an indication of no-go areas, an indication of non-localized traces (such as smells and sounds) and the ability to add spoken annotations, to free up the use of their hands more. Concerned for crime scene contamination, users stressed that the tool should be very easy to decontaminate and clean (by wrapping it in a transparent plastic sleeve), and a handlebar should be added for a more natural grip. Incorporating a flashlight would improve the tool's performance even in low lighting conditions.

5 Discussion and Conclusions

This paper presented the design and evaluation of a prototype Augmented Reality annotation tool for crime scene investigation. This tool added virtual tags to evidence traces, based on position information from feature tracking software. The trace list could be digitally completed, reviewed and exported to a database, contrary to current paper-based evidence lists. With twelve forensic professionals, the added value and impact on the forensic process was qualitatively evaluated in a user walkthrough and discussion. In general, end-users were positive about the expected impact of the tool on forensic processes. They considered AR annotation to speed up the trace collection process, reduce the amount of time needed for documentation and found adding and reviewing virtual tags usable and intuitive. However, because the tool was a prototype, the weight and required slow movements (due to feature tracking and processing) constituted technical drawbacks. In addition, flexibility of tag manipulation and visualization should be improved. Functionalities should be extended to include speech recognition software, indication of no-go areas, and a chronological timeline. Finally, the tool should be easier to hold and easier to clean and decontaminate.

These results confirm the added value of AR annotation hypothesized in earlier work [2, 8, 9]. On a broader note, this evaluation again raised the question whether to use a handheld platform or a platform that keeps the hands free (i.e. head-mounted display). While the participants disagreed on this issue, they stressed the need to be able to see the crime scene unmediated for the collection of evidence traces, e.g. without a display in front of their eyes. This would favor the current handheld solution over an HMD. Following hardware developments, the technical limitations in weight and processing are likely to be overcome in the next few years. Moreover, future opportunities of using this AR tool include creating an accurate 3D model representation of the crime scene, complete with features, textures, traces and tags, using the same feature tracking software. Such a model could be used for crime scene reconstruction, and even for juridical purposes in a court of law.

Forensic professionals considered that this AR tool will have added value in terms of speed of trace collection and transfer of information for coordination and collaboration. This evaluation limited itself to collecting qualitative data after a short usage period. Further research over longer periods of time and in real forensic settings should show whether implementing such relatively simple AR tools for crime scene investigation will actually improve performance and collaboration.

Acknowledgements. We thank Jurrien Bijhold and Thanh Le (Netherlands Forensics Institute) for their support in the evaluation in the Forensic Field Lab. We also thank all participants in the evaluation. This research was funded by "Pieken in de Delta", Provincie Zuid-Holland, and Gemeente Den Haag and is part of the CSI The Hague project (www.csithehague.com).

References

1. Azuma, R., Yohan, B., Reinhold, B., Steven, F., Simon, J., Blair, M.: Recent advances in augmented reality. Comp. Graphics and Applications 21(6), 34–47 (2001)
2. Wither, J., DiVerdi, S., Höllerer, T.: Annotation in outdoor augmented reality. Computers & Graphics 33, 679–689 (2009)
3. Kruijff, E., Swan, E., Feiner, S.: Perceptual issues in augmented reality revisited. In: IEEE Proc. of Mixed and Augmented Reality (ISMAR). IEEE (2010)
4. Kooi, F., Toet, A.: Visual comfort of binocular and 3D displays. Displays 25, 99–108 (2004)
5. Wang, R., Popovic, J.: Real-time hand-tracking with a color glove. ACM Transactions on Graphics 28(3), 63 (2009)
6. Schall, G., Mendez, E., Kruijff, E., Veas, E., Junghanns, S., Reitinger, B., Schmalstieg, D.: Handheld Augmented Reality for underground infrastructure visualization. Pers. & Ubiq. Computing 13, 281–291 (2009)
7. Kim, J., Jun, H.: Vision-based location positioning using augmented reality for indoor navigation. IEEE Transactions on Consumer Electronics 54(3), 954–962 (2008)
8. Gee, A.P., Escamilla-Ambrosio, P.J., Webb, M., Mayol-Cuevas, W., Calway, A.: Augmented crime scenes: virtual annotation of physical environments for forensic investigation. In: ACM International Workshop on Multimedia in Forensics, Security and Intelligence, pp. 105–110. ACM (2010)
9. Poelman, R., Akman, O., Lukosch, S., Jonker, P.: As if Being There: Mediated Reality for Crime Scene Investigation. In: Proceedings of ACM Conference on CSCW, pp. 1267–1276 (2012)
10. Olsson, T., Kärkkäinen, T., Lagerstam, E., Ventä-Olkkonen, L.: User evaluation of mobile augmented reality scenarios. J. of Ambient Intelligence and Smart Env. 4, 29–47 (2012)
11. Yusoff, R., Zaman, H., Ahmad, A.: Evaluation of user acceptance of mixed reality technology. Australasian J. of Educational Technology 27, 1369–1387 (2011)
12. Streefkerk, J.W., Van Esch-Bussemakers, M.P., Neerincx, M.A., Looije, R.: Evaluating Context-Aware Mobile Interfaces for Professionals. In: Lumsden, J. (ed.) Handbook of Research on User Interface Design and Evaluation for Mobile Technology, pp. 759–779. IDEA group (2008)

The Virtual Reality Applied in Construction Machinery Industry

Yun-feng Wu[1], Ying Zhang[2], Jun-wu Shen[2], and Tao Peng[2]

[1] School of Mechanical Science & Engineering,
Huazhong University of Science & Technology, Wuhan 430074, China
[2] IT Department, Sany Heavy Industry CO., LTD., Changsha 410100, China
yunfengwu@hust.edu.cn

Abstract. Nowadays, the competition in the construction machinery industry is increasingly fierce. So how to realize the fastest speed to market, best quality, lowest cost, and best service are key factors for enterprises to win the market and users. Advanced technology can do help. It is in this context that VR (Virtual Reality) steps in, applied by manufacturers as a sharp weapon. Sany Heavy Industry is China's largest and the world's sixth construction machinery manufacturer. As a full-scale enterprise, Sany always emphasizes digitization and information technology for the construction of a modern enterprise. VR has been integrated into the enterprise's workflow. In R&D department, VR is commonly used in digital prototyping. But this paper mainly introduces Sany's own virtual reality roaming platform: *VR Flier,* developed to make significant applications in simulation of plant layout and process planning. And service and marketing departments have also received a significant effect via VR&AR applications.

Keywords: construction machinery industry, Virtual Reality application, digital manufacturing, marketing and services.

1 Virtual Reality in Construction Machinery Industry

1.1 VR Technology Owns Its Place in Digital Manufacturing

First, it must be mentioned that the development and growth of the digital manufacturing technology is very significant in the construction machinery industry. With the in-depth application of IT technology, as well as digitization, informatization and modern enterprise operating management concept onto a new step, digital information is no longer the exclusive means of the R & D department. Meanwhile, digital manufacturing has become a high-efficiency tool on which construction machinery companies have focused to pursue best performance in production and management.

Many large enterprises: Manitou (France), CNH, Caterpillar, Terex [1], John Deere (U.S.), Volvo (Sweden), etc., for new products to meet the customers' changing needs more quickly into the market, will do the digital process validation before production, including the balance of the production line, the device, workers, tools, operation and

R. Shumaker (Ed.): VAMR/HCII 2013, Part II, LNCS 8022, pp. 340–349, 2013.

other factors of the entire production line, which would improve the overall manufacturing process of high efficiency and accuracy. VR technology as a means of 3D verification is used in the process design and validation phase, playing an important role in digital manufacturing. In addition, the rise of the 3D operation instructions, using 3D interactive way to extend the usage of 3D product design data, helps a lot prompting traditional operating standards in form, content and effect [2].

1.2 VR Technology Adds More Enchantment and Interest in Sales and Services

Virtual reality technology with its full three-dimensional, immersive, multi-angle interactive experience, combined with taking the Internet's advantages of rapid propagation and low cost, is more and more chosen by enterprises as an important product display and propaganda channel. The large corporations in construction machinery industry at home and abroad have gradually introduced VR technology to create new network marketing and service models for their products. Here are some examples below, many are not listed in this paper, like Liebherr (Germany) [3]. It could be said that digital Marketing services are covering the full lifetime cycle of the product.

- Hyster's warehouse simulator software helps customers choose the right configuration of trucks for their warehouse(Fig.1.a).
- HAMM's interactive PDF documents supply the self-selection products (Fig.1.b) [4].
- Zoomlion presents AR technology on their new product launch stage (Fig.1.c).
- Xugong's online three-dimensional products exhibition hall (Fig.1.d).
- Caterpillar's virtual training system(VTS) for training program (Fig.1.e) [5].
- Ritchie Bros. online secondhand auctioneers (Fig.1.f).

The common characteristics of these applications are to provide complete digital solutions supporting a wide range of products, which promote customer experience greatly. What's more, many companies appeared and specifically focused on doing digital solutions. The applications and simulators they created and developed could be fitted on various models of similar products from different companies& brands in the market. The A1A company's online "3D lift plane" is a lift planning and crane selection application, which helps find the most economical crane configurations for the lift (Fig.2). It can also simulate the entire lift in 3D Lift Plan to save time and increase efficiency on the jobsite. This application's database has a crane library of over 900 cranes on the market which contains accurate crane dimensions and load charts, and provide hundreds of custom 3D objects.

As a branch of the VR technology, Augmented Reality (AR) technology, also called mixed reality, with the help of computer technology, applies the virtual information into the real world, making real environments and real-time virtual objects superimposed coexist in the same screen space. AR technology in recent years is widely used in marketing services, new product releases, and show stage, bringing audience surreal sensory experience, deep impression and visual impact. Also there are AR exhibition interactive publicity, AR marketing service manuals, AR products experience online, AR technical support for maintenance, etc.

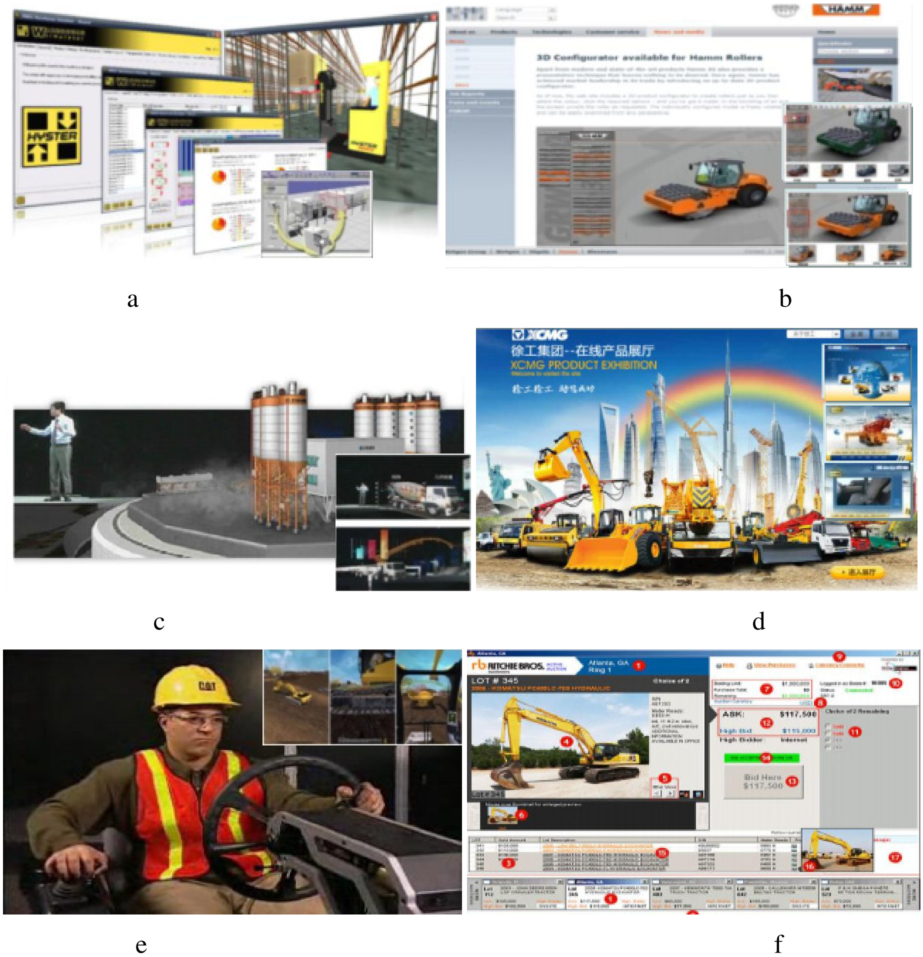


Fig. 1. VRApplications in Construction Machinery Industry

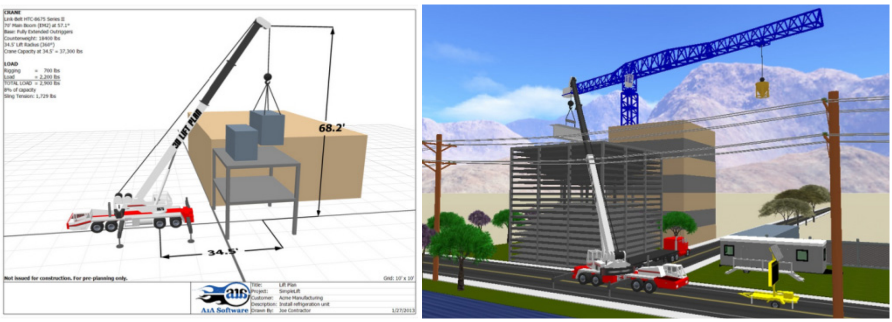


Fig. 2. A1A Software's "3D lift plane"

2 Virtual Reality in SANY

Sany Heavy Industry as an emerging enterprise has developed into China's biggest and the world's sixth international construction machinery manufacturer in less than two decades. It has also accumulated a lot of experience in the digital manufacturing and VR application tools since more than four years ago. VR technology is playing an important role in the development and growth of Sany, and Sany accordingly provides VR technology with a broad experimental space and scale applications in the construction machinery industry. Therefore below this paper will introduce the application of VR technology in Sany, primarily from four aspects: plant planning, process training, marketing and services, and VR training.

2.1 Digital Factory and Plant Planning

The concept Digital Factory is put forward to solve the "divide" between product design and product manufacturing, and improve various aspects of the product life cycle, including design, manufacturing, assembly, logistics function. Via advanced inspection in a totally virtual environment, it can reduce uncertainty from the design to manufacturing, and improve product reliability. In 2009Sany introduced the concept of the digital factory and decided: for all the new or to be rebuilt plants and production lines, it must go through a three-dimensional digital validation, and make sure that the layout and process are qualified before starting the projects, which is also called reality comes after virtualization, commence after validation.

Sany has introduced a variety of professional logistics and process planning simulation software, such as Dassault's DELMIA, QUEST, etc., to help improve the process of planning and logistics simulation. In aspect of plant planning, Sany has more than a dozen domestic and foreign industrial parks, so the planning simulation for layout of factories and production lines has unique and systemic requirements. Therefore, Sany developed a three-dimensional simulation platform named as VR Flier, and concluded a fixed set of plant planning solution based on this platform.

VR Flier Platform. Based on the OSG, VR Flier provides general application software platform architecture to help users with physical simulation, user interface, scene and sound's management and drive. With object-oriented design, VR Flier is modular, extensible, supports a variety of VR peripherals, and has friendly interface [6]. Figure 3 shows that VR Flier's software platform architecture, which consists primarily of the basal class library and application building blocks.

Modeling Preparation. At the beginning, the first job is to prepare all the needed 3D models. Model categories include buildings, equipment, machinery, products and other things in the actual plant. Model format should be converted through VR Flier converter into the required format, and then all the models can be imported into the platform. The final virtual scene in VR Flier will display each element of the actual plant. The initial stage of this work looks very cumbersome, because many 3D models need to be measured and built. But once this modeling process has been done for the first time, it will accumulate to form a models library with clear classification and rich content, which builds a good foundation for future use.

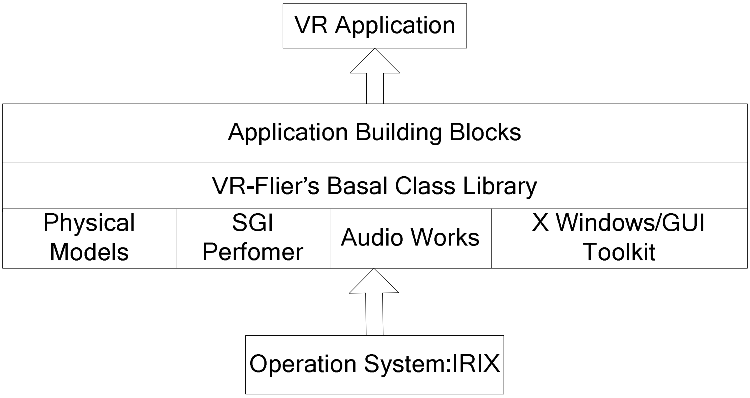


Fig. 3. VR Flier's Software Platform Architecture

Static Layout. After all the required models have been put into the platform, users can transform and rotate the models to place them in the appropriate position. The overall layout of a plant should be formed in this step, with models of production lines and facilities in right place (Fig.4).

Dynamic Simulation. For dynamic simulation to reflect the state of operation in the plant, VR Flier provides scripts to drive models in virtual scene, showing the flow of the production process and transportation logistics. Also, it has embedded interface which can be used by some professional logistics simulation software. For example, data result from QUEST (logistics Simulation software by DassaultSystèmes) can be directly input into VR Flier to drive virtual equipment to run.

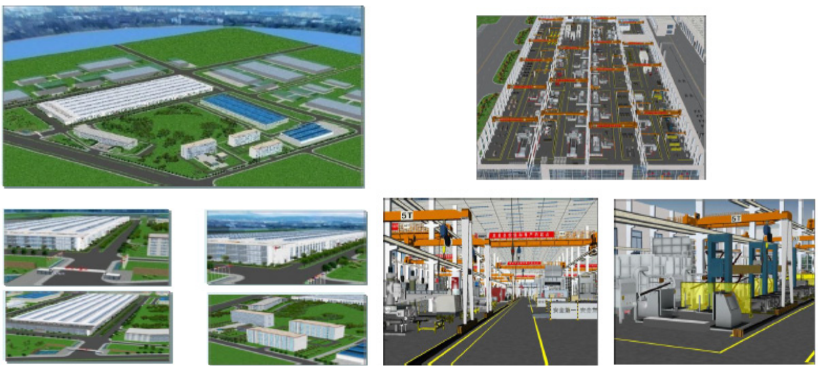


Fig. 4. Layout of Whole Park and Production Lines in Virtual Scene

Result. According to dynamic simulation results, static layout will be adjusted, and then validated by dynamic simulation again. This circular validation can ultimately determine the static layout and process logistics program. In the same project, users can save multiple sets of layout options for policy-makers to watch and evaluate (Fig.5). These optimization before construction can reduce overall cycle time and cost for a planning and construction project.

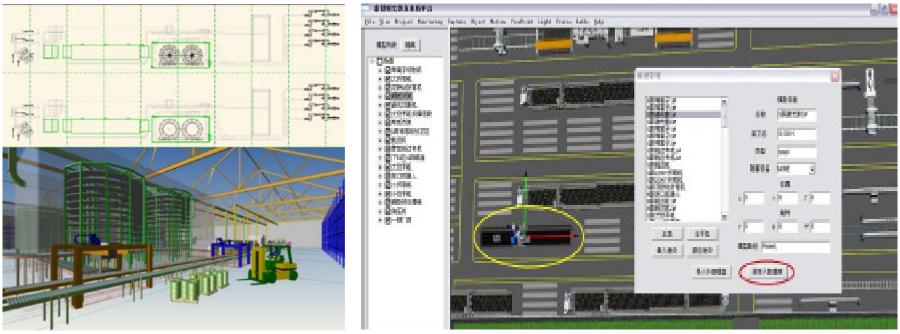


Fig. 5. MultipleSets of Layout in VR Flier

NingxiangIndusty Park began in 2009

In 2009, NingxiangIndusty Park (Fig. 6) was planned to build, covering an area of 367,332m^2, construction area of 93,867m^2 and plant area of 91,940m^2, all of which are presented in virtual world. In the second phase of the construction, it saves a total investment of 8% including equipment and infrastructure, increases equipment space occupancy rate from 16% to 30%, cuts material inventory occupancy rate falling to 18.8%, and shortens infrastructure cycle in half: from 1year to 6 months.

Fig. 6. The Virtual Ningxiang Industry Park (partial)

Green concrete mixing station park planning in Wuhu City

Wuhu City plans to build an eco-friendly concrete mixing station industrial park focus on the supply and distribution of concrete for whole city. The 3D simulation helps planners with industrial park layout and logistics and transport scheme in city. The project's report to the municipal government of Wuhu finally was well received with 3D case display (Fig.7).

Fig. 7. Green Concrete Mixing Station Park Planning

2.2 3D Operation Standard (3D OS) for Process Guiding and Training

Process control is very important throughout the product manufacturing. In the assembly line, workers usually rely on thick paper operation standard to learn the knowledge of assembly process. Animations with3D models and annotations are embedded in the document template, instead of drawing and text description, forming3D interactive electronic operating standards. In Sany, 3D OS has covered a number of products' whole assembly line process.

Fig. 8. 3D Operation Standard and Workers Browsing in MES Machine

The electronic 3D operation standard is put in the MES machines right beside the assembly line in the workshop. Workers can conveniently look for what they want, watching guidance animation, inquiring on material codes, and so on (Fig.8). Manufacturing errors have been reduced by 38.5%. According to data from the workers' training center, training time has been shortened by 56%.

2.3 Digital Sales and Services with VR&AR

Sany also uses mature Virtual Reality and Augmented Reality technology combined with its own business and product characteristics, to develop unique products and services with VR&AR.

- Fast configuration simulator for concrete mixing station (Fig.9.a).
- 3D products exhibition hall online, containing all product series (Fig.9.b) [7].
- Virtual training for product maintenance (Fig.9.c), and operation evaluation system(Fig.9.d).
- AR product brochures show products for customers. The mark and AR display for a pump truck in Fig.9.e&f.

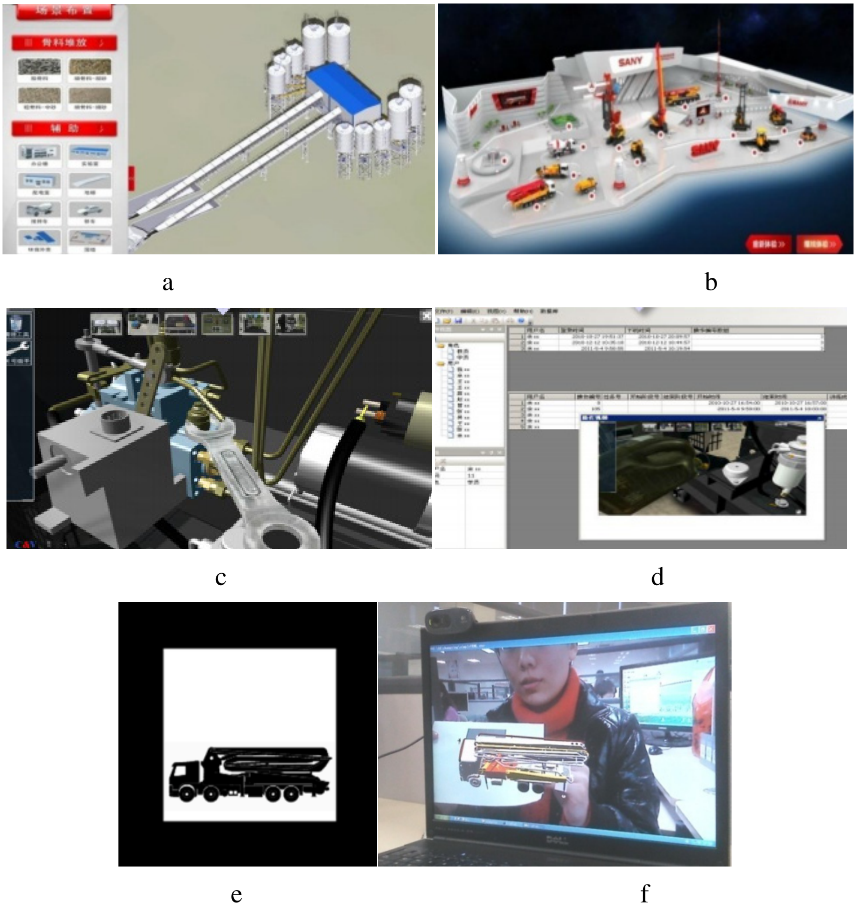


Fig. 9. Digital Sales and Services with VR&AR

2.4 Virtual Training System

Construction machinery is characterized by bulky equipment and complicated working conditions, thus the driver and operator's training is difficult, and takes long time cycle. Virtual training system used in the construction machinery driving operation, replacing most of the hardware by software, has greatly reduced the cost and time, but still resulted a good training effect, and also protected the safety of the operators during training.

With visual simulation by software, as well as real driving environment of hardware, virtual training system can complete a full set of users’ training. This application covers products of crawler crane (Fig.10), crane, and pump truck and so on.

Fig. 10. Virtual Training System for Crawler Crane

3 Conclusions

Generally speaking, VR technology is mainly applied in two aspects in construction machinery industry. On the one hand, co-products using VR technology, such as simulation system software for construction machinery products’ lectotype and training, electronic manuals, operating simulator, have been on market as ancillary and proliferous products which for consumers to use and enhance enterprises' competitiveness and service in training, marketing, and maintenance. On the other hand, VR technology used as a tool in the layout planning and construction of industrial park, workshop and lines, provides a three-dimensional digital experience vecification means to reduce errors, cost and time. In a word, with increasing development towards 3D digitization, the construction machinery enterprises would use VR technology in more and more applications, to create high-return value.

Acknowledgement. In this paper, the cases involved are from the relative companies' official website. Cases about Sany are from interior of the company. The research work was funded by the Ministry of Science, in the National Science and Technology Support Program (2012BAF12B20).

References

1. http://www.gottwald.com/gottwald/site/gottwald/en/system-solutions/simulation.html (Terex simulation and emulation for ports terminal system)
2. Li, S., Peng, T., Xu, C., Fu, Y., Liu, Y.: A Mixed Reality-Based Assembly Verification and Training Platform. In: Shumaker, R. (ed.) VMR 2009. LNCS, vol. 5622, pp. 576–585. Springer, Heidelberg (2009)
3. http://www.liebherr.com/CC/en-GB/region-(asia)/products_cc.wfw/id-17998-0/measure-metric/tab-12111_2228 (Liebherr virtual product display)
4. http://www.hamm.eu/en/aktuelles-und-presse/news/2011/3dconfigurator.html (3D configurator available for Hamm Rollers)
5. https://www.catsimulators.com/ (CAT virtual training simulator)
6. Fan, X., Zhang, X., Cheng, H., Ma, Y., He, Q.: A Virtual Experiment Platform for Mechanism Motion Cognitive Learning. In: Shumaker, R. (ed.) Virtual and Mixed Reality, Part II, HCII 2011. LNCS, vol. 6774, pp. 20–29. Springer, Heidelberg (2011)
7. http://sanygroup.com/showroom/SD/index.htm (Sany online products exhibition hall)

Part IV
Culture and Entertainment Applications

On the Use of Augmented Reality Technology for Creating Interactive Computer Games

Chin-Shyurng Fahn, Meng-Luen Wu, and Wei-Tyng Liu

National Taiwan University of Science and Technology
Department of Computer Science and Information Engineering
Taipei 10607, Taiwan, R.O.C.
{csfahn,d10015015,m9915029}@mail.ntust.edu.tw

Abstract. In this paper, we design interactive games systems that adopt augmented reality (AR) technology. By virtue of a conventional webcam for capturing source images, we develop real-time visual tracking techniques based on edge detection and make 3D virtual objects display on our defined markers that are within the source images in the field of view (FOV) of the webcam. Two kinds of gaming interfaces are created for example: one is an AR based Monopoly game, and the other is an AR based fighting game. There are five classic human computer interface design methods considered to create the above AR based game systems. In the example of Monopoly games, we demonstrate how a traditional table game can be turned into an interactive computer game using the AR technology. We also list the advantages of a marker based approach and state why it is suitable for the interactive computer game. Further, the existing popular game consoles with different gaming interfaces are compared to the two AR based game systems. The comparison results reveal that our proposed AR based game systems are lower in cost and better in extensibility.

Keywords: Augmented reality, human computer interface, AR based game system, interactive computer game, marker recognition.

1 Introduction

AR is an interaction technology which combines the real world and virtual objects in the FOV of a camera, as shown in Figure 1. Because of the combination of real and virtual worlds, it is ideal for entertainment, medical science, education, human-robotic system, and so on. Currently, most of AR interfaces are only based on cameras. Previous literatures usually focused on improving the accuracy of object detection, but entire user interfaces are ignored.

AR is realized by the technology of computer vision and computer graphics. In the computer vision, there are three phases, including object detection, tracking, and recognition. In the detection phase, we detect the regions in which objects are located. Second, we track each of the objects using spatial temporal relations. Third, we recognize each of them by means of a classifier. In the computer graphics, we apply the geometry transformation by the relation of a detected object (in the 3D marker

R. Shumaker (Ed.): VAMR/HCII 2013, Part II, LNCS 8022, pp. 353–362, 2013.

coordinate system), the FOV of a camera (in the 2D screen coordinate system), and the camera (in the 3D camera coordinate system), as Figure 2 shows. After the transformation, a 3D virtual model will be displayed at the position of the detected object (the marker) in the FOV.

Fig. 1. Illustration of AR based interaction technology [1]

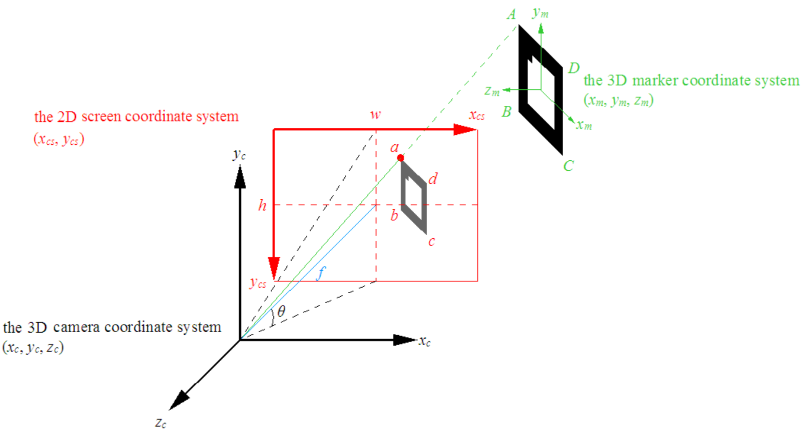


Fig. 2. The relation among the 3D camera coordinate system, the 2D screen coordinate system, and the 3D marker coordinate system

There are two categories of AR technology: marker based and markless based approaches. In the marker based approach [2], we need to design binary patterns for identification. Of such a binary pattern, the corner feature is clear because only black and white colors exist. In the markless approach, we need not design binary patterns for identification, but the computation cost is very high. Usually the algorithm is more difficult to design, and the accuracy is lower in most cases, especially influenced by the variation of the environment lighting conditions. In our proposed AR based game systems, real-time processing is an important factor, and the binary patterns can be regarded as game cards. Therefore, we choose the marker based approach, and design some markers, each of which is associated with a binary pattern printed on a game card for identification. Our design method can provide more than sixty thousand identities that give a good variety of markers for users to apply them. An example of the marker based AR technology is shown in Figure 3.

(a)

(b)

Fig. 3. An example of the marker based AR technology: (a) distinguish binary patterns printed on different game cards in our proposed system; (b) the corresponding 3D virtual model appears on the binary pattern of a game card.

In this paper, two computer game scenarios based on different user interfaces are developed for the AR based environment. We create an easy-to-use interface that is a graphical human interface developed through cross-platform APIs. In our game systems, users can interact with computers via a very intuitive way. Our proposed method enables computers to sense multiple markers in the sight of a camera, and which identity associated with a marker possessing the type, position, and pose is identified. With this information, computers can superimpose 3D virtual models on the specific markers respectively, so that users can interact with the 3D virtual models simultaneously through the markers.

2 Designing AR Based Gaming Interfaces

There are five classic methods to design man-machine interfaces called heuristic evaluation, observation, interviews and questionnaires, logging actual use of users, and user feedback [3],[4]. What follows introduces the details of these five design methods and how they are applied to AR based game systems.

1. Heuristic evaluation:
   Heuristic evaluation can make the research and development team quickly and easily identify usability problems in the man-machine interface products. In the process of heuristic evaluation, based on the control list of usability

guidelines provided by researchers, the evaluators check whether the form of man-machine interface products is right, and identify which items violate usability guidelines. The advantage is that the individual usage problems can be found and the needs of expert users can be listed. The disadvantage is that because the real thoughts of users are not included, unexpected demands of experts cannot be found.

2. Observation:
   Observation is employed to analyze and inquire the actions of users in the study phase, which usually requires three or more users. The advantage is to possess ecology force and to indicate users' work clearly. The disadvantage is that without participant control, the result will be hard to handle.
3. Interviews and questionnaires:
   An interview is used to the analysis phase of operation. In general, it needs five people to participate in. The advantage is that flexible and thorough point of view together with investigation experiences can be required. The disadvantages are that it takes quite a long time and the results are difficult to analyze. Questionnaire is applied to operational analysis and reviews of researches. At least thirty people are required. The advantages are that subjective preference of users can be found. Besides, repeating the whole process is easy. The disadvantage is that in order to avoid misunderstanding, a previous test is needed.
4. Logging actual use of users:
   The method needs more than twenty participants to keep track of full results, like link analysis, layout analysis, and hierarchical task analysis methods. The advantage is the patterns of high usage or less usage can be found. The disadvantage is that it requires a lot of data to analyze contents, which may violate users' privacy.
5. User feedback:
   The method needs hundreds of participants to join in for a long period of time. It is applied to research reviews. The advantage is that it can continuously track whether requests and viewpoints of users are changed. The disadvantage is that it needs stable implementation according to the questions from a specific organization, data retrieval, and so on.

In the following, we create two AR based game systems. Three design methods are chosen to implement them, including observation, interviews and questionnaires, and user feedback.

3 The Design Flow of AR Based Game Systems

In this section, we will elaborate the design flow of the two AR based game systems. One is an AR based Monopoly game, which is a static table game. The other is a joystick controlled AR based fighting game, where the user input is intensive. Because the characteristic of the two games is not the same, the design flow of each of them will be depicted separately as follows.

3.1 An AR Based Monopoly Game

The main inputs of the AR based Monopoly game are different kinds of markers. Figure 4 illustrates the architecture of this game system, and Figure 5 shows a screenshot of the system where many 3D virtual models are located in a real world. There are three input modes in this game. The first one is a marker dice, which is a replacement for real dices used in the traditional Monopoly table game. The second one is a marker cover, which is used for selecting items. As for miscellaneous operations, mouse and keyboard are employed.

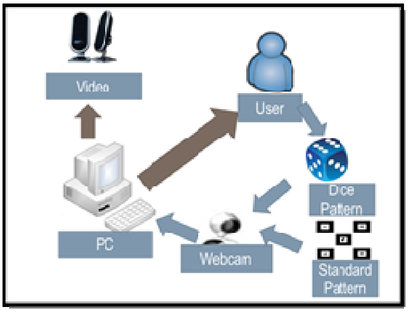


Fig. 4. The architecture of our proposed AR based Monopoly game

Fig. 5. A screenshot of our proposed AR based Monopoly game

Marker Dice

Throwing a dice is an important action in a Monopoly game; therefore, the interface of manipulating this action needs to be designed carefully. To accomplish it, we choose the "interviews and questionnaires" design method depicted in Section 2. Five members are assigned to carry out interviews and questionnaires with thirty people. These thirty people have experiences in playing table games. Several requirements are that not only the dice must be touchable and able to throw, but also the rolling process must be visible in monitors. In [5], it provides an idea for a marker dice, which is in form of a cube with a marker on each side. We adopt such a marker dice in our proposed AR based Monopoly game, where the stop rolling detection of the dice is essential, as shown in Figure 6. The main algorithm of this detection is based on determining whether the parallelism between the orientations of two markers occurs. If the marker on the top side of the dice is parallel to the marker on the table, the marker dice will stop rolling.

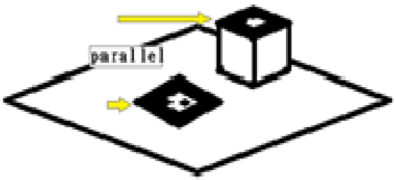


Fig. 6. Stop rolling detection of a dice by checking the parallelism between the orientations of two markers

Marker Cover

There are many conditions that we need to do “select” actions in a Monopoly game, such as deciding which land to buy and determining where to build a house. The design of this interaction is on the basis of user feedback. First, some prototypes of the system are made, and we ask one hundred players to try each of them. After the trials, players send their feelings and suggestions back, and gradually finish the overall system.

The marker cover is the solution to accomplishing a “select” action. The marker cover mechanism is realized using a two-phase approach. In the first phase, when a player wants to choose an item on the table, he or she applies the marker cover to mask an item. In the second phase, to prevent a wrong selection, the action in the previous phase must be confirmed. Users must employ another marker cover to mask the previous cover again. There are two kinds of this cover: one is a “confirm” cover, and the other is a “cancel” cover.

3.2 An AR Based Fighting Game

Generally, the inputs to the system are very intensive in a fighting game, so they are not possible to adopt markers to control the game. The marker here is used to identify the orientations of 3D virtual models. Figure 7 illustrates the architecture of such a fighting game system and Figure 8 shows a screenshot of the system, which is extended from the traditional fighting game scenario. In the AR based game system, users can have their virtual characters displayed in a real world, not limited to a pre-designed virtual background. In addition, the virtual characters are attainable to interact with a real background.

When playing the AR based fighting game, a marker is put in the FOV of the camera. The virtual characters are shown at the related positions to the marker. According to both the “heuristic evaluation” and “interviews and questionnaires” design methods described in Section 2, users tend to play fighting games using joysticks. This is because button-based peripherals are the most responsive input devices, which coincides with the requirement for the characteristic of fighting games.

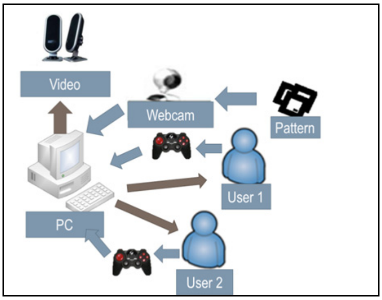


Fig. 7. The architecture of our proposed AR based fighting game integrated with joysticks

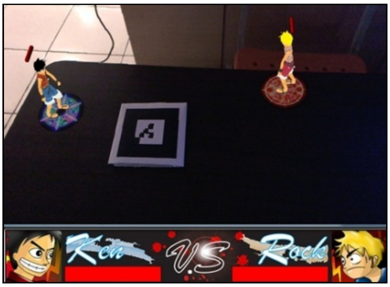


Fig. 8. A screenshot of our proposed AR based fighting game

4 Comparisons of the Existing PC Gaming Interfaces

In this section, our proposed AR based game systems are compared to the existing popular game consoles with different gaming interfaces. The first proposed system is an AR based Monopoly game, and the comparison result between the AR based and traditional Monopoly games whose respect scenarios are shown in Figure 9 is stated below.

(a) (b) (c)

Fig. 9. Three types of Monopoly games created by different technologies: (a) a traditional Monopoly table game; (b) a traditional Monopoly computer game; (c) our proposed AR based Monopoly game.

The traditional Monopoly game on PCs has been popular because the computer treats most trivial and boring parts of the game where users can focus on their strategies to win the game. However, this traditional Monopoly game deprives of the real interactions in a table game. By virtue of the AR technology, our proposed game system still keeps real interactions, when users assign the trivial and boring parts of the game to a computer. The detailed comparison of the three types of Monopoly games is listed in Table 1.

Table 1. Comparison of Monopoly Games Created by Different Technologies

Item compared \ Game type	Traditional Monopoly table game	Traditional Monopoly computer game	Our proposed AR based Monopoly game
Process controlled by a computer	No	Yes	Yes
Interaction on a table	Yes	No	Yes
Interface	Dice	Keyboard & Mouse	Printed patterns
Equipment	Card	PC	PC & Webcam
Viewpoint	Not adjustable	Not adjustable	Adjustable
Resume function	No	Yes	Yes
Multiple players	Yes	Yes	Yes

Next, our proposed AR based fighting game is compared to the existing ones, such as traditional 2D and 3D fighting computer games. The screenshot of each type of fighting games mentioned above is illustrated in Figure 10.

(a) (b) (c)

Fig. 10. Three types of fighting games created by different technologies: (a) traditional 2D fighting computer game; (b) traditional 3D fighting computer game; (c) AR based 3D fighting game

In traditional 2D fighting computer games, users can move the characters only in four directions: up (jump), down (squat), left, and right. In the traditional 3D fighting computer games, users can move the characters freely in a 3D space. In this situation, besides both jump and squat actions, users can move the characters in a 2D space on the ground; thus, the users have more moving strategies to adopt. However, in either the traditional 2D or 3D fighting computer games, their gaming backgrounds are pre-designed. Users can only choose limited scenes in these types of computer games. But, in the AR based fighting game, users need not choose which scene to play in; they simply select a real world as a gaming background arbitrarily. Moreover, it is possible for the characters appearing in the game to interact with the real world, such as the cups colliding with each other on a desk and falling from the edges of a table. Therefore, the extensibility of our proposed AR based fighting game is higher than those of traditional 2D or 3D fighting computer games. Table 2 lists the overall comparisons between our proposed AR based fighting game and the two traditional ones.

Table 2. Comparison of Fighting Games Created by Different Technologies

Item compared \ Game type	Traditional 2D fighting computer game	Traditional 3D fighting computer game	Our proposed AR based fighting game
Game scenes	Limited (Usually 8~20)	Limited (Usually 8~20)	Infinite
Interface	Joystick	Joystick	Joystick
Degrees of freedom	Two	Three	Three
Game extensibility	Normal	Normal	High
Interaction with the real world	No	No	Yes
Object render type	2D image	3D virtual model	Adaptive 3D virtual model
Viewpoint	Fixed	Changeable	Changeable

Finally, two major game consoles, Kinect and Wii, are compared with our system. Both of them have special gaming interfaces. The Kinect exploits a depth sensor to capture the body motions of users; however, the cost of the depth sensor is relatively high [6]. The Wii takes a gyroscope and an accelerator to sense the moving trajectory of the controller. In contrast to the Kinect, the computational cost of the Wii is lower and the hardware cost is cheaper, but the Wii lacks of an AR support [7]. Our proposed game system is based on the AR technology, and its main equipment is a webcam that captures the scenes of a real world into games. Compared to the Wii and Kinect, the price of our proposed AR game system is the cheapest. The computational cost is also lower since we adopt markers as game cards in our proposed system, which are apt to identify the virtual characters and fighting weapons. The comparison result is listed in Table 3.

Table 3. Comparison of the Kinect, Wii, and Our Proposed System

Items compared \ Platform	Kinect	Wii	Our Proposed System
Peripherals	Depth sensor	Sensor and keypad	Camera and joysticks
Equipment cost	Highest	Medium	Low
Interface	Kinect sensors	Wii remotes	Webcam and joysticks
Object render type	3D Virtual model	3D Virtual model	3D Virtual model and real world
Interaction mode	Body motions	Wii remote motions	Moving game cards

5 Conclusions and Future Works

In this paper, an AR based game design flow is presented, and the comparison among different platforms is also made. First, in developing the AR based Monopoly system, we evaluate many user interface design methods. The design of the marker dice is according to the outcome of interviews and questionnaires, and the design of the marker cover is based on the result of user feedback. Second, in the AR based fighting game, the virtual characters are put into the real world by the aid of the marker. Third, the comparison results reveal that our proposed AR based game systems are better than the traditional ones in many aspects. Additionally, the cost and the extensibility of the former systems are lower and higher than other popular gaming interfaces used in the Kinect and Wii.

If the performance of the collision detector for finding virtual characters colliding with real objects can be increased, the extensibility of our proposed AR based game systems can be enhanced. In the future, a framework for such a collision detector will be proposed. This can be a solution to simplifying the applications in most of the AR based systems.

Acknowledgement. The authors thank the National Science Council of Taiwan for supporting this work in part under Grant NSC101-2221-E-011-140-MY2.

References

1. AR Defender, (December 31, 2012), `http://www.ardefender.com/`
2. Kato, H., Billinghurst, M.: Marker tracking and HMD calibration for a video-based augmented reality conferencing system. In: Proceedings of the 2nd IEEE and ACM International Workshop on Augmented Reality, San Francisco, California, pp. 85–94 (1999)
3. Koh, R.K.C., Duh, H.B.-L., Gu, J.: An integrated design flow in user interface and interaction for enhancing mobile AR gaming experiences. In: Proceedings of the IEEE International Symposium on Mixed and Augmented Reality - Arts, Media, and Humanities, Seoul, South Korea, pp. 47–52 (2010)
4. Kirner, C., Zorzal, E., Kirner, T.: Case studies on the development of games using augmented reality. In: A Practical Guide to Usability Testing. Intellect Ltd, Bristol (1999)
5. Colvin, R., Hung, T., Jimison, D., Johnson, B., Myers, E., Blaine, T.: A dice game in third person augmented reality. In: Proceedings of the IEEE International Workshop on the Augmented Reality Toolkit, Tokyo, Japan, pp. 3–4 (2003)
6. Lange, B., Chang, C.-Y., Suma, E., Newman, B., Rizzo, A.S., Bolas, M.: Development and evaluation of low cost game-based balance rehabilitation tool using the Microsoft Kinect sensor. In: Proceedings of the IEEE Engineering in Medicine and Biology Society, Boston, Massachusetts, pp. 1831–1834 (2011)
7. Komlodi, A., Jozsa, E., Hercegfi, K., Kucsora, S., Borics, D.: Empirical usability evaluation of the Wii controller as an input device for the VirCA immersive virtual space. In: Proceedings of the 2nd International Conference on Cognitive Infocommunications, Budapest, Hungary, pp. 1–6 (2011)

A 3-D Serious Game to Simulate a Living of a Beehive

José Eduardo M. de Figueiredo, Vera Maria B. Werneck,
and Rosa Maria E. Moreira da Costa

Universidade do Estado do Rio de Janeiro
IME – Mestrado em Ciências Computacionais,
Rua São Francisco Xavier 524- 6º andar – Bl. B,
CEP 20550-013 - Rio de Janeiro – RJ – Brazil
prof_edu@oi.com.br, {rcosta,vera}@ime.uerj.br

Abstract. Computational tools are increasingly supporting the learning process in several areas. They open new opportunities for teachers to teach contents and interact with their students. This group of tools includes simulations based on multi-agent systems. This work aims to present a simulation game to study the population growth of a beehive. System variables can be changed in order to analyze different results. Aspects such as duration and time of flowering can be manipulated by the student. The multi-agent approach in Distributed Artificial Intelligence has been chosen to automatically control the activities of the application. Virtual Reality is used to illustrate the behavior of the bees that in general, are not able to be seen in the real world or through mathematical simulation.

Keywords: Simulation, Virtual Reality, Multi-Agents Systems, Serious Games.

1 Introduction

Computer simulations are used to recreate real-life processes in computing environments. They allow us to check how the modeled system works under certain conditions and assess the consequences for the overall behavior of the system [1].

The biology area has many phenomena that can be simulated, which include the process of cell division [2] and predator-prey relationship [3]. Another possibility to simulate the real world in a digital system is to reproduce the activities of social insects. In this case, several agents, the insects, act as their experience and interact with other agents, making decisions in order to maintain the unity of the group. This group is a swarm [4].

This paper aims to present a 3-D game based on a simulation environment which supports the study of the population growth of a society of bees. One of the aspects involved in the beehive growth is the need to collect and produce nectar to ensure energy supply for the entire hive. The solution adopted was based on the areas of Distributed Artificial Intelligence and Swarm, exploring Multi Agents Systems

R. Shumaker (Ed.): VAMR/HCII 2013, Part II, LNCS 8022, pp. 363–371, 2013.

concepts. The simulation environment allows the configuration of different variables involved in the hive development process, such as intensity of rain, humidity, wind speed and floral resources (nectar and pollen). The game manages the performance of bees in the collecting tasks. The virtual game offers some aspects of this process that cannot be visualized by mathematical simulation.

To reach these targets, the paper discusses some general related concepts as the use of simulations in the educational process; the Virtual Reality technology and the Distributed Artificial Intelligence. Next, we describe the game characteristics and the results of an experiment to analyse the software usability. Finally we present the conclusions and the reference list.

2 General Related Concepts

2.1 Education and Simulations

The construction of artificial worlds allows students and teachers to experiment environments with high power to support the teaching-learning process [3]. However, in the current model of education the teacher takes the role of the content host. Thus, students are not given the chance to make discoveries, test hypotheses, discuss results, using their own experiences in order to build their knowledge [5].

The simulation of a biological phenomenon in computers motivates learning, since the graphical interface and the ability to explore the environment variables serve to stimulate student. Hence, in this context, the simulations open new horizons in teaching practice as they allow experiments to be reproduced and repeated as often as needed by students [6].

In addition to the simulations, the animations and more recently, the Virtual Reality technology, have the ability to give information in different formats, where the evolution of various phenomena can be seen over time [7].

2.2 Virtual Reality and Games

The Virtual Reality (VR) is an "advanced interface" to access computational applications, with some features: the visualization of three-dimensional (3-D) environments in real time and the interaction with their elements. [8].

Thus, it can be considered as an important resource in facilitating learning. It turns the virtual experiments more realistic. There are 3 features that make VR environments attractive to students: interaction, immersion and presence. The integration of these features creates a sense of reality, captivating the attention of the student, and interfering positively, motivating him to learn [9].

The educational tools must present characteristics that offer the possibility of the students to build their own knowledge, and so enable the teacher to conduct this teaching-learning process in the context of knowledge construction. Among their key features, we can mention: Interactivity, simulation of real aspects; and representation of information in various ways, such as text, graphics, and animations [10].

In this sense, some initiatives have been developed to explore the potential of simulations in the educational context. The PhET is a web environment that has multiple simulations for the areas of Physic and Chemistry. The simulations have graphical interface, good usability and they do not use VR or Artificial Intelligence technologies. It is an initiative of researchers from the University of Colorado [11]. The Cell Biology Animation has several animations that simulate contents of Cell Biology. The animations have ambiance and low graphical interactivity with the user [12]. The NetLogo Hares & Lynx Model [13] consists of a simulation of the predator-prey relationship between lynxes and hares. It offers a graphical environment where we can choose some initial values and we can see a simple animation of these populations' evolution.

In general, these systems have a low level of interaction; they do not explore intelligent strategies neither Virtual Reality technologies.

Currently, a 3-D computer game with the goal of education and/or construction of concepts can be considered as "Serious Games" [9]. In general, Serious Games allow the simulation of real-world situations, providing activities that stimulate the learning process.

2.3 Agents

The Distributed Artificial Intelligence (DAI) is an Artificial Intelligence area that addresses complex problems. It explores the idea of cooperation between software or hardware agents to perform a task. The DAI area includes the multi-agent approach, i.e., multi-agent systems (MAS). The MAS provide computational mechanisms using autonomous entities, the agents, which interact in a shared environment [14]. Unlike traditional approaches, the multi-agent approach has as its main focus the study of collectivity and not the individual entity. In this sense, the MAS approach can be used for the simulation of decentralized and collaborative environments. Agents are entities that can solve these minor problems to achieve an overall goal.

An agent can be considered an autonomous system seeking different ways to reach pre-established goals in a real or virtual environment [14]. Their behavior is the result of their knowledge about the environment and from the exchanged messages with other agents.

Agents have some important features: An environment of operation, a continuous cycle of life, sensors seeking changes in the environment, actuators acting on the environment and autonomy, that is, independent of user action. Agents have continuity over time and are continuously active. Among all these features, the ability to manage its internal state and its actions without human involvement is a consensus [14]. The agent model used in this work was the reactive one. Reactive Agents are simple agents based on simple event-response model, reacting to environmental changes. These agents have no memory and thus are unable to plan future actions [15]. The idea of this architecture is that a global intelligent behavior is achieved by the interaction of several simple behaviors.

In general, simulations with agents recreate an environment with autonomous entities that interact with each other. They evolve over time and undergo adaptations that depend on the conditions imposed by the changes in the environment.

Next, the game that integrates all these concepts is presented.

3 The Game

The 3-D game is composed by a scenario with some different plantations and a beehive. The goal of the game is to maintain the stock of honey in the beehive. A bee must go to some plantations to get nectar and take it back to the hive. There are some obstacles that can hinder the success of this action: a bird can eat the bee or the wind can take the bee out of its good way.

To start the game, the user must choose an option on the first screen: to play the game or to study the bee morphology.

If the game play is chosen, the player chooses the difficult level: beginner, intermediate and advanced; enter a name to register the scores; and chooses the climatic conditions: sunny day, cloudy day or evening. Next, the user receives information about his scores, the nectar and the honey stored in the hive, and the number of the bees that are in the hive. After that, a scene that shows a forest with a beehive opens up (Figure 1).

Then the user takes the position of the bee and flies in searching of flower fields. There are seven different plantations distributed in the environment. Each field has a flowering period.

Over time the player earns points but loses honey, due to consumption by the bees in the hive. When the bee collects nectar and takes to the hive to produce honey, the user gets points. Every 100 points he earns a lifetime. When the stored honey finishes, the player loses a life. If he has no more lives, the game ends.

The game has two cameras, one provides the vision that the bee has and another that shows the bee in flight (Figure 3).

On the other hand, if the user chooses to study the morphology of the bee, a screen containing the 3-D bee model is shown (Figure 2). In this case the parts of the bee can be viewed from different angles. The bee model can be rotated in various directions and if clicking on an area of its body, a brief description of their functionality is shown.

3.1 The Game Implementation

The simulation system and the game development process explored specific methodologies for agent modeling and implementation. They were modeled using a goal orientation method, the Multi-agent System Engineering-MaSE[16]. The JADE framework supported the agents' implementation. In order to integrate all particularities of these components we used the NetBeans platform [17].

Blender modeling environment was chosen for the creation of 3-D models [18]. To create the game environment and the configuration of the system interactivity the Unity 3D engine was used [19].

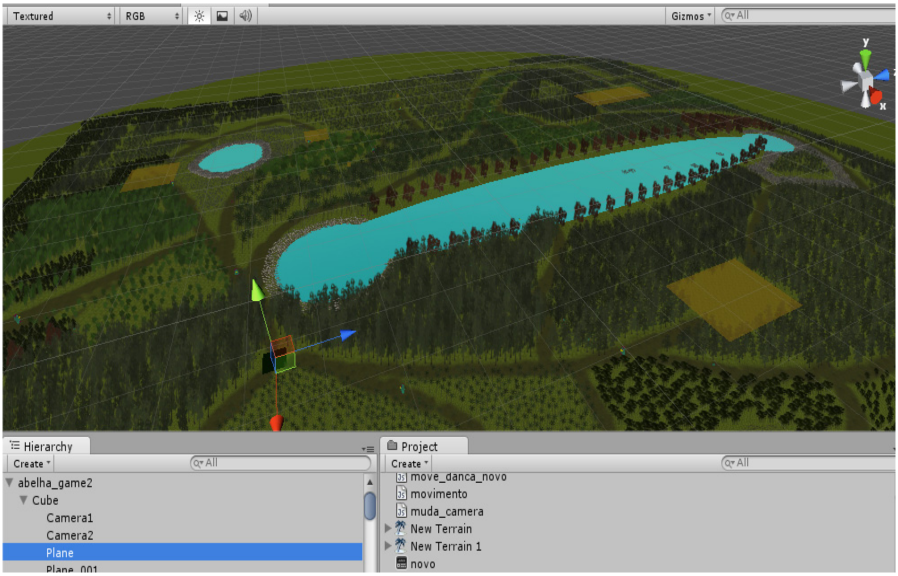


Fig. 1. The visualization of the game scene

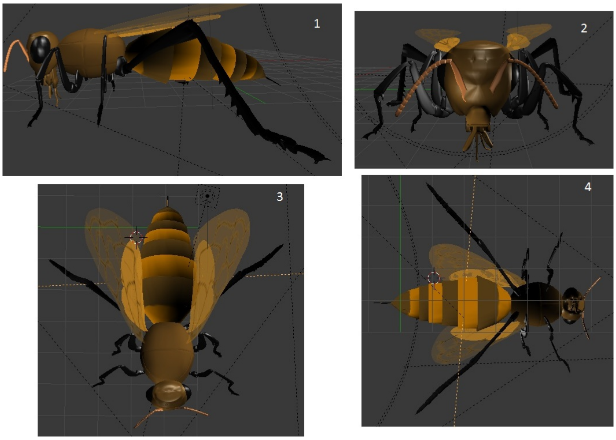


Fig. 2. The morphological study of a bee

Fig. 3. The two cameras vision

In the game, the behavior of bees and the hive are similar to real. To support that requirement we used mathematical models that simulate the hive's population growth, the hive's energy balance; climatic variations and flowering cycle of different plants. Next, these models are briefly described.

3.2 Simulation Models

A quantitative model of bees' population dynamics was applied to calculate the population growth curve of bees [20]. In this mathematical model the eggs originate larvae that become the young bees, which carry out tasks in the hive. Then they become forage bees that work for a period and die. The egg laying held by the queen is influenced by the beehive size. The lack of sufficient food for all beehive individuals can raise the mortality rate of bees causing in the long run, the collapse of the beehive.

The beehive energy balance was based on other work [21] that quantified the pollen and nectar collecting and the honey and beeswax production by bees from a beehive. A beehive with approximately 50,000 bees produces 125 kg of honey. And this requires about 250 kg per year of nectar. For the pollen they collect about 25kg per year. On average, the beehive consumes about 90% of the honey produced to run their activities. Thus, approximately 15 kg of honey are stored in a honeycomb unmanaged, i.e. in the wild.

In general, a bee carries a charge of about 40 mg of nectar. As for the pollen they carry about 20 mg [21]. Since the nectar and pollen loads carried by a bee are known, it becomes necessary to discover how many bees are foraging in a beehive per hour. From the knowledge of the average percentage of foraging bees that make the collection of pollen and nectar at a certain time, we can calculate the amount of nectar and pollen transported in a day.

All presented calculations on the energy balance do not consider the climatic influences and possible seasonality in bloom. Thus, we had to search for data that can be used to simulate these factors.

The Department of Ecology of the Biology Institute in the University of Sao Paulo (USP) has a bee Laboratory that provides a table with the timing of flowering plants visited by the *Apis melífera* bees [22].

Beyond flowering, that information provides the resources to ensure the beehive energy balance, and the climatic factors that affect the flight activity of bees, such as humidity, rainfall and wind speed.

4 Tests

Some usability tests were applied with elementary and secondary school students to observe some aspects of the game usability. According to Nielsen [23], there are five attributes that define usability: Ease of learning, Efficiency of use, Ease of use, Error Rate, and Satisfaction.

The usability test was conducted with five users. Nielsen [23] stresses that this number of users is sufficient to meet about 85% of the usability problems of a system. The test results are shown in Table 1.

Table 1. Usability tests results

Questions	Users					Average
	1	2	3	4	5	
1. The option “help” clarifies all doubts about playing.	5	5	4	5	5	4,8
2. The option “help” shows clearly the game goal.	5	5	4	4	5	4,6
3. The displacement in the environment is easily done.	4	4	4	4	4	4
4. The information on the screen help the player to achieve its goal.	5	5	5	5	5	5
5. It's easy to identify the scene elements.	5	5	5	5	5	5
6. The layout of the navigation keys facilitate its use.	5	4	5	5	5	4,8
7. The beginner level allows players to learn the game.	5	5	4	4	5	4,6
8. The colors of the scenes elements are adequate.	5	5	5	4	5	4,8
9. The game clearly displays the score and game time.	5	5	5	5	5	5
10. The game encourages the player to play at a higher level.	5	4	4	5	5	4,6
11. It's easy to pause the game and return without losing data.	5	5	5	5	1	4,2
12. The fact the game has a rating of record holders encourages the player to try again.	5	5	5	4	5	4,8
13. The information shown on the screen providing easy visualization.	5	4	4	5	5	4,6

This test had a great importance for the development of the system. It allowed us to detect problems in game control. From that, we changed some navigation controls and the format the information was presented to user. In general, we considered that the students had a high level of interest in the game activities.

Now we are starting a more formal experiment with a group of students to observe the gain in the learning processes.

5 Conclusions

In general, it is a hard task to teach some complex concepts that were not easily observed in the real world. The integration of three-dimensional visualization with simulation systems can facilitate the study of these topics.

Specifically, in the area of biology there are some software that offer diversified materials to support learning of complex concepts. However, they do not offer options for working with various learning strategies, because they address general topics.

Aiming to overcome these limitations, this paper presented some initial results of a project that has two objectives. The first one is associated with the technical and educational questions related to the development of a multi-agent simulation system to simulate the behaviour of a beehive, which is not exactly the focus of this paper. The other, aims at finding the combination of technologies to support the integration of intelligent agents within a 3-D serious game to open new possibilities for learning this theme in a more motivational way. The game depends on the simulation system because the bee behaviour in the game is based on the simulation models implemented in the first one.

The main contribution of this work in relation to others reported in the literature is the possibility to play a game to learn biology concepts associated to the bees' lives, supported by a simulation intelligent system. In this case, the game considers real situations and combines climatic variables and with the behavior of swarms. Another important issue is the innovative integration of specific technologies and languages for agents' modelling, development and their integration in a 3-D environment.

References

1. Bachen, C.M., Hernandez-Ramos, P.F., Raphael, C.: Simulating Real Lives. Simulation Gaming 43(4), 437–460 (2012)
2. Mulinari, M.H., Ferracioli, L.: A utilização da tecnologia no ensino de Biologia: um experimento com um ambiente de modelagem computacional. Revista Brasileira de Ensino de Ciência e Tecnologia 1(1), 98–115 (2008) (in Portuguese)
3. Pereira, A.S.T., Sampaio, F.F.: Avitae: desenvolvimento de um ambiente de modelagem computacional para o ensino de Biologia. Ciências & Cognição 13(2), 51–70 (2008) (in Portuguese)
4. Dorigo, M.: Editorial. Swarm Intelligence, vol. 1(1), pp. 1–2. Springer (2007)
5. Lima, L.: Ensino de Conceitos Biológicos: a Relação entre Aprendizagem Significativa e Objetos Educacionais Digitais Anals of XX Simpósio Brasileiro de Informática na Educação (2009) (in Portuguese)

6. Carson, J.: Introduction to modeling and simulation. In: Winter Simulation Conference, vol. 1, pp. 9–16 (2004)
7. Lowe, R.K.: Animation and learning: selective processing of information in dynamic graphics. Learning and Instruction. The Journal of the European Association for Research on Learning and Instruction 13(2), 157–176 (2003)
8. Burdea, G., Coiffet, P.: Virtual Reality Technology, 2nd edn. John Wiley & Sons (2003)
9. Moraes, R.M., Machado, L.S., Nunes, F.L.S., Costa, R.M.E.M.: Serious Games and Virtual Reality for Education, Training and Health. In: Cruz-Cunha, M.M. (Org.) Handbook of Research on Serious Games as Educational, Business and Research Tools, vol. 1, pp. 315–336. IGI Global, Pennsylvania (2012)
10. Rezende, F.: As novas tecnologias na prática pedagógica sob a perspectiva construtivista. Ensaio Pesquisa em Educação em Ciências 2(1), 1–18 (2002) (in Portuguese)
11. PhET. Interactive Simulation. Visited in (January 2013), `http://phet.colorado.edu/`
12. Kyrk, J.: Cell Biology Animation. Visited in (January 2013), `http://www.johnkyrk.com/index.html`
13. McNelis, E.: NetLogo Hares & Lynx Model. Visited in (January 2013), `http://www.shodor.org/refdesk/BioPortal/model/NLpredatorPrey`
14. Wooldridge, M.J.: An Introduction to Multi-Agent Systems, 2nd edn. John Wiley and Sons Limited, Nova Jersey (2009)
15. de Abreu, P.F., de Carvalho, L.A.V., Werneck, V.M.B., da Costa, R.M.E.M.: Integrating Multi-agents in a 3D Serious Game Aimed at Cognitive Stimulation. In: Shumaker, R. (ed.) Virtual and Mixed Reality, Part II, HCII 2011. LNCS, vol. 6774, pp. 217–226. Springer, Heidelberg (2011)
16. Wood, M.F., DeLoach, S.A.: An Overview of the Multiagent Systems Engineering Methodology. In: Ciancarini, P., Wooldridge, M.J. (eds.) AOSE 2000. LNCS, vol. 1957, pp. 207–221. Springer, Heidelberg (2001)
17. Netbeans.(November 2012) (Visited in), `http://netbeans.org`
18. BLENDER. (November 2012)(Visited in), `http://www.blender.org/`
19. UNITY-3D. (November 2012) (Visited in), `http://unity3d.com/`
20. Khoury, D.S., Myerscough, M.R., Barron, A.B.: A Quantitative Model of HoneyBee Colony Population Dynamics. PLoS ONE 6(4), 1–6 (2011)
21. Southwick, E.E., Pimentel, D.: Energy Efficiency of Honey Production by Bees. BioScience 31(10), 730–732 (1991)
22. IAG - USP, Seção Técnica de Serviços Meteorológicos.(November 2012) (Visited in), `http://estacao.iag.usp.br`
23. Nielsen, J.: Why You Only Need to Test With 5 Users. (November 2012) (Visited in), `http://www.useit.com/alertbox/20000319.html`

Presentation of Odor in Multi-Sensory Theater

Koichi Hirota[1], Yoko Ito[2], Tomohiro Amemiya[3], and Yasushi Ikei[4]

[1] Graduate School of Interdisciplinary Information Studies, The University of Tokyo
[2] Graduate School of Frontier Sciences, The University of Tokyo
[3] NTT Communication Science Laboratories, NTT Corporation
[4] Faculty of System Design, Tokyo Metropolitan University
k-hirota@k.u-tokyo.ac.jp

Abstract. This paper discusses an approach to implement and evaluate odor display, with the goal of using it in multi-sensory theaters. A display system that mixes odors with an arbitrary ratio was developed, and a sensor system that is capable of measuring the concentration in a relatively short time period using a sample and hold function was devised. Experiments clarified the time delay and attenuation of the concentration in the transmission of an odor from the display to a user, and the feasibility of utilizing a quantitative mixture of odors was confirmed.

Keywords: Underlying & supporting technologies: Multimodal interfaces, Olfactory Display, Multi-sensory Theater.

1 Introduction

Many investigations have been conducted with the aim of integrating various sensations into a virtual environment. Such investigations have included the feedback of haptic and vestibular sensations, in addition to visual and auditory sensations. The presentation of these sensations is expected to increase the similarity to a real environment and hence improve the reality of a virtual environment. The sensation of odor is one of the sensations that we use in daily life. The sensation informs us of both preferable and abhorrent situations such as the presence of flavorful food or smoke from a fire, and it contributes to our safety and enriches our lives.

The authors are working on a project involving the development and application of a multi-sensory theater. The goal of this project is to attain a higher sensation of reality by integrating haptic, wind, odor, and body motion presentations. The study reported in this paper is part of our effort to develop an olfactory display for the theater environment.

A difficulty when dealing with the sensation of odor from an engineering viewpoint is that the measurement of odor is not necessarily easy. Although semiconducting gas sensors are commercially available at a moderate price, they are often difficult to use because no quantitative calibration method has been established, and their response is relatively slow. This study also investigated fundamental techniques to use such sensors.

R. Shumaker (Ed.): VAMR/HCII 2013, Part II, LNCS 8022, pp. 372–379, 2013.

2 Related Work

The sensation of odor is a chemical sensation involving the perception of chemical substances. An artificial presentation of this sensation is realized by emitting odorants that are perceived by the receptors. However, at present, it is impossible to synthesize arbitrary substances in real time. An easy approach is to use substances prepared in advance; it is expected that a wide variety of odors could be generated by mixing smaller sets of odorants. However, the relationship between a set of odorants and the possible variations in the sensation of odor is unclear.

Many studies have dealt with the presentation of odor by mixtures of odorants. Pioneering research in this field was done by Hiramatsu et al., who examined the transmission of odor by estimating the mixing ratio of elemental odorants [1]. Tanikawa et al. developed a device that mixes vaporized odorants using air flow control [2]. Yokoyama et al. implemented a portable odor display, and carried out an experiment on a searching task using odor[3]. Yanagida et al. proposed a method of conveying odor using the vortex rings generated by an air cannon[4]. In a project by NICT, downsizing an odor display was investigated, and a device that was 20 × 20 × 20 mm in size was developed[5]. Kadowaki et al. investigated the characteristics of human odor perception, including the adaptation to stimuli, and applied the knowledge gained to the development of an odor display[6].

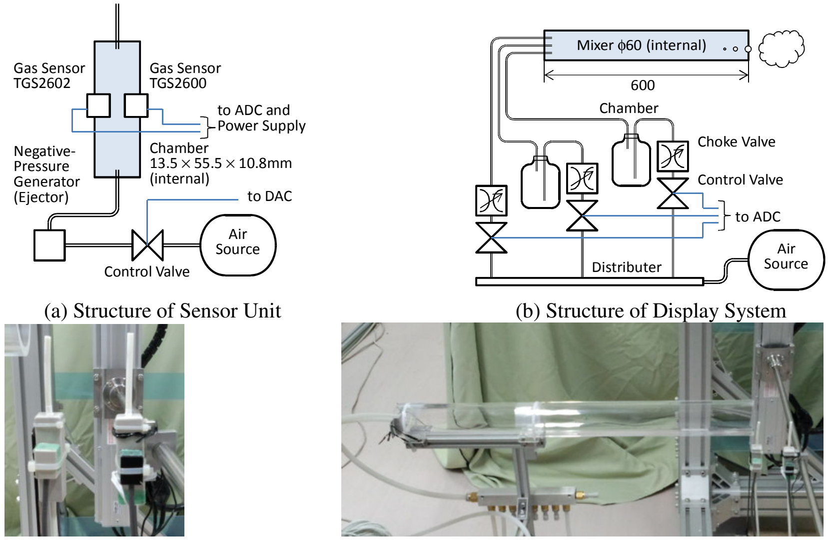
Most of the devices described above were designed to present odors directly to the nose of a user. However, in the case of a multi-sensory theater, it is not necessarily appropriate to attach or wear a device around the user's face because this could hinder the free motion of the user. Therefore, our study investigated an approach to convey odor to the user by an air flow. This approach has the drawback that the diffusion and time delay in the conveying process are not negligible. In addition, to evaluate this approach, precise measurement of the odor concentration in the transient state is necessary.

3 Experiment System

This section describes the sensor and display systems that were implemented in the following experiments.

3.1 Sensor System

A drawback of applying a semiconducting gas sensor to a human interface is its relatively slow response. This makes it difficult to measure the intensity of an odor in real time. Our study proposed a sample-and-hold type sensor unit that samples the air with the odorant and holds it in a small chamber until the outputs of the sensors become stable. A negative-pressure generator (i.e., ejector) was used to introduce the external air into the chamber. This generator was turned on and off from the computer by switching the air supply (Figure 1 (a) (c)).



(a) Structure of Sensor Unit (b) Structure of Display System

(c) Sensor Units (c) Mixing Chamber of Display System

Fig. 1. Sensor and Display Systems

The volume of the chamber was approximately 8 mL, and the flow of the suction was approximately 100 mL/s. Two sensors of different types were attached to the chamber, whose metal cover was removed to improve the response. The resistance of the sensor device decreases according to an increase in the concentration of the gas (or odor) around the device; the resistance R and concentration C have approximately the following relationship:

$$\log(C/C_0) = K \log(R/R_0), \tag{1}$$

where suffix 0 indicates the values in plain air.

Two sensor units were created; in the following experiments, they are referred to using the letters A and B. Moreover, the types of sensors are represented by the numerals 00 and 02 for TG2600 and TG2602, respectively. For example, the TG2602 sensor on unit A is described as A02.

3.2 Display System

A display system that mixes odorants was developed (Figure 1 (b) (d)). This system was designed for use in a multi-sensory theater. However, it could also be used as an odor generator for the evaluation of the sensor units. The principle of the system is similar to that of the device by Tanikawa et al.[2]; the odorants are vaporized in chambers to achieve saturated state, and the air in each chamber is pushed out by infusing air into the chamber. Then, the extruded air is mixed with plain air. The air flows are controlled using pressure-control valves and choke valves. Generally, the choke valves have the characteristic that the amount of flow Q is proportional to $\sqrt{P}$, where P is the differential pressure before and after the valve. Because the air flow is relatively small, the loss of pressure after the valve is negligible. Hence, it can be assumed that the pressure at the exit of the valve is equal to that of the atmosphere.

The plain air that is mixed with the extruded air is also used to convey the odor to the user, in addition to attenuating the concentration. Because the flow of the plain air has a much larger volume, the change in the total volume by mixing in the extruded air is negligible. Hence, the concentration of the odor presented to the user, C, is considered to be proportional to the flow volume Q.

In the following experiments, the pressure control range was limited to 0.1–0.4 MPa. The choke valve was tuned so that the concentration of the output odor did not exceed the measurement range of the sensors. Because no quantitative measure for the odor concentration was available, we defined our own standard based on the display system, where the concentration of the odor at the output of the mixing chamber at the maximum air pressure (0.4 MPa) was defined as $C = 1.0$.

3.3 Odor Source

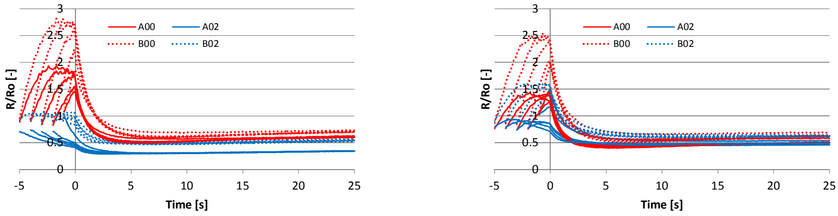
Two odor sources were used: "Queen Rose" (denoted as QR) and "Mix Berry" (denoted as MB). These are essential oils that are commercially available.

4 Measurement of Odor

The measurement method using the sample-and-hold-type sensor unit was investigated.

4.1 Sample and Hold Characteristic

The suction duration and hold time required for the sensors to become stable were important factors for the sensor unit. To determine these factors, the changes in the sensor outputs during the sample and hold operations were observed. The suction duration was changed from 1 to 5 s. Figure 2 shows the results; all of the data were aligned at the end of the suction on the temporal axis. From these results, it is clear that there were large changes in the values during the suction, which were probably caused by a temperature change in the sensor device. The sensor values became stable after about 10 to 20 s; moderate changes in the values are considered to be caused by the odor leaks from the chamber. During this stable period, the value differences caused by the effect of the suction duration were relatively small. Considering these features, the ideal suction duration and hold time were determined to be 1 s and 20 s, respectively. This measurement sequence was used in all of the subsequent experiments.



(a) Queen Rose (QR) (b) Mix Berry (MB)

Fig. 2. Changes in Sensor Values during Suction Operation

4.2 Sensor Calibration

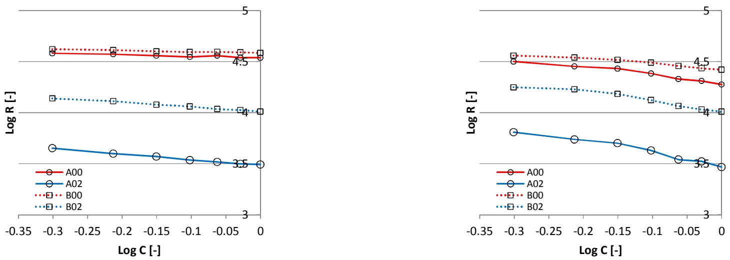
It is known that the sensors used in our implementation have different sensitivities depending on the gas (i.e., odorant). In addition, the individual differences between sensor devices, even with the same model number, are not negligible.

Hence, the measurement parameters for every combination of sensors and odors had to be measured and recorded. The following relationship between the concentration of odor C and the sensor resistance R was assumed:

$$\log(C) = a \log(R) + b, \tag{2}$$

where a and b are parameters that take different values for different combinations.

The measurement results for the sensor resistance while changing the air pressure P are shown in Figure 3. As stated above, the odor concentration was assumed to be proportional to $\sqrt{P}$. Through fitting of lines to the plot using the least squares method, parameters a and b were obtained, as given in Table 1.



(a) Queen Rose (QR) (b) Mix Berry (MB)

Fig. 3. Responses of Sensors to Odor Concentration

Table 1. Calibration Parameters

Sensor	Queen Rose (QR)		Mix Berry (MB)	
	a	b	a	b
A00	-0.147	4.53	-0.760	4.29
A02	-0.545	3.49	-1.153	3.49
B00	-0.119	4.58	-0.483	4.43
B02	-0.438	4.01	-0.870	4.02

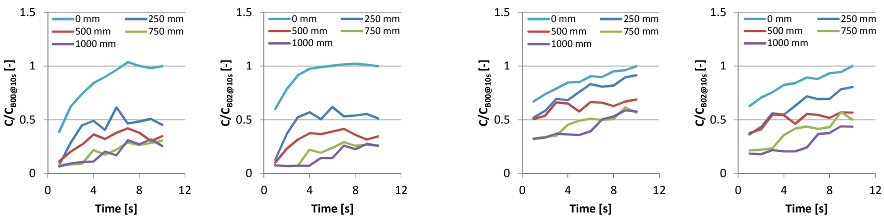
5 Presentation of Odor

Controlling the timing and concentration is essential for the presentation of odor in a multi-sensory theater. To clarify the characteristics of the display system, the following two experiments were carried out.

5.1 Time and Distance

It takes time to convey an odor to a user because the user is located at some distance from the display system. In addition, the odor diffuses in the air during the conveyance. The delay and attenuation were evaluated by measuring the transitional changes in the concentration at various distances. The distance from the exit of the display system to sensor unit A was changed to 250, 500, 750, and 1000 mm, whereas sensor unit B was placed at the exit (i.e., 0 mm). The air was sampled at 1 to 10 s in steps of 1 s. Actually, because the sensor unit requires a hold time of 20 s, the same emission sequence was repeated for every sample time and distance combination.

The concentration values were computed using the parameters obtained in Section 4.2. Because this experiment had the goal of determining the transmission in air, the values were normalized using the concentration at the exit; the value from sensor unit B at 10 s was used as the standard for this normalization. The results are shown in Figure 4. They suggest that at 1000 mm from the display, the timing of the increase in concentration was delayed by about 6 s, and the concentration was attenuated by from 1/2 to 1/5.



(a) Queen Rose (QR) (b) Mix Berry (MB)

Fig. 4. Changes in Concentration Depending on Time and Distance

5.2 Mixture of Odors

As stated previously, the display was designed with the ability to mix odors at any given ratio. This mixing function was confirmed by an experiment using the sensors. The relationship of Equation 2 was not applicable to the case of a mixed odor. Hence, a different model was assumed, as follows.

$$D = D_0 + h_{\mathrm{QR}} C_{\mathrm{QR}} + h_{\mathrm{MB}} C_{\mathrm{MB}}, \tag{3}$$

$$\log(R/R_0) = k \log(D/D_0). \tag{4}$$

The reduction in the resistance of the sensor device is caused by the reducing character of the gas (or odor). In the case of a mixed gas under a low concentration condition, the reducing ability is considered to have an additive nature. In the above equation, D_0 and D are the reducing abilities of the plain air and the air with the gas, respectively; h_{QR} are h_{MB} are parameters that represent the reducing abilities of the QR and MB odors, respectively. All of the parameters are considered to have different values for individual sensors.

In the experiment, the extruding air pressure for QR and MB was changed to 0.1, 0.2, 0.3, and 0.4 MPa, and the resistances of sensors A and B were recorded. Using the resulting data, the parameters in Equations 3 and 4 were estimated. The estimation error was defined as

$$E = \log(R/R_0) - k \log(D/D_0), \tag{5}$$

and the least squares method was used to minimize the square error. The resulting parameters are listed in Table 2. Using these parameters, the measurement results can be mapped into a concentration space whose axes represent the intensities of QR and MB (i.e., C_{QR}-C_{MB} space). The plot is shown in Figure 5. It should be noted that the additivity magnitude relation is preserved to some extent; especially in the area where C_{QR} is not large, there is no fold back in the plotted mesh.

Table 2. Fitting Parameters

Param.	Sensor	
	A00	A02
h_{QR}	9.66×10^{-5}	5.22×10^{-4}
h_{MB}	1.66×10^{-3}	6.31×10^{-4}
D_0	3.58	2.61×10^{-3}
k	476.26	1.55

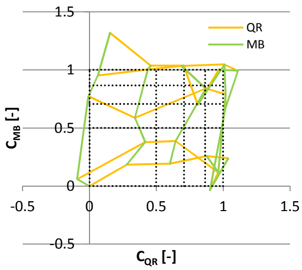


Fig. 5. Mapping into Concentration Space

6 Conclusion

This paper discussed our investigation for the presentation of odor in a multi-sensory theater. A display system that mixes odors and conveys them to users using air flow was developed. A sensor system with a sample and hold function was implemented for the evaluation of the display. Experiments clarified the characteristics of the time delay, concentration attenuation, and quantitative mixture of the odors.

Future work will include a more precise investigation into the spatial distribution of odor. Because a user is allowed to move his/her head in a theater, the distribution must be designed to cover the area of this head motion. Other tasks will involve further improvements in the sensor unit and sensing method. In particular, a reduction in the hold time is required for efficient measurement.

Acknowledgement. This research was conducted as part of a project that is supported by the National Institute of Information and Communication Technology (NICT).

References

1. Hiramatsu, T., Nakahira,Y., Nakamoto,T., Moriizumi, T.: Study of Odor Recorder using Active Odor Sensing System. IEICE Technical Report, CPM2000-32, 17-22 (2000)
2. Tanikawa, T., Hirota, K., Hirose, M.: Olfactory Field Simulation by using Wearable Olfactory Display. Trans. VRSJ 9(3), 265–274 (2004) (in Japanese)
3. Yamada, T., Yokoyama, S., Tanikawa, T., Hirota, K., Hirose, M.: Wearable Olfactory Display: Using Odor in Outdoor Environment. Proc. IEEE Virtual Reality 2006, 199–206 (2006)
4. Yanagida, Y., Kawato, S., Noma, H., Tomono, A., Tesutani, N.: Projection Based Olfactory Display with Nose Tracking. Proc. IEEE Virtual Reality 2004, 43–50 (2004)
5. NiCT press release: Development of multi-sensory interaction system integrating four sensations (March 01, 2013) (in Japanese), `http://www2.nict.go.jp/pub/whatsnew/press/h21/091104/091104.html` (retrieved)
6. Kadowaki, A., Sato, J., Bannai, Y., Okada, K.: Presentation Technique of Scent to Avoid Olfactory Adaptation. In: Proc. ICAT 2007, vol. 2007, pp. 97–104 (2007)

Using Motion Sensing for Learning: A Serious Nutrition Game

Mina C. Johnson-Glenberg and the EGL Lab group at Arizona State University

Learning Sciences Institute and Department of Psychology, University of Arizona, Tempe
Mina.Johnson@asu.edu

Abstract. A mixed reality game was created to teach middle and high school students about nutrition and the USDA *My Plate* icon. This mixed reality game included both digital components (projected graphics on the floor) and tangible, physical components (motion tracking wands that were handheld). The game goal was to feed the alien the healthiest food item from a pair of items. Students learned about the amount of nutrients and optimizers in the digital food items and practiced making rapid food decisions. In the final level of the game players interacted with *My Plate* and each food item filled the appropriate quadrant in real time. Nineteen 4th graders played through the game in one 1.5 hour session. Significant learning gains were seen on a pretest and posttest that assessed nutrition knowledge, paired $t_{(18)} = 4.13$, $p < .001$. We support the need for call for more embodied games that challenge children to practice making quick food choice decisions and we explore how motion capture games can affect engagement, health behaviors, and knowledge outcomes.

Keywords: Applications: Education, Mixed Reality, Nutrition and Exer-Games.

1 Introduction

The rigorous design of health games and assessment of learning gains associated with such games is in its infancy. We present here a study that uses a mixed reality game platform and also explores metrics for assessing learning gains in such games. Mixed reality means a scenario has both digital components (projected graphics on the floor) as well as tangible, physical components (in this case students held a motion capture wand). To date, there has been little controlled, empirical support for the inclusion of games in formal education settings. A recent meta-analysis revealed some learning effects for exergames (Young, et al., 2012), but more research is required. Exergames are a natural extension of the new interfaces that have been popularized by the videogame entertainment industry (especially Wii FitTM and KinectTM platforms).

While there is increasing interest in exergames, current research shows mixed results for the use of systems to motivate behavior change. For example, although there are some data to suggest the possibility of promoting physical activity through systems like Wii Fit [Graves et al. 2010], a recent longitudinal three month study of the Wii Fit showed no significant health outcomes among its participants [Owens

R. Shumaker (Ed.): VAMR/HCII 2013, Part II, LNCS 8022, pp. 380–389, 2013.

2011].) However, Souter (2008) observed a number of factors that motivated students to participate in the exergames. For example, students who were reluctant to participate in other physical activities or other social events were able to comfortably play the Wii with a group and be successful. That success then led to increased self-confidence, and some students were even motivated to try real sports after playing the Wii games. There is some evidence that games overcome a student's hesitancy and this overrides other factors like a child's gender or weight. Epstein, Beecher, Graf, and Roemmich (2007) studied overweight and nonoverweight children who exercised in a variety of different ways using dance dance revolution (*DDR)*. The results showed that the children were more motivated to play *DDR* with the interactive dance pad over the other non-dance options because of the interactive nature of the game and that the increased motivation did not depend on gender or weight status.

Video games for nutrition training. Less research has been done to attempt to integrate both healthful eating education and exergaming principles. To date, the closest research to explore this area has occurred within the realm of mobile health, particularly for weight loss. For example, SmartDiet (Lee, Chae, Kim, Ho, & Choi, 2010) features a thin, normal weight, or overweight avatar, depending on the user's self-reported weight changes. While this study influenced fat mass, weight, and body mass index, the possibility for unintended consequences (e.g., increased prevalence of eating disorders) was not explored. The app was targeted toward overweight adults and uses self-reported diet and PA to calculate necessary caloric intake and expenditure for weight control. It also included a game that teaches users about nutrition and PA and features an avatar that reflects the user's weight change via the avatar's changing body shape (Lee et al., 2010).

Video games offer potential behavior change channels by embedding functional knowledge and change procedures such as goal setting, modeling, and skill development activities into a meaningful, entertaining, and immersive game environment (Baranowski, Buday, Thompson, & Baranowski, 2008). One reason might be because these games create environments where players' actions and decisions can have immediate effects. This rapid feedback cycle coupled with the power of collaboration within games could be efficacious for altering students' decisions about food choices. Developing age-appropriate game-based intervention requires "substantial formative research (e.g., focus group discussions, intensive interviews, observations) with the targeted demographic group on story and character concept, story arc, personality and visual representation, and alpha testing on fun and functionality of the interactive components" (Baranowski, Buday, Thompson & Baranowski, 2008. Further advice on game creation for this space comes from Rahmani and Boren (2012), who suggest that the interesting pieces of knowledge should be "hidden in the gameplay" and that story, graphics, and music be tightly designed and integrated in the play.

We are in the early stages of creating a nutrition game, but wanted to share the first study results because one of our goals is to create shared knowledge regarding the magnitudes of the effect sizes that researchers might expect from health game interventions. In addition, the field is in need of rapid-choice engaging, and valid nutrition assessments for the targeted age group.

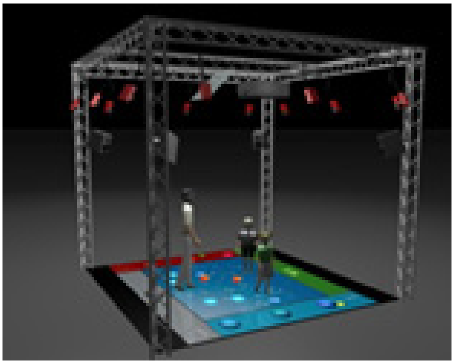


Fig. 1. Schematic of *SMALLab*

Learning *in Mixed Reality Environments.*The game in this study melded exercise with instruction on nutrition. There are several desktop games to teach children about nutrition, which are created on the model of one-child-to-one-screen. As such they are not embodied and they are not collaborative. The goal of our game was to engage an entire class in the instruction of nutrition while allowing two students to be active at a time. *SMALLab* (Situated Multimedia Arts Learning Lab) is a mixed reality educational platform that engages the major modalities (i.e., the sense systems including visual, auditory, and kinesthetic) that humans use to learn. *SMALLab* uses 12 infrared OPTITRACK motion tracking cameras either mounted in a ceiling (when in a classroom) or on a trussing system if in the research lab where this study took place. See Figure 1 for a schematic. The cameras send information to a computer about where a trackable object or "wand" is being held in the floor-projected environment. The space is called mixed reality because it meshes the digital virtual world with hands-on manipulables (Milgram & Kishino, 1994). The floor space is 15 X 15 feet and the tracked space extends approximately eight feet high; wands are tracked with X, Y, Z coordinates up to one millimeter in precision. Using the wand (a rigid body trackable object with retro-reflective markers on the ends), the physical body can now function like a 3D cursor in the interactive space. For example, by dipping down with the wand a virtual object projected on the floor can be "grabbed" by the wand and moved to other locations for immediate scoring and feedback – similar to a click and drag action on a computer screen. Some notable differences between this platform and a traditional desktop one are that students can locomote through the immersive environment and be active; further many peers can partake in observational learning and real time discourse.

SMALLab allows for multiple students (up to four) to be tracked simultaneously. With turn-taking, entire classrooms with 30 students are able to physically experience a learning scenario in a typical class period. Students outside the active space sit around the open periphery and collaborate by discussion and whiteboard activities. The lab at ASU has researched the efficacy of the mixed reality platform in several content domains including language arts (Hatton, Birchfield, & Megowan, 2008), science, technology, engineering, and mathematics [STEM] - physics (Johnson-Glenberg, Birchfield, Savvides, & Megowan-Romanowicz, 2011; Tolentino, Birchfield, Megowan-Romanowicz, Johnson-Glenberg, Kelliher & Martinez, 2009) - geology (Johnson-Glenberg, Birchfield, & Usyal, S., 2009) and Disease Transmission (Johnson-Glenberg, Birchfield, Tolentino, & Koziupa, in revision). This range suggests that embodied learning in mixed reality is not content dependent.*Design principles in game.* The lab follows several principles of design as it co-creates content with teachers and subject matter experts. Content always incorporates components of embodiment with gestural congruency and collaboration – we expand on these constructs in the following sections. The game begins with the narrative of the player (you) finding a lost and hungry alien under your bed. Coincidentally, the alien's body functions similarly to a human's and you must feed him and get him back to health so he can get home. The first student in the space takes the role of the selector and the other student is the transporter. Figure 2 shows an example of the floor projection. A forced choice task is shown in the top right corner. The selector hovers the motion tracking wand over a food item and on the top left of the space the nutrients present in that item begin to glow. We highlight three nutrients (protein, carbohydrates and fatty acids) and two 'optimizers' (fiber and vitamins/minerals; note these were chosen based on expert consultation from a registered dietician). The selector "picks" the virtual item up and with the wand brings it to the alien's mouth, where a chewing animation follows. At that point, feedback is shown regarding the quality of the choice. When the better choice is picked (in Figure 2 that would be the bowl of blueberries with less fat and more fiber), the alien stays green and smiling; had the piece of pie been chosen the alien would display the beginning of five stages of decline into sleep or lower alertness. In the final stage, the alien is a sallow yellow with drooping antennae, closed eyes and no smile.

Fig. 2. Example of the floor projection

Now, the transporter player needs to grab the glowing nutrients at the top of the play space and physically "run" them down to the tissue at the bottom of the space. All nutrients that are lit up must be transported before the next food pair can appear. Thus, the transporter is encouraged by his/her peers on the sidelines to hustle to get the nutrients down to the tissue area to continue the game. We specifically wanted to design in high levels of activity for at least one of the players.

The five "nutrients." We wanted children to have practice rapidly choosing food they might be exposed to in a day at home, stores, or school. Two subject matter experts (i.e., a registered dietician and a nutrition professor) were consulted extensively during the design of the game. We did not want to focus on simple calorie counts because children do not seem to make rapid decisions this way and because, overall, we were more interested in emphasizing food quality over caloric quantity. Our goal was to get children to think about foods that satiate, and were "more" nutritious in a comparative manner. They need practice making quick choices about what "to grab in a cafeteria line" or in a store during a busy day. The use of the alien avatar allowed for a human physique that would react immediately to the foods but not get "fat." We did not think this would be appropriate for classes that would surely contain a spectrum of body sizes and shapes. The decision was made to give feedback on nutritional value via drowsy sleep states.

We needed to keep the number of nutrients tractable, so no more than five are listed with four levels of magnitude: 0 = outline, 1 = dim, 2 = bright, 3 = bright and spinning. We chose these nutrients as these represent broad nutritional domains important for understanding food quality. With regard to food quality metrics, a balance was sought between providing enough information to aid children in making better decisions, while not complicating matters too much, thereby reducing retention of the learning objectives.

2 Methods

Participants. Nineteen 4th graders were recruited from a K-8 school in an urban city. The class was 80% Caucasian, 10% African American and 10% Hispanic. Though it was the end of the school year, they had not yet received a unit on health and nutrition.

Procedure. One week before students came to the University lab to play the videogame, two experimenters went to the school to administer the two pretests. The following week, the students arrived and played the game at the same time of day (before lunch). Immediately upon finishing their turn, the pair of students was taken into another room to complete the two invariant posttests (this took about 15 minutes). When the students finished the posttests, they were escorted back into the main play area of *SMALLab* to observe their peers in the game. Each pair was assigned an ordinal pair number so order of play effects could be tested. The hypothesis was that the more a student was able to observe the gameplay (i.e., the higher the pair number), the higher the gains would be by posttest.

Gameplay. The first author led the gameplay session. Student pairs were quasi-randomly chosen. Volunteers were sought and enthusiasm played a role as to the order they were called upon, as well as the desire to create mix gender pairs. Each pair played approximately eight minutes, in the middle of play, the pair switched roles of selector and transporter. Again, one student (the selector) selected the food with the wand, verbally checked with the transporter for agreement on choice, and then carried it to the alien's mouth. When the alien started to chew, the second student (the transporter) brought the now-active nutrients down to the tissue at the bottom of the playspace. The three nutrients were protein, carbohydrates and fatty acids, and the two 'optimizers' were fiber and vitamins/minerals. After the end of the first full play session, two of the practice levels were skipped since all students had seen those with earlier players. At a later level, we introduced the *My Plate* icon. The experimenter went over the names of the food groups and kept instruction brief, saying, "In general, you should try to eat more fruits and vegetables and fewer grains and proteins." The circle for dairy that some templates include was not used because the dairy requirement is considered by some to be controversial. When the selector chose a food item the appropriate section or quadrant of the *My Plate* icon would begin to fill. We mapped the item to one serving - or a portion that a 12 year-old male would eat for a lunch so the grilled fish was deemed one serving and it filled up one third of the protein section of the *My Plate* template. An apple would fill one fourth of the fruit section. (These portions will be refined in future iterations.) The figure above shows the icon filled except for the vegetables.

Nutrition and Food Choice Test – This experimenter-designed test was a mixture of forced choice and open-ended items. While we did find several potential measures to use related to nutrition knowledge (e.g., see Townsend, Johns, Shilts, & Farfan-Ramirez, 2006; Wall, Least, Gromis & Lohse, 2012), we could not find a measure that was in line with the educational content specifically focused upon within this test, particularly for this age group. This new measure was piloted with a focus group of 9 year-olds before the study. It includes 31 items and a blank *My Plate* template that students were asked to fill in. The maximum score on the test was 100, this was never reached. The highest score at posttest was 95.

3 Results

Students were assessed with pretests one week before play and then immediately after. The Descriptives and effect sizes (Cohen's *d*) are listed in Table 1.

Table 1. Descriptives for Alien Health-*SMALLab*

Test	Pretest M and SD	Posttest M and SD	Difference	Effect Size
Nutrition Test				
Total Test	70.87 (11.32)	77.95 (13.66)	7.08	.57
My Plate Sub Items	7.74 (4.48)	11.17 (5.01)	3.43	.72

4 Discussion

The results indicate that one 1.5 hour session in a mixed reality game environment can have positive impacts on knowledge about nutrition. Statistically significant gains were seen in explicit knowledge regarding food choices on a nutrition test. There is evidence of transfer of knowledge of general principles related to nutrition because over half of the items on the nutrition test were new items that students had not seen in the game. In addition, students learned about *My Plate* categories and proportions by playing the game, on this subtask they displayed a large effect size in learning of .72 by the end of play. This is interesting because students received only cursory instruction on *My Plate*, they were told it replaced the Food Pyramid icon. The experimenter went over the four food group labels only once and said, “In general, you should try to eat more fruits and vegetables and fewer grains and proteins.” The icon for dairy was not used because it was deemed somewhat controversial by our expert nutritionist panel. This suggests that exposure to and short interactions with the icon during play may be instructive and result in significant knowledge change, though additional research is required to fully delineate the active mechanism of knowledge transfer.

Implications. We contend that instructional gameplay in a mixed reality environment holds great promise for nutrition education and whole class participation. Active whole class participation is difficult to achieve with more traditional multimedia designs when each child is viewing an individual screen. Students reported being “very interested in” and engaged with the alien storyline and showed strong evidence of transfer to new nutrition facts and to the *My Plate* icon. We are currently researching how to facilitate an entire class exercising with the active players in the game and are creating a vertical, on-the-wall version using the Microsoft *Kinect*™ as the input device. Our goal is to have the game used in a PE and/or nutrition classes.

Limitations. As this was a first-generation game study focused on determining the feasibility of a multimodal game to promote nutrition knowledge, there were several limitations. First, there was no control condition. The school we had access to had two fourth grade teachers, but the second teacher did not want to be part of the study since she “had no plans” to cover nutrition the entire school year. It did not seem prudent to have a “business as usual” control condition if the class would not focus on nutrition for 1.5 hours. Further, there were too few students in the experimental class to do a split class waitlist design, which we have done for other *SMALLab* studies. For the next iteration of studies we will bring the mobile KINECT camera to the classroom and recruit several schools.

Future Directions. The advent of cost effective skeletal tracking systems like the Xbox Kinect could have a positive effect on exercise in home and classroom setting. We are currently designing a game based on many of the lessons learned with this mixed reality game. We want to get all the passively viewing member of the class to participate even if they are not being tracked by the system. The nutrition or PE teacher could encourage those who are getting automated feedback to do projected exercises on the sideline. Xbox Kinect may prove to be a ''game changer'' in the exergames domain (Boulos, 2012).

Future iterations of the game should be more precise about proportion of food that a child should eat in a day, especially in reference to the *My Plate* quadrants. With the Kinect we are able to track the size of the child and make the plate portion fill accordingly. For example, a short child with a slim outline would require fewer servings of protein in a day, this information can be instantly translated into gameplay. We are currently researching the pros and cons of making the avatars take on the body morphology of the active student. (This may only be appropriate for at home play because of children's social discomfort.)

Conclusions. Games hold much promise for instructing young people in nutrition, in addition games can be key motivators in getting students active during the school day or in afterschool care. We present a mixed reality game that uses a large open space and allows the non-active students to also learn about nutrition and My Plate. Significant gains were seen on a nutrition test after 1.5 hours of play for an entire class. One goal is to give students practice making fast food decisions that may carry over into real life. We are working to revise an engaging food choice test using replica food items to better assess this type of transfer.

Acknowledgements. Special thanks to Eric Heckler, Maureen Zimmerman, Tatyana Koziupa, Caroline Savio-Ramos, Chris Dean, Ben Luke, Sean Griffin, Sarah Cargill and Punam Ohri-Vachaspathi.

References

1. Baranowski, T., Buday, R., Thompson, D.I., Baranowski, J.: Playing for Real: Video Games and Stories for Health-Related Behavior Change. American Journal of Preventitive Medicine 34(1), 74–82 (2008)
2. Barsalou, L.W.: Grounded Cognition. Annual Review of Psychology 59, 617–645 (2008)
3. Boulos, M.: Xbox 360 Kinect Exergames for Health Games. Research, Development, and Clinical Applications 1(5) (2012), doi:10.1089/g4h.2012.0041
4. Davis, C.L., Tomporowski, P.D., Boyle, C.A., Waller, J.L., Miller, P.H., Naglieri, J.A., Gregoski, M.: Effects of aerobic exercise on overweight children's cognitive functioning: a randomized controlled trial. Research Quarterly of Exercise and Sport 78(5), 510–519 (2007)
5. Dustman, R.E., Emmerson, R.Y., Steinhaus, L.A., Shearer, D.E., Dustman, T.J.: The effects of videogame playing on neuropsychological performance of elderly individuals. J Gerontol 47(3), 168–171 (1992)
6. Dustman, R.E., Ruhling, R.O., Russell, E.M., Shearer, D.E., Bonekat, H.W., Shigeoka, J.W., Bradford, D.C.: Aerobic exercise training and improved neuropsychological function of older individuals. Neurobiology of Aging 5(1), 35–42 (1984), `http://dx.doi.org/10.1016/0197-4580`
7. Epstein, L.H., Beecher, M.D., Graf, J.L., Roemmich, J.N.: Choice of interactive dance and bicycle games in overweight and nonoverweight youth. Ann. Behav. Med. 33(2), 124–131 (2007), doi:10.1080/08836610701307835

8. Gee, J.: Good Video Games and Good Learning: Collected Essays on Video Games, Learning and Literacy (New Literacies and Digital Epistemologies). Peter Lang Pub. Inc. (2007)
9. Graves, L.E.F., Ridgers, N.D., Williams, K., et al.: The Physiological Cost and Enjoyment of Wii Fitin Adolescents, Young Adults, and Older Adults. Journal of Physical Activity and Health, 393–401 (2010)
10. Grunbaum, J.A., Kann, L., Kinchen, S.A., Williams, B., Ross, J.G., Lowry, R., Kolbe, L.: Journal of School Health 72(8), 313–328 (2002), doi:10.1111/j.1746-1561.2002.tb07917.x
11. Hatton, S., Birchfield, D., Megowan-Romanowicz, C.: Learning metaphor through mixed-reality game design and game play. In: Paper presented at the Proceedings of the 2008 ACM SIGGRAPH Symposium on Video Games, Los Angeles, California (2008)
12. Jenkins, H., Clinton, K., Purushotma, R., Robison, A., Weigel, M.: Confronting the Challenges of Participatory Culture: Media Education for the 21st Century: MacArthur Foundation (2006)
13. Johnson-Glenberg, M.C., Birchfield, D., Savvides, P., Megowan-Romanowicz, C.: Semi-virtual Embodied Learning-Real World STEM Assessment. In: Annetta, L., Bronack, S. (eds.) Serious Educational Game Assessment, pp. 241–257. SensePublishers (2011)
14. Johnson-Glenberg, M.C., Birchfield, D., Usyal, S.: SMALLab: Virtual geology studies using embodied learning with motion, sound, and graphics. Educational Media International 46(4), 267–280 (2009)
15. Lee, W., Chae, Y.M., Kim, S., Ho, S.H., Choi, I.: Evaluation of a mobile phone-based diet game for weight control. J Telemed Telecare,16(5), 270–275 (2010), doi: jtt.2010.090913 [pii] 10.1258/jtt.2010.090913
16. Maillot, P., Perrot, A., Hartley, A.: Effects of interactive physical-activity video-game training on physical and cognitive function in older adults. Psychology and Aging 27(3), 589–600 (2012)
17. Milgram, P., Kishino, F.: A Taxonomy of Mixed Reality Visual Displays. IEICE Transactions on Information Systems, E77-D(12)(1994), doi: citeulike-article-id:477597
18. Owens, S.G., Garner, J.C., Loftin, J.M., van Blerk, N., Ermin, K.: Changes in Physical Activity and Fitness after 3 months of Home Wii Fit use. Journal of Strength & Conditioning Research 25(11), 3191–3197 (2011)
19. Souter, J.: Nintendo Wii as an inclusive learning tool. Disability Support Services Unit Online Newsletter 7(3) (2008)
20. Thomas, H.: Obesity prevention programs for children and youth: why are their results so modest? Health Education Research 21(6), 783–795 (2006), doi:10.1093/her
21. Thompson, R., Jago, R., Griffith, M.J., Thompson, D., Baranowski, T., Buday, R., Baranowski, J.: Development of a Serious Video Game. Simulation Gaming 41, 587 (2010)
22. Tolentino, L., Birchfield, D., Megowan-Romanowicz, C., Johnson-Glenberg, M.C., Kelliher, A., Martinez, C.: Teaching and Learning in the Mixed-Reality Science Classroom. Journal of Science Education and Technology 18(6), 501–517 (2009), doi:10.1007/s10956-009-9166-2
23. Townsend, M., Johns, M., Shilts, M.K., Farfan-Ramirez, L.: Evaluation of a USDA nutrition education program for low-income youth. Journal of Nutrition Education and Behavior 38, 30–41 (2006)
24. USHHS. Physical Activity Guidelines for Americans (2009), `http://www.govtrack.us/congress/bills/111/hr3851`

25. Wall, D.E., Least, C., Gromis, J., Lohse, B.: Nutrition education intervention improves vegetable-related attitude, self-efficacy, preference, and knowledge of Pennsylvania fourth grade students. Journal of School Health 82(1), 37–43 (2012)
26. Watts, K., Jones, T.W., Davis, E.A., Green, D.: Exercise Training in Obese Children and Adolescents: Current Concepts. Sports Medicine 35(5), 375–392 (2005)
27. Young, M.F., Slota, S., Cutter, A.B., Jalette, G., Mullin, G., Lai, B., Yukhymenko, M.: Our Princess Is in Another Castle: A Review of Trends in Serious Gaming for Education. Review of Educational Research 82(1), 61–89 (2012), doi: citeulike-article-id:10312339, 10.3102/0034654312436980

AR'istophanes: Mixed Reality Live Stage Entertainment with Spectator Interaction

Thiemo Kastel, Marion Kesmaecker, Krzysztof Mikolajczyk, and Bruno Filipe Duarte-Gonçalves

St. Poelten University of Applied Sciences, St. Poelten, Austria
Institute for Creative\Media/Technologies
Research Group Video Production
{thiemo.kastel,marion.kesmaecker,krzysztof.mikolajczyk, bruno.goncalves}@fhstp.ac.at

Abstract. Mixed Reality and Augmented Reality for live stage productions have been used ever more frequently by artists over the past few years. AR'istophanes is an experimental stage production aimed at bringing the new technical possibilities of Mixed and Augmented Reality to the stages of this world. This document describes the first phase of pre-production from 2011 to 2012 and demonstrates the possibilities of integrating motion capturing and 3D animation. This also includes the use of Smartphone Apps and real-time rendering. Audience interaction is a key focus in this production – which means technical approaches are demonstrated and opinions were collected from potential viewers.

Keywords: Mixed Reality, Augmented Reality, Interaction, Theatre, Live Entertainment, Optical See-Through Glasses.

1 Introduction

Nowadays, technology offers quite a number of options for exploring live performance activities. Theatre, in particular, offers new ways of performing with different approaches. Interaction between a real person and a virtual one can lead into 3D environments and deliver to spectators new experiences of content creation. Stage designers extend their traditional props to interactive ones or combine them into Mixed Reality environments. Traditional audiences often have to learn how to participate in these new scenarios and sometimes need special hardware to consume the Mixed Reality scenarios. Augmented Reality, in combination with Smartphones and high-end computer hardware, is already quite a powerful tool for supporting new Mixed Reality approaches with very detailed real-time character animation and interaction. Over the past few years many stage companies and tech-driven researchers have developed new theatre experiences often known as "Digital Theatre".

Based on their own research the authors propose a unique method of performance and stage setup:

R. Shumaker (Ed.): VAMR/HCII 2013, Part II, LNCS 8022, pp. 390–399, 2013.

AR'istophanes [1] is a Mixed Reality play which was developed in 2011 by Thiemo Kastel. The story is based on the Aristophanes play Eirene – The Peace. Preproduction of the 3D content like virtual characters with motion capture and lip-sync animations and the overall control of different Smartphone and computer Apps were developed and largely all produced with a group of students in July 2011 and July 2012.

The stage itself is subdivided into several layers, where the play happens. Unlike classical theatre productions AR'istophanes offers more than one space where the play can be acted out. Beside the main stage with a holographic projection of the virtual characters, the play is also performed on the spectator's optical see-through glasses [2]. AR'istophanes enables spectators to interact live and in real-time with the director of the play. This includes the use of so-called personal optical see-through glasses in conjunction with Smartphones.

What is real or virtual depends on the director's genius during the live performance. The director's creativity is directly linked via a network to the actor's wearable information system. This information system is the same as the spectator's. Based on the spectators' live feedback via Smartphone the director can decide during the play whether a virtual or real character should enter the next scene. This leads to a direct involvement of the spectators and from day to day each performance of the play becomes an independent and exclusive piece of entertainment. The metadata of each play can then be recorded and made available on the Facebook social media platform for other theatre directors. Via touchscreen devices spectators would be able to share screenshots and comments with their smart community on Facebook. The smart community could also include contacts to the actors, director and production crew of the play enabling the audience to gain deeper insights into their artistic work.

Augmented Reality requires a trigger to play on Smartphone-stored content or loaded content from the Internet. Typical triggers for Smartphones are camera or position i.e. GPS-based. This overall spectator experience is only possible in a smart theatre environment (Smart Venue) with Future Media Internet products and services.

2 Digital Theatre Productions

2.1 Live Entertainment Productions with Spectator Interaction

Active audience involvement in stage productions can create exciting experiences. The actors or director do not know in advance how the audience will react when their collaboration is requested. In his book Digital Performance Steve Dixon formulates four types of interactive art and performance [3]. These are:

Navigation, Participation, Conversation and Collaboration

In our view, the opportunities of participation and collaboration are particularly interesting when Smartphones are used for this at live productions due to the possibilities they provide for interacting at "Smart Venues".

2.2 Augmented Reality Used at Theatre Productions

Even if the general public is still fairly unfamiliar with Mixed Reality productions, quite a few projects have already taken place aimed at connecting the audience and the performance thereby taking it to a new level.

One of the first theatre productions that used AR was *Everyman: The Ultimate Commodity* [4]. The production was performed in Singapore (2006) and Toronto. Daniel Jernigan's theatre group referred to the first definitions of AR by Ronald Azuma and G. Bishop when developing this performance. As they themselves state, this is the first time the option of moving AR markers and holding the actors in a station position is used. Typical black/white tracking markers were held in front of the actors' faces and these were then filmed live. A computer generated the AR by giving the actors different faces. The result was then projected onto a screen. The audience then saw the projection and in front of it other actors acting alongside the projected ones. The artists' summary of these attempts is formulated in relatively neutral terms:

"In the final analysis, however, we believe that our project serves as a reasonable example of how any attempt to integrate technology into theatre can be a double-edged sword, as technology can simultaneously be both supportive of – and disruptive to – the themes and aesthetics of a particular production."

3 Eirene – The Peace by Aristophanes

Created in 421 BC the comedy *Eirene* (The Peace) was staged by the poet Aristophanes at the great Dionysia of that year. It depicts the problems caused by the turmoil of the Athenian-Spartan War, in particular those affecting hard-hit farmers who take a delegation to the Olympians and successfully deliver the goddess Eirene from exile. The fairly simple plot based on the daily news of the war is enlivened by a wealth of fantastic and unreal events and contains detailed parodist allusions to Euripides' *Bellerophon* [5].

Aristophanes was born in Athens around 445 BC. At that time Athens was the capital of Attica and the largest city in Greece next to Sparta. At Aristophanes' birth a period of peace and prosperity prevailed in Athens under Pericles (peace with the Spartans and also with the Persians already in 449 BC) which was to end with the 20-year Peloponnesian War (431BC-404 BC) [6]. Aristophanes' idea of the world was shaped more conservatively and his political and philosophical inclinations (preference for aristocracy and older philosophy over democracy and the Sophists) are always present in his plays causing him to clash with prominent contemporaries (Euripides, Socrates and others).

4 Stage Setup for AR'istophanes

Since time immemorial mankind has always tried to elude its own perception – from cave paintings to 3D cinema man always been fascinated by the power of the image. In a period where technology offers numerous possibilities we decided to take a step forward and propose something that will change our perception of reality, offering an experience that many will want to repeat.

The stage will need equipment that can deliver with quality a believable illusion, good enough to deceive the eye of an ordinary citizen. In order to provide a complete and convincing illusion we must be sure to use state-of-the-art technologies in video-projection and holograms for Mixed Reality experiences. The idea is to use a layer for a virtual character. The technology for projecting holograms created by Musion seems to fit our needs [7].

Musion has a hardware setup called "Musion Eyeliner" that features a high-quality projector, placed on the ceiling of the structure pointing down, projecting to a mirror that will reflect the image to a 6 by 4m (standard), very thin and almost invisible foil. The foil is rotated 45-degrees relative to the ground. If the projection sent the image directly to the foil spectators would see some of the projector light reflected at them and this would spoil the illusion. With 45-degree angle the light sent by the projector is not reflected back to the audience, creating a more realistic hologram. The design is briefly explained in Figure 1:

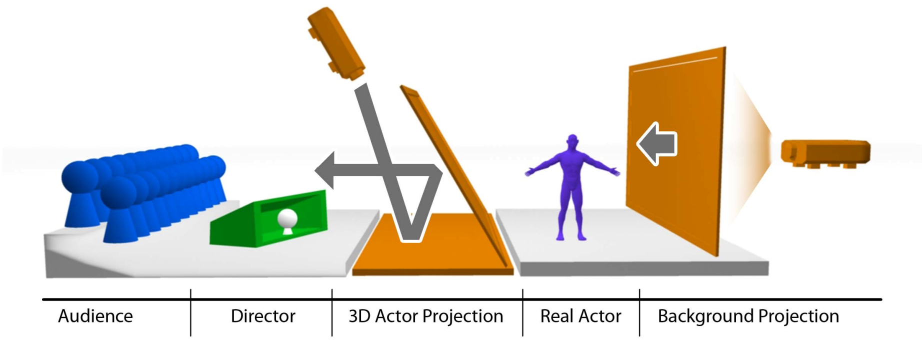


Fig. 1. Side perspective of the Stage Setup for AR'istophanes

In order to establish a connection between all participants, all computers for the projections and Smartphones/tablets must be on the same network (Fig. 2).

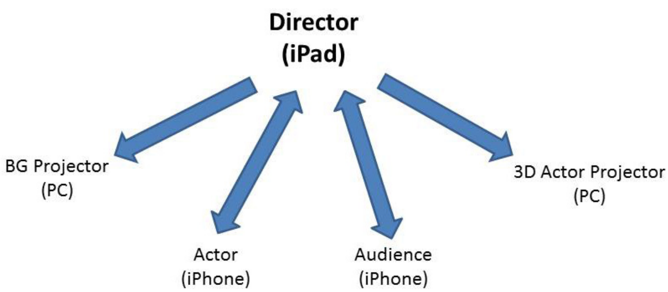


Fig. 2. Data schematic – one-way arrows represent only output from the director and two-way arrows represent input/output from the director

5 3D Content Creation for AR'istophanes

5.1 Low-Cost Motion Capturing

> *"Motion capture is the process of recording a live motion event and translating it into usable mathematical terms by tracking a number of key points in space over time and combining them to obtain a single three-dimensional representation of the performance. In brief, it is the technology that enables the process of translating a live performance into a digital performance."* (Menache 2000, p.1) [8]

Motion capture is a process which allows us to record actions quickly and to import the information in 3D software to animate the characters. Thanks to motion capture it is possible to animate characters in a really realistic way, also with complicated actions. It is indeed hard to achieve these kinds of realistic movements using key frames or other animation techniques. The problem of motion capture has been that it was only affordable for large companies or big-budget film productions. Fortunately, Microsoft's Kinect Xbox reached the market in 2010 [9]. Initially thought of as a way to revolutionise video games, this hardware became a target for programmers who adapted it to all kinds of other uses. The Kinect is used in games for detecting the player's movements and controlling virtual characters. In the same way, it is possible to use this depth information to capture the motion of an actor and use it to animate a 3D character. For this purpose, the right software is required. For the AR'istophanes project we chose to use ipiSoft software. ipiSoft is a piece of software that can record a video using the ipiRecorder and record the depth information of a person in a room. It can then transfer this information in ipiStudio to a track and save the information in the format needed for the rest of the animation [10]. Figure 3 shows the setup of this capturing system.

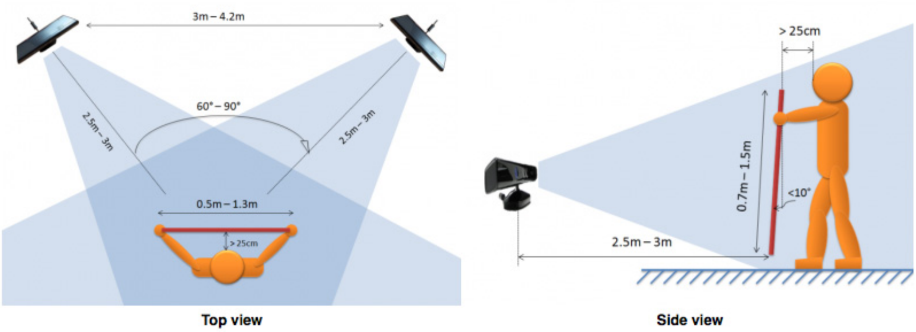


Fig. 3. Positioning for calibration of the Kinect for a good depth-capture

5.2 Character Design for Real-Time Rendering

Unity3D

Unity is a high-performance rendering engine permitting the development of computer games and interactive 3D graphics applications. The development environment runs in Microsoft Windows and Mac OSX. Potential target platforms here, alongside PCs, are also game consoles, mobile devices and web browsers. For AR'istophanes the only platforms used are iOS, Android and Microsoft Windows.

Generate 3D Objects for Unity
3D objects for Unity can be generated with any 3D software. As an export format .FBX is mainly used. When generating 3D objects for Unity3D the following aspects need to be taken into account:

(a) Optimised Mesh Division

The extent to which you allocate divisions to the 3D objects in the modelling phase is crucial here. For instance, too many polygons on a 3D object will unnecessarily compromise the performance of a PC or other medium (mobile device). Figure 4 depicts two fairly complex 3D objects.

The house on the left contains 1523 polygons. Its optimised version on the right has three times fewer polygons. For a game engine like Unity3D a 3D object with an optimised number of polygons is best suited as fluid representations are also possible even on lower-performing end devices.

(b) Real Size of the 3D Object

It is very practical to model 3D objects in real scale. This means you can keep real proportions and avoid additional adaptive scaling of the objects in Unity (Fig. 5).

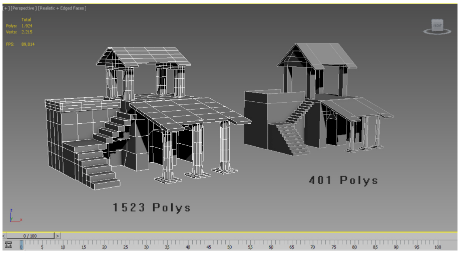


Fig. 4. Optimization of the Number of Polygons

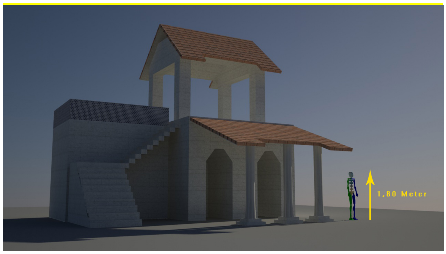


Fig. 5. Correct Proportions when Modelling

Animated 3D Objects in Unity

In Unity it is possible to move objects via pre-defined paths, scripts or physical paths. Character animation can be realised using a skeleton system. Using a skeleton, individual character body parts are animated. Head animation is usually generated separately. One face animation technique is morph animation. Character animations are prepared within and imported from external programs like 3DS Max.

Creation of 3D Heads using Facial Studio
Facial Studio (Microsoft Windows Edition) is a software program used to create 3D heads easily and simply. More than 500 parameters are available here for the generation and deformation of a head. You can import your own photo in two shots (front and side) and create a head with an automatic texture using control points (Fig. 8).

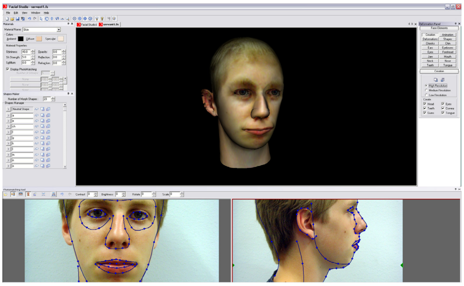


Fig. 6. Facial Studio (Windows Edition)

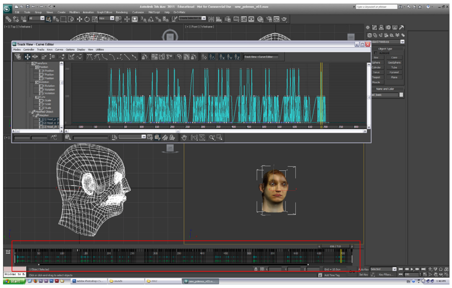


Fig. 7. Key Frames after Lip-Synching

By adapting these control points to a photo you automatically create a 3D geometry with the image texture of imported photos. When the head is finished you can create different facial expressions individually and save them as morph targets. These different facial expressions can then be animated using an external 3D program. You can import a 3D head like this as an .FBX file – for instance in Autodesk 3D Studio Max.

Voice-O-Matic for Lipsynching

The imported head has different facial expressions as morph targets and also those that correspond to speech phonemes. Using this information and the Voice-O-Matic plug-in (3D Studio Max Edition) it is possible to achieve quick and automatic lip-synching with a language for any desired audio file. In a few seconds key frames are created on a timeline. The individual key frames can then also be manipulated and optimised manually in an animation editor. Voice-O-Matic saves a great deal of time on lip-synching (Fig. 9).

Exporting 3D Characters for Unity

Once lip-synching is completed you link a 3D head to an animated body. For Unity it is vital a 3D character does not have unnecessary divisions in the mesh as this can very much compromise the performance of a computer. To import a 3D character in Unity3D one other factor needs to be taken into consideration: the 3D head is purely a morph animation.

6 Mixed Reality Development

6.1 Augmented Reality for Spectators

"Augmented Reality (AR) technology for digital composition of animation with real scenes is to bring a new digital entertainment experience to the viewers. Augmented Reality is a form of human-machine interaction. The key feature of the Augmented Reality technology is to present auxiliary information in the field of view for an individual automatically without human intervention. The effect is similar to composing computer-animated images with real scenes."[11]

This reference describes accurately the intention we have planned for the audience: Interaction between the audience and the play, the feeling they are actively participating in the play. For this purpose optical see-thorough AR Glasses were chosen. Unfortunately, these glasses are state-of-the art technology and not yet in circulation. For this reason, we have only performed preparatory work and will do the tests as soon the glasses arrive.

6.2 Augmented Reality for Actors

In AR'istophanes actors are to be equipped with the same technical system as the audience. This allows very fluid interaction as some information can be represented quickly via the glasses – for instance using simple symbols.

7 Spectators Interaction with the Director and Actors

7.1 Interaction Control System

The interaction will be an ever present element with every participant, whether on stage or in the audience. The director is the only agent interacting directly with each individual as all other individuals only interact with each other indirectly. This is how it works:

The director can ascribe different behaviours to the virtual actor: acting for the story, as an idle listener to the real actor or one waiting for the results of the audience survey. This kind of improvisation of a character from the virtual world can make people relate to the character, since it behaves in a human manner; reacting to events instead of having scripted behaviour.

Actors will also use Augmented Reality glasses to obtain information about the play, to know what happens next or any kind of information the director might find relevant to share; for example, survey results, an upcoming event or a change in mood or attitude of the character being played. The spectators will be able to answer queries sent by the director. These queries can be about small things – details like the colour of a character's clothing, or more important things like changing the course of the plot. The director can work with the information provided by the audience and can decide what is best for the play to keep it interesting.

More interesting details can be added to the play's environment. Let us suppose the intrigue in the play becomes denser. Here the director can add dynamic environment effects like rain, snow, de-saturate colours or dim the lights to create a more fitting environment. This is particularly interesting as it gives different directors to opportunity to test their creative abilities, delivering to the audience and actors an "ever new" and dynamic experience. This system is being developed in Unity 3D, a game creation platform that can deliver the adequate mechanisms for the realisation of all those features listed so far. All participants must have installed the AR'istophanes App on their Smartphone (in the case of spectators and actors) or a tablet in the case of the director.

7.2 Spectator Experience

Thanks to Augmented Reality the audience is no longer passive. Spectators take part in the story directly. The Augmented Reality glasses allow spectators to control the show. At any time they can ask for more information about the play (this information will be chosen by the director before) giving them a better understanding of the show. If the director so wishes spectators can have a direct influence on the story. In this way it would be possible for spectators to choose between different story endings. There is still one question that remains unanswered: Is the audience interested in performances of this kind? A show like this brings to the audience an interactive dimension like never before. But will people be keen to take part in the show, or is this a change they are not willing to take on board?

The next chapter will present the results of a survey carried out from November 2012 to February 2013 to discover what people expect from this kind of show and how much involvement they are prepared to provide.

7.3 Results of the Survey

The first thing to note here is that this kind of performance is aimed more at a younger audience. Indeed, the younger generation are accustomed to using Smartphones for every possible reason which makes it easier for them to imagine a play where they would be involved directly with their Smartphones. Accordingly, it is harder to appeal to older audiences with Augmented Reality theatre because they would first need a "course" to learn how to use the application. However, with the help of assistants this audience group could also participate in an interactive performance.

As mentioned earlier in the paper, this kind of performance is still relatively unfamiliar to the general public and only 37% have already experienced a theatrical production with audience interaction. 72% would not wish to use the opportunity of interaction at an interactive theatre production.

Concerning the use of optical see-through Augmented Reality glasses, 68% are willing to use them directly, 26% want to give it a try first and only 5% do not want to use them at all. The audience is relatively positive about this AR experience. However, some are still not sure whether wearing AR glasses might be bothersome as these glasses are still not yet available on the market.

Most respondents (60%) agree to paying more because of the advanced technology a show of this kind would offer. However, most of them (51%) are not willing to pay more than Euro 2 extra for this technology. 40% of the audience are still not prepared to pay more than a normal ticket despite the added technology. Regarding the application, 78% of respondents are ready to download the free App to use on their cell phone. The remainder of audience expect organisers to provide Smartphones at the entrance for people to use. On a positive note for this kind of performance, 69% of people already own a Smartphone and 20% who currently do not have one wish to own one in the near future. By February 2013 34 people from 7 countries had taken part in the public survey. 88% of those polled were aged 16 to 35.

8 Further Work

As we approach our project's final goals, we continue to test all parameters and attempt to overcome any obstacles we might encounter. Sometimes is not possible to incorporate all features designed for a project, but creativity enables us to navigate problems and sometimes develop new ideas and methods for improving our product. The next step in our work will be to test stress Unity with more demanding tasks. We will gauge its ability to handle long morph animations so as to avoid the need to split sequences of animations from the same character. Improving the quality of the 3D meshes and textures is also something we aim to improve without compromising global performance of the game engine. Due to the rapid technological development of mobile end devices and real-time rendering software the quality of the 3D character is sure to improve and will take up a large proportion of our work. We are confident that optical see-through Augmented Reality glasses with reasonable display sizes will be available at consumer prices. There is certainly still a great deal of work to be done before a real production is ready to be presented on the stages of this world.

References

1. Online Documentation, `http://mixedreality.kastel-net.com`, `http://webtv.fhstp.ac.at/aristophanes`
2. Keynote Day 1; Project Glass: Live Demo at Google I/O, San Francisco Google announced the Glass Explorer Edition of their AR glasses for early, This will only be for a few developers and will not be a consumer product (2012), `http://youtu.be/VuC0i4xTyrI`
3. Dixon, S., Smith, B.: Digital performance: a history of new media in theater, dance, performance art, and installation. MIT Press, Cambridge (2007)
4. Jernigan, D., Fernandez, S., Pensyl, R., Shangping, L.: Digitally augmented reality characters in live theatre performances. International Journal of Performance Arts and Digital Media 5.1, 5(1), 35–49 (2009)
5. Aristophanes section of the Aphoristiker-Archiv (May 12, 2012), `http://aphoristiker-archiv.de/A9.html`
6. Sharpley, H.: The Peace of Aristophanes, Edinburgh and London, William Blackwood and Sons, p. 1 (1905)
7. Musion Eyeliner (January 01, 2013), `http://www.musion.co.uk`
8. Menache, A.: Understanding Motion Capture for Computer Animation and Video Games. Morgan Kaufmann Publishers (2000)
9. Microsoft Kinect (January 01, 2013), `http://www.microsoft.com/en-us/kinectforwindows/`
10. iPi Soft (January 01, 2013), `http://www.ipisoft.com/`
11. Lyu, M.R., King, I., Wong, T.T., Yau, E., Chan, P.W.: ARCADE: Augmented Reality Computing Arena for Digital Entertainment. In: IEEE Aerospace Conference, pp. 1–9, 5-12 (March 2005)

System Development of Immersive Technology Theatre in Museum

Yi-Chia Nina Lee[1], Li-Ting Shan[2], and Chien-Hsu Chen[3]

[1] Department of Industrial Design, National Cheng Kung University
[2] Titania Design Int'l
[3] Department of Industrial Design, National Cheng Kung University
jellynina@gmail.com, tinashan@titania.com.tw,
chenhsu@mail.ncku.edu.tw

Abstract. Varieties of museum theatres include historical characters, puppetry, movement, music, etc. Visitors can experience the storyline of knowledge about history, science and technology in the theatre which creates immersive environment and engages experiences, such as those in performance, games and simulation. With this kind of experience, knowledge learning in museums becomes more effective and interesting. In addition, it requires multiple disciplines to accomplish the multi-sensory experience provided in the theatre environment. This article focuses on the design process and the system development of the Immersive Theatre in a systematical method. There are three phases in the process: Design, Configuration Negotiation, and Implementation.

Keywords: immersive technology, museum theatre, Configuration negotiation.

1 Introduction

Interactive technology exhibition is becoming a stage for government, engineers, designers, and artists collaborating together. Meanwhile people can learn about the life in the future from interactive technology exhibition which usually take immersive technology as a tool. Immersive technology refers to technology that blurs the line between the physical world and digital or simulated world, thereby creating a sense of immersion.

Museum theatre is the theatre and theatrical techniques being use in a museum for educational, informative, and entertainment purposes. It can also be used in a zoo, an aquarium, an art gallery, at historic sites, and so forth [1].

The theatre is generally performed by professional actors in addition. Varieties of museum theatres include historical characters, puppetry, movement, music, etc. Visitors can experience the storyline of knowledge about history, science and technology in the theatre which creates immersive environment and engages experiences, such as those in performance, games and simulation. With this kind of experience, knowledge learning in museums becomes more effective and interesting. In addition, it requires multiple disciplines to accomplish the multi-sensory experience provided

R. Shumaker (Ed.): VAMR/HCII 2013, Part II, LNCS 8022, pp. 400–408, 2013.

in the theatre environment [2]. The design of this kind of theatre has been studied by painters, theatre directors, scenic designers, lighting designers, filmmakers, producers, artists and control and image engineers. These fantastic experiences rely on the collaboration both of designers and engineers. The process of this collaboration is quite complicated and messy; also, it needs an organizational management [3].

2 Case Study

In this section, three cases of immersive theatre will be reviewed. Two cases bring out culture and educational meaning with immersive technology. Each case is with different purpose and outward appearance. However, this study expected to find a general path and elements of these immersive technologies

2.1 Pavilion of Dreams

Pavilion of Dream [4] is one of the exhibition sections of 2010 Taipei International Flora Exposition. The Pavilion of Dreams is an interactive exhibition hall with art and technology integrations. It is a grand platform of ITRI technologies fused with creativities exuded by Taiwanese artists, whereby technologies are granted with new application opportunities.

Through continuous communications and attempts, engineers and artists have come together to solve problems and overcome obstacles, and together with the production team, all parties have worked as a team to build the Pavilion of Dreams and to turn it into a magical space full of imagination and amazement.

The team of the Pavilion of Dream selected five technology: FleXpeaker [5], Multi-View Naked-Eye Stereoscopic Display [6], Smart Controllable Liquid Crystal Glass, Non-Contact UWB Physiological Sensing Technology, and Ultra High Frequency RFID Technology

Audience will received a RFID bracelet which will record the data during the interactive journey, also with the RFID bracelet people can play a trigger role of the special event in the scenario. The way to experience the immersive technology theatre of Pavilion of Dreams is complete free, audience follow the path designed without any time limitation. In 360 square meter real-time interactive environmental theater, the display will follow the behavior of the audience have different animated.

In the process of development, the 2010 Taipei International Flora Exposition decided the main theme, then they planed the series topic of areas. The designers constructed the view of the Pavilion of Dream with looking into the technology provided by ITRI (Industrial Technology Research Institute). After the vision of the Pavilion of Dream was completed, designers communicated with artists and engineers about how to work together. The Artist put esthetic feeling into the scenario, and the engineers need to face to the technical problem and find out the solution with designers and artist. Finally, the Pavilion of Dream would be ready for visiting, then the operators of 2010 Taipei International Flora Exposition have to make sure the experience of audience would run as the working team expected.

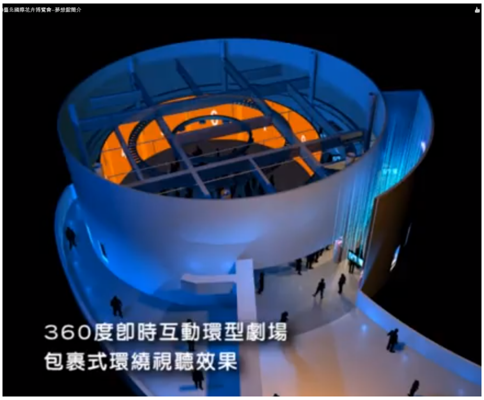


Fig. 1. 360 square meter real-time interactive theatre [4] (加圖片)

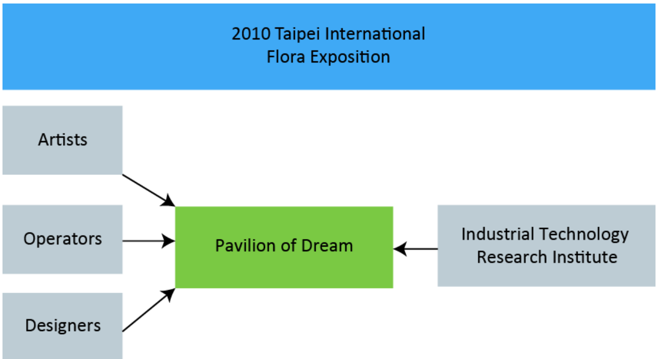


Fig. 2. Involving organizations of Pavilion of Dream

2.2 Discovery Theater in Discovery Center of Taipei

The "treasure" within the Discovery Center of Taipei [7] is the 360 square meter Discovery Theater with its rotating screen. It is different from other theaters because it is an anthropocentric and dynamic theater reflecting daily life. The rotating screen, concrete models, twelve solid projectors, and special music and light generate a unique and futuristic vision.

This kind of experience is completely immersive, because there is no boundary of the stage. The view angle is dynamic with the rotating seat, and with the physical model, the experience emphasize the blur of the reality and visual world.

Difference from Pavilion of Dream, the audience of Discovery Theater sit still in the middle of the theatre and surrounded by the 360 square meter display. There are different films for audience to choose to look. Hence that, the collaboration will be very long term if the client (Taipei City Government) decides to add a new topic of film.

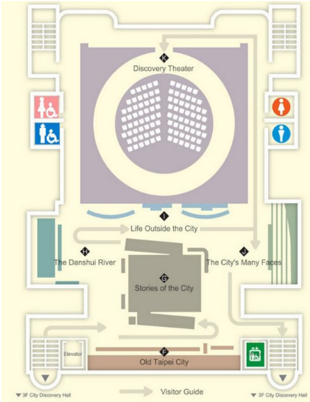


Fig. 3. A plan view of Discovery Theater in Discovery Center of Taipei [7]

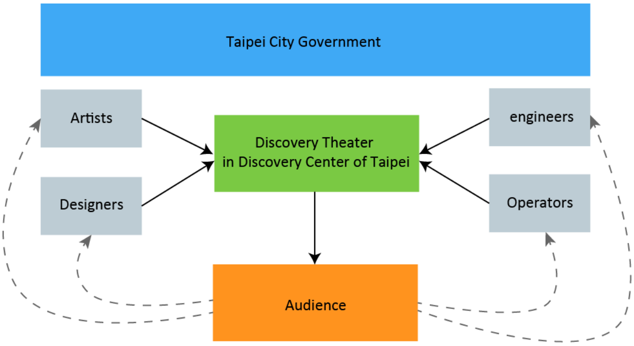


Fig. 4. Involving organizations of Discovery Theater in Discovery Center of Taipei

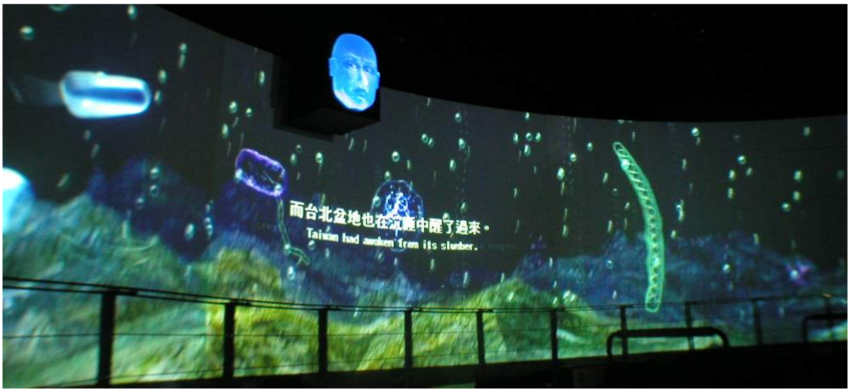


Fig. 5. A view of Discovery Theater in Discovery Center of Taipei

In this case, the audience will play an important role to the immersive technology theatre. The content of the experience will update, and the designers and artist will consider the feelings of common from audience when they get into a new film. Therefore, audience become a factor of the theatre.

3 Methodology

The design of this kind of theatre has been studied by painters, theatre directors, scenic designers, lighting designers, filmmakers, producers, artists and control and image engineers. These fantastic experiences rely on the collaboration both of designers and engineers. The process of this collaboration is quite complicated and messy; also, it needs an organizational management.

Based on the above cases of cooperation, we can be summarized in the following ITT's basic cooperation as fig. The clients provide the requirement of the theatre, also most of the limitation such as the budget, physical space and the special request, is from the clients. The designer is the one who draw the big picture of the immersive technology theatre, and he merges all the elements (sounds, visual effect, and human factor) together. The engineer brings the fancy scenario to real. When the immersive theatre is well developed.

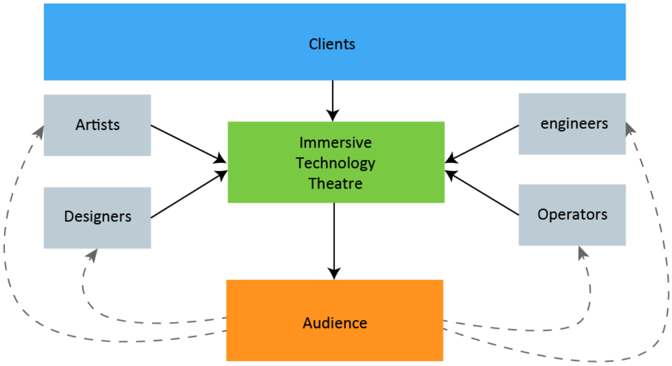


Fig. 6. Involving organizations of Discovery Theater in Discovery Center of Taipei

Fig.7 shows the relationship of the collaboration in making an immersive technology theatre. The clients make the offer of an immersive technology theatre, then the designers will draw a picture of the immersive technology theatre, planning the way of the experience, and listing the elements and tools needed. The engineers provide the technical provide professional technical to make the fancy scenario come true. The artist is the one who put the meaningful content (which with a culture or an educational meaning). Finally, an immersive technology theatre will be ready to construct, the operator have to make sure the theatre can be operated.

3.1 Phase 1 Design

Once an Immersive theatre is needed to be composed, designers and engineers have to consider the requirement of the immersive technology theatre.

In Design phase, designers brain storm the view of what will the theatre look like; then designers and engineers have to work together about the solutions of these ideas.

The first thing to think of is the scenario and atmosphere in the theatre including the storytelling to present. According to the cases reviewed, the main elements needed to concerned are in following: View angle, audience orientation, Life performance, and special effect.

View Angle

In display technology parlance, viewing angle is the maximum angle at which a display can be viewed with acceptable visual performance [8].

The ideal picture size is defined for movie theatre for a horizontal viewing angle of 45 degrees at the prime seat. The prime seat is (mostly agreed) a center seat two-thirds to the rear of the house [9]. THX recommends having a 36 degree viewing angle from the farthest seat in the auditorium. Prior to construction, THX advises cinema designers and architects to accommodate for the 36 degree horizontal viewing angle. And, to make sure that every seat has an unobstructed view, THX often recommends either elevating or lowering the entire floor to adjust the seating location [10].

Audience have to be completely immersed in the atmosphere of the immersive technology theater, as the result, view angle have to be concerned based on the requirements in the beginning of the design. The main field is produced based on the view angle decided.

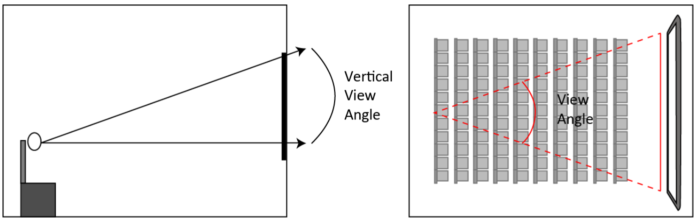


Fig. 7. View angle of theatre

Audience Orientation

Audience orientation [11] is the direction of the audience view, and the main stage will follow the audience orientation. Audience Orientation can be dynamic, so it is important to consider about audience orientation when designing the stage which mainly present.

Live Performance
Without live performance, the experience of an immersive technology will be nearly the same. Live performance is not definitely refer to drama act with actors at stage, but it also could be the real time interaction between the storytelling content and the audience. Therefore, there will be plenty of possibility during the experience.

Special Effect
Special Effect is refer to the effect which create the atmosphere, including lighting, video display, working model, and immersive technology. In an immersive technology theatre, there will be a lot of programs, dynamic mechanical, and audio effect filling in the space of the theatre. Therefore, it become a hug issue to make the arrangement of all the equipment.

3.2 Phase 2 Configuration Negotiation

After the whole view of the immersive theatre is constructed, it is stepping into the second phase: Configuration Negotiation. In the phase of Configuration Negotiation, designers and engineers need to face the limitation of the physical space, or the request of the business entrepreneur, and the limitation of legislation. The elements discussed above will need to be adjusted base on the reality limitation.

Orientation will be an important matter in Configuration Negotiation phase. Multiple technology are applied in the theatre, and they need to be merged into the physical space, meanwhile, it is important to make sure that the mechanical or electronic equipment will not harm the audience. The shows in the theatre should be checked again and again. Designers and Engineers can have a clear view of the details in the shows with a script.

3.3 Phase 3 Implementation

Finally, the theatre will start building in implementation phase. In this phase plenty technical support will be needed. All the parts have to work together, and it is necessary to prepare a check list of works, tools, and equipment. Moreover, the schedule should break down all the detail of works, so there will be no mistake, and wasted time.

4 Discussion

With the three phases mention above, as well as with various experts in the field of professional required, the study sorted out the construction of a journey map as Fig. 8. Developing an immersive technology theatre an intensive iterative design process.

Frist, an offer is brought by clients, which usually will be government units. Clients narrate the purpose of the immersive technology theatre, and let design and engineer team know how many recourse can be used. Then designers and engineers will get together to draw a picture of the theatre. In some case, artist will be invited to add some culture and educational meaning into the theatre. For wasting no time, it is very important for the team to settle all the details before construction.

The human factor is an important issues in an immersive technology theatre, because the immersive environment will be a deadly mechanic experience. The human factor here is not just for the user part but all the participants of all the process of the theatre. The communications is the main topic in this issue, including culture between technologies, clients between producing team, government between people, and audience between the theatres. In the last part of the journey map (Fig. 8) is the “Operations”, however is not the end of an immersive theatre. Because the culture and the technology keep moving forward, as a present of the ideal future living and a role of spreading knowledge and culture, an immersive technology theatre in museum should keep refreshing and moving forward with the society.

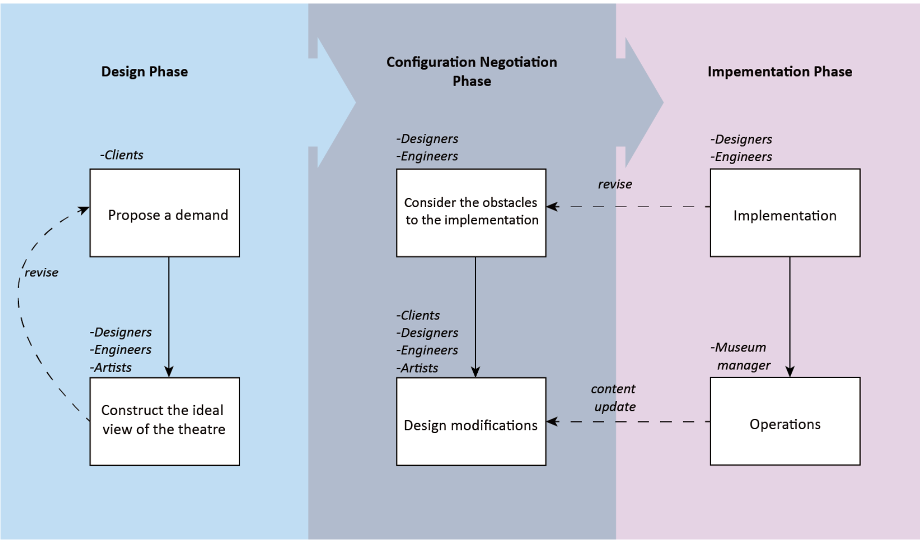


Fig. 8. Process of immersive technology theatre development

5 Conclusion

The immersive technology theatre in museum shows people the ideal of the future living, and it is a stage of professionals in various fields. With a large number of ideas for different kind of views, the team have to communicate very carefully, and the leader, which could from any fields, needs to have a clear mind that which stage is it. Although between each immersive theatre and each museum, the topic, content, and the technology implied will be very different, the rules and process is the same. The three phases is the big structure of the develop process, in the future, there will be more cases, so that the research team can break down the big phases into more details steps for development.

References

1. Darroch, M.: Intermedial Theatre:±Technology (2007)
2. Gruzelier, J., et al.: Acting performance and flow state enhanced with sensory-motor rhythm neurofeedback comparing ecologically valid immersive VR and training screen scenarios. Neuroscience Letters 480(2), 112–116 (2010)
3. Vanhoutte, K., Wynants, N.: Performing Phenomenology: Negotiating Presence in Intermedial Theatre. Foundations of Science 16(2-3), 275–284 (2011)
4. Government, T.C.: Pavilion of Dreams, `http://www.dreams21.tw/`
5. Chen, Y.-C., et al.: Modeling and experimental verification on directivity of an electret cell array loudspeaker. Journal of Intelligent Material Systems and Structures (2012)
6. Jeng, T.-R., et al.: New 3D Image Technologies Developed in Taiwan. IEEE Transactions on Magnetics 47(3), 663–668 (2011)
7. Government, T.C.: Discovery Center of Taipei, `http://www.discovery.taipei.gov.tw/web/index.html`
8. Wikipedia. Angle of view, `http://en.wikipedia.org/wiki/Angle_of_view`
9. INFO, A. Viewing Angle, `http://www.bnoack.com/index.html?http&&&www.bnoack.com/audio/thx.html`
10. THX, THX Certified Cinema Screen Placement
11. Weaver, D.H.: Audience need for orientation and media effects. Communication Research 7(3), 361–373 (1980)

An Immersive Environment for a Virtual Cultural Festival

Liang Li[1], Woong Choi[2], and Kozaburo Hachimura[3]

[1] Ritsumeikan Global Innovation Research Organization, Ritsumeikan University, Japan
[2] Department of Information and Computer Engineering, Gunma National College of Technology, Japan
[3] College of Information Science and Engineering, Ritsumeikan University, Japan

Abstract. This paper describes the development of a virtual reality (VR) system and the use of an immersive environment for a traditional Japanese virtual cultural festival. With the development of computer graphics (CG) and VR technologies, extensive researches have been carried out on digital archiving of cultural assets. The goals of our Virtual Yamahoko Parade project are to record and preserve digitally the Yamahoko Parade of the Gion Festival, an intangible traditional grand scale cultural event, as well as to open the product to the public. Therefore, not only the quality of the VR contents but also the display and demonstration are important to reproduce the atmosphere of the festival. The proposed system combines vision, sound, immersive display, and real time interaction, which enables the users to feel as if they are actually participating in the parade.

1 Introduction

Digital archives, which measure, record, and preserve tangible and intangible cultural assets using digital information technologies, have attracted increasing attention in the last two decades [1, 2]. In recent years, applications of virtual reality on cultural heritages have been attempts to cover not only individual tangible and intangible cultural assets but also large scale intangible cultural events, such as traditional festivals and behaviors of participants in cultural events [3–5]. This raises a problem that how to present these virtual cultural assets in an immersive environment to the audiences. Pre-rendered contents and traditional displays limited the options of interactions, which cause the users can not be involved in the virtual environment. Because the whole atmosphere in a cultural event is crucial, so a real time interactive environment is required.

The goals of our Virtual Yamahoko Parade project are to record and preserve digitally the Yamahoko Parade of the Gion Festival, an intangible traditional grand scale cultural event, as well as to open the product to the public. The Gion Fetival is one of Japan's greatest and biggest event, which is founded on Japanese history and culture (Fig. 1). Every year on July 17, the festival culminates in a parade of the Yama and Hoko, floats known as "moving museums." Every year the Parade attracts over 150,000 spectators from all over the world.

R. Shumaker (Ed.): VAMR/HCII 2013, Part II, LNCS 8022, pp. 409–415, 2013.

In our previous work, We generated “Virtual Yamahoko Parade,” which includes CG models and animations of the floats, crews, and spectators, as well as the music and ambient sounds heard at the time of the parade [6, 7]. We used a large scale cylindrical screen as the display and a gamepad as the input device to control the viewpoint. We received positive feedbacks from most of the users. However, according to the feedbacks from some users, the gamepad-controlled system is more like a video game other than an immersive festival experience. Gamepad indeed increased the flexibility of user control but reduced the immersive experience as well. In this study, we use an immersive and hands-free system to reproduce the Virtual Yamahoko Parade. The new system provides real time interaction and more realistic immersive atmosphere.

Fig. 1. Yamahoko Parade of the Gion Festival in Kyoto

2 Constructing the VR Contents of the Virtual Yamahoko Parade

We constructed the virtual parade and combine the motion and acoustics of the floats, crews, and spectators using a virtual reality toolkit called Vizard.

The components of the VR contents were created as follows.

(1) Virtual Kyoto: The street model of “Virtual Kyoto” [8] was developed using various technologies and materials, such as geographic information system (GIS) data, cadastral maps, aerial photos, street photos, and landscape paintings. We used the model of part of Shijo Street (approximately 550 meters), which is extracted from Virtual Kyoto. It is reproduced along with the buildings and arcades on both sides of the street.

(2) CG floats: Four CG models of the floats (Fune-hoko, Naginata-hoko, Kanko-hoko, and Kitakannon-yama) were created and introduced into the virtual environment for use. The CG models of Fune-hoko and Kanko-hoko were

built from the data obtained by measuring their miniatures (on a scale of one to eleven) using 3D shape measurement. As for the CG models of Naginata-hoko and Kitakannon-yama, their 3D models were created out of measured drawings. Figure 2 shows the created CG models of the Yama and Hoko floats.

(3) CG crews: Four kinds of CG parade crews belong to Fune-hoko were created: Hikikata who pull the float; Ondotori who direct the float; Kurumakata who control the traveling direction and the start-stop movement of the floats; and Hayashikata who play ohayashi music on traditional Japanese instruments on the hayashibutai, a stage on the upper part of the float. The body motion data given to each character model were obtained from actual actions using a motion capture technique. We captured a variety of these motions performed by a Fune-hoko crew who participated in our experiments. The textures of the costume for each character were created from the photographs of the costumes that are actually used. Figure 3 shows the created CG models of the parade crews of Fune-hoko.

(4) Crowd simulation: We have so far arranged 770 CG models of spectators on both side of the street to regenerate the atmosphere of the event. Idle motions were randomly added to these characters.

(5) Music and sound: We have employed a multi-point sound measuring technology to record and reproduce the acoustic environment. We also collected acoustic data by recording audio sources such as creaking sounds of the wheels of the float, speaking voices of the spectators, and noises made by the crowds, using the same technology [9].

Fig. 2. CG models of the Yama and Hoko floats: from left to right: Fune-hoko, Kanko-hoko, Kitakannon-yama, and Naginata-hoko

Fig. 3. CG models of the parade crews of Fune-hoko: from left to right: hikikata, ondotori, kurumakata, ohayashikata

3 Immersive Environment and Interactive Control

In our previous study, we displayed our Virtual VR contents on a large cylindrical screen and used a gamepad as the input device to control the viewpoint. In this study, as a new step, we introduced a more immersive and interactive display system to allow the users to have a more realistic virtual experience.

The new environment is based on a head-mounted display (HMD), an orientation sensor and a Kinect motion sensor, which are provided by Wirks of WorldViz. This system is the interplay between sensor, tracker, display, software, and content.

The Virtual Yamahoko Parade can be displayed in 3D using the HMD by separately rendering scenes for each eye. The orientation sensor attached to the HMD gives a feedback of which direction the user is facing, in combination with the head position information acquired from the Kinect to update the correct location and looking direction for HMD display. The motion tracking sensor of Kinect is connected with Vizard via a middleware called Flexible Action and Articulated Skeleton Toolkit (FAAST), which enables full-body control and provides a user friendly interface. Kinect allows us to track the user's position and movement in our virtual environment and can be used as a hands-free input device to control the viewpoint and carries out other interactions using gestures. The composition of the system is shown in Fig. 4.

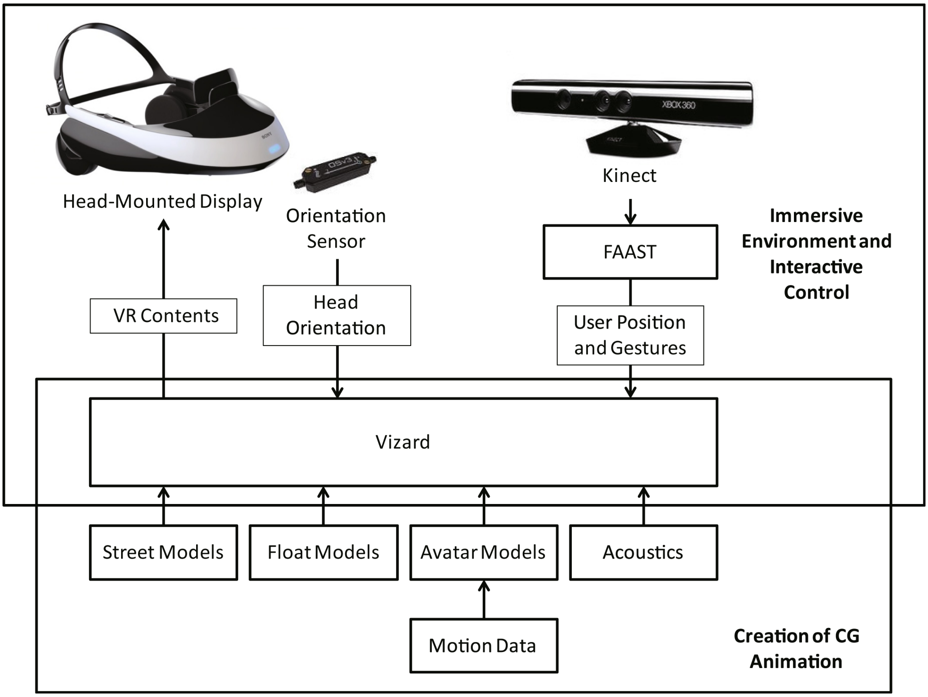


Fig. 4. Composition of the system

4 Results and Discussion

We set four different viewpoints. By gestures of "left hand left", "right hand right", "both hands up", and "crouch", the user viewpoint will be navigated to left crowd, right crowd, bird view, and close float view. At each viewpoint, the view will be linked to and dynamically follows the user's position and looking direction. We received positive feedback from the subjects of our pilot experiments. Figure 5 shows a subject who's using our system.

In current step, the interaction only limited to the viewpoint control. As our next step, we are testing a real time interaction technique that allows the users to interact with the spectators and the float crews in the virtual environment.

We also testing the use of the new interaction technique for our virtualk Yamahoko Parade experience system [10], which allows the user to experience the atmosphere of the parade from another viewpoint, that of the parade crews.

Fig. 5. Virtual Environment using HMD. The synthesized background is the VR contents of the Virtual Yamahoko Parade, which is displayed on the HMD.

Acknowledgement. This study is funded by the "Research and Development of Demonstration Experiment System for the Advancement of Digital Museums" by the Ministry of Education, Culture, Sports, Science and Technology, Japan.

References

1. Furukawa, K., Woong, C., Hachimura, K., Araki, K.: CG restoration of a historical noh stage and Its use for edutainment. In: Zha, H., Pan, Z., Thwaites, H., Addison, A.C., Forte, M. (eds.) VSMM 2006. LNCS, vol. 4270, pp. 358–367. Springer, Heidelberg (2006)
2. Hachimura, K.: Digital archiving of dance by using motion-capture technology. In: Kawashima, A., Akama, R., Yano, K., Hachimura, K., Inaba, M. (eds.) New Directions in Digital Humanities for Japanese Arts and Cultures, pp. 167–181. Nakanishiya Publishing (2008)
3. Magnenat-Thalmann, N., Foni, N., Papagiannakis, G., Cadi-Yazli, N.: Real time animation and illumination in ancient Roman sites. The International Journal of Virtual Reality 6, 11–24 (2007)
4. Papagiannakis, G., Magnenat-Thalmann, N.: Mobile augmented heritage: enabling human life in ancient Pompeii. The International Journal of Architectural Computing 2, 395–415 (2007)
5. Virtual Vaudeville, `http://vvaudeville.drama.uga.edu/`
6. Choi, W., Fukumori, T., Furukawa, K., Hachimura, K., Nishiura, T., Yano, K.: Virtual Yamahoko Parade in virtual Kyoto. In: SIGGRAPH (2010)

7. Li, L., Choi, W., Hachimura, K., Nishiura, T., Yano, K.: Presentation of Japanese Cultural Event Using Virtual Reality. In: Park, J.-I., Kim, J. (eds.) ACCV Workshops 2012, Part II. LNCS, vol. 7729, pp. 72–82. Springer, Heidelberg (2013)
8. Yano, K., Nakaya, T., Isoda, Y.: Virtual Kyoto: Exploring the past, present and future of Kyoto. Nakanishiya, Kyoto (2007)
9. Fukumori, T., Nishiura, T., Yamashita, Y.: Digital archive for Japanese intangible cultural heritage based on reproduction of high-fidelity sound field in Yamahoko Parade of Gion Festival. In: 13th ACIS International Conference on Software Engineering, Artificial Intelligence, Networking and Parallel/Distributed Computing, pp. 549–554 (2012)
10. Li, L., Choi, W., Hara, Y., Izuno, K., Yano, K., Hachimura, K.: Vibration reproduction for a virtual Yamahoko Parade system. IEEE Virtual Reality 2012, 129–130 (2012)

Mission: LEAP

Teaching Innovation Competencies by Mixing Realities

Christopher Stapleton[1], Atsusi "2C" Hirumi[2], and Dana S. Mott[1]

[1] Simiosys, Orlando, Florida, USA
{cstapleton,StoryChaos}@simiosys.com
[2] University of Central Florida, Orlando, Florida, USA
Atsusi.Hirumi@ucf.edu

Abstract. Mixed Reality (MR) melts more than boundaries between realities. MR also melts boundaries between disciplines to stimulate innovation. In a project originally sponsored by NASA, the authors of this paper discuss the case study Mission:LEAP, a Mixed Reality Experiential Learning Landscape. In achieving the core objective of building innovation competencies in youth, we had to expand Space STEM education to include the Arts, Media, Design and Humanities to teach innovation competencies. By play-testing a full-scale mock-up, the process also revealed the value of MR in experiential learning landscapes and defined new aspirations and requirements for innovative ways of how we interface with MR environments in free-choice learning venues.

Keywords: Innovation, Mixed Reality, Informal Education, Interplay, Phydgital InterSpace.

1 Inspiring the Next Generation of Innovation

NASA is historically an internationally recognized leader in innovation. The creative leaps in Science, Technology, Engineering and Mathematics (STEM) accomplished from space exploration have inspired the world across all industry sectors. NASA did not become innovative; it was born innovative out of necessity. It was created to radically innovate the US capability to compete in the world quest of space exploration. The Cold War fear to "keep up" soon transformed into the impassioned hope of "human endeavor" to achieve unfathomable feats of innovation. NASA continues to innovate with ambitious goals to accomplish, "what has never been done before; to go where no man has gone before," which makes it an ideal case study for teaching the core competencies of innovation.

NASA's celestial ambition inspired a nation of innovative visionaries who spawned remarkable economic growth. The rapid rise of companies such as Apple, Google and Facebook because of their innovation is equal only to the surprisingly near extinction of venerable American companies such as IBM, Kodak and Chrysler for their lack of foresight of innovation. The consistent and predictable "operational

R. Shumaker (Ed.): VAMR/HCII 2013, Part II, LNCS 8022, pp. 416–425, 2013.

excellence" that helped companies to last in the past, will be their systemic downfall in the future [1]. The future, with the anticipated constant change of radical innovation, requires bottom-up and agile processes for the intellectual workforce to be creative, passionate, independent-minded, and enterprising so that they not only anticipate change, but drive change [2]. Innovation skills are not just for the leaders, but for the entire workforce [3]. In a new consumer-driven marketplace, the users play a critical role in the innovation process [4]. Where invention can take just one person, innovation requires adoption by large groups of consumers. This adoption requires an innovation culture willing to change and adapt to new means, methods and unforeseen challenges [5].

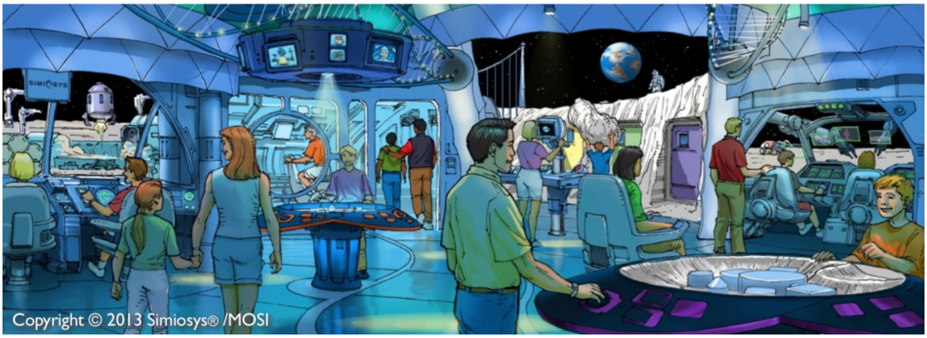


Fig. 1. Concept Rendering, Mission: LEAP (Learning Expedition for Astronaut Pioneers)

2 Is STEM Education Enough to Drive Future Innovation?

How do you prepare today's youth to passionately dive into the unknown and unpredictable future of innovation? Can NASA inspire the next generation innovators? Our goal was to create a Mixed Reality (MR) Experiential Learning Landscape (ELL) that would enlighten the public on the importance of how deep space exploration can teach core innovation competencies. A public that is more innovative can stimulate the speed and ease of the innovation flow that the economy needs for survival [6].

In the 2011 Atlantic Century Report, the United States ranked fourth among industrialized nations in innovation-based competitiveness and ranked last in improvement in international competitiveness and innovation capacity over the last decade (Atkinson & Andes, 2011). Apparently, "the scientific and technological building blocks critical to economic leadership have been eroding at a time when many other nations are actively laying strong foundations in these same areas" (Dept of Commerce, 2012, p. v). With such reports, the current emphasis and funding of STEM education is understandable, but is it enough? While STEM education provides the foundation of invention, the core competencies of innovation rely on soft skills and attitudes, such as fostering creativity, imagination and passion to imagine the possibilities rather than just developing the capabilities. These are competencies that are aligned more with the Arts, Media, Design and the Humanities more than STEM [7]. The economy does not benefit until inventions are transformed into marketable innovation.

The problem is while STEM education is solidifying the foundations for innovation, it is not necessarily fostering the passion and creativity necessary to innovate that is critical for market diffusion. The creative expression of ***art***, the critical problem-solving of ***design***, the compelling impact of marketing (***media***) and the understanding of the human experience (***humanities***), drive the success of products. While Federal and state mandates continue to pressure K12 educators to teach-to-the-test STEM knowledge, they cut the creative skills training of the Arts, Media, Design and Humanities (AMDH) that are critical to achieving the economic rewards of innovation [8]. In addition, corporate culture and training that often focuses on eliminating error, reducing variance and increasing operational excellence, particularly in highly regulated industries, also inhibit creativity and innovation [1]. New approaches to education and professional development are necessary to prepare the next generation of innovators essential for economic success.

3 Convergence of STEM and AMDH to Learn Innovation

NASA's unique position spans civic, academic and commercial interests while tapping both technical and creative fields to produce future innovative solutions. NASA has also become an expert of the larger innovation process, managing both the invention and adoption of the innovation. To complete the cycle, NASA's technology-transfer program bridges innovation back to civilian life where it is designed to stimulate the future economy. So what distinguishes Space STEM from other STEM professionals? Deep space exploration is dependent on a state of constant change driven by persistent innovation that, in turn, depends on innovation competencies that support the legacy of discovery. With this perspective, the space innovation experience can inform the learning experience and the preparation of our overall future workforce.

When interviewing Space STEM Professionals about their role, we were struck by the fact they had the enthusiasm of creative artists. NASA space professionals emphasized their dreams, passions, collaboration, creativity and imagination. In response, our creative design staff were just as passionate and fascinated with the possibilities of space STEM. These shared professional interests and passions reinforced the cross-disciplinary synergy that is at the core of innovation collaboration [9]. The interdependencies between STEM and AMDH drove the design of our learning landscape. STEM and AMDH are in fact, "two sides of the same coin." To guide the design, we crafted a series of antithetical aphorismic stanzas describing transdisciplinary collaboration that was inspired by Pablo Picasso's quote, "Art is a lie, that reveals the truth."

Science *is the truth that unveils the myths;* ***Art*** *is the myth that unveils our truths.*
Technology *imagines the extension our abilities;* ***Media*** *is our ability to extend our imagination.*
Design *is the dream of the achievement;* ***Engineering*** *is achieving our dream.*
Mathematics *is the expression of insight;* ***Humanities*** *is insight of expression.*

Simiosys.com/Transdisciplinarian

4 From Innovation Expert to Novice to Future Expert

We asked Subject Matter Experts (SME) from the highest levels of research at JPL to operational staff at Kennedy Space Center about their pathway to Space Exploration. Many shared a similar frustration of their youth in following their life's dream because, at the time, there were no explicit paths to become an astronaut.

Every space professional we interviewed remembered a similar time when they were first inspired as a child, whose youthful experience still drives their passion today. How can we capture that experience to inspire kids today and motivate learning about their role in innovation? Each of those different experiences of NASA personnel consistently described three characteristics that defined their passion today:

REAL (hands-on): The childhood experience represented a tangible engagement with the reality of STEM/AMDH principles to achieve a satisfying result (taking apart an appliance, watching birds, designing something or blowing it up, etc.). Such experiences emphasized the physical, hands-on nature of learning innovation through curiosity, creativity and experimentation.

RELATIONAL (heart-felt): The experience was with a mentor or peer group (family, class, group leader, etc.) who stoked their curiosity and guided a memorable experience they couldn't have had on their own. Such experiences emphasized the important role of family learning, passions and group (or team) interaction in experience-based venues.

RELEVANT (head-strong): The experience opened their eyes (and mind) to so excite them about a concept that they could imagine themselves doing it for a career ("people can do this for a living?") providing them with a lifetime of wonder. Such experiences emphasized the importance of role playing and the sense of belonging to a larger cause with a meaningful purpose and experience.

A community, family-based learning center is an ideal place for bringing together these aspects of experiential learning. However, so many museum exhibits are one-time experiences that do not span the breadth of experience, nor the knowledge of space or the imagination of the child. These experiences don't change with the time, nor expand beyond the walls of the museum. The experiences are over-simplified, pre-programmed, individual experiences that don't invite that child-parent interaction or problem solving skills [10]. It's not the knowledge that will excite that child about a career in space, but it is stoking their imagination and providing the tools of the trade for them to jump in and have a real, relational and relevant experience. We need to make a memorable experience similar to the childhood memories of the space professionals to make a lifetime impact. This experience challenged even the best practices of museum exhibitory forcing us to innovate to teach innovation.

5 Transporting Visitors into Their Own Innovative Future

To inspire visitors about the future of innovation, we find ourselves where science meets fiction. Mission: LEAP, a Learning Expedition for Astronaut Pioneers

(Fig 1 & 2) is a simulated colony set in the future to spark the imagination of young kids with a vision of their future life as an astronaut. Visitors arrive at a fully operating, simulated habitat to make the experience real, relational and relevant. They then explore opportunities to initiate different levels of innovation by honing their creative skills.

To start, we picked the year 2075 when living in space may be commonplace and astronauts will be more than scientists, pilots, engineers or doctors. Venturing into long-term, deep-space exploration, astronauts will be chefs, artists, media producers, therapists, farmers, entrepreneurs and architects. This future involves creating colonies that will require us to reinvent life as we know it to survive (Fig 2). Without the pristine ecosystem of earth's atmosphere, we will rely on new inventions of STEM to become Techno-Species, as James Cameron describes future space explorers, who are dependent on technology for survival.

The Mixed Reality Experiential Learning Landscape becomes a hybrid of a visceral theme park and a dynamic video game [6] with social networking capabilities. The science-based fantasy challenges the participant's creative imagination and problem solving skills beyond their current experience and abilities. Our attempt is to create the hardest fun the students ever had. As frustrating and ambiguous that difficult creative problems can pose, using the interplay of story structure, game mechanics and play testing makes each challenge more motivating by being fun and social [11].

6 Innovation PlayTank: Transforming Idea to Innovation

Museums offer an ideal venue to augment a person's formal education and to foster the development of innovation competencies with creative hands-on activities. The challenge lies in telling such an infinite story in such a finite space. The typical museum exhibit falls short of providing the breadth and depth of experience to teach such a broad spectrum of innovation competencies. Instead, many museums seem to re-emphasize school standards in a limited, "hit and run" level of interactivity. An exhibit about future innovation should embody a living example of innovation and constantly change and challenge visitors with new learning opportunities. But like many other institutions, when they sense the need for fundamental change, "museums remain enmeshed in a network of interdependencies that undercut their ability to innovate" [15].

Taking a cue from industry and academia, we set up a rapid prototyping facility (Fig 2) called the Innovation PlayTank (versus Think Tank) to make creative leaps in demonstrating innovation. By taking the messy creativity and unpredictable experimentation out of the arena of "operational excellence," the PlayTank avoided defeating criticisms and resistance to change. To satisfy the museum, designers and client (NASA SMEs), we produced four iterations of the full scale exhibit collecting feedback to make improvements on each monthly iteration using subjects of all ages. A series of exploratory play-tests entailed expert observations, survey responses and a focus group.

Fig. 2. Traditional hands-on phenomenon based experiences were combined with virtual game-based activities to extend play, encourage collaboration and deepen learning

Over-thinking and over-planning of complex, creative ideas such as experience design can kill the most promising of experiments. The best way to validate entertainment value and usability is to try it out on an audience sooner rather than later. The practice of "play-testing" is to create experiments much like a child creates play on the playground-- expending the least effort for the most impact and immediate response. The practice was as intense and diverse as a "think tank," but the full scale mock-up was more like a playground. This process became to be known as the Innovation PlayTank (versus think tank).

The Innovation PlayTank took form at the business incubator at the UCF Institute for Simulation and Training. In a 3000 square feet, high bay laboratory the exhibit concept design was fully blocked out with the exhibit floor design and with minimal lighting, audio and scenic elements created a sense of place on the moon with the earth setting over the crater wall. Each station at first used free online games to represent the simulation. A Skype connection represented a social network between stations. Live inter-actors provided the over arching premise and story. Placed between computer stations were simple physical exhibits such as sandbox demonstrating lunar cratering; a salad bowl and air ducting on foam core to represent a space suit; water bottles and ping-pong balls representing physical monitoring stations. In each iteration, the role-playing became more sophisticated, the floor plan evolved, the activities became refined and the virtual content became more specific and and physical. Beyond the extensive creative development and exchange, there were critical insights and "lessons learned" (below) into new requirements for experiential venues that were required for us to improve the Mixed Reality interface.

Be social to be successful: Challenging experiences done alone became frustrating for participants, while the same experience with somebody joining them, became "fun." The aspect of being relational became the difference between success and failure. Ease-of-use concepts to eliminate the frustration would actually kill the fun. Instead, we needed to create not an ***interface*** between computer and human, but an ***InterSpace*** between people and their activities to emphasis the face-to-face interaction between people.

Make Mixed Reality as ONE world. The use of a single virtual display in a physical venue consumes the attention of a single user into the virtual presence beyond. This single screen decreasing their awareness of their interpersonal or physical surroundings. This virtual-real chasm in their sense of presence prevented cross-over activities between real and virtual spaces detracting from the effectiveness of mixing realities. Leveraging multiple screens provided emphasis to the space between the displays, versus in the display themselves. This allowed for a more spatial awareness enhancing the InterSpace between people. We needed to start with the physical world and embed or project the virtuality in or on the reality to make sure we did not lose the social in lieu of the virtual.

Make it real. The more realistic the interaction, the more the user enjoyed, engaged and believed in the experience. Allowing the guests to engage with real tools of the trade, was dramatic in its ability to transport audiences. Virtualizing real equipment should take priority over trying to make equipment virtual. The physical haptics engaged the learning experience more.

Make it unreal. The power of experiential venues is the ability to “be there” with peripheral stimulus with all the senses. Just a small amount of peripheral and tangential perception can stimulate the imagination with sounds, sights, smells and feelings. Leaving more to the imagination directed with well orchestrated tangential, multi-sensory stimulus can better transport their presence to places that you could never fabricate real or virtual. However, no effort to the design immersive sensory illusion can be detrimental to the experience.

Smiling is the best learning assessment tool: While video games can have multi-players, the typical game design becomes co-active, versus socially interactive event. To stimulate the relational aspects of the learning event we emphasized the face-to-face visual contact across the playing field or via teleconferencing. This provided critical cues for the mentor or docent to exchange joy, inspiration, comprehension or frustration.

The best part of reality are the living. The coolest displays of virtuality and exhibitory were ignored after the first few minutes. To stimulate longer term engagement and inter-zone interaction, took significant effort technically. However, the efforts of a single live inter-actor to make connections, stoke exploration, excite challenges transformed the experience. We soon needed to develop a networked system that allowed a single interactor to more readily instigate group interaction, generate a colony-wide emergency, puppet a virtual robot, or monitor all displays via teleconferencing, providing a powerful and salient experience. Technology could never replace a live “experience conductor,” but technology can significantly augment their abilities. Without the live interactor, other exhibits equated to mere vending machines of learning with little engagement, depth or duration.

Competition is a double edge-sword. Game points dramatically increase competitive engagement and emotional investment. However we saw, if competition and point making is not uncheck or monitored, competition can quickly lead to cheating, spiteful aggression and panic that can dramatically interfere with the fun and the learning. In a face-to-face experiential venue, cooperative play as in the case of an emergency response or competing against a non-present virtual competitor increased motivation, stimulated learning and providing more enjoyment.

The GUI gets in the way. Every example of a “Drag and drop,” or “point and click,” GUIs needed to be replace with “cause and effect” responsive simulations. This is because the abstraction of the GUI required too many instructions, demonstration and rule sets within a physical simulation venue, which led to decreased understanding, fun and engagement. Guests wanted immediate interaction and stimulus for “trial and error” feedback loop or “play.” Virtual simulations and physical “sandbox” type of phenomenon-based experiments (with intuitive and naturalistic interaction and immediate tangibility) significantly out performed the more push-button computer graphic games.

In creating a prototype of this off-world colony experience to a high level of engagement, we had to challenge the common physical interfaces of traditional museums and question the validity of the common GUI interfaces used in virtual interaction. The series of iterative play-testing experiments provided critical new requirements of the next generation mixed reality experience for informal education facilities.

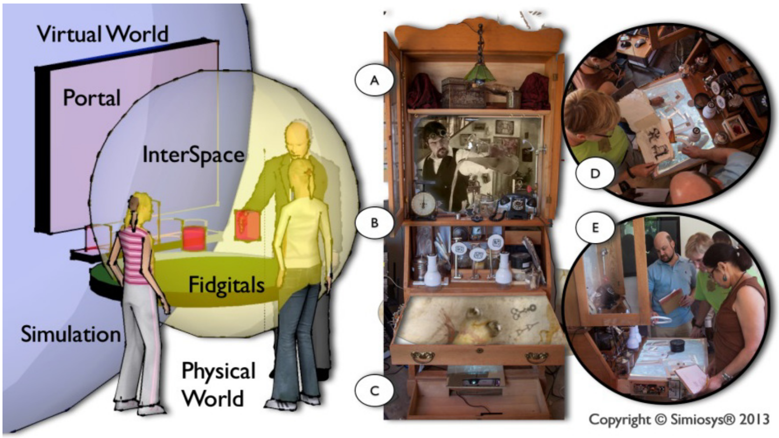


Fig. 3. Phydgital InterSpace prototype play-testing at an Innovation PlayTank

7 Phydgital InterSpace Melting Boundaries with Mixed Reality

The Mission: LEAP mock-up provided critical requirements for innovating immersive Mixed Reality interfaces for museums. Museums are experiential venues whose entertainment value comes mostly from social face-to-face interactivity with hands-on engagement. The use of a physically dynamic objects in between people is the most rapid, intuitive way to invite participation, escalate engagement and increase the desire to challenge users. A ball, a sandbox or a water fountain are simplest examples that achieve this state of engagement with kids. With these examples, there is an immediate attraction, intuitive interaction and the sensory feedback that provides satisfying and repeatable social play. We needed to achieve the same in a interactive station.

Our solution was a Phydgital InterSpace that starts with the physical object and environment (Fig 3) and embeds physical artifacts with the virtual life so that they can be intuitively manipulated within a shared space between participants and virtual worlds. This object becomes the "Phydgital," the physical/digital object that you fidget with. This works in tandem with the shared tangible space in between participants and virtual portals that we call the "InterSpace." Once we used multiple screens and physical objects, the tendencies of the users was to not be consumed by the virtual presence beyond any one virtual portal interface. This concept combines the intuitive interaction of physical reality and leverages the dynamics of the virtuality.

The first prototype of the Phydgital InterSpace was play-tested in an experiment to make a curio cabinet come alive for use in a history museum or science center (Fig 3). In the cabinet a fictional, virtual biologist was creating genetic mutants and came to us within an image of the mirror that capture his presence from years before (A). His experimental biology tools became Phydgitals (B) that could be picked up and used in his vernal pool to feed, kill or infect swimming creatures (C), which was a virtually projected in a drawer representing a watery ecosystem of the biologist's creation. Other physical artifacts, like his journals provided secrets and clues as to the mystery of this curious display. Both virtual displays, one computer generated and the other video captured represented extensions of the physical space. The vertical reference portal provided character interaction through video while the horizontal simulation portal provided immediate cause and effect, trial and error interaction and feedback. Designed as a time-based puzzle, participants gathered around the vernal pool to experiment with the phydgitals to determine if they should save the creatures or euthanize to prevent contamination. No instructions were provided. However, guests naturally and instinctively went through a loose scientific process to figure things out. The frustration and ambiguity of the open-ended problem solving (there was no "right" answer) seemed to increase the entertainment value, because the participants could choose their own creative solution. Objects were so curious and responsive, while the social interaction prevented the challenge from becoming to frustrating.

The simulations provide repeatable play and intuitive interaction that can provide an opportunity for creativity. This proved to create an escalation of inquiry based learning that naturally followed the scientific process. These tests will inform the next generation of Mission: LEAP and push the standards of museum design.

8 Conclusion

Our future economy is dependent upon the competitive edge of creative innovation. To that end, we can't over-emphasize Science, Technology, Engineering and Math at the cost of the loss of the Arts, Media, Design and the Humanities. In inspiring the next generation of innovator, neither can we myopically assess the retention of knowledge without evaluating the application of knowledge with skills creative and attitudes. We can't think that virtuality can do everything that reality or imaginality can achieve in learning [16] and need to innovate new paradigms in interface design in mixing realities. This act of innovation to help teach innovation emphasizes the importance of experience-based education venues as real world laboratories of learning and experimentation need to evolve and innovate with the times.

References

1. Foster, R., Kaplan, S.: Creative Destruction, Why Companies that are Built to Last Underperform the Market – How to Successfully Transform Them. Currency Books, New York (2001)
2. Herbig, P.: The Innovation Matrix, Culture and Structure Prerequisites to Innovation. Quorum Books, West Port (1994)
3. Howkins, J.: The Creative Economy, How People Make Money from Ideas. Penguin Books, New York (2001)
4. Johnson, S.: Where Good Ideas Come From, The Natural History of Innovation. Riverheads Books, New York (2010)
5. Brown, T.: Change by Design, How Design Thinking Transforms Organizations and Inspires Innovation. Harper Collins Books New York (2009)
6. Ramo, J.C.: The Age of the Unthinkable, Why the New World Disorder Constantly Surprises Us and What We can Do About It. Little, Brown & Company, New York (2009)
7. Pink, D.H.: A Whole New Mind, Why Right-Brainers Will Rule the Future. The Berkeley Publishing Group, New York (2005)
8. Kelly, T., With Liftman, J.: The Art of Innovation. Currency, New York (2001)
9. Schrage, M.: Shared Minds, The New Technologies of Collaboration. Random House, New York (1990)
10. Jonassen, D.H.: Learning to Solve Problems, An Instructional Design Guide. John Wiley & Sons, Inc., Pfeiffer Publishers, San Francisco, CA (2001)
11. Stapleton, C.B., Hughes, C.E.: Emerging Technologies of Augmented Reality: Interfaces & Design, Making Memories. In: Science Meets Fiction Creating Compelling Entertainment with Augmented Reality. IDEA Group Inc. Publishers (2008)
12. Lindsey, L., Berger, N.: Experiential approach to instruction. In: Reigeluth, C., Carr-Chellman, A. (eds.) Instructional-Design Theories and Models: Building a Common Knowledge Based, vol. 3, pp. 117–142. Routledge, New York (2009)
13. Stapleton, C.B.: &Hirumi, A. Interplay instructional strategy: Learning by engaging interactive entertainment conventions. In: Shaughnessy, M. (ed.) Pedagogical Implications for Online Instruction. Hindawi Publishing Corporation (2009)
14. Hirumi, A., Stapleton, C.B.: Integrating Fundamental ID Tasks with Game Development Processes to Optimize Game-Based Learning. In: Miller, C. (ed.) Games in Education (2008)
15. Rounds, J.: The Best of Practices, The Worst of Times. In: McLean, K., McEver, C. (eds.) Are We There Yet? Conversations about Best Practices in Science Exhibition Development (2004)
16. Stapleton, C.: Imagination, The Third Reality. In: International Symposium on Mixed and Augmented Reality (ISMAR) Proceedings (2011)

ChronoLeap: The Great World's Fair Adventure

Lori C. Walters[1], Darin E. Hughes[1], Manuel Gértrudix Barrio[2], and Charles E. Hughes[1,3]

[1] Institute for Simulation and Training, University of Central Florida, Orlando, Florida, USA
{lcwalter,dhughes}@ist.ucf.edu
[2] Faculty of Communication, Universidad Rey Juan Carlos, Madrid, Spain
manuel.gertrudix@urjc.es
[3] Department of Electrical Engineering and Computer Science, University of Central Florida, Orlando, Florida, USA
ceh@cs.ucf.edu

Abstract. *ChronoLeap: The Great World's Fair Adventure* utilizes the educational potential of immersive 3D virtual venues for children and early adolescents between 9 and 13. Virtual reality environments transport the mind beyond the 2D bounds of text or photographs; they engage the imagination and can be a powerful tool for conveying educational content [1]. *ChronoLeap* leverages these innate qualities and weaves together the individual threads of single disciplines into a multi-disciplinary tapestry of web-based exploration through the 1964/65 New York World's Fair. Through their myriad of pavilions and exhibits, World Fairs offer links to science, technology, engineering, mathematics, art and humanities topics. *ChronoLeap* provides an immersive 3D environment with highly accurate and detailed models, and merges it with games and themes designed to provide users an educational STEAM environment. The project is a collaborative effort between the University of Central Florida, Queens Museum of Art and New York Hall of Science.

Keywords: STEAM, STEM, Immersive Education, virtual environments, virtual heritage, interdisciplinary, 1964/65 New York World's Fair.

1 Introduction

The roots of the 1964/65 New York World's Fair (NYWF) are entwined with those of a Fair held a generation earlier. The 1939/40 NYWF rose from the "Valley of the Ashes" depicted in F. Scott Fitzgerald's The Great Gatsby to transform over 1200 acres of blight to that of public parkland [2]. Visitors to the 1939/1940 Fair found an ideal future world carefully crafted to provide escape from the economic depression and global unrest that dominated the 1930s. When New York welcomed the world back to Flushing Meadows in 1964/65, Fairgoers experienced more than a celebration of post World War II American prosperity and Space-Age wonder as the Fair also reflected the changing domestic social landscape and Cold War fears.

World's Fairs are a centralized celebration of discovery. The 1964/65 New York World's Fair was the last 'great' Fairs held in the United States with over 51 million

R. Shumaker (Ed.): VAMR/HCII 2013, Part II, LNCS 8022, pp. 426–435, 2013.

attendees. For the price of a $2.00 admissions ticket, a Fairgoer could discover art from around the world, glimpse into the nation's history, view advances in science and technology and peer into the future, all in one compact locale. Many of the technologies we enjoy today, or the foundations that made them possible, were showcased at the 1964/65 NYWF. The 21" color televisions on display beckoned viewers to make the costly transition to a new color set. At numerous pavilions, presentations explained why the computer was a technology to embrace not fear. General Electric's Carousel of Progress allowed everyone to journey through a 100-year evolution of home technology. The Bell System provided visitors a close-up look at the Telstar satellite and a glimpse into the communications revolution.

The 1964/65 New York World's Fair environment that *ChronoLeap* builds upon is unique in that it was an actual place with strong science content that possessed a futuristic world appearance. Our user testing has found that this visionary world has strong appeal with the youth of today. Evaluation of the NYWF environment of users from 9 to 12 revealed the content "resonated with their past experiences and/or current interests" and was capable of sustaining continued interest within that demographic [3]. *ChronoLeap* examines the foundations of the technologically driven world we currently occupy and forges links that promote intellectual curiosity and engaged exploration.

Fig. 1. New York State Pavilion and avatar

2 Relevance to STEAM Learning

The virtual world of *ChronoLeap: The Great World's Fair Adventure* provides a true interdisciplinary learning environment, within which users can not only explore a multitude of STEM topics, but can also come to understand the impact arts and humanities had on this showcase of scientific and technological innovation. *ChronoLeap* provides a rare opportunity for users to break out of the siloed view of education. The environment demonstrates how the humanities influence STEM discoveries, how STEM discoveries have inspired critical events in the humanities and incorporates the artistic aspects of technology and engineering innovations pertaining to the 1960s and their evolution to today.

As there was following the Soviet Union's launch of Sputnik in 1957, there is again considerable concern within the education community of the state of science, technology, engineering and mathematics education. Recently there has been an effort to educate the public on the value of art and now the letter A has been added to STEM to make STEAM [4]. This stresses the value of creative/artistic skills in maintaining the competitiveness of our youth. However, the value of the humanities in this new wave of educational thought is being underplayed. The humanities provides not only a review of what has been, but an understanding of the evolution of STEM, the forces that led to where we are and the skills to understand how STEM topics can influence where we are going as a society.

3 Experiences

The world of *ChronoLeap* is explored through a series of “Time Bubbles.” Each bubble is a self-contained level that focuses on a specific STEM theme, linked together through STEM artifacts found within a virtual representation of our partner, the Queens Museum of Art. As the user embarks on a time bubble journey, they find themselves in a world where exploration, inquiry, and interaction are fostered through a variety of mechanisms – a virtual data pad, STEM stops, games and content connections that link seemingly disparate experiences by a thread of shared knowledge. The storyline of restoring a time paradox drives the user to complete each level and continue to move from bubble to bubble. Within each bubble the user finds secondary quests or games that enhance their understanding of how technology influences society and its evolution. The completion of a time bubble permits the user to return to the Queens Museum of Art and embark on a new mission. As with commercial video game titles, the user can unlock playful elements within the virtual world. *ChronoLeap* also serves as a central location to embark on in-depth interactive experiences that help bring understanding and relevance to the advances in technology and humanities/art content that were showcased at the NYWF.

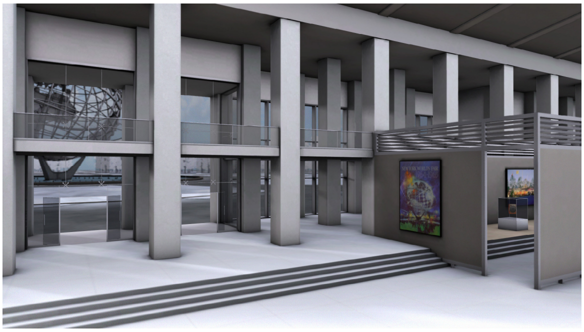


Fig. 2. Interior Virtual Queens Museum of Art

Within *ChronoLeap,* the collaboration between various disciplines such as music, cultural heritage, science and technology permits users to address the content holistically. Thereby, users can make rich connections, investigate the complexity of history, understand the artistic achievements of humanity, and apply these learned base skills to different STEAM learning experiences. These experiences are found within the larger World's Fair environment as a series of independent in-depth games. Hosted on the project's website, our independent games have a direct connection to a pavilion and address specific STEAM topics. Each game can be downloaded and several played online allowing for individual users to explore content areas that align with their own interests – enriching and supplementing lessons learned in the Fair while also providing a unique opportunity to discover new ideas.

3.1 Pythagoras' Music Challenge

At the Greece Pavilion *Pythagoras' Music Challenge* is a full interactive 3D game that blends art and mathematics. The experience provides the user with an understanding of the mathematical basis for musical intervals – the main component of Western scales and chords – while also allowing for free-play and original composition. Embedded within the visual design are the physics principles behind sound waves wherein each note emits a particle effect representative of its relative frequency. The 12 notes of a C Major scale are represented by bells laid out before the user. "Hitting" a bell generates a note corresponding to the letter displayed on the bell.

Fig. 3. Pythagoras' Music Challenge

3.2 Sea Hunt

Accessed from the Travel and Transportation Pavilion, *Sea Hunt* is a 3D, first-person "eater" from the view point of an anglerfish on a deep sea mount in the "midnight zone" (3000 – 13,000 feet below sea level). This game teaches users about deep sea ecology, the predator/prey relationships of that ecosystem, bioluminescence, hydrothermal vent ecology, and the historical and engineering accomplishments of the Bathyscaphe Trieste (a submersible that reached the bottom of the Mariana Trench in 1960 – a feat that has only recently been matched). In this experience, the player swims around in the dark waters encountering bioluminescent creatures, some of which they can eat, while others, such as the colossal squid and larger toothfish, will attack and eat them. The user encounters giant tube worms, sea cucumbers, jellyfish, copepods, lantern fish, Beroe forskalii, arrow worms, and other creatures of the deep. The game is completed after the player has encountered all of the various creatures, eaten enough to keep its health up, and avoided being eaten.

Fig. 4. Sea Hunt Title Card

3.3 Spain Pavilion

Exhibiting works of art from El Greco, Goya, Dali, Picasso, the Spain Pavilion was often referred to as the "Jewel of the Fair." The exhibits within its walls provided a glimpse into Spain's heritage and its influence on the New World. It explored how the processes of cultural change, through communication, assimilation, acculturation, and transculturation, generated rich cultural systems. In Spain's case, the exemplar is cultural evolution through folklore music heritage, and more specifically, through its internationally known folk expression, Flamenco.

In *ChronoLeap's* Spain Pavilion, the planned experience is told through the eyes of Paco, a 12 year old assistant of an adult anthropologist. Initially Paco introduces the breadth of Pavilion activities in the general context of Fair. He then enlists the aid of our player to help him organize various cultural elements found within the Pavilion. The user must help Paco investigate three levels of discovery that ultimately culminate in a game addressing Flamenco. From the perspective of STEAM, the experience demonstrates a novel relationship between the knowledge of arts, humanities and science, and it builds bridges between different forms of access to knowledge. This relationship has been previously studied and defended from several points of view: the development of creativity through arts [5], the benefits of integrating the arts into STEM study [6], the role of music and arts in STEM initiatives [7] and the economic impact of cultural and creative industries.

4 Learning Theories

Studies have addressed the capacity of immersive environments to improve and enhance the education process. Immersive environments can provide multiple perspectives into a particular topic thereby allowing a user to view a real world scenario from a variety of disciplines. This kind of perspective-taking offers unique opportunities that might be difficult, dangerous, or even impossible in everyday life (e.g. learning about deep ocean ecology through the perspective of an angler fish) [8] [9] [10]. Additional studies have repeatedly indicated the relevance in the present context of integrate immersive environments and games in education to improve the motivation and engagement of youths. Successful learning strategies in games are when the users have an active participation and interact in the experience of learning [11]. Research has demonstrated the ability of educational gaming to improve the engagement of students both in formal and informal learning situations. Some theorist consider educational gaming as a kind of "empirical application of constructivist theory" [12] [13].

Specifically within *ChronoLeap*, educational content is delivered using a wide variety of modalities. We have utilized the expansive topical world of the actual Fair to serve as a mechanism to deliver content-appropriate knowledge associated with with particular pavilion. A "learning by" set of educational metaphors was adopted during the overall experience design: learning by seeing/hearing, doing, playing, creating, and sharing. Each of the various STEAM delivery mechanisms utilizes one or more of these metaphors maintain a wide range of learning strategies and most importantly levy the advantages and disadvantages of any individual approach.

Within the *ChronoLeap* environment, users encounter "STEM Stops" that adopt the learning by seeing/hearing approach. The most traditional of methods within the Fair, STEM Stops provide the user with a high level of content but a low level of interactivity. Text, image, video, and sound clips appear in pop-up windows that are collected as the user explores the various pavilions within the virtual Fair. Each of these STEM Stops can be revisited at any time by the user. This method provides educational information in a passive format that is similar to that found in books or websites.

Learning by doing is experienced by "hands-on" games similar to the Pythagoras Music Challenge game described above. These are intended to provide the 'if this – then that' experience that a student might find in a "lab" class. In these modules there are no right or wrong answers, the focus is on real-time interaction and experimentation. Our user is presented with a simulation based on scientific models and dynamic data sets that can be manipulated with sliders and buttons.

The time travel adventure aspect of *Chronoleap* focuses on freeplay learning. The user undertakes the various themed missions within the time bubble levels collecting items and completing missions. This is designed to be an age appropriate version of popular video games such as *LA Noire*. The emphasis on freeplay allows users to embark on a voyage of discovery and exploration. This method differs from many traditional approaches in education in that the path to knowledge is not predetermined by the designer/instructor but rather the designer/instructor must design the game in such a way that the user feels compelled to explore.

Learning by creating is a method in which the user generates content given the context provided by the simulation. Just as an educator might assign a student to write a short story about a historical event, in this context, the user might be asked to create a digital interpretation of a particular painting, event, or concept. This method of learning empowers the student to give their own impression of an idea or concept. By giving the user the ability to take control of the content, the subject matter can become more personal and meaningful. We have adopted this thought in our Sketch and Tell exerience. Once a user has completed *ChronoLeap* and defeated the time paradox, they are presented with a pure user-generated content experience. Building on the future visions presented at the 1964/65 Fair, in Sketch and Tell we ask the user to consider their home, automobile or city of the future. They physically draw their item of the future and our software captures the visuals and, most importantly, an audio explanation of these visions. Here we are stimulating their ability to observe, interpret and create – providing a true STEAM experience.

Finally, we examine learning by sharing of selective materials through online forums. This method ultimately becomes the showcase for user-generated content, enabling the users to teach and learn from each other. We view this as a digital version of the traditional show-and-tell.

5 Discussion of Results

Educational experiences in immersive environments have demonstrated improved learning processes in different aspects: engagement, contextualization of learning, transfer capability and collaborative work [14] [9], and especially in the arts and music [15] [16]. However studies have also established some of its current limitations [17]. In addition, other studies have evaluated the importance of sound and music in an immersive environment to improve communication processes [18] or the development of skills such as memory [19].

The New York World's Fair environment underwent significant evolution from when the project began to its current beta *ChronoLeap* version. Formative testing in the second half of 2011 revealed the desire by users for a more compartmentalized world than the massive sandbox environment initially envisioned. We found that user feedback continually requested a level approach. Users were uncomfortable with freely navigating 660 virtual acres and effectively making the STEM connections across the expansive environment. This also enabled the introduction of a strong story line to assist users, as we found this to be the second inquiry during testing – the all-important "why am I doing this task?"[12]. Subsequent testing has revealed the users enjoyed the time travel story line coupled with the mission based level approach. As each mission focused on an overriding STEM theme it became easier for the user to make the connections between the subjects.

Formative testing conducted on all twelve of the independent games yielded positive results both in terms of general enjoyment and transference of knowledge. Specific to Sea Hunt, the perspective-taking approach to education was employed wherein the player, takes the "point-of-view" of an angler fish. This metaphor provoked a discussion about what it might be like to play the game through the eyes of several of the other creatures and experience the predator/prey relationships through a variety of modalities. Another player commented that he would like to be able to "see himself" (as in a third-person perspective) suggesting that the player had embraced the metaphor of player-as-fish. Furthermore, several participants commented on the "mysteriousness" of the game and felt this quality made the game more exciting – perhaps indicating an increased sense of immersion [20].

The children in this study were unanimous in their belief that the game was compelling and enjoyable. When asked if they believed if the game was a good way to learn about deep ocean ecology, all students answered "yes." Many of the participants communicated that they were unaware of bioluminescence, tube worms, hydrothermal vents, and most of the creatures in the game. They referenced many of the organisms by name in the post surveys indicating that the students took the time to read the pop-ups and were able to actively discuss the various elements of the ecosystem. The participants also commented that they wished there were more levels and they wanted the game to be longer [20].

In addition to the formative evaluation, both Sea Hunt and Pythagoras' Music Challenge have been played by large groups of middle school students over the course of the last 12 months within the research laboratory. It has been observed, informally, by both faculty and staff that these and our other independent games foster collaboration among groups of school children wherein the students gather around the computer while one person plays the game. The observers often read-aloud the educational pop-ups offering advice and discussing the content together.

6 Interdisciplinary Collaboration

So what happens when you bring together computer scientists, artists, historians, musicians, and variety of science disciplines? The project benefitted from a highly interdisciplinary collaboration between the University of Central Florida, Queens Museum of Art and the New York Hall of Science. Each organization brought a unique perspective as to how the 1964/65 New York World's Fair could be interpreted as a virtual learning environment, thus providing a solid interdisciplinary perspective at the early planning stages. Of equal importance was the real world experience both the Queens Museum of Art and New York Hall of Science possessed with the target age group. The New York World's Fair collector community was also most helpful in providing assistance in securing the historical materials required to create the pavilions within *ChronoLeap*. We highly recommend securing partnerships beyond the confines of the academic community.

7 Conclusion

ChronoLeap: The Great World's Fair Adventure provides engaging exploration and opportunities to learn about foundational themes, events and individuals from the 1960s and their continuing relevance today. Users are transported to the 1964/65 New York World's Fair – an event whose pavilions and exhibits bring together an expansive range of learning points in a single environment. The virtual world that recreates this event is more than a simple architectural recreation of the Fair; it is a destination for examining the historical foundations of the world we all now occupy and forging interdisciplinary links of wonder and exploration for the user.

Educational researchers have frequently cited the potential for virtual environments to promote learning [21] [9]. Virtual environments provide learners with access to artifacts and contexts that are otherwise not available to them in the real world (e.g., expensive or rare technologies, historical settings such at NYWF). Such environments can encourage exploration and discovery that are consistent with constructivist pedagogies [22] [23]. Virtual environments can also immerse people within new perspectives and new identities that can stimulate learning [24] [25], [9] [26] and lead users to think more like experts [27] Given their strong potential for learning, the virtual environment will continue to develop for education in STEAM and the humanities, including teaching literature [28], art and design education [29], and history education [30].

Acknowledgements. The authors wish to acknowledge the contributions the Queens Museum of Art and New York Hall of Science and for the creative talents of Michelle Adams, Nicolas Beato, Steven Braeger, Eleana Eribiceau, Michael Hopper, Eric Imperiale, Edward 'Chip' Lundell, Daniel Mapes, Arjun Najendran, Julian Orrego, Remo Pillat, Evan Shafran, Kelly Vandegeer, Yiyan Xiong, and Alex Zelenin. The contents of this paper are based upon work supported by the National Science Foundation under Grant No. DRL-0840297 and the National Endowment for the Humanities under Grant No. HD5020707. Any opinions, findings, and conclusions or recommendations expressed in this material are those of the authors and do not necessarily reflect the views of the National Science Foundation or the National Endowment for the Humanities.

References

1. Fernie, K., Richard, J. (eds.): Creating and Using Virtual Reality: a Guide for the Arts and Humanities. Oxbow Books, Oxford (2003)
2. Fitzgerald, F.S.: The Great Gatsby. Charles Scribner's Sons, New York (1925)
3. Kitalong, K.S. (2012). Annual Report of Project Evaluator: June 1, 2011 – May 31, 2012 Interconnections Revisiting the Future. Technical Report. Date Submitted: (June 14, 2012)
4. Fantauzzacoffin, J., Rogers, J.D., Bolter, J.D.: From STEAM Research to Education: An integrated Art and Engineering Course at Georgia Tech. In: Proceedings of 2nd IEEE Integrated STEM Education Conference, ISEC 2012, Trenton NJ, pp. 1–4 (2012)
5. Shapiro, D.: Reaching Students Through STEM and the Arts. National Science Teachers Association (2010), http://www.nsta.org/publications/news/story.aspx?id=56924
6. Crayton, J.: STEAM Art Education Movement: a Creative Approach to Eduction in Innovation for the 21st Century (2011), http://janedapain.net/files/STEM_Art_Education_Movement_crayton.pdf
7. Platz, J.: STEM to STEAM: How do you turn STEM into STEAM? Add the Arts (2007), http://www.oaae.net/index.php?option=com_content&view=article&id=58&Itemid=114
8. Aldrich, C.: Learning Online with Games, Simulations and Virtual Worlds: Strategies for Online Instruction. Jossey-Bass, San Francisco (2009)
9. Dede, C.: Immersive Interfaces for Engagement and Learning. Science 323, 66–69 (2009)
10. Kapp, K.M., O'Driscoll, T.: Learning in 3D: Adding a New Dimension to Enterprise Learning and Collaboration. Pfeiffer, San Francisco (2010)
11. Johnson, L., Levine, A., Smith, R.: The, Horizon Report. Austin, Texas: New Media Consortium (2009)
12. Sorathia, K., Servidio, R.: Learning and Experience: Teaching Tangible Interaction & Edutaiment Procedia, pp. 265–274. Social and Behavioral Sciences (2012)

13. Özkal, K., Tekkaya, C., Çakıroğlu, J., Sungur, S.: A Conceptual Model of Relationships among Constructivist Learning Environment Perceptions, Epistemological Beliefs, and Learning Approaches. Learning and Individual Differences 19, 71–79 (2009)
14. de Freitas, S.: Learning in Immersive Worlds: A Review of Game-based Learning (2006) Bristol: JISC, http://www.jisc.ac.uk/media/documents/programmes/elearninginnovation/gamingreport_v3.pdf
15. Sebald, D.: The CAMP: Teaching and Learning Music Technology Through Virtual Environments. In: Resta, P. (ed.) Proceedings of Society for Information Technology & Teacher Education International Conference 2012, vol. 2012, pp. 893–895. AACE, Chesapeake (2012)
16. Gertrudix, F., Gertrudix, M.: Music in Virtual Worlds. Study on the Representation Spaces.Comunicar, vol. XIX, pp. 175–181 (2012), http://eprints.rclis.org/16742/1/en175-181.pdf
17. Dunleavy, M., Dede, C., Mitchelle, R.: Affordances and Limitations of Immersive Participatory Augmented Reality Simulations for Teaching and Learning. Journal of Science Education and Technology 18(1), 7–22 (2009)
18. Patel, N., Hughes, D.: Studying Listener Comprehension of Sonifications through Visual Replication. In: Proceedings of the Interservice/Industry Training, Simulation & Education Conference (I/ITSEC), Orlando, FL (2011)
19. Fassbender, R., Richards, D., Bilgin, A., Thompson, W., Heiden, W.: VirSchool: The Effect of Background Music and Immersive Display Systems on Memory for Facts Learned in an Educational Virtual Environment. Computers & Education 58(1), 490–500 (2012)
20. Kitalong, K.S.: Sea Hunt Evaluation Report: June 2011 – Interconnections Revisiting the Future. Technical Report. Date Submitted: (July 19, 2011)
21. Bailenson, J.N., Yee, N., Blascovich, J., Beall, A.C., Lundblad, N., Jin, M.: The Use of Immersive Virtual Reality in the Learning Sciences: Digital Transformations of Teachers, Students, and Social Context. The Journal of the Learning Sciences 17, 102–141 (2008)
22. Dalgarno, B., Lee, M.J.W.: What are the Learning Affordances of 3-D Virtual Environments? British Journal of Educational Technology 41(1), 10–32 (2010)
23. de Freitas, S., Neumann, T.: Exploratory Learning for Supporting Immersive Learning in Virtual Environments. Computers and Education 52(2), 343–352 (2009)
24. Ross, J., Tomlinson, B.: How Games can Redirect Humanity's Cognitive Surplus for Social Good. ACM Computers in Entertainment 8(4), Article 25, 4 pages (2010)
25. Bers, M.U.: Identity construction environments: Developing personal and moral values through the design of a virtual city. The Journal of the Learning Sciences 10(4), 365–415 (2001)
26. Squire, K.: From content to context: Videogames as designed experiences. Educational Researcher 35(8), 19–29 (2006)
27. Lindgren, R.: Generating a Learning Stance through Perspective-Taking in a Virtual Environment. Computers in Human Behavior 28(4), 1130–1139 (2012)
28. Webb, A.: Teaching Literature in Virtual Worlds. Immersive Learning in English Studies. Routledge, Taylor & Francis Group, Florence (2011)
29. Hanrahan, S., de Pietro, P., Brown, L.H., Haw, A., Malins, J., Milojevic, M., et al.: Interface: Virtual environments in art, design and education. Arts and Humanities in Higher Education 8(1), 99–128 (2009)
30. Prasolova-Forland, E., Ørjan Hov, O.: Eidsvoll 1814: Creating Educational Historical Reconstructions in 3D Collaborative Virtual Environments. International Journal of Web Portals 3(4), 1–14 (2011)

The Electric Bow Interface

Masasuke Yasumoto and Takashi Ohta

Tokyo University of Technology, Japan
yasumoto@stf.teu.ac.jp

Abstract. The research intends the establishment of the cognitive space where the force feedback, the sense of immersion and existence are implemented in interactive space functioned as the input of physical movements. The Electronic Bow Interface system has been developed based on a Japanese archery with consideration to apply the application of games such as FPS(First Person Shooting) into it; furthermore, contents are also developed with the Electronic Bow Interface system. This research also attempts the actualization of the realistic forced feedback and the interface where physical movement can be reflected.

Keywords: interface, game, interactive art.

1 Introduction

A game pad, a mouse and keyboards with a fingertip were the mainstream of inputs in the territory of interactive contents including video games. The game devices such as Wii Remote[1], PlayStation Eye[2] and Kinect[3] have made possible to the input of various physical movements. Especially, Kinect is possible to process inputs by taking body movement into image analysis, while holding nothing on hands[4]. The spread of these new devices make possible to implement Natural Interaction even other studies apart from video games. However, several issues are caused when these inputs are applied to the standard contents of video games.

Firstly, it is difficult to implement realistic physical movement with high flexibility. The input of an interface has been developed in the way that body movement as input information can take into a computer. However, Output still depends on the fixed image output. In other words, the players still need to pay attention to the fixed images when the players move around their body. Problems of images relation can be happened in video games accompanying no physical movement. 3D contents are functioned with the rotation of images in order to represent the rotation of sights, but it is not functioned with human body movement itself. In other words, human body such as the movement of muscles and centrifugal force does not react to them. Thus, it is necessary to develop the sense of sight motions with high flexibility of realistic physical movement.

Moreover, another difficulty is to give the force feedback to a camera utilized for the reorganization of physical movement when nothing on hands. In case of human body, they do not recognize manipulation only with sights and hearing, but they also recognize the condition of touch. The sense of a fingertip, pressure to muscles and

R. Shumaker (Ed.): VAMR/HCII 2013, Part II, LNCS 8022, pp. 436–442, 2013.

temperature are useful information. The information of them is significant to give the sense of immersion and the sense of existence to the players.

This issue also rises up questions to interactive contents production which physical movement information is utilized. Therefore, in this research, the contents, the issues of sights and an interface are contemplated to overcome technical difficulty in order to implement realistic physical movement in the Electronic Bow Interface device. We would like to introduce “The Light Shooter”, this work contains an interactive content, and the Electronic Bow Interface device by mean of a controller a design and construction of exhibition space.

2 Related Works

In this research, the construction of contents utilized with force feedback is significant. In many cases, an actuator as the feedback of a hand hold type of a controller brings vibration to the players. It is suitable for transmitting the condition of a road surface in a race game. However, since it only conveys simply vibration, it is not versatile. The examples of high versatile researches are as follows. One is in regard to counterforce and impulsive force represented by controlling a spindle [5]; the other is regarding to the sense of being pulled over gives to a user by controlling the rotation movement [6]. However, the amount of strain is limited comparing to realistic physical movement, and it is pseudo and not correct. The impression from uses is different from the manipulation of an original one. It is hard to adopt human physical ability into a virtual space such as games.

The method that versatile devices imitate the faction of original objects is implemented, the game with the shape of a handle and rifle is also available. Their abilities remain poor from a feedback point of view. On the contrary, they can give the sense of reality in motion of a player since the same position of a hand movement is applied to them.

Moreover, although there is no versatility for this, but another case study has been confirmed. One is the interface which includes the feature of a stuffed toy by inserting a sensor into its toy [7]. The other case is a Balance Ball Interface [8], which the back is centralized to adjust the whole body movement. However, they are not reacted by the operation of the players in the force feedbacks. Thus, it is still difficult to feedback from contents randomly. Since objects in reality are used in the interface, a feedback itself brings realistic sensation to the players. These stuffed toy and balance balls can be rebounded and change their shapes with a sensor or without, if pushed. On the contrary, their texture is almost as same as objects in reality, but these are initially made for another purpose, not for operation.

Therefore, the bow interface is developed for making possible to utilize a force feedback repetitively. A player can pull over a bow without an arrow, which means that the position can generate an interface. Moreover, a bowstring is precisely adjustable and controllable by a player.

A bow is a weapon as the same as a rifle, due to their uses and images, violent impression tends to apply for a bow as well. To avoid this impression, simple Biological

Motion [9], adapted in ”Dot Man, Line Man” [10] in “Transform yourself Exhibition” [11], is utilized for simple graphics in the application of the Electronic Bow Interface.

3 Overview of the Electric Bow Interface

3.1 Concept

The Concept intends that the sense of touch and temperature contained in objects and a life are expressed in the force feedback.

The reason of a bow being adapted in this research as follows; firstly, a bow can aim the target at 360 degrees and from a distance. Secondly, it is hand- portable. Thus, it can apply flexible physical movement and the accordance to directions of sights and images into expressions.

A bow is different from a rifle in terms of distance and speed being controllable to a certain extent. The strength of arm is necessary in order to shoot an arrow for a distance and fingers holding a bowstring are also pressured more in proposition. When an arrow is let off, it generates vibration and impact on a bow. Namely, these functions are significant for the force feedback to distinguish its operation as a player wishes. From these reasons, the archery interface is possible design realistic physical motion with the information of physical motion in reality.

An arrow relatively takes time to reach to the target. The construction of archery does not allow rapid fire, which means that it can provide time to a player to consider.

In addition, Contents, urging a player consideration, can be constructed as well.

3.2 Overview of the Electronic Bow Interface System

We have produced the Electronic Bow Interface based on the material and structure of Japanese traditional archery. This system is composed of a player, a projector, PC and a screen (Fig. 1). The information of a bow is acquired, disposed to the Electronic Bow device itself and correspondent to PC, thus, it is possible to use at any location. Moreover, real time calculation is possible due to the advanced calculation of the position of a bow, a player; distance; and the fixation of screen.

The Electronic Interface device is based on the structure and material of Japanese bow. Thus, the touches, sensation, impact of lunching are almost the same as an original Japanese bow. Therefore, a player can apply a bow and Japanese bow experiences to this operation. Similar to archery operation in reality, it is necessary to pull the both side of arms strongly to shoot an arrow, and if an arrow is aimed at longer distance, a bowstring needs be dragged stronger and the strength needs to be added to their hands to prevent a bow to slide due to its reaction. For these reasons, a player does not need to learn new operation procedure with a certain experience of archery due to the availability of the same force feedback in human reactions; a player can start playing immediately.

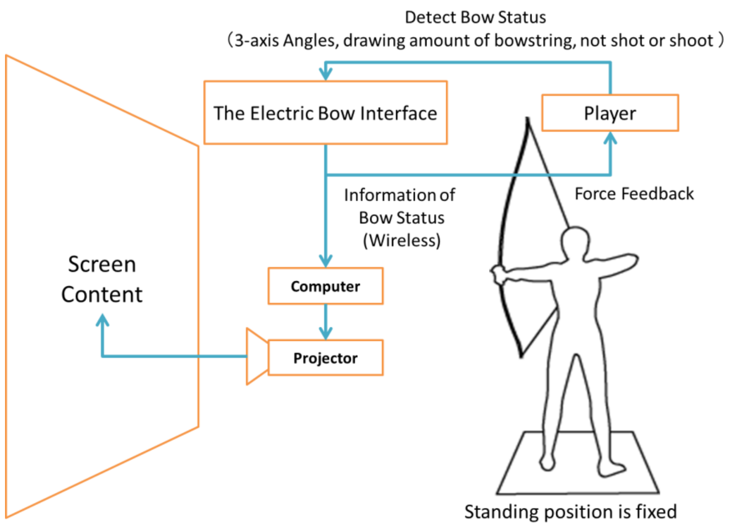


Fig. 1. Overview of the Electronic Bow Interface System

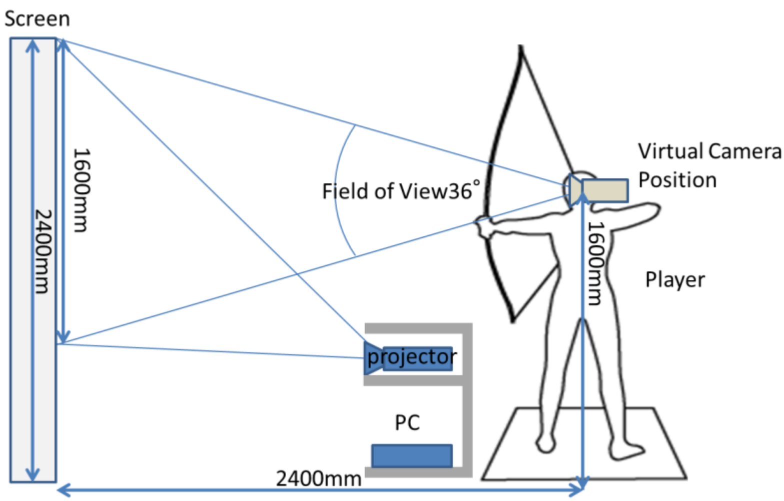


Fig. 2. Positional relationship between the screen and the virtual camera and the player

The position of a player and a screen are fixed. Therefore, the sight angle of a player can correspond with the angle of contents. As a result, a player can acquire immersive experience in the electronic bow (Fig. 2).

3.3 The Electronic Bow Interface Device

The Interface Archery Device is based on the structure and material of Japanese bow. It is utilized by holding a grip attached to the center of a bow with the left hand (Fig. 3).

All electronic devices are contained in its grip. The ArduinoFio [12] is utilized as a Micro Computer to process information sent from each sensor (Fig. 4). Then the Xbee [13] Module is equipped to correspondent to PC. The battery is also equipped in the grip. The three axis accelerate sensor is to gain the slant of archery; the Terrestrial Magnetism Sensor is to get the direction of archery; and the strain gauge is to make sure if an arrow is shot to the target or not. These are attached to it. Moreover, the small size laser pointer is also installed in front. It is for the calibration of directions the bow aims at.

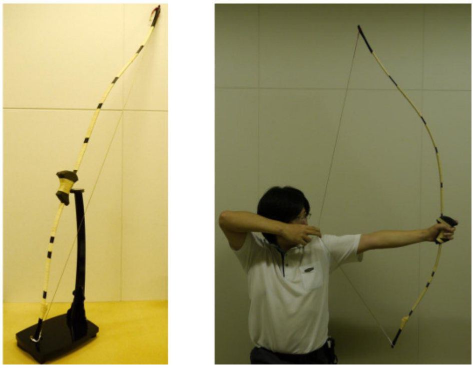


Fig. 3. The Electronic Bow Interface Device

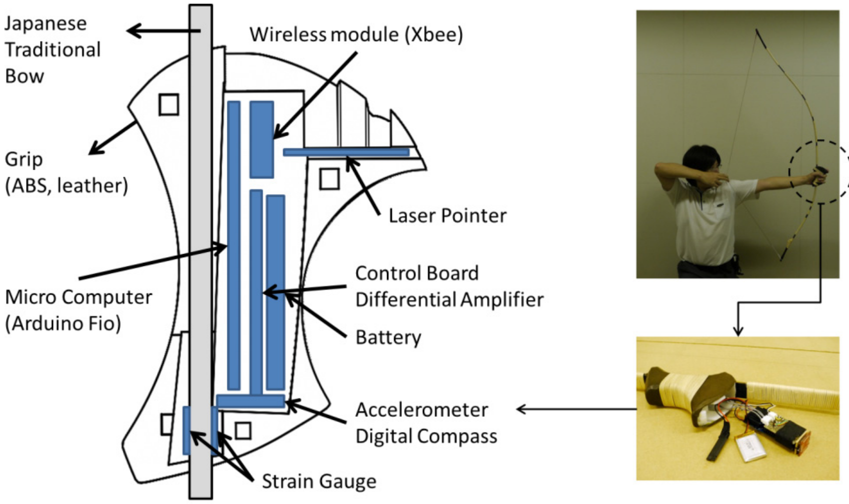


Fig. 4. Detail of The Electronic Bow Interface Device

The grip is produced with ABS resin to protect from the shock of shooting and hand strength. The part of grip is covered with cowhide and the other whole parts are covered with rattan peel used also for traditional Japanese bow. The strength and accuracy of the Electronic Bow device has been confirmed due to the examination by nearly 1000 examiners in the Tokyo Game show 2012.

4 Application

The game application is produced for the experiment to the Electronic Archery interface. First idea is that blocks on screen are simply targeted and destroyed (Fig. 5). For this experiment and entertainment, the device is reproduced in the way that a player can shoot light into darkness where targets are more difficult to be found.

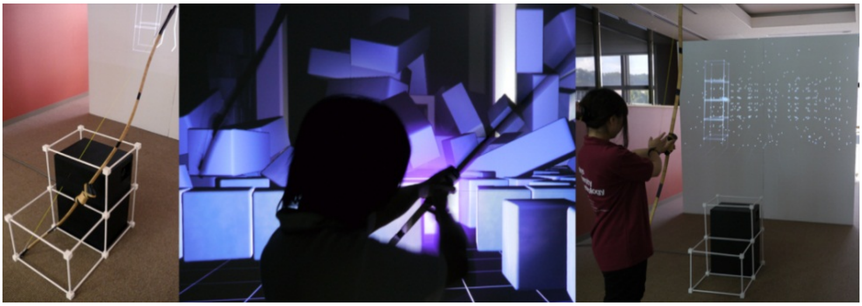


Fig. 5. Applications of The Electronic Bow Interface

4.1 The Light Shooter

The Light Shooter is the interactive art work where the Electronic Archery Interface is equipped, and the function as a game is highlighted (Fig. 6). Firstly, Darkness is shown on the screen, secondly, once the electronic archery is targeted on screen, the target area is expanded, and white dots appears on the screen. These processes utilize the phenomenon of human reactions called, the Biological Motion, which humans

Fig. 6. The Light Shooter

recognizes a life even with a little information such as a few dots. With this phenomenon, the players can distinguish white dots composed of a life enemies and obstacles and they can shoot enemies appearing on the screen for defense. When an arrow is shot to the target, an arrow turns to be a light and it flashes the white lines of the edge composed of a life and obstacles.

In the exhibition, this work is displayed with a mat black color wall, floor and screen. A speaker, a computer and a projector, are all cover up not to disturb well focus on the work.

5 Conclusions

The Light Shooter had been shown in Tokyo Game Show 2012 and SENSE OF WONDER NIGHT 2012 and other exhibitions. We have found that the visitors of the archery experienced operate well than one without. Thus, this device can be more suitable for the experienced. We will continue the research including this issue. Furthermore, we would like to develop the mobile projector, where three dimensions configuration data is included is attached to the electronic archery device itself. With this development, the electronic archery will provide 360 degrees directions mobility to the players.

References

1. Wii Remote, `http://wii.com/`
2. PlayStation Move, `http://us.playstation.com/ps3/playstation-move/`
3. Kinect, `http://www.xbox.com/kinect/`
4. Shotton, J., Fitzgibbon, A., Cook, M., Sharp, T., Finocchio, M., Moore, R., Kipman, A., Blake, A.: Real-Time Human Pose Recognition in Partsfrom Single Depth Images. In: CVPR (2011)
5. Yoshino, K.: `http://www.ipa.go.jp/about/jigyoseika/07fy-pro/mito/2007-0736d.pdf`
6. Amemiya, T., Maeda, T.: Directional Force Sensation by Asymmetric Oscillation from a Double-layer Slider-crank Mechanism. ASME Journal of Computing and Information Science in Engineering 8(3) (2008)
7. Sugiura, Y., Lee, C., Withana, A., Makino, Y., Inami, M. IGARASHI, T.: Pinoky: a ring that animates your plush toys. In: SIGGRAPH Asia 2011 Emerging Technologies, SA 2011, vol. 14, pp. 14:1–14:1. ACM, New York (2011)
8. Yasumoto, M.: Balance ball interface. In: ACM SIGGRAPH ASIA 2008 artgallery: emerging technologies (2008)
9. Johansson, G.: Visual-perception of biological motion and a model for its analysis. Percept Psychophys (1973)
10. Yasumoto, M., Ishikawa, M., Sato, M.: Dot Man, Line Man. NTT InterCommunication Center (2008)
11. Sato, M., Kiriyama, T., Ishikawa, M., Saito, T., Yasumoto, M., Idaka, K., Katsume, Y.: ICC Kids Program 2008 Transform yourself. NTTInterCommnucationCatner (2008)
12. ArduinoFio, `http://www.arduino.cc/`
13. Xbee, `http://www.digi.com/xbee/`

Author Index