Journal Subline

LNCS 6250

Xiaopeng Zhang
Kevin Wong
Guest Editors

Transactions on Edutainment IV

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

Springer

Lecture Notes in Computer Science 6250

Zhigeng Pan Adrian David Cheok
Wolfgang Müller Xiaopeng Zhang
Kevin Wong (Eds.)

Transactions on Edutainment IV

Springer

Editors-in-Chief

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

Adrian David Cheok
National University of Singapore, Singapore
E-mail: adriancheok@mixedrealitylab.org

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

Guest Editors

Xiaopeng Zhang
Chinese Academy of Sciences, IACAS, Beijing, China
E-mail: xiaopeng.zhang@ia.ac.cn

Kevin Wong
Murdoch University, Perth, WA, Australia
E-mail: k.wong@murdoch.edu.au

Library of Congress Control Number: 2010930658

CR Subject Classification (1998): I.2.6, I.3, H.5.2, K.4, H.5.1, H.4

ISSN 0302-9743 (Lecture Notes in Computer Science)
ISSN 1867-7207 (Transactions on Edutainment)
ISBN-10 3-642-14483-7 Springer Berlin Heidelberg New York
ISBN-13 978-3-642-14483-7 Springer Berlin Heidelberg New York

springer.com

Printed in Germany

Typesetting: Camera-ready by author, data conversion by Scientific Publishing Services, Chennai, India
Printed on acid-free paper 06/3180

Preface

E-learning and digital entertainment techniques, tools and systems are becoming popular and can be found in many real-world educational applications in many countries. The driving force behind these technologies is the rapidly growing requirements of edutainment, especially from the perspective of the users. This has led to the increasing interest of researchers in this area. The articles in this issue give a rich overview of how edutainment technologies can be creatively used for training and education purposes.

The first 12 articles of this issue represent a selection of outstanding contributions from Edutainment 2010, the 5th International Conference on E-learning and Games, held in Changchun, China, in August 2010. The main purpose of this event is to facilitate the discussion, presentation, and information exchange on the scientific and technological developments in the emerging field of digital entertainment techniques and systems for education. These 12 papers concentrated on three aspects: e-learning system and applications, game techniques for learning, and virtual reality techniques for entertainment. They are closely related to the topics of this journal.

The first four papers cover e-learning systems and applications: "The Study and Design of Adaptive Learning System Based on Fuzzy Set Theory," "Modeling Personalized Learning Styles in a Web-Based Learning System," "An Emotional Agent in Virtual Learning Environment," and "Lunar Surface Collaborative Browsing System for Science Museum Exhibitions." The next four papers are on game techniques for learning: "Towards a Structural Model for Intention to Play a Digital Educational Game," "Case Study of FISS: Digital Game-Based Learning for a Broad Range of Ages," "Woodment: Web-Based Collaborative Multiplayer Serious Game," and "Learning with Virtual Reality: Its Effects on Students with Different Learning Styles." The subsequent four papers are about virtual reality techniques for entertainment: "Automatic Motion Generation Based on Path Editing from Motion Capture Data," "Exploration of Metaphorical and Contextual Affect Sensing in an Intelligent Agent," "Synchronizable Objects in Distributed Multimedia Applications", "Anisotropic Cloth Modeling for Material Fabric."

The last ten papers in this issue are regular papers, focusing on two main topics: virtual reality and game systems for learning, and animation and interaction for entertainment. In "A Virtual Reality Simulator Prototype for Learning and Assessing Phaco-sculpting Skills," Kup-Sze Choi presents a virtual reality-based simulator prototype for learning phacoemulsification in cataract surgery. In "An Augmented Reality Nanomanipulator for Learning Nanophysics: The 'NanoLearner' Platform," Florence Marchi et al. focus on the development and evaluation of an augmented reality nanomanipulator, called the "NanoLearner" platform, to be used as an educational tool in practical work for nanophysics. In "Fast Prototyping of Virtual Reality-Based Surgical Simulators with PhysX-Enabled GPU," Wai-Man Pang et al. present their efficient approach in prototyping of a series of important but computation-intensive functionalities in surgical simulators based on newly released PhysX-enabled GPU. In "Dance-Based ExerGaming User Experience Design Implications for Maximizing Health Benefits Based on Exercise Intensity and Perceived Enjoyment," Alasdair Thin et. al. indicate that the game play mechanics and skill demands of the dance-based ExerGames limited the subjects' level of physical exertion over the period of study. In "Learning Ultrasound-Guided Needle Insertion Skills Through an Edutainment Game," Wing-Yin Chan et el. present a novel training system with the integration of game elements in order to retain the trainees' enthusiasm.

In "Sketch-Based 3D Face Modeling for Virtual Characters," Wei Jiang et al. propose a mapping mechanism on the basis of contour lines for the use of sketch-based interface in 3D face modeling and focus on how to map 2D sketchy features onto a 3D model. In "A Framework for Virtual Hand Haptic Interaction," Xiaoxia Han et al. propose a framework for virtual hand haptic interaction, in which a virtual hand model simulates natural anatomy in its appearance, motion and deformation, and the feature of force feedback data gloves is reflected. In "Phone, Email and Video Interactions with Characters in an Epidemiology Game: Towards Authenticity," Muriel Ney et al. show how four challenges faced by the designer of authentic games have been addressed in a game for an undergraduate course used in a medical school. In "A Real-Time Interactive System for Facial Make-up of Peking Opera," Feilong Cai et al. offer a vector-based free form deformation (FFD) tool to edit patterns for facial make-up and, based on editing, the system creates automatically texture maps for a template head model. In "Design of Educational Games: A Literature Review," Ting Wei et al. point out the balance and integration between educational characters and playfulness in educational games, and suggest limited research areas in the existing literature and possible further research in educational game design.

The papers in this issue represent a large number of techniques and application examples of edutainment. These verify the potential and impact of digital entertainment technologies on the education and training domain. We would like to express our appreciation to all who contributed to this issue. They are the authors of the papers, the reviewers, and the International Program Committee members of Edutainment 2010 for recommending high-quality articles for this issue. Special thanks to Yi Li, Ruwei Yun and Qiaoyun Chen from the Editorial Office in Nanjing Normal University, and to Wei Wang in Northeast Normal University. They put a lot of effort into contacting authors, managing the reviewing process, checking the format of all papers, and co-collecting all material.

May 2010

Xiaopeng Zhang
Kevin Wong
Zhigeng Pan
Adrian David Cheok
Wolfgang Müller

LNCS Transactions on Edutainment

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

Editorial Office:

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

Table of Contents

The Study and Design of Adaptive Learning System Based on Fuzzy Set Theory

Bing Jia[1], Shaochun Zhong[1,2,3,4], Tianyang Zheng[1], and Zhiyong Liu[1,4]

[1] Ideal Institute of Information and Technology in NENU, JiLin ChangChun, 130024
[2] Engineering & Research Center of E-learning, JiLin ChangChun, 130024
[3] E-learning laboratory of Jilin Province, JiLin ChangChun, 130024
[4] School of Software in NorthEast Normal University (NENU), JiLin ChangChun, 130024
{Bing_jia,356981}@qq.com, sczhong@sina.com,
zhengty1984@hotmail.com

Abstract. Adaptive learning is an effective way to improve the learning outcomes, that is, the selection of learning content and presentation should be adapted to each learner's learning context, learning levels and learning ability. Adaptive Learning System (ALS) can provide effective support for adaptive learning. This paper proposes a new ALS based on fuzzy set theory. It can effectively estimate the learner's knowledge level by test according to learner's target. Then take the factors of learner's cognitive ability and preference into consideration to achieve self-organization and push plan of knowledge. This paper focuses on the design and implementation of domain model and user model in ALS. Experiments confirmed that the system providing adaptive content can effectively help learners to memory the content and improve their comprehension.

Keywords: Adaptive Learning System; fuzzy set theory; domain model; user model.

1 Introduction

Along with the development and mature of network, It is accepted and approved by people to carry on the study using E-learning system to accept day by day. But most of E-learning system [1] generally has some problems at present, such as Piled-up courses, mixed learning resources, lack of effective and sensitive response and feedback, lack of testing and evaluation system, lack of effective mechanisms for Guidance, low intelligence of the system. The raise of Adaptive Learning System [2] [3] is an effective solution about these matters. But most of intelligent E-learning system focus on the adoption of learner's behaviors, interests and habits to provide personalized E-learning services[4], and commonly neglect to consider if learner's ability are matched to the content and difficulty level of learning resources. Unsuitable resources may cause learner's cognitive overload or disorientation [5]. This paper proposes a new Adaptive Learning System (ALS) based on fuzzy set theory. It can effectively estimate the learner's knowledge level by test according to learner's target. Then take

Z. Pan et al. (Eds.): Transactions on Edutainment IV, LNCS 6250, pp. 1–11, 2010.

the factors of learner's cognitive ability and preference into consideration to achieve the organization and push plan of knowledge to meet with the need of the learners.

In the view of the above mentioned analysis, in this study, We analysis the theoretical basis of ALS in section 2, introduce the design of ALS based on fuzzy set theory in section 3, give the construction method of domain process of LM in section 4, explore the related technologies to achieve the model and the application in section 5 and summarize the study and forecast the future in section 6.

2 Related Work

2.1 Adaptive Learning System

In the fact, the core of adaptive learning thought was already discussed quite a lot before. It is a dream of educators and distance instructing experts to meet the individual demand and teach students according to individual character [6]. Adaptive learning system is to provide a learning support for individuals suitable for their individual characteristics about their differences in the learning process (due to people, due to time).

The system will integrate the theories and technologies such as artificial intelligence, network communication, item response theory and adaptive test all together. The system will be a real distance adaptive hypermedia learning system, which can organize and present the proper content most suitable for the student's current level of knowledge and ability according to the individual characteristics (the data comes from the student's cognitive model).

2.2 Fuzzy Set Theory

Fuzzy set theory [7] is the extension of conventional (crisp) set theory. It handles the concept of partial truth (truth values between 1 (completely true) and 0 (completely false)). It was introduced by Prof. Lotfi A. Zadeh of UC/Berkeley in 1965 as a mean to model the vagueness and ambiguity in complex systems.

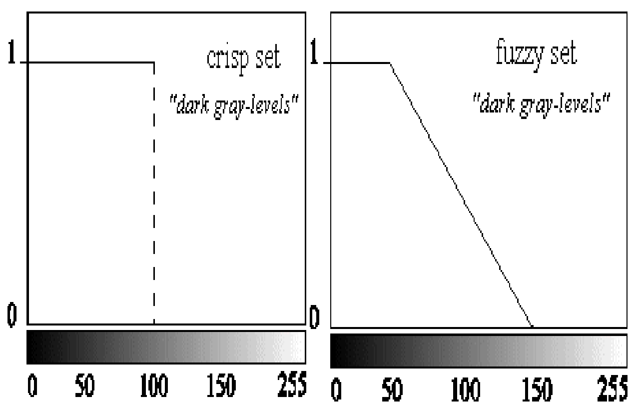


Fig. 1. Representation of "dark gray-levels" with a crisp and a fuzzy set

The idea of fuzzy sets is simple and natural. For instance, we want to define a set of gray levels that share the property dark. In classical set theory, we have to determine a threshold, say the gray level 100. All gray levels between 0 and 100 are element of this set; the others do not belong to the set (left image in Fig.1). But the darkness is a matter of degree. So, a fuzzy set can model this property much better. To define this set, we also need two thresholds, say gray levels 50 and 150. All gray levels that are less than 50 are the full member of the set, all gray levels that are greater than 150 are not the member of the set. The gray levels between 50 and 150, however, have a partial membership in the set (right image in Fig.1).

3 The Design of Adaptive Learning System Based on Fuzzy Set Theory

As the level of knowledge of learners inherently ambiguous, this paper presents the adaptive learning system based on fuzzy set theory to estimate the learner's level of knowledge, so as to provide the basis for the right resources to recommend. ALS is mainly constituted by two models (learner model, domain model), three processor (inference engine (IE), resources recommendation engine (RRE), information collection and extraction engine (ICEE)) and four databases (knowledge base, test base, resources base, and learner information base). System's main flow shows in fig.2.

3.1 Processors' Major Function Descriptions

- **Func 1:** Arrange the disorder KP in LKLS which educed by LM according to the relations of KP in DM. The requests of KPs are stored in knowledge library. Judges KP one by one, and deletes the KP that Meet the requirement from the set. Finally obtains KP sequence to be learning.
- **Func 2:** Select learning resources which correspond with the KP sequence. At the same time, define the match conditions according to learner’s cognitive ability and preferences to find the best-fit resources.
- **Func 3:** Record the information of learning process and testing process. Include log in information, learning history, testing history, Wrong question information, breakpoint information, etc. Through carrying on the mining and the extraction to this information, reappears for learner's characteristic information. Causes the learner characteristic information to realize the renewal.
- **Func 4:** Complete learner's Q/A demand, searches the corresponding answer through the key word match's way in the answer library.

Processors' Major Function Descriptions

- Step1: The new user needs to register. The learner’s personal information, academic information and other information are recorded in the learner info base.
- Step2: After user log in, need to choose learning goals if it is the first time the learner use the system. The learning objectives can be a chapter, a section or a knowledge point. Relations between them are included. That is, a chapter formed by a number of sections, and a section also formed by a number of knowledge points. Based on the learner’s goal, choose knowledge points as goal knowledge points set in the knowledge base under the goal, Express this set with target-knowledge.

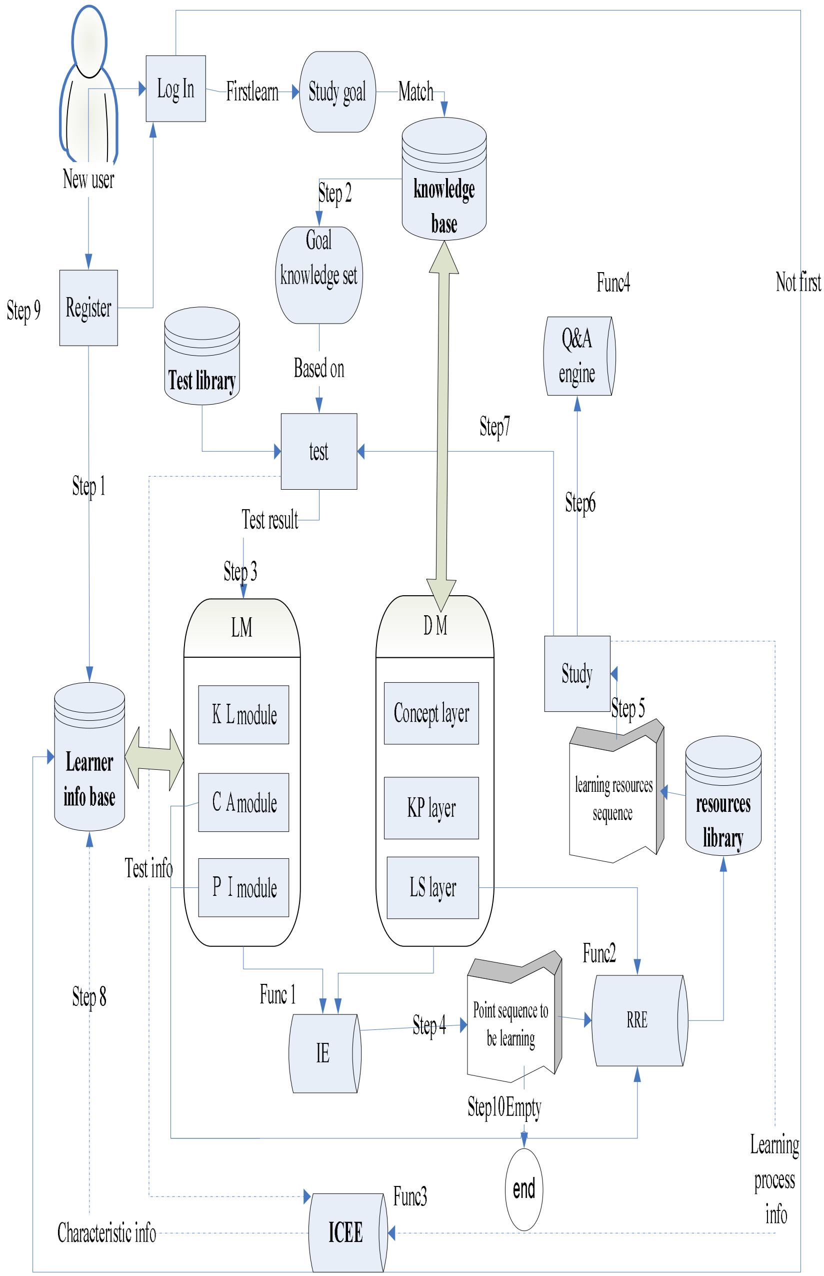


Fig. 2. ALS's main flow

- Step3: Design the test questions according to this set, and the test data need to feed back to LM.
- Step4: Educe the sequence of the knowledge points that to be learning for the learner from the relevant data between LM and DM.
- Step 5: The system shows the selected learning resources to users sequentially.
- Step6: Learners can get the answers from the Q&A engine when they are in trouble.
- Step7: Learners can test at any time when they are learning.
- Step8: Record learners' learning process information in learner information base.
- Step9: If the learner has used the system, it directly shows the resources that the learner should learn according to his learning record.
- Step10: If the sequence is empty, then end. Learners need to choose a further target.

4 The Construction of Domain Model and Learner Model

In this system, the domain model is the foundation, the learner model is the core and the foundation, all knowledge push plan is the knowledge structure which and the relations describes based on the domain model in acts according to the learner model learner's characteristic information formulation, to understand fully learner's status messages, the system carries on the judgment to learner's state-of-art and the cognitive capacity conforms to its ability level learning program according to the judgment result for its formulation, then conforms to its request resources form by chance according to the learner information for its push.

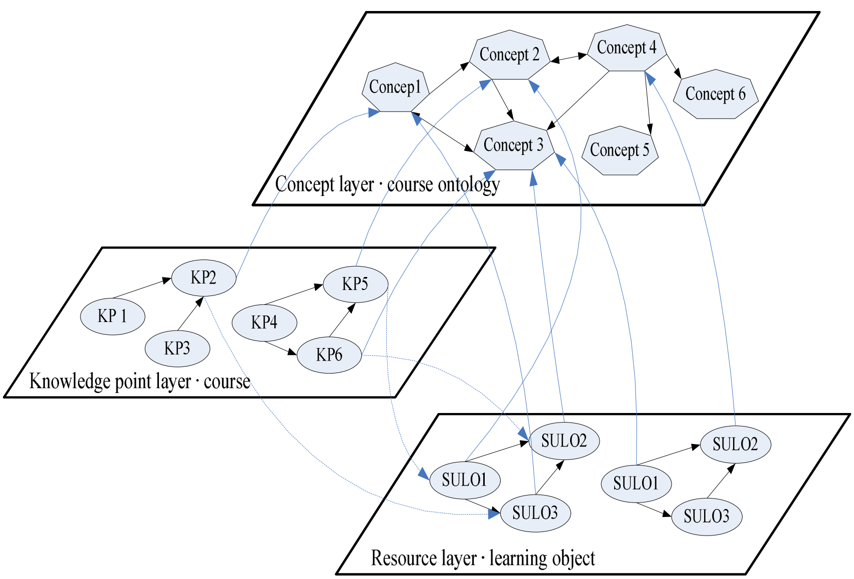


Fig. 3. DM based on ontology

4.1 Domain Model

The domain model may regard as is a professional field knowledge library, it is containing the fact knowledge and Yuan knowledge. How is the system auto-adapted height decided to a great extent by organizes the knowledge, the expression knowledge and the utilization knowledge.DM is a graphical representation of domain knowledge, and is an abstract of each constituent of domain knowledge[8][9] . As shown in Fig.2, the relations of learning objects are from the course ontology. System realization is divided into three tiers, namely concept tier, knowledge-point (KP) tier and resources tier. The concept tier in Fig.3 is the abstract description of course ontology, and the knowledge-point in the Figure 2 is an atomic unit, associated with learning objects, and cannot be separated. The resource tier presents the specific learning objects.

The relationship between knowledge-points depends on the relationship of course concept, In other words, the relationship between learning objects depends on course ontology, so we need to determine the course ontology and its semantic relationship first. To make this clearly, here, with the course "java programming language" as an example, we construct the curriculum ontology and its semantic relationship, as shown in Fig.4.

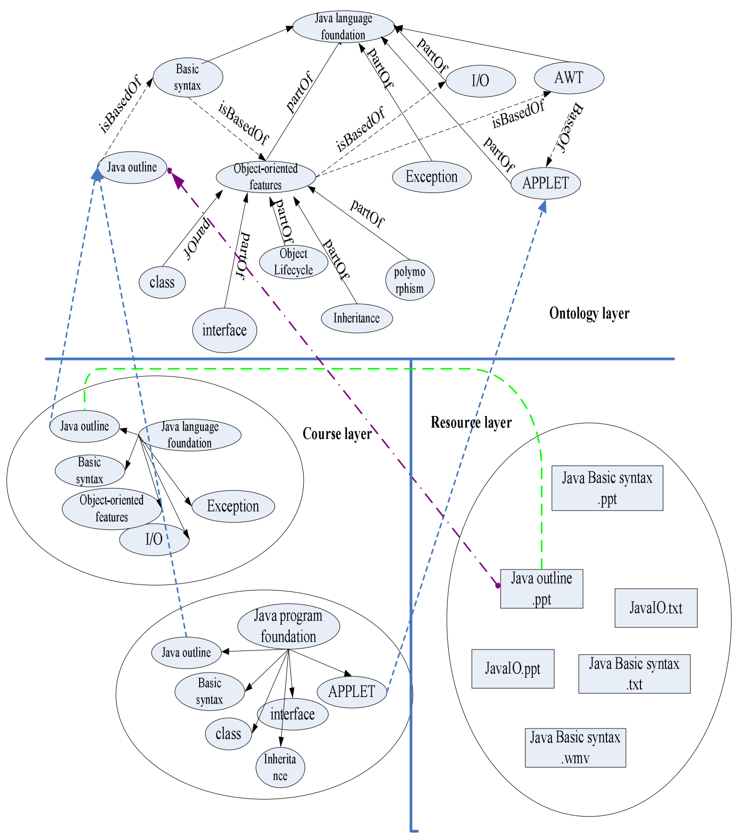


Fig. 4. Courses ontology

First identified courses ontology and their relationship, then construct the knowledge-point and their relationship, and ultimately linked the courses to specific learning resources.

4.2 Learner Model

Learner model refers to produce one reliable way of expression to demonstrate what the student understood and could do, what he does not understand and could not do, what he wants to do as well as he should do[10]. In this study, propose a new LM consist of Knowledge Level (KL) module, Cognitive Ability (CA) module and Preference Information (PI) module.

Knowledge Level (KL) Module

KL module can give a set of learner's KL based on the learner's goal through testing, that is, give "Knowledge-how" to IE. Use Knowledge-how to denote learner's knowledge level set (LKLS), The main function of the knowledge level module is to gain a set of learner's existing knowledge under a certain goal by test, use "Knowledge-how" to represent the set. Definition1: Knowledge-how ={(k1,h1),…,(ki,hi),…,(kn,hn)} where ki represents the ith knowledge point(KP), n represents the total of KPs, hi represents learner's mastering of the ith KP, hi∈H, H={0,1,2,3,4,5,6}, where "0" represents the situation that the learner have no idea of the KP, 1~6 respectively represents educational objectives proposed by Bloom[11] (knowledge, comprehension, application, analysis, synthesis , evaluation). We will introduce the steps to determine "Knowledge-how" in the follow.

- Obtains a set of learner's goal knowledge: Target-knowledge=$\{t_1, t_2,\dots,t_m\}$.
- Discovers the pre-KPs corresponding with the m KPs in the "Target-knowledge" set according to domain model. The set of pre-KPs: Pre-knowledge=$\{p_1,p_2,\dots,p_n\}$.
- Judges the learner's mastering of pre-KPs one by one (marked with 0~6). so the set of learner's pre-KPs level: Pre-Knowledge-how =$\{(p_1,h_1),\dots,(p_i,h_i),\dots,(p_n,h_n)\}$, $h_i \in H$.
- Records KPs whose mastering degree between 1 and 6 in "Knowledge-how". Set p_j as target-KPs and turn to step b when h_j=0.
- Achieves the disorder set "Knowledge-how".

But how to Estimate the degree of learner's mastering of a KP. We mainly introduce the application of fuzzy set theory in estimating the degree of learner's mastering of a certain KP [12]. According to Definition 1, we use these seven ranks (0~6) to express the degree of learner's mastering of a certain KP. The membership of each rank is expressed as μk(i), i=0~6. Express the degree of learner's mastering of a certain KP by fuzzy set K, so $K=\{i \mid (i, \mu_k(i)), i \in H\}$, where, $0 \le \mu_k(i) \le 1$, $\sum \mu_k(i) \le 1$. To a KP, the highest level of membership is the level that the learner has achieved currently about this KP.

The basic principle of changing rule: when learners do a certain rank of test question correctly, adopt up rule to the rank and the following, and adopt down rule to the above. When learner do test question wrongly, the membership does not change.

Table 1. Changing Rules of Membership

If	Rule	Type
i<m	μk(i)= μk(i)- μk(i)q+μk(i-1)q	up rule
i=m	μk(i)= μk(i)+ μk(i-1)q+ μk(i+1)q	up rule
i>m	μk(i)= μk(i)- μk(i)q+ μk(i+1)q	down rule

From the "down rule" in table 1, using "Δ" to represent the variation of $\mu_k(i)$, $\Delta=(\mu_k(i+1)-\mu_k(i))q$, so parameter "q" decides the scope of membership changing , It decides the declining speed of learner's mastering of KP. The determination of parameter "q" should consider the following two aspects: the difficulty of the question and the degree of learner's familiarity with the question. The difficulty of the question is expressed as "d", where 0<d<1.the degree of learner's familiarity with the question will be obtained by the speed of the learner to answer the question correctly. We stipulate two time sections: normal time and longest time. If the learner answer the question correctly in normal time range, indicate that he is familiar with the knowledge. If he answers the question correctly between the normal time and the longest time, indicate that he is not very familiar with it. If beyond the longest time, indicate that he is not familiar with it.

As the degree of learner's familiarity with the question is fuzzy, we use S membership function to express it. Because S function increases monotonically along with the increasing of x, but the degree of learner's familiarity with the question is descending along with time. We need to make t some small modifications to S function. So, assume that α is the normal time, γ is the longest time, β=(α+γ)/2.The function of the degree of learner's familiarity with the question F(t) is defined as: (where α, γ of each question can be different).

$$F(t)=1-S(t;\alpha,\beta,\gamma)=\begin{cases}1 & \text{if } t\le\alpha \\ 1-2(\frac{t-\partial}{\gamma-\alpha})^2 & \text{if } \alpha\le t\le\beta \\ 2(\frac{t-\gamma}{\gamma-\alpha})^2 & \text{if } \beta\le t\le\gamma \\ 0 & \text{if } \beta\le t\le\gamma\end{cases}$$

So that q = c^d * F (t), c is a constant larger than 1, 0<d<1. When the learner does the question correctly in the normal time range, F(t)=1, q become large. When the learner does the question correctly exceed the normal range, F(t) becomes small, q become small too. It proves that the learner's mastering degree of KP is changing slower When exceeds the longest time, F(t)=0, q=0. Similarly, q increases with increasing d. When the learner does a difficulty bigger question correctly, It proves he can learn this content well, also proves his mastering degree of KP is rising faster.

Cognitive Ability Module

Cognitive ability(CA) is individual's ability possessed while reconstructing and employing knowledge. We can describe the level of learners' ability by depicting the cognitive status of learners[13]. These abilities include induction ability, memory ability, observation ability, analysis ability, abstract ability, deductive ability, mathematic ability, association ability, imagination ability, logic reasoning ability and so on. We use Ability-level to represent the set of learner's CA, so,

Ability-level $=\{(a_1,l_1),\ldots,(a_i,l_i),\ldots,(a_n,l_n)\}$,where a_i is the i_{th} CA, l_i is the level of the i_{th} CA, n is the total of CAs. We can obtain the learner's CA by testing, using G(i) to express the ability type tested, T(i) to represent the question type, η to represent the coefficient of difficulty of question type, $0<\eta<1$, λ to represent the coefficient of difficulty of question, $0<\lambda<1$. φ to represent the correct answer to the question. Q to represent the content of the question. Then the form of every test is : Test(i) = (A(i),T(i),η,λ, Φ:Q), A series of test questions constituted a test paper. TEST= $\{t_1=(A_1,T_1,\eta_1,\lambda_1,\Phi_1:Q_1),\ldots,tn=(A_n,T_n,\eta_n,\lambda_n,\Phi_n:Q_n)\}$. For example, in order to obtain the level of induction ability (G_1) and memory ability(G_2). The series of test questions are as follows:

Test(1)= (A(1), T_1, η_1, λ_1, $\Phi_1 : Q_1$).
Test(2)= (A(1), T_2, η_2, λ_2, $\Phi_2 : Q_2$).
Test(3)= (A(2), T_3, η_3, λ_3, $\Phi_3 : Q_3$).
Test(4)= (A(2), T_4, η_4, λ_4, $\Phi_4 : Q_4$).

Then obtain the answers of the test questions. The set of answers : Answer =Answer $=\{answer(t_1, \varphi'_1), \cdots, answer(t_i, \varphi'_i), \cdots, answer(t_n, \varphi'_n)\}$.So, we can definite the level of the i_{th} CA as follow: Ability-how (A_i)=

$$\sum_{j=1}^{n}\left[\left(\phi_j'\cap\phi_j \middle/ \phi_j\right)\times\eta_j\times\lambda_j\right] / \sum_{j=1}^{n}(\eta_j\times\lambda_j)$$

where $(\varphi'_j\cap\varphi_j)/\varphi_j$ is the accuracy rate of the learner's answers to the j_{th} question, using R_j to express the accuracy rate, so $0\le R_j\le 1$. n is the total of the questions which the student replied to test the i_{th} kind of ability. Let l_i =Ability-how (A_i), and we can obtain the set of learner's CA "Ability-level".

Preferences Module

Preferences are defined as demonstrated interest, hobby and other information in learning. Use a five-tuple "Preferences-set" to represent it , so Preferences-set=$<P_b, P_p, P_s, P_t, P_f>$,where P_b represents background material preferences(sport, entertainment, education, etc.), P_p represents resource presentation preferences (text, audio, animation, video), P_s represents learning strategies preferences(teaching type, inquisition type, cooperation type), P_t represents study time preferences(time period), P_f represents system function preferences(using the system frequency-related features). P_b , P_p, P_s, P_t, P_f are five concept sets , A concept is a pair of (c, σ),where c is a concept ,and σ is a coefficient denotes the preference degree of the concept. The system identifies the similarity of learning resources among learner's preferences according to the angle between the characteristics vector and the vector space . The most similar learning resources will be returned to the learner.

5 Technologies and Applications

The system is implemented in MVC mode. It mainly uses JAVABEAN to enclose business logical layer. The presentation of interface adopts JSP and XML technology.

The database is SQL server. Besides above-mentioned technologies, the system also consults JavaScript, DOM and others. The used software includes Jena, JBuilder, Dreamweaver, SQL server etc[14]. This work has been applied in junior high school for assisted learning. Through experiments it proves to be an efficient method for adaptive learning.

6 Conclusion

Adaptive learning is an effective way to improve the learning outcomes, that is, the selection of learning content and presentation should be adapted to each learner's learning context, learning levels and learning ability. Adaptive Learning System (ALS) can provide effective support for adaptive learning. But most of intelligent E-learning system focus on the adoption of learner’s behaviors, interests and habits to provide personalized e-learning services, and commonly neglect to consider if learner’s ability are matched to the content and difficulty level of learning resources. Unsuitable resources may cause learner’s cognitive overload or disorientation. In this paper, This paper proposes a new ALS based on fuzzy set theory and detailed description system's flow and main function. ALS can effectively estimate the learner’s knowledge level by test according to learner’s target. Then take the factors of learner’s cognitive ability and preference into consideration to achieve self-organization and push plan of knowledge. This paper also focuses on the design and implementation of domain model and user model in ALS. In addition, this system has the following two points to study further:1) The system in serves for the user in the process, needs the user to supply many authentic information, therefore when the system realizes must consider the security and the secrecy fully 2）In order to enable the learner model to achieve from the renewal, needs to choose a more appropriate algorithm to carry on the collection and the excavation to the learner learning process information, with the aim of reappearing for learner's characteristic, thus satisfies system's compatibility.

Acknowledgment

We would like to thank professor Shaochun Zhong for his insightful comments in reviewing drafts of this paper. We also would like to thank National Science and Technology Innovation Fund(04c2622211081) for the sponsor to our study.

References

1. Xian, J.F., Zhao, X.X., Zhong, Y.Z.: The Status Quo and Trend of Development of Multimedia Technique. Journal of Guangxi Radio and TV University 13(3), 10–12 (2002)
2. Wang, H.B., Zhu, Y.P., Wang, J.: Design and Development of the Adaptive Testing System Base-on Internet. Research and Exploration in Laboratory 24(5) (2005)
3. Liu, H., Cao, H.J.: Adaptive Learning System Research and Design. Wuhan: Software Guide (6), 39–41 (2006)

4. Fu, X., Budzik, J., Hammond, J.K.: Mining navigation history for recommendation. In: Proceedings of the 2000 international conference on intelligent user interfaces, New Orleans, pp. 106–112 (2000)
5. Chen, C.M., Duh, L.J.: Personalized Web-based Tutoring System based on Fuzzy Item Response Theory. Expert Systems with Applications 34(4), 2298–2315 (2008)
6. Yu, S., He, K.: Reserch on the Adaptive Learning System Based on Internet. In: Proceeding of 35th Annual Conference of Frontiers in Education Conference 2005, vol. (3), pp. 18–23 (2005)
7. Wikipedia, the free encyclopedia, `http://en.wikipedia.org/wiki/Fuzzy_set`
8. Zhao, W., Liu, X., Qiu, B.-S.: Construction of Domain Ontology in Adaptive Learning System Based on Semantic Web. Journal of Jilin University(Information Science Edition), 514–518 (2008)
9. Zhang, F., Zhou, X.: The field of Knowledge Representation and UML modeling. The Journal of Computer Engineering and Design (03) (2004)
10. Lai, X., Wu, Y., Cheng, W.: The Design and Implementation of Multimedia Intelligent CAI in Mining. Computer Systems & Applications (02) (1997)
11. Bloom., B.S., Engelhart., M.D., Furst., E.J., Hill., W.H., Krathwohl, D.R.: Taxonomy of Educational Objectives. In: handbook I, Cognitive domain, London, Longman (1956)
12. Zhuang, Z.: Estimation method of student's cognitive level in self-adaptive distance tutoring system. Beijing: Computer Engineering and Applications 43(3), 220–223 (2007)
13. Tang, S.Q., Zhong, Z.: Evaluation of Cognition of Students. Nanning: Journal of Academy of Sciences 18(4), 281–284 (2002)
14. Zhong, Y.J., Liu, J., Zhong, S.C., Zhang, Y.M., Cheng, X.C.: Programming of informatized instrctional design platform for physics. In: Proceedings of Technologies for E-learning and Digital Entertainment, pp. 171–177 (2006)

Modeling Personalized Learning Styles in a Web-Based Learning System

Chia-Cheng Hsu, Kun-Te Wang, and Yueh-Min Huang

Department of Engineering Science, National Cheng Kung University, Taiwan
chia.cheng.hsu.tw@gmail.com, taito@easylearn.org,
huang@mail.ncku.edu.tw

Abstract. An innovative learning mechanism for identifying learners' learning styles to improve adaptive learning is proposed. Hypermedia-learning tools are highly interactive to learners in web-based environments that have become increasingly popular in the field of education. However, these learning tools are frequently inadequate for individualize learning because accessing adaptive learning content is required for learners to achieve objectives. For predicating adaptive learning, a neuron-fuzzy inference approach is used to model the diagnosis of learning styles. Then, according to the diagnosis results, a recommendation model is constructed to help learners obtain adaptive digital content. The proposed approach has the capability of tracking learning activities on-line to correspond with learning styles. The results show that the identified model successfully classified 102 learners into groups based on learning style. The implemented learning mechanism produced a clear learning guide for learning activities, which can help an advanced learning system retrieve a well-structure learning unit.

Keywords: Adaptive Learning, Learner Diagnosis, Neuron-fuzzy, Learning Style.

1 Introduction

Online learning technology has changed the traditional way of teaching and learning. Since learners control the hypermedia-learning tools to access rich digital content, adaptive learning has become learner-oriented [1, 2]. However, with the easy access to information in web-based learning, information overload can be a problem. Therefore, teachers should create adaptive content, and the learning systems should index the recommended content and provide suggestions to learners.

Carroll proposed a model to account for school learning [3]. His major premise was that school learning is a function of time. To be more specific, Carroll proposed the degree of learning f = (time spent / time needed). Kolb described the learning cycle model [4]. People learn in various ways, so introduced experiential learning theory (ELT) and learning styles inventory (LSI). The LSI is a 12-item assessment tool that identifies four learning styles based on ELT: (1) diverging, (2) assimilating, (3) converging, and (4) accommodating, as shown in Fig. 1.

Z. Pan et al. (Eds.): Transactions on Edutainment IV, LNCS 6250, pp. 12–21, 2010.

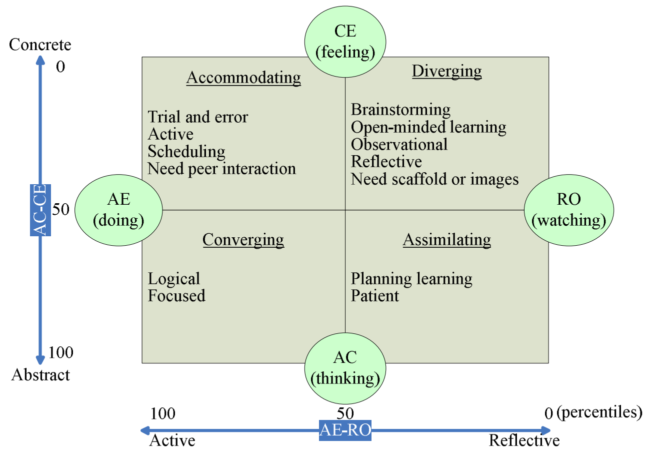


Fig. 1. Features of learning styles according to Kolb's LSI model

The present study proposes that the learning styles are closely correlated to their learning activities (e.g., learning time). In the analysis, the factors of learning styles from LSI are a function of variables such as learning schedule, quiz, discussion interaction, and number of reviews. With these linguistic variables, a fuzzy inference system (FIS) is used to develop a network for identifying the learning style of e-learners using LSI [5]. The recommended learning content is provided to learners by the recommendation model. The content is packaged as a learning object (LO) according to the Sharable Content Object Reference Model (SCORM) specification [6].

The learning mechanism uses the diagnosis approach to engage learners in adaptive learning. The mechanism has two stages: (1) the preprocessing of an inquiry into learning activities and (2) the delivery of categorized content from the recommendation model. This study confirms that the proposed mechanism is feasible and practicable.

2 Method

The proposed learning controls are shown in Fig. 2. The learning style of e-learners is identified using the FIS approach and the association control approach is used for providing recommended content to learners.

2.1 Mechanism of Learning Controls

Learners have many learning activities in web-based courses. Their learning activities proceed from LO to LO in a learning sequence created by learning management system (LMS). The completed activities are monitored over time by a specific agent. Then, the fuzzy inference rules and neuron network are implemented in the model to determine the learning style. In the proposed learning mechanism, if the learner cannot pass the given threshold, the recommended LO is delivered. Therefore, the learner is assisted by the model.

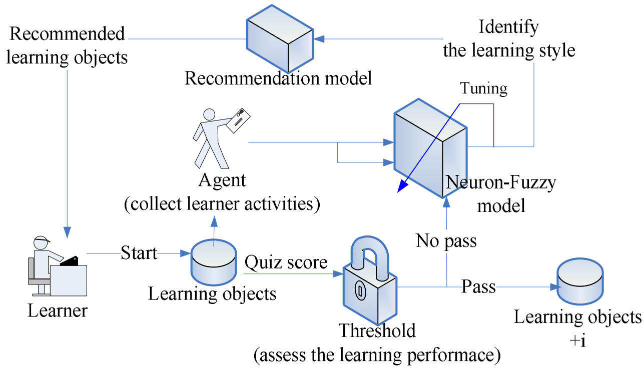


Fig. 2. Mechanism of learning controls consists of a neuron-fuzzy model and a recommendation model in a web-based learning environment

2.2 Fuzzy Inference Rules for Learning Activities

In order to predict the learner's activities, the records are transformed to a fuzzy set *B*, where each response is timed and recorded, as shown in Fig. 3. The responses are learning time and frequency on a course, activities on forums, quiz reaction speed, and frequency of topic review.

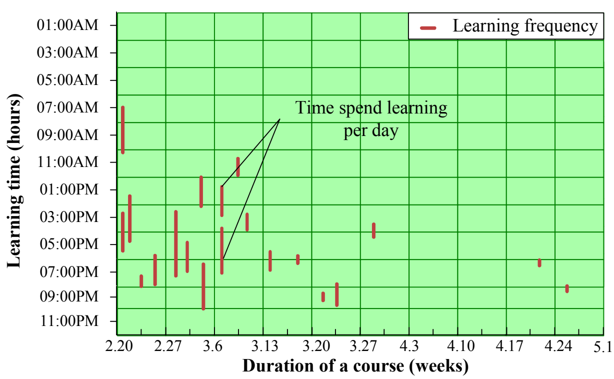


Fig. 3. Learning activities for an active learner

Depending on these characteristics, the records are given a linguistic value *v* as follows: learning schedule (learning time and frequency) = {always, sometimes, seldom}, forum participation = {infrequent, normal, frequent}, quiz reaction speed = {fast, normal, slow}, and frequency of topic review = {often, sometimes, seldom}. The value of evaluated activities is summarized into *k* groups. Then, each $B_i = \{v_1, v_2, \ldots, v_m\}$ has a *v* term that consists of *m* linguistic values. The set $T(B_i) = \{B_1, B_2, \ldots, B_k\}$, where B_i $(i = 1, 2, \ldots, k)$ represents a sentence describing the *i*th type of response of a learner's activities. For example, consider the set $T(B_i)$ = *T*(time on course) = {short, normal, long}, where the linguistic variable $m = 3$. In our study, four fuzzy sets B_i ($i = 1, \ldots, 4$) and the terms v_m ($m = 1, \ldots, 3$) are used to formulate the input to the identifying process.

The fuzzy variables and IF-THEN rules are used to infer the consequences of the four styles of LSI.

$$\begin{aligned}
&IF\ B_1\ is\ v_{11}\ AND\ B_1\ is\ v_{12}\ AND \ldots B_1\ is\ v_{1m}\ AND \\
&B_2\ is\ v_{21}\ AND\ B_2\ is\ v_{22}\ AND \ldots B_2\ is\ v_{2m}\ AND \\
&\ldots \\
&B_k\ is\ v_{k1}\ AND\ B_k\ is\ v_{k2}\ AND \ldots B_k\ is\ v_{km}\ AND \\
&THEN\ Y_1\ is\ CE\ AND\ Y_2\ is\ AC\ AND\ Y_3\ is\ RO\ AND\ Y_4\ is\ AE
\end{aligned}$$

2.3 Tuning Learning Styles Using FIS Model

The fuzzy inference rules are used to construct a neural network for identifying the relationships between learning activities and learning styles. Based on the output, the desired data D_i ($i = 1, \ldots, 4$) in training the set represents the results of LSI (CE, AC, RO and AE) $T(D_i) = \{D_1, D_2, \ldots, D_k\}$ is used to provide supervised data in the identification network. If the outputs of the network Y_i are not the desired variables D_i, the network performs backward calculations to minimize the error function $E = \frac{1}{2}(Y - D)^2$ for tuning the results of LSI. As shown in Fig. 4, the fuzzy logic system with a singleton fuzzifier, a max-min composition, and a center average deffuzzifier uses the following equations (1) and (2):

$$Y = \frac{\sum_{l=1}^{k} \overline{y}_l \{\prod_{i=1}^{n} \exp[-(\frac{x_i - \overline{x}_{li}}{\sigma_{li}})^2]\}}{\sum_{l=1}^{k} \{\prod_{i=1}^{n} \exp[-(\frac{x_i - \overline{x}_{li}}{\sigma_{li}})^2]\}} \quad (1)$$

where $\overline{y}_l$ is the center of membership function B_l, σ_{li} is inflection point, and $\overline{x}_{li}$ is the highest point of membership function B_l.

$$Z_e = \exp[-(\frac{x_i - \overline{x}_{li}}{\sigma_{li}})^2] \quad (1.1)$$

$$a = \sum_{l=1}^{k} \overline{y}_l Z_e \quad (1.2)$$

$$b = \sum_{l=1}^{k} Z_e \quad (1.3)$$

The Gaussian membership function is:

$$u_{Ali}(x_i) = \exp[-(\frac{x_i - \overline{x}}{\sigma_{li}})^2] \tag{2}$$

where $\overline{x}$ is the center of x_i.

Therefore, the FIS network has to adjust three variables, $\overline{y}_l, \sigma_{li}, \overline{x}_{li}$ to minimize the error function E. In addition, the network in Fig. 5 uses the steepest descent algorithm given by:

$$\begin{aligned} \overline{y}_l(k+1) &= \overline{y}_l(k) - \alpha \frac{\partial E}{\partial \overline{y}_l}\Big|_k \\ \overline{y}_l(k+1) &= \overline{y}_l(k) - \alpha \frac{(Y-D)}{b} Z_l \Big|_k \end{aligned} \tag{3}$$

$$\begin{aligned} \overline{x}_{li}(k+1) &= \overline{x}_{li}(k) - \alpha \frac{\partial E}{\partial \overline{x}_{li}}\Big|_k \\ \overline{x}_{li}(k+1) &= \overline{x}_{li}(k) - \alpha (Y-D) \cdot \frac{\overline{y}_l - Y}{b} \cdot Z_e \frac{2(x_i - \overline{x}_{li})}{\sigma_{li}^2}\Big|_k \end{aligned} \tag{4}$$

$$\begin{aligned} \sigma_{li}(k+1) &= \sigma_{li}(k) - \alpha \frac{\partial E}{\partial \sigma_{li}}\Big|_k \\ \sigma_{li}(k+1) &= \sigma_{li}(k) - \alpha (Y-D) \cdot \frac{\overline{y}_l - Y}{b} \cdot Z_e \frac{2(x_i - \overline{x}_{li})^2}{\sigma_{li}^3}\Big|_k \end{aligned} \tag{5}$$

Updating these variables using equations (3), (4), and (5) adjusts the neural-fuzzy network to evaluate learner's activities. The membership function shown as equation (1.1) describes a revisable learning curve. The learning activities can thus be identified without a questionnaire.

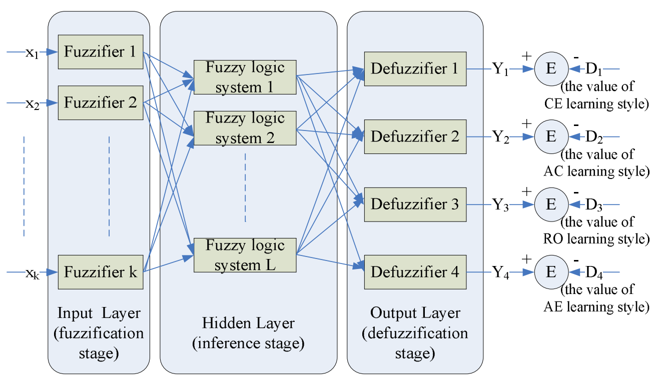


Fig. 4. Schematic of the inference model used to identify learning styles (CE, AC, RO, and AE)

2.4 Constructing the Ontology of Recommended Learning Content

After the learning style is determined, the recommendation model is triggered when a learner does not pass the threshold score set by the teacher. The recommendation model finds the relationships between learning style and LO on the structured ontology, as shown in Fig. 5. The ontology consists of two domains of knowledge: the learning style domain and the LO domain. LOs are represented as an aggregation that consists of learning documents, knowledge forums, quizzes, multimedia materials, examples, and lectures. There is an association between the domains of learning style and LO that is used to provide machine-assessable semantic content.

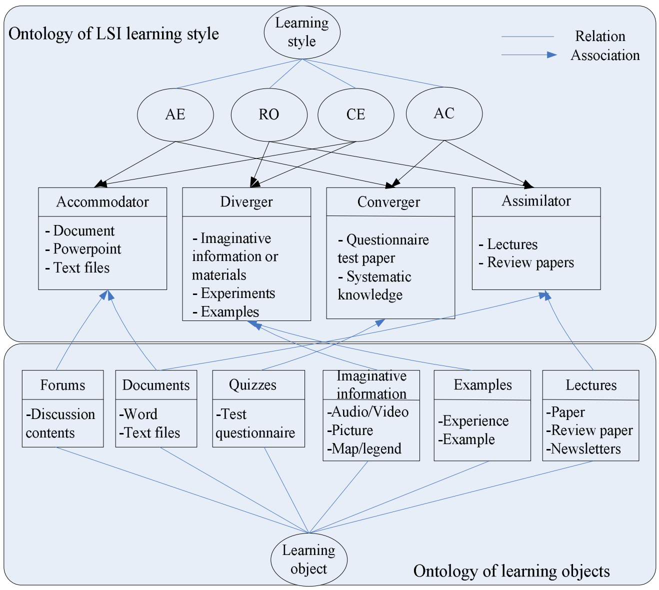


Fig. 5. Ontology of learning style and learning object is used to form the interoperable associational relationship of content

3 Experimental Results and Discussions

3.1 Monitoring of Learning Activities in Web-Based Learning

The web-based learning environment consists of a learning portal and the learning control mechanism described in Section 2. The course is developed according to the specification of SCORM, as shown in Fig. 6. Learners then do the selected activities and the learning agent records completed activities.

3.2 Tailoring the Neuron-Fuzzy Model

In order to obtain correlation between learning style and learning activities, 102 learners finished two questionnaires. First, in the experiments, the learning style of the 102

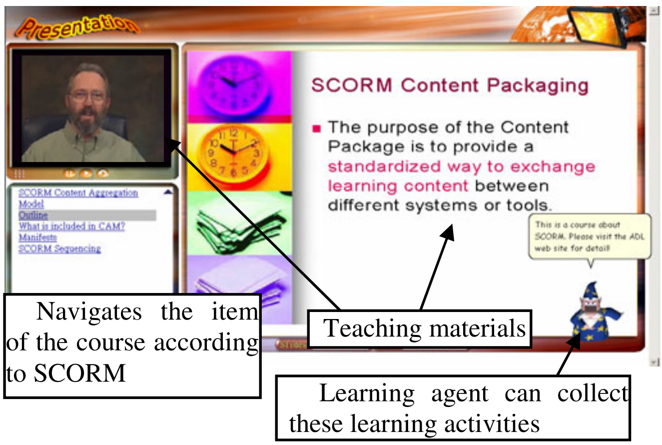


Fig. 6. Navigation of a SCORM-based course based in web-based learning environment, and the collection of learning activities by a learning agent

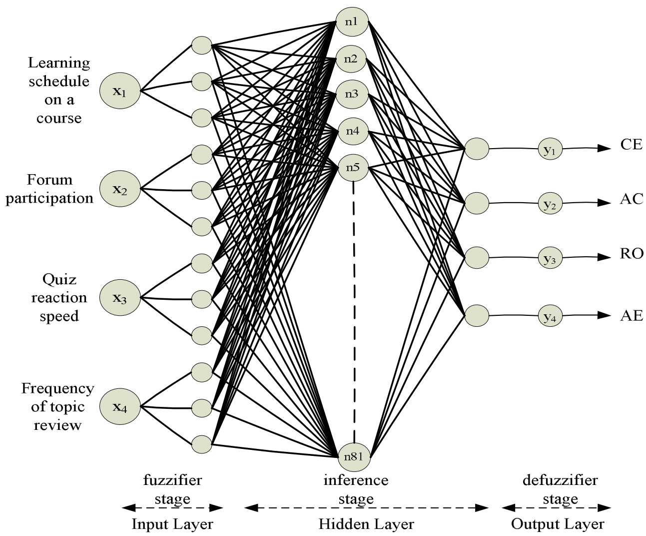


Fig. 7. Network implemented to identify the learning style by monitoring the four types of learning activity

learners was surveyed using LSI. Then, the learners completed an activities preference questionnaire (APQ). The APQ is an 8-item assessment tool that understands the activities of a learner based on four modes: learning schedule, quiz, discussion interaction, and number of reviews. The samples using questionnaire were regarded as a training set in the database to supervise the output of the network to adjust the error function *E* (see Section 2.3). The implemented neuron-based fuzzy model is shown in Fig. 7. The model has four inputs x_1, x_2, x_3, and x_4; each input has three linguistic variables in the input layer with 81 (i.e. 3x3x3x3) possible combinations of rules. The output of the identification process was described with four linguistic values $T(y_i)$ = {CO, AC, RO, AE} (see Fig. 1). The membership function was bound to $0 \le x_1, x_2, x_3, x_4 \le 1$.

The training data were substituted into the neuron-fuzzy network, which was run 150 times for tuning network parameters that related to the membership function and fuzzy model. The results of the experiment showed that the convergence of network for the times of training was increase but the root mean squared error was gradually decrease about 100 training times. The error is minimized between 0.1 and -0.1 with the real data $T(D_i)$, as shown in Fig. 8.

The objective at this stage is to adjust the parameters of the initial model, via a training procedure, for best performance. Of note, the initial model structure was not optimized. The parameters need to be optimal in the sense that the antecedent Gaussian membership functions have a uniform standard deviation, and then the parameters are only optimal with respect to the membership functions using an FIS.

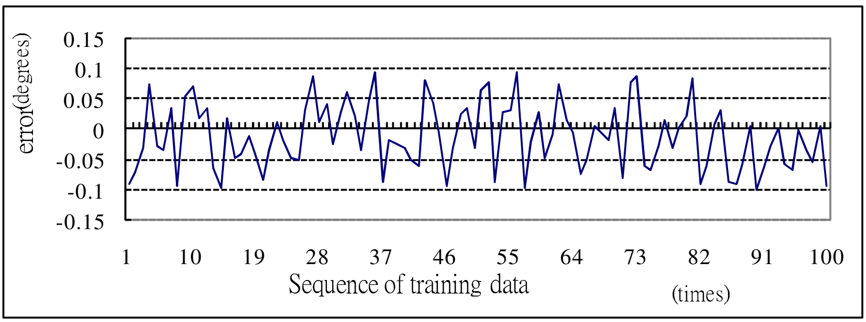


Fig. 8. Error rate of training network between -0.1 and 0.1

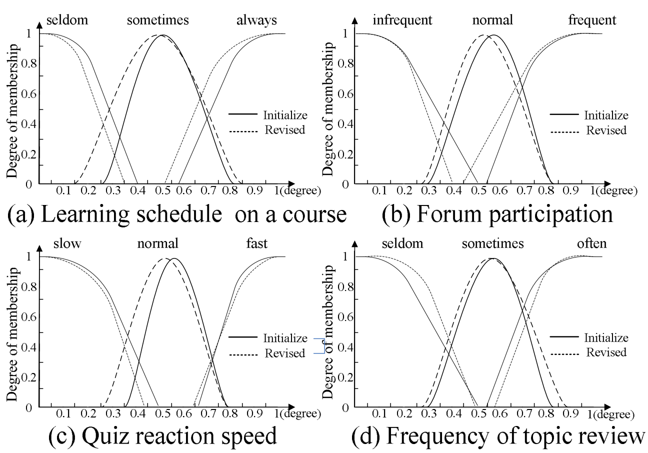


Fig. 9. Initialized and revised membership functions for four linguistic variables after the training of the FIS network

3.3 Evaluating the Neuron-Fuzzy Model

In order to evaluate the performance of the rule-based neuron-fuzzy model, three test sets, each with 34 sampled learners with predefined linguistic values in the linguistic variables of their learning activities, and with predefined membership degrees for these values, were generated. Then, the input values $X = \{x_1, x_2, x_3\}$ of each pattern of the test sets were processed by the trained neuron-fuzzy model and classified into one of the linguistic values of the term set {CE, AC, RO, AE}. In addition, the three sets of learners were classified into one of the linguistic values of the term set {CE, AC, RO, AE} by a group of experts. The group classifications were compared against the neuron-fuzzy model classifications. The overall average in diagnosis accuracy was 90%; it was 93%, 91%, and 86% for each of the three data sets, respectively.

3.4 Evaluating the Quality of the Learning Mechanism

In order to assess the proposed approach when the learning style is available, the results of pretests and posttests were used to evaluate learning performance, which contains test scores for three subjects. Fig. 10 shows the progress for each subject. The results indicate that the learning mechanism improves self-learning.

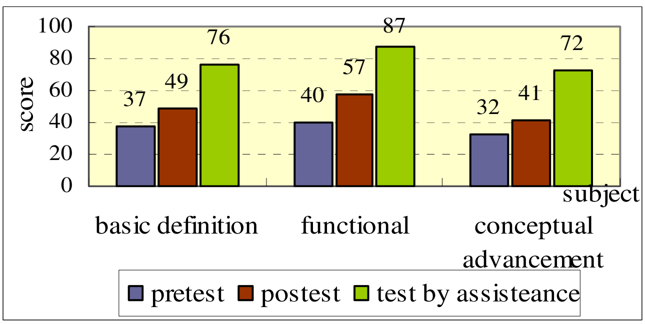


Fig. 10. Performance of e-learners for pretest, posttest, and test

4 Conclusion

Although adaptive learning has become an important issue that researchers have found individuals learned with ease on style when the learning environment is compatible with their learning style [7]; however, most of the teaching strategies do not consider accommodating the learning style when teachers produced a curriculum.

A neuron-fuzzy model for the identification process was proposed for determining the learning rules of learners. The method is convenient for diagnosis.

The learning mechanism can be integrated into a learning management system for each learner. The mechanism clusters learners into groups based on learning style, and recommends learning content for each group. The learning system does not need to evaluate the LSI. In future work, the learning actives can be measured more accurately to develop more individualized computer-assisted learning.

Acknowledgements. This work was partially supported by the National Science Council of Taiwan (R. O. C.) under grant NSC 97-2511-S-006-001-MY3 and NSC 98-2631-S-024-001.

References

1. Felder, R.M., Brent, R.: Navigating the bumpy road to student-centered instruction. College Teaching 44(2), 43–47 (1996)
2. Razzaq, L., Heffernan, N.: Towards designing a user-adaptive web-based e-learning system. In: Extended Abstracts Proceeding of the 2008 Conference on Human Factors in Computing Systems, pp. 3525–3530. ACM, Florence (2008)
3. Berliner, D.C.: Telling the stories of educational psychology. Educational Psychologist 27, 143–161 (1992)
4. Smith, M.K.: David A. Kolb on experiential learning. The Encyclopedia of Informal Education (2001)
5. Stathacopoulou, R., Grigoriadou, M., Samarakou, M., Mitropoulos, D.: Monitoring students' actions and using teachers' expertise in implementing and evaluating the neural network-based fuzzy diagnostic model. Expert System with Applications 32, 975–995 (2007)
6. Advance Distributed Learning (ADL): SCORM specification: Shareable Content Objects Reference Model (2004)
7. Wang, T.I., Wang, K.T., Huang, Y.M.: Using a style-based ant colony system for adaptive learning. Expert Systems with Applications 34(4), 2449–2464 (2008)

An Emotional Agent in Virtual Learning Environment

Ailiya[1], Zhiqi Shen[2], and Chunyan Miao[1]

[1] School of Computer Engineering, [2] School of Electrical Electronic Engineering,
Nanyang Technological University, Singapore
{ai0001ya,zqshen,ascymiao}@ntu.edu.sg

Abstract. Integrating emotional agent into virtual learning environment is a flourishing research topic to assist the pedagogical applications. However, most of existing emotional agents still lack of a straightforward and understandable computational strategy for both educational experts and developers. In this paper, we integrate an emotional agent into a pedagogical project which uses OCC-based rules to do cognitive appraisal and involves Fuzzy Cognitive Map to do the numerical computation. In this way, the system can not only provide agents flexible reactions in both qualitative and quantitative aspects, but also bring designers an easy-to-use approach.

Keywords: Emotional Agent, Virtual Learning Environment, DINO project, OCC-Based rules, Fuzzy Cognitive map.

1 Introduction

Emotion plays a very important role in people's everyday life. It influences our attention persisting, memory retrieving, decision making, as well as problem solving. Specifically from the perspective of education, emotion also influences the entire learning process. The emotionality of instructors can affect their teaching strategies so that they may provide students different cognitive scaffoldings. Moreover, the interaction quality between learners and instructors/pedagogical systems highly depends on the involved emotional factors. Since education does not merely focus on the final teaching results but the entire teaching process, we should enhance the agent-augmented pedagogical system with emotional agents.

In several existing pedagogical systems [1-3], emotional agent has already been applied. However, these systems still have limitations such as the one-size-fits-all design and the lack of precise computational approach to provide flexible reactions and protean behaviors for agents. To solve these problems, this paper proposes an emotional pedagogical agent which can qualitatively elicit various emotions and quantitatively do the computation in a simple and straightforward approach.

Typical emotional behaviors of a pedagogical agent in a virtual learning environment consist of two categories. One is related to agent's internal feeling due to its own perceptions over environment changes, and the other is related to agent's sympathetic emotions to learners as a feedback of student's performance. The two categories functionally serve different parties, but their emotion eliciting process can be unified. To elicit the emotions, first we use OCC-based rules to analyze conditions and do the

Z. Pan et al. (Eds.): Transactions on Edutainment IV, LNCS 6250, pp. 22–33, 2010.

qualitative cognitive appraisal, and then we use Fuzzy Cognitive Map (FCM) [4] to represent the causal relations and do the quantitative computation. This mechanism can qualitatively analyze what type of emotion needs to be generated, and quantitatively compute the specific intensity of emotions under a certain situation. With the help of this emotional approach, the emotional agent will work as an emotionally believable learning companion for learners during a learning process.

The rest of this paper is organized as follows. In Section 2, we first introduce a pedagogical system developed by our research group called "DINO" and discuss why we need to involve emotional agents in this project. Section 3 specifically describes the emotion qualitative elicitation strategy in our system, and the corresponding quantitative computational approach is presented in Section 4. Section 5 briefly explains how an emotional agent based on the approaches in Section 3 and 4 is integrated into the DINO project. The final section concludes this paper and discusses some future work.

2 Related Work and the Project DINO

Concerning the generation of emotions in agent system, many researchers have been looking for answers from the psychological models of emotions. These models normally use the cognitive conditions to elicit emotions such as the Ortony Clore Collins (OCC) model [5], the Roseman's cognitive model [6], the Plutchik Robert's three-dimensional circumplex model [7], and Russell's circumplex model [8]. As OCC provides a more concise classification of the cognitive generation of emotions, and is easier to be applied in the virtual environment, OCC is more common to be used in this field. However, with all the advantages of OCC it is still hard to be directly integrated in agent design, because there is a gap between the real operation of agent system and the psychological theories.

To bridge this gap, many extensions and applications emerged. In summary, there are five types of extensions to the OCC model: 1) the rule-based emotion inference [9] which directly transforms the elicitation process of OCC into hard coded rules to control agent, 2) the fuzzy logic approach [10] to map elements in OCC, 3) the dynamic decision networks [11] to depict the relationships in OCC, 4) the PAD (Pleasure-Arousal-Dominance) emotional state [12] to measure emotions, and 5) the logical formulization [13] of OCC. Except the first work, the rest research work all can do the quantitative computation, and they used different ways to represent OCC model. However, these models do not have a straightforward representation for educational designers, and if the design has some mistakes it is hard to find out where the problem is. As the research of emotional pedagogical agents is related to cross-disciplinary experts of education and computer science, an easy understanding and simple use of models are very important. For this concern, we propose using FCM to represent OCC model and realize the elicitation of agents' emotions. Its straightforward formulation and simple computation can meet the designing requirements of a pedagogical agent.

Our proposed approach will be used in a virtual learning environment, the Dino project. The full name of this project is "Serious Immersion and Embodied Learning: Traces of Dinosaurs in Earth System Science"[14]. The main purpose of this project is

Fig. 1. Screen shots of Dino project

to conduct a research and develop an interactive digital media to support learning of geography for secondary school students in Singapore. In the form of an immersive game, interactive virtual worlds with dinosaur agents are created for users to explore. The game is in a 3D virtual world and provides first-person-view for students. A screen shot is shown in Fig. 1.

The learners can enter the present virtual world to collect dinosaur fossils and traverse the "time tunnel" to past world to see the ancient creatures and play with dinosaurs. During performing a series of tasks, students can learn earth knowledge in an open-ended problem-solving space.

In order to trigger students' learning interest and make them engage more time in the earth knowledge study, a learning companion is involved which is an emotional agent as a dinosaur, to interact with students. The emotional reactions of the dinosaur include its own emotional reactions according to the environment (such as fear for the earthquake) and having sympathy for the student (such as feel happy when the student dose well and feel sorry when does not). The benefits of involving emotional agents as a learning companion include three aspects:

1. First, the interesting and attractive emotional responses of the agent can increase the believability of the virtual world and make students have more immersive experiences which attract their interests and bring them deeper remembrances.
2. Secondly, when the emotional agent expresses sympathetic reactions to students, the students can have the feeling of being cared and being thought of. Such experience will encourage them to inquire and explore.
3. Thirdly, since the learning companion usually comes with a cartoon appearance, when it has fancy emotional expressions students may be easily released from learning burden and be more willing to learn.

The details of how to build the emotion generating system will be introduced in the following two sections.

3 OCC-Based Emotion Elicitation

In this paper, we are particularly interested in OCC model, which concisely depicts the classification of emotions under different eliciting conditions. According to the cognitive structure of OCC, we will analyze this process with a rule-based formulization. After this qualitative analysis, we will continue to do the quantitative computations in Section 4.

3.1 Rule-Based Emotion Elicitation

In this part, we will explain how to build up a rule-based emotion elicitation model based on OCC upon a scenario in a virtual learning environment. The OCC model assumes that emotions arise from the evaluation of three aspects of the world, events, agents and objects. The three aspects regard the different valenced reactions of the emotional agent to situations, and are naturally overlapping with each other. Therefore at the beginning of resolving a scenario, the first procedure is to identify from which perspective the causation of emotion is aroused. Subsequently from the corresponding perspective, one continues to decide whether the valenced reaction is focused on consequences for the other or for oneself. This is the general logic flow of OCC to differentiate emotions.

The causation of emotion, denoted as $Emo\_Causation$, belonging to the scenario is an entity which initiates an emotion, an $Emo\_Causation \in \{Event, Agent, Object\}$, where $Event$, $Agent$ and $Object$ are the sets of all possible events, agents and objects across the environment. Among the three groups, event-based emotions are the most complicated, yet they are the ones which most frequently occur in a virtual learning environment. Since virtual worlds are commonly modeled as event-driven (implemented by event listeners), and a virtual agent's behavior is usually considered as a certain reaction to virtual world events, in this paper we only consider cases where $Emo\_Causation \in Event$. However, the eliciting processes of both agent-based and object-based emotions are resembled closely to the proposed approach, and thus are left for the interest of space.

An event at time t is defined as $event_t = \langle event\_content_t, event\_endurer_t \rangle$, where an event with $event\_content_t$ happens to $event\_endurer_t$, and $event\_endurer_t$ belongs to the set of virtual agents. An event with multiple event endurers can always be considered as multiple events with a single endurer, and therefore in this paper a single agent as the event endurer is discussed. The factor, desirability, applies to all event-based emotions, which does not only distinguish an emotion from its opposite within a subgroup, but also influences the computation of emotion intensity. Function $desirable(event_t, Goal) \in \{True, False\}$ returns a boolean value to specify whether an event is desirable in the light of endurer's goal, where $Goal = goalof(event\_endurer_t)$.

Moreover, a virtual agent who is experiencing an emotion is known as Emo_Holder. Note that $event\_endurer_t$ can be the Emo_Holder (consequences for the self) or other agents (consequences for the other). If $event\_endurer_t$ is not the Emo_Holder, subgroup *fortunes-of-others* is raised. Then we need to access the will of Emo_Holder towards $event\_endurer_t$, and define $will(Emo\_Holder, event\_endurer_t) \in \{good\_will, ill\_will\}$ as a function. In this subgroup, emotion is *happy-for* if *desirable* returns true and *will* returns well; *pity* if *desirable* false and *will* good; *resentment* if *desirable* true and *will* ill; and *gloating* if *desirable* false and *will* ill.

If $event\_endurer_t$ is the Emo_Holder, we need to differentiate subgroups according to $prospect\_relevant(event_t) \in \{True, False\}$ which means whether prospects are relevant or irrelevant. If $prospect\_relevant$ is false, subgroup *well-being* is raised, and further emotion is *joy* if desirable or *distress* if undesirable.

By contrast, if $prospect\_relevant$ is true, subgroup *prospect-based* is raised, and include emotion *hope* or *fear* if *desirable* returns true or false. Four more emotions can be elicited by the status of *hope* or *fear* which may be confirmed or disconfirmed: emotion is *satisfaction* if *hope* is confirmed; *disappointment* if *hope* disconfirmed; *fears-confirmed* if *fear* confirmed; and *relief* if *fear* disconfirmed.

Using above primitives and functions, the process of eliciting an emotion of a virtual agent can be implemented, for which inputs are $Emo\_Causation_t \in Event$ and Emo_Holder, and output is the generated current emotion of the underlying emotion holder, $emotion_t$. A sample fragment of rule-based emotion elicitation is illustrated in Fig. 2.

```
queue ← Emo_Causation_t ∈ Event
while (queue != null)
    causation_t = pop − out(queue);
    set Goal = goalof(event_endurer_t);
    get desirablity = desirable(causation_t, Goal);
    if  desirablity = true
        ⎡if event_endurer_t = Emo_Holder
        ⎢    ⎡if prospect_relevant(event_t) = True
        ⎢    ⎢   ⋮
        ⎢    ⎣else  emotion_t = Joy;
        ⎢else
        ⎢    ⎡if will(Emo_Holder, event_endurer_t) = good_will
        ⎢    ⎢         emotion_t = Happy − for;
        ⎣    ⎣else     emotion_t = Resenment;
    else
      ⋮
```

Fig. 2. Sample fragment of rule-based emotion elicitation

3.2 Illustration

To illustrate the soundness of the proposed implementation, we take one scenario as an example to illustrate how to elicit the dinosaur companion's emotion. The following example is related to the dinosaur's sympathetic emotion about the student. In this case, the event does not directly happen to dinosaur. Rather, it is the situation where dinosaur perceives what happens to the learner and generates a type of sympathetic emotion.

Example: Given that a student (ID: studentA) gets a bonus for 100 scores when he/she successfully completed a learning task (event_content ID: event_content001).

Inputs:

1. $Emo\_Causation_t = event_t = \langle event\_content001, studentA \rangle \in Event$
2. $Emo\_Holder = avatarA$, and it is the agent dinosaur.

Details:

1. Get $Goal = goalof(studentA)$ which is $'obtain\_highest\_score'$.
2. Check $desirable(event_t, Goal)$ which returns $True$.
3. Check $prospect\_relevant(\langle event\_content001, studentA \rangle)$ which returns $False$, because getting a bonus has already happened, so it is prospect irrelevant.
4. Since $event\_endurer_t \neq Emo\_Holder$, we need to check the will of the emotion holder to the event endure. As we presume the learning companion dinosaur is a friend of the student, the dinosaur always holds a good will to the student, and $will(Emo\_Holder, event\_endurer_t)$ returns $good\_will$.

Output: $emotion_t = happy\_for$ that the agent is happy for the student getting a bonus.

Quick check: Comparing the event (*event_content ID: event_content001*) and the goal 'obtain_highest_score', we know that the event is desirable. Moreover, the emotion that the dinosaur will have is not related to its internal feeling due to its own perceptions over environment change, but is regarding a learning companion's attitude to the student's performance. As the dinosaur is always pleased to see the student performing well, it will feel happy for the learner when he gets a high score.

From this simple example, we can find that the information required by the eliciting process is non-trivial. But fortunately in a virtual learning environment, a lot of data can be easily collected by the system and used for analysis. However, the rule-based elicitation only returns qualitative results. In order to get the quantitative results for precise agent control, we need another phase—using FCM to compute the intensity of emotions.

4 Computing Intensities of Emotions

Using the OCC-based elicitation rules described in the above section, we are able to find out what emotions the agent will generate according to the current situations in virtual world. However, only identifying what emotions are likely to be is not enough. We also need to compute the specific intensity of emotions to achieve the precise control of agent behaviors. In this section, we will illustrate how to provide a convenient way to transform the OCC-based rules to a computable causal graph, which flexibly depicts the dynamic features and the interactions between emotional elements.

4.1 Using FCM as a Computational Tool for Emotion Modeling

Fuzzy Cognitive Map (FCM) is a fuzzy-graph structure which can simulate the complex systems in the world through causes, effects, and the causal relationships in between. FCM as an efficient fuzzy tool can be defined as a trio (C, R, W), in which $C = \{C_1, C_2, ..., C_n\}$ is the Concept set. Each element, concept, is represented as a node in FCM graph, and the causes and effects are all defined as concepts in this set. $R = \{R_1, R_2, ..., R_m\}$ refers to the Causal Relation set. Each element $R_k = \overrightarrow{C_i C_j}$ refers to the causal relation between concepts C_i and C_j. The relations are represented as arcs in FCM graph. Each causal relation has a weight to depict the influential degree from the former concept to the latter one, and it is defined as $W = \{W_1, W_2, ..., W_m\}$. All weights of the causal relations can be also compactly represented as an element of connection matrix $\mathbf{W} = [w_{ij}]$.

FCM as a fuzzy tool not only has the potential to describe the causal elicitation process of emotions based on emotional theory, but also can do mathematical computations to guide the emotional agent. The advantage of using FCM involves three aspects as follows,

1. FCM represents knowledge in a symbolic manner. The relationship between each concept can be directly signed by inter-linkages. The graphical representation of FCM is convenient for representing the elicitation rules mentioned previously.
2. Besides graphical representation FCM also provides the mathematical way to analyze the problem. Each concept can be defined as a fuzzy set. In addition, the causal strength and the interactive relations can be depicted by weighted values for each connection. The symbolic and numeric transformation in FCM is straightforward.
3. For building a dynamic system, FCM is efficient to describe a complex dynamic process, because FCM makes the complex operation of the whole system transparent through defining the causal relationships merely within each concept pair, while preserving the complex dynamics by iterative calculations. Therefore, no matter how complicated the elicitation process of emotions will be, FCM is a powerful tool for designers to model such processes.

4.2 Integration

In order to use FCM to analyze the dynamic system of emotions and compute the emotion intensities, we need to do three phases of integration. First phase is to map the inputs, outputs, and related concepts of the OCC-based rules into the concept set C of FCM. This phase is to model the rule-based entities as a collection of concepts, and these concepts will then be used to simulate the dynamic process through the causal relationships between them. Therefore, the concepts collected to set C should contain those factors of which numerical values are necessary for computing the intensity of emotions. Based on the OCC theory, the factors (local factors in OCC) affecting the intensity of event-based emotions include

1. the degree of desirability of an event,
2. the degree of belief that an anticipated event will occur (likelihood),
3. the degree to which resources were expended in obtaining or avoiding an anticipated event (effort), and

4. the degree to which an anticipated event actually occurs (realization).

Besides these factors, some other important concepts also need to be drawn into the FCM. These are

1. the impact of the causational event,
2. the impact of reactions of the emotion holder, and
3. the intensity of emotions

where the first factor can reflect the environmental changes of the system, and the second and the third factors reflect the behaviors of the emotional agent.

The second phase, after collecting all the concepts, is to find out the causal relations between these concepts and represent them to set R . In FCM the causal relations are divided into two categories, *positive* and *negative*. The positive causal relation refers to relations for which the increase of the starting concept value may cause the increase of the ending concept value. Conversely, the negative causal relation refers to those by which the increase of starting concept value will cause the decrease of the ending concept value. To precisely describe the quantitative degree of causal relations, each relation or each element in set R is associated with a weight or weight function. The weight values and the concept values generally are real numbers, but sometimes they may use predefined functions to achieve stability or dynamics which provides various feasibilities for modeling a dynamic system.

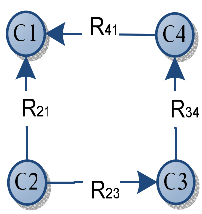


Fig. 3. An Example of FCM

$$W = \begin{bmatrix} 0 & 0 & 0 & 0 \\ w_{21} & 0 & w_{23} & 0 \\ 0 & 0 & 0 & w_{34} \\ w_{41} & 0 & 0 & 0 \end{bmatrix}$$

Fig. 4. The Weight Matrix

In the third phase, we draw the concepts as nodes and connect all the concepts in light of their relations with arcs as shown in Fig. 3. Moreover, we represent the weights of all the relation arcs in a matrix format. If there is no causal relation between two concepts, we define the weight as zero. The example weight matrix for Fig. 3 is given in Fig. 4. With the weight matrix we can use simple linear algebra to do the computation iteratively as $C_{i+1} = C_i \bullet W$. The details will be explained with an example in the case study.

5 Case Study

In this section, we will illustrate how to analyze a scenario with concrete environmental conditions and how to use it to do the intensity computation. We will introduce a simple example which has been used in Dino project.

Example: Given that a learning companion dinosaur called Dilong (ID: avatarA) notices that a giant carnassials dinosaur is coming (event_content ID: event_content002), and Dilong realizes that maybe he is in danger of being eaten by the giant dinosaur.

Analyzing by the rule-based elicitation process we have,

Inputs:

1. $Emo\_Causation_t = event_t = \langle event\_content002, avatarA \rangle \in Event$
2. $Emo\_Holder = avatarA$ and it is the agent dinosaur.

Details:

1. Get $Goal = goalof(avatarA)$ which is $'avoid\_giantDino'$.
2. Check $desirable(event_t, Goal)$ which returns $False$.
3. Check $prospect\_relevant(\langle event\_content002, avatarA \rangle)$ which returns $True$, because being eaten by the carnassials dinosaur is the "prospect" of this learning companion dinosaur.
4. Since $event\_endurer_t = Emo\_Holder$, and because of the undesirability of this event, we straightforwardly come to the conclusion.

Output: $emotion_t = Fear$ that Dilong feels fear when the giant dinosaur is near.

After the qualitative analysis, we generate the corresponding FCM to map the rules. First, collect relevant concepts of FCM. The concepts affecting the intensity of event-based emotions include *desirability*, *likelihood*, *effort* and *realization*. To simplify the situation, we only concern the first two factors, the Dilong's desirability about the event, and the degree of Dilong's belief of being eaten by giant dinosaur. The impact of the causational event is defined as the distance between Dilong and the giant dinosaur. The output concept Emotion is "fear", and the concept Action is a sequence of Dilong's frightened behaviors which controls Dilong's reaction in the application level. To sum up, all the concepts are listed in Table 1.

Table 1. List of FCM Concepts for Dilong

Concepts	Definition
C1 Event	Distance between Dilong and the giant dinosaur
C2 Desirability	Dilong's desirability
C3 likelihood	The degree of Dilong's belief of being eaten
C4 Emotion	Dilong's emotion "fear"
C5 Action	Dilong's frightened actions

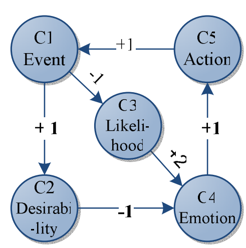


Fig. 5. Simplified FCM for Dilong being eaten situation

According to the distance between the giant dinosaur and Dilong C1, Dilong's desirability C2 is changed correspondingly. At the same time, C1 also affects Dilong's belief of being eaten, the likelihood of being eaten C3. Furthermore, the desirability C2 and the likelihood C3 will together determine Dilong's emotion C4. Emotional state will influence the agent's reaction to the event, and thus Dilong's emotional state C4 will regulate its reaction C5. Finally, this reaction directly changes the distance between the giant dinosaur and Dilong in return. The weights define the degree of the causal effect. Based on the relations between each concept, the FCM is drawn as Fig. 5.

After the graphical FCM is represented, we need to focus on how to do the intensity computation. In light of OCC, this model is related to the emotion "fear". By using the formulation of emotion fear [15], we have $Fear = 2 \times Likelihood^2 - Desirability$. As we have mentioned in Section 4.2, each concept or weight can use a predefined function, and we adopts three functions for concepts C1, C3 and C5 to simulate the causal relations of fear emotion. $C_k^{'}$ denotes the original concept values before passing through the concept function. For C1, the function is $f_{C_1} = C_1^{'} / d_{\max}$, where $d_{\max}$ is the maximum distance defined for cases in the virtual world. By divided by $d_{\max}$, each real distance value between Dilong and giant dinosaur in the system is normalized within 0 and 1.

For C3, the function is defined as $f_{C_3} = \begin{cases} (1-|C_3^{'}|)^2, & C_3^{'} \in [-1,0) \\ C_3^{'}, & C_3^{'} \in [0,1] \end{cases}$. Because the likelihood of being eaten is decreased with the increasing of the distance, we use $(1-|C_3^{'}|)^2$ to keep the value positive.

Concerning C5, it refers to the reaction of Dilong. In this model, Dilong's emotion C4 directly influences its reaction C5. We simply define the variation of reactions is Dilong's running speed

$$v_{Dilong} = (emotion / emotion_{\max}) v_{Dilong\max} = (C_4 / C_{4\max}) v_{Dilong\max}$$

Then, we use function f_{C_5} to define the effect of Dilong's speed on the changes of distance. We assume, during the period of Dilong making running decision, the speed of the giant dinosaur remains the same. Thus we have,

$$f_{C_5} = C_1 - (v_{giantdino} - v_{Dilong}) / d_{\max}$$

With all these functions and the weight matrix, this approach can do the matrix computation with an initial concept vector. There is only one input in the vector, that is, the distance between Dilong and giant dinosaur C1, which can be read from the game engine directly. Given the distance d, the input vector of the five concepts is

$$\begin{array}{ll} & \text{C1 C2 C3 C4 C5} \\ \vec{C_1} = & [\, d \quad 0 \quad 0 \quad 0 \quad 0 \,] \end{array}$$

The matrix computation is an iterative process, and computations for iteration i are summarized as follows,

$$\begin{cases} \vec{C_i} \times \mathbf{W} = \overrightarrow{C'_{i+1}} = \begin{bmatrix} C'_{1(i+1)} & C_{2(i+1)} & C'_{3(i+1)} & C_{4(i+1)} & C'_{5(i+1)} \end{bmatrix} \\ C_{1(i+1)} = f_{C_1}(C'_{1(i+1)}) \\ C_{3(i+1)} = f_{C_3}(C'_{3(i+1)}) \\ C_{5(i+1)} = f_{C_3}(C'_{5(i+1)}) \\ \overrightarrow{C_{i+1}} = \begin{bmatrix} C_{1(i+1)} & C_{2(i+1)} & C_{3(i+1)} & C_{4(i+1)} & C_{5(i+1)} \end{bmatrix} \end{cases}$$

We used Matlab to simulate this computation, and got the results as shown below. Given $v_{giantdino} = 10$, $v_{Dilong\max} = 8$, $d_{\max} = 80$ and the initial vector $\vec{C_1} = [0.9 \;\; 0 \;\; 0 \;\; 0 \;\; 0]$, the computation results are shown as Fig. 6. The figure shows how the distance between giant dinosaur and Dilong changes (blue line), and how the fearful feeling of Dilong changes (red line). When it runs to the 13th round, the distance is 0, and the whole computation terminated meaning Dilong has been caught. If we assume $v_{Dilong\max} = 20$, the corresponding results are given in Fig. 7. It shows that if we change the maximum speed of Dilong as 20, the distance between giant dinosaur and Dilong will reach a steady state at the 34th step. In this way, Dilong will never be caught, and its emotion of fear will remain at a constant level.

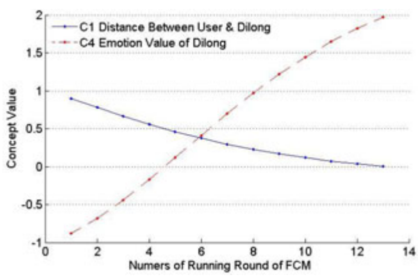


Fig. 6. Simulation1

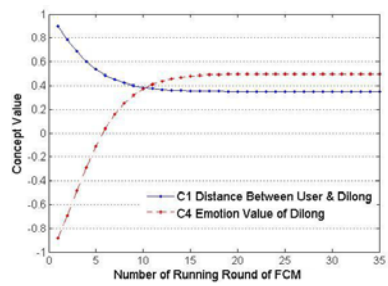


Fig. 7. Simulation 2

From the simulation results we find the relationship between distance and Dilong's fearful feeling, and these emotion values can control the Dilong's emotional expression in virtual world. We can also alter the value of $v_{giantdino}$ in real time and do the computation similarly. In this way we can achieve the adjustment of Dilong's fearful emotions in the virtual world with dynamic changing variables.

From this case study, we can see that this quantitative approach is simple and fast, and owing to FCM's symbolic representation it is very convenient to modify if having any design mistake.

6 Conclusion and Future Work

This paper proposed a practical approach which uses FCM to represent OCC-based emotional rules to bridge the gap between psychological theory and the real operation of agent system. We also examined how this model can achieve the requirements from the qualitative and the quantitative concern. FCM as an efficient fuzzy tool can not only bring a system logical formulization in a straightforward way, but also conduct the simple computation with matrix calculation. For future work, to provide more natural and believable agent behaviors the situation of mixed emotions especially the conflicting mixed emotions should be investigated deeply, and the agent behaviors controlled by the computational results should be made more creative and interesting.

References

1. Johnson, W.L., Rickel, J.W., Lester, J.C.: Animated pedagogical agents: Face-to-face interaction in interactive learning environments. International Journal of Artificial intelligence in education 11(1), 47–78 (2000)
2. Zakharov, K., Mitrovic, A., Johnston, L.: Towards emotionally-intelligent pedagogical agents. In: Woolf, B.P., Aïmeur, E., Nkambou, R., Lajoie, S. (eds.) ITS 2008. LNCS, vol. 5091, pp. 19–28. Springer, Heidelberg (2008)
3. Ortiz, A., Oyarzun, D., Puy Carretero, M.: ELEIN: E-Learning with 3D Interactive Emotional Agents. Springer, Heidelberg (2009)
4. Kosko, B.: Fuzzy cognitive maps. International Journal of Man-Machine Studies 24(1), 65–75 (1986)
5. Ortony, A.A., Collins, A.A.: The cognitive structure of emotions. Cambridge University Press, Cambridge (1988)
6. Roseman, I.J.: Cognitive determinants of emotion: A structural theory. Review of Personality & Social Psychology (5), 11–36 (1984)
7. Plutchik, R.: The nature of emotions. American Scientist 89(4), 344–350 (2001)
8. Russell, J.A.: A circumplex model of affect. Journal of personality and social psychology 39(6), 1161–1178 (1980)
9. Jaques, P.A., Vicari, R.M.: Infering emotions and applying affective tactics for a better learning. In: Agent-based tutoring systems by cognitive and affective modeling, pp. 135–155 (2008)
10. El-Nasr, M.S., Yen, J., Ioerger, T.R.: Flame fuzzy logic adaptive model of emotions. Autonomous Agents and Multi-Agent Systems 3(3), 219–257 (2000)
11. Zhou, X., Conati, C.: Inferring user goals from personality and behavior in a causal model of user affect. ACM, New York (2003)
12. Jiang, H., Vidal, J.M.: From Rational to Emotional Agents. University of South Carolina (2007)
13. Steunebrink, B.R., Dastani, M., Meyer, J.J.C.: A formal model of emotions: Integrating qualitative and quantitative aspects. IOS Press, Amsterdam (2008)
14. Zhang, H.L., et al.: Emotional agent in serious game (DINO). In: International Foundation for Autonomous Agents and Multiagent Systems (2009)
15. Price, D.D., Barrell, J.E., Barrell, J.J.: A quantitative-experiential analysis of human emotions. Motivation and Emotion 9(1), 19–38 (1985)

Lunar Surface Collaborative Browsing System for Science Museum Exhibitions

Kikuo Asai[1,2], Tomotsugu Kondo[1,2], Akira Mizuki[3], and Mark Billinghurst[4]

[1] Center of ICT and Distance Education, The Open University of Japan, Chiba 261-8586, Japan
[2] The Graduate University for Advanced Studies, Miura 240-0193, Japan
[3] TransferOrbit Corporation, Matsudo 271-0087, Japan
[4] Human Interface Technology Laboratory, Christchurch 8140, New Zealand
asai@code.u-air.ac.jp

Abstract. The lunar surface collaborative browsing system was developed for exhibitions at science museums. It visualizes multimedia data using the geographic locations in the area of the lunar surface explored by the NASA Apollo missions, providing visitors with a collaborative-learning environment through networked interactive functions. We designed scenarios based on real episodes during exploration activities and assumed a quasi-role-play with children acting as astronauts and their parents acting as mission commanders. Children manipulate a rover on the lunar surface and view a landscape at a viewpoint on the lunar surface. Parents instruct their children to find objects or information to complete tasks. Our contribution is to create a collaborative-learning environment by integrating map-based and virtual environments to view the lunar surface from exocentric and egocentric viewpoints. The system has capabilities that encourage children and their parents to get together to learn.

Keywords: Collaborative Learning, Virtual Environment, Map-based Augmented Reality, Lunar Surface, Science Museum.

1 Introduction

Exhibits at museums have usually been organized based on real objects, providing both education and enlightenment to their visitors. Replicas are often used if the real objects cannot be exhibited or are not ideally suited for display [1]. Multimedia presentations and computer graphics are widely used for visualizing various types of data from digital archives [2]. However, these exhibits encourage passive behavior by visitors, and they spend little time on these exhibits [3, 4]. Such glance-overs are insufficient for knowledge to become established in their memories.

Interactivity is an important factor in constructing knowledge and concepts and to motivate visitors to learn. Interactive experience provides children with an environment where they can enjoy learning. User interfaces based on conventional WIMP (window, icon, menu, and pointer) environments for controlling multimedia content are not well suited to public exhibits because not all visitors may be familiar with using them. Our group has taken the approach of using map-based augmented reality

Z. Pan et al. (Eds.): Transactions on Edutainment IV, LNCS 6250, pp. 34–43, 2010.

(AR) to provide users with an interface [5]. Map-based AR presents geographically embedded information that is overlaid on an actual map. It creates a tangible user interface that intuitively supports interactivity with virtual objects by enabling users to manipulate physical objects that correspond to data elements [6, 7].

However, we have not considered another important feature of visitor attributes, which is that visitors often come to museums with their families. Typical visitors to science museums are children and their parents. Their cooperative learning may facilitate mutual construction of knowledge and concepts through communication, if a constructivist approach [8] to learning is taken, theoretically assuming that knowledge and concepts are created through experience and are obtained by interacting with information as well as through activities. Computer-supported collaborative learning is also considered to work as constructivist learning that can be applied to virtual worlds [9, 10].

Here, we tried to create a collaborative-learning environment where children and their parents could learn together at science-museum exhibitions. We targeted an age group on the upper grade students of elementary school. Therefore, we upgraded the lunar surface browsing tool to a collaborative browsing system by combining map-based AR with virtual environments. The system enables children and their parents to cooperatively browse the lunar surface and engage in exploration activities in the area explored by the NASA Apollo 17 mission [11]. To make the situation more authentic, we arranged a quasi-role-play where astronauts and mission commanders communicated remotely. Children play the roles of astronauts who carried out exploration activities on the lunar surface, and parents play the roles of mission commanders at the earth control center. In practical situation, tasks to be solved for game-play aspects may be given them in order to keep engaging their attention.

The lunar surface collaborative browsing system was designed to describe scenarios at science-museum exhibitions, interaction techniques to control the visualizations, and useful functions to make exhibits more attractive.

2 Related Work

Applying AR technology to museum exhibitions has recently become popular. Most museums use AR to augment real objects with computer graphics or add detailed explanations with annotations. For example, an exhibit at the Allard Pierson Museum in Amsterdam presented an old photograph of the Roman Forum that was superimposed with a 3D reconstruction [12]. The AR experience installed in the Nga Taonga gallery at the Canterbury Museum in New Zealand makes the legends behind the artifacts as compelling and authentic as the animated stories [13]. The AR guidance system by Metaio GmbH and the Louvre-DNP Museum Lab assists visitors with information on directions and explanations about the exhibits [14]. Museum experiences have also been enhanced with AR in tabletop environments enabling more social and tangible interactions [15, 16].

Augmented Maps [17], in which AR has been applied to cartography, have enabled the visualization of 3D digital geographic data to be overlaid on printed maps using a video see-through head-mounted display. Augmenting maps with information that is projected overhead facilitates tangible interactions in tabletop environments [18].

A PDA device was used as a tangible user interface, enabling intuitive access to information linked to locations on a printed map. The Illuminating Clay system [19], a GIS (geographic information system) interface for landscape design and analysis, uses a clay model instead of a printed map. These systems provide users with a collaborative environment by allowing them to share a common physical space, which enables communication behaviors to become easier and more natural and social [20].

There have been several navigation tools for visualizing lunar-surface-exploration data and for displaying images of lunar-surface reliefs. One of these is World Wind [21], which is a navigation system developed by NASA to visualize geographical data using zoom-in and zoom-out functions. The moon in Google Earth [22] enables users to view high-quality images and landscapes of the lunar surface using the data obtained by spacecraft.

Our browsing system is similar in concept to these map-based visualization systems. However, a collaborative-learning environment is created by integrating map-based AR and virtual worlds, so that users can interactively browse geographically embedded information on exploration activities and view realistic landscapes of the lunar surface from an egocentric viewpoint.

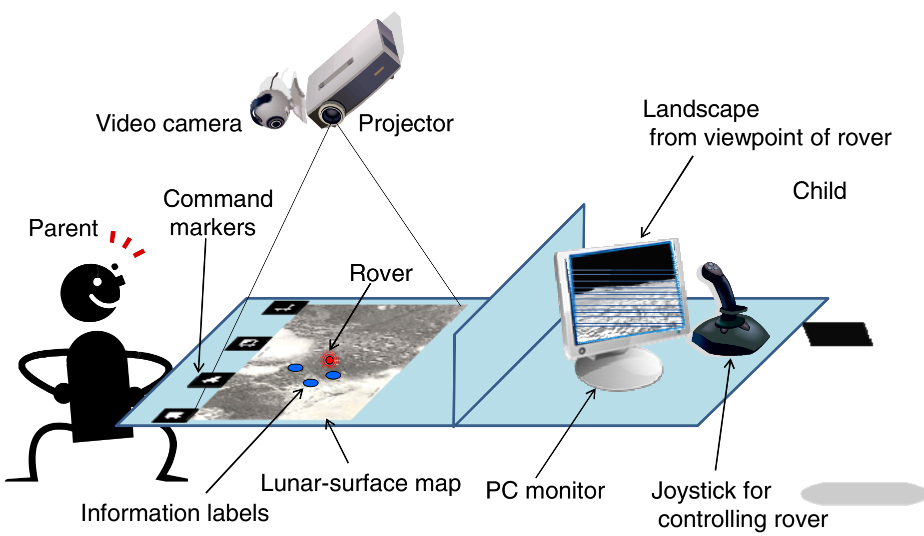


Fig. 1. Schematic of setting in collaborative browsing system. Child and his parent learn about the lunar explorations together. Child is looking at landscapes from viewpoint of lunar rover and parent is looking at whole area of lunar-surface map.

3 System Design

Here, we describe the design of the lunar surface collaborative browsing system, which was divided into three parts, i.e., scenarios for using the system at science-museum exhibits, interactive techniques for controlling visualization, and useful functions for making exhibits more attractive. Figure 1 is a schematic of a setting in the collaborative browsing system. A child and his parent are learning about the lunar-surface explorations together. The child is looking at landscapes from the viewpoint of the lunar rover and the parent is looking at the whole area of the lunar-surface map and at the additional information that is projected onto the map. Communication is assumed to be accomplished by their use of transceiver equipment, so that the situation can be more realistic in simulating telecommunications during the explorations.

3.1 Scenarios

As the Earth's only natural satellite, the moon is one of the most familiar astronomical objects. It has been extensively explored, particularly during the NASA Apollo missions, and much scientific data has been collected, including images of the lunar surface and audio recordings of the communication between the astronauts themselves and between the astronauts and their ground commanders. Scientific data on the lunar surface have been used in many science-museum exhibits.

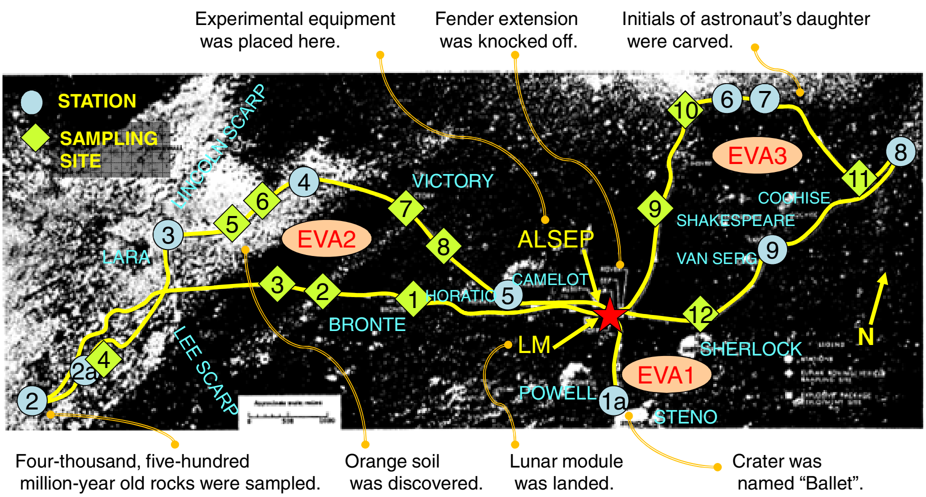


Fig. 2. Navigation area around Apollo 17 landing site. Real episodes during three separate lunar surface excursions were embedded into scenarios.

The lunar module (LM) of Apollo 17 landed in a valley in the Taurus-Littrow region (Fig. 2). Three separate lunar surface excursions were conducted during its stay. The astronauts conducted various scientific experiments during the excursions, spending roughly 22 hours outside the LM and traversing about 35 km in the lunar rover. They collected nearly 110 kg of rock and soil samples from various areas of stratigraphic interest. The Apollo 17 preliminary science report [11] contains a huge amount of data, which is neither easy nor practical for laymen to comprehend in entirety. We selected several remarkable experiments and samplings, and used them as episodes for learning about lunar-surface explorations.

A quasi-role-play was introduced to the communication between children and their parents so that the situation would simulate lunar-surface explorations. Children play the roles of astronauts who moved around on the lunar surface on a lunar rover. Children can find some geographically embedded information at locations where real astronauts found it or where they did something during the excursions. The information is presented with multimedia data such as photos, sounds, and text annotations. Parents play the roles of mission commanders who gave astronauts instructions on exploration activities and proposed solutions to problems during the excursions.

Children and their parents are given tasks of answering questions based on real episodes. This way of relating tasks to real episodes may have a better impression on their memories. To perform tasks, parents ask their children to find information that is required to answer questions, which virtually creates an environment for meaningful role-play. Since time limitations are also imposed, the role play has game-like characteristics.

One of questions, for example, is that "what is the purpose in placing the lunar rover 150 m away from the lunar module before leaving the lunar surface?" The tips on answering the question may be located around the landing site. Then, a parent instructs his/her child to drive the rover to the landing site. The child may find information of the photo, annotations, and narration related to the question. They can learn the lunar surface activities through answering the question with the realistic landscapes and real records.

We chose real episodes that were not just planned activities but also various episodes, e.g., where a fender extension was knocked off, a crater was named "Ballet", the initials of an astronaut's daughter were carved in moon rock, orange soil was discovered, and four-thousand, five-hundred million-year-old rocks were sampled.

3.2 Visualization, Interaction, and Communication

Figure 3 overviews the collaborative browsing system, where the landscape is being viewed by children from a viewpoint on the lunar surface at a PC monitor. The locations of geographically embedded information were labeled with colored signs, but some signs could not be seen from the viewpoint on the lunar surface because of geographic relief. Children move around on the lunar surface to find the locations of information that they require answering questions based on instructions from their parents. A joystick is used to control the viewpoint of the landscape, because the real lunar rover was manipulated with a joystick-like controller instead of the steering wheels we see in cars.

A tabletop environment is arranged to provide a bird's-eye view of the lunar-surface map viewed by parents who acted as mission commanders. Since mission commanders in the real situations were on earth, they could not obtain details on situations surrounding astronauts. They could just piece together the situation through video-camera images mostly captured by the lunar rover. The position and orientation of the lunar rover is projected onto the lunar-surface map by a projector, and is expressed as a colored arrow. The landscape from the viewpoint of the lunar surface, which is the same as the one viewed by children, is displayed at low resolution in a small window.

As previously described, information that astronauts have had is quite different from that that mission commanders have had during exploration activities. Although photos and annotations are presented to children, only annotations are presented to

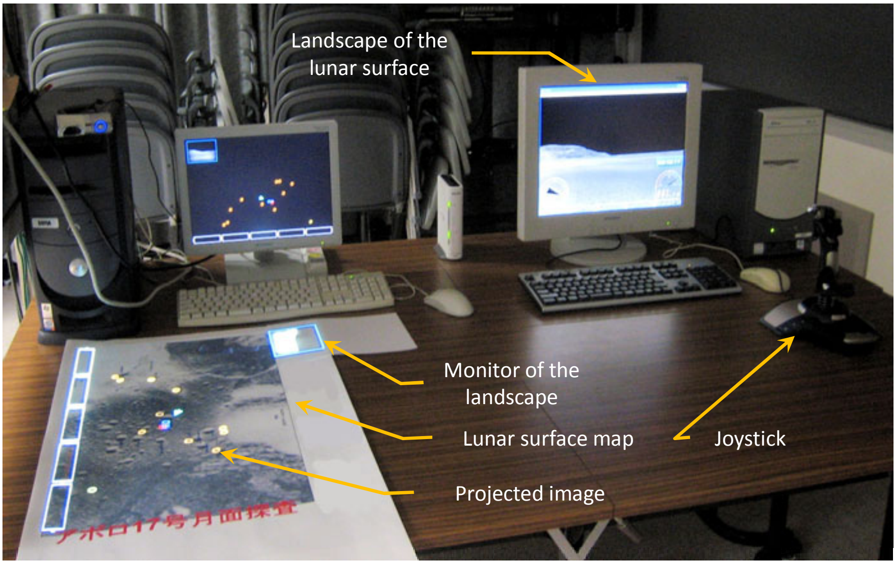
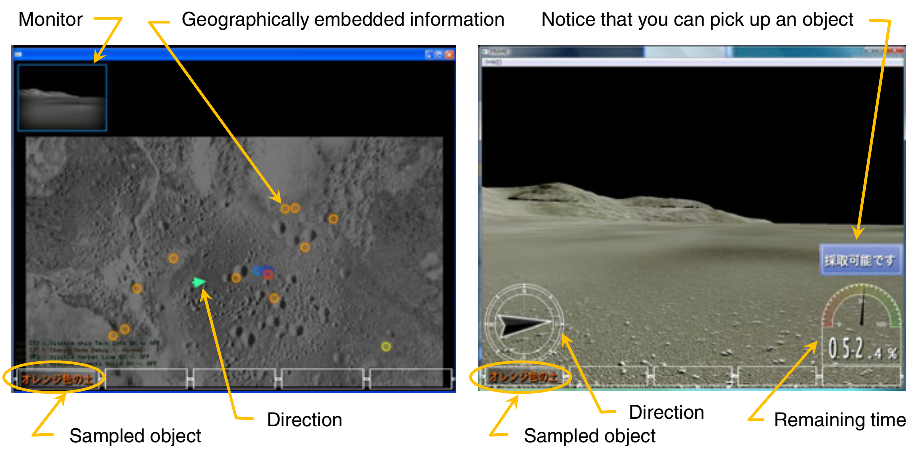


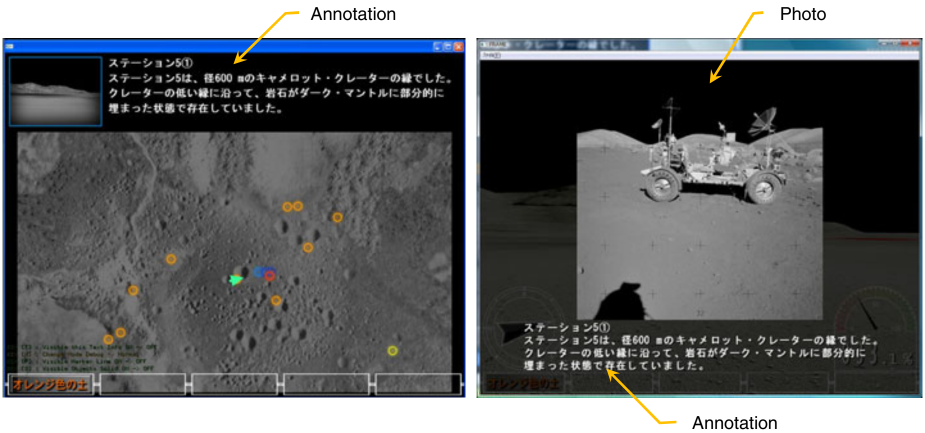
Fig. 3. Overview of lunar surface collaborative browsing system



(a) (b)

Fig. 4. Example snapshot of lunar surface browsing system being used; (a) Bird's-eye view of lunar surface viewed by parents who monitor astronaut excursions and (b) landscape on lunar surface viewed by children who drive lunar rover during role play

parents, so that the role-play simulates a situation where astronauts were engaged in exploration activities on the lunar surface and the mission commander obtained information through communication with them. Therefore, when parents need specific information to answer questions, they can ask their children about this. Transceiver devices are used by children and parents for communication, which creates an atmosphere that they are communicating some distance from each other. Figures 4 and 5 have example screenshots of the lunar surface browsing system being used.



(a) (b)

Fig. 5. Presentation of the geographically embedded information; (a) just the annotations are presented for parents. To perform a task, parents ask their children about details in the photo, and (b) a photo and the annotations are presented with narration for children.

A paper chart was used as the lunar surface map, so that parents could note tips to answer questions. The chart would also be a souvenir of their visit to the science museum and keep encouraging them to learn home more about lunar explorations and the moon. The geometric relation of the video camera, projector, and map is fixed, to prevent the presentation of information from becoming unstable due to map tracking.

Although mission commanders could not obtain details on the situations around astronauts, they did have huge amounts of geographic information about the lunar surface. Visualization functions were prepared to present various kinds of information on the lunar-surface map, such as contour lines for altitude, locations of planned stations, geographic data, grids for scale, and rover tracks. Pattern markers were placed beside the lunar-surface map, enabling them to be used as a user interface for control of the visualization functions. Hiding one of the pattern markers with a finger works as a command switch to change the information that was projected onto the lunar surface map.

Children have the chance to obtain some objects such as the initials, hammers, and the orange soil related to real episodes. The objects are placed at certain locations on the lunar-surface map. When a lunar rover driven by children enters a certain area involving an object, a notice appears with a text sign that asks them whether they want to pick up. When the object is picked up by the children, it is noticed to the parents in their window. Object collection can be used as a task for them to comprehend where the real episodes have occurred in the excursions.

When a lunar rover driven by children turns the south-southwest, the earth can be viewed in space. Since the moon presents an almost steady facade when viewed from the earth, the earth should be viewed from a regular elevation on the lunar surface. If the earth and moon are exhibited at museums, they allow us to see their behavior from an exocentric viewpoint. The egocentric and exocentric views give children an advantageous environment enabling them to comprehend the geometric relations between the earth and moon.

3.3 Implementation

Figure 6 shows software architecture of the lunar surface collaborative browsing system. It consists of a server and a client because of the load distribution. The client presents landscapes at a viewpoint of the lunar rover and multimedia data based on the position of the lunar rover. The position data of the rover is transferred from the client to the server. The server presents an overview of the landscapes that are the same as the ones at the client, and presents just annotations based on the position of the lunar rover. The server also works an image-processing machine for recognizing control commands. Socket communications were used for data exchange between the server and the client.

ARToolkit [23], a set of image-processing libraries was used to detect control command cards in the server, which are composed of patterns with square frames. The Direct3D was used for producing graphics of the lunar surface landscapes in the client. Since the observational texture data of the lunar surface was not so high in resolution, the area close to the viewpoints in the landscapes did not look realistic. The artificial texture data were superimposed in order to make the area more realistic in the lunar surface landscapes. The DirectMusic was used for producing sounds of narration in the server. It takes a few seconds to produce the sounds in the server after the client presents the geographically embedded information, which also simulates the long distance between the earth and moon.

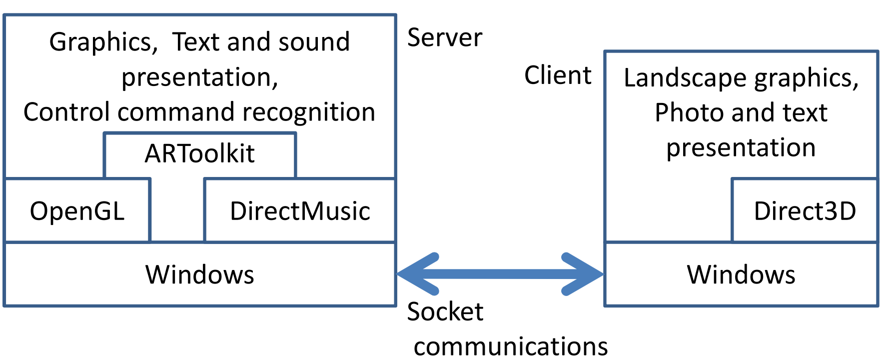


Fig. 6. Software architecture of lunar surface collaborative browsing system

4 Summary

The lunar surface collaborative browsing system that we developed was described, which was aimed at being used in exhibitions at science museums. It was designed to enable children and their parents to learn together in a collaborative-learning environment. We created a scenario where children move around acting as astronauts on the lunar surface by driving a lunar rover and their parents navigate them acting as mission commanders who viewed the whole area on a lunar-surface map. Future work includes evaluating the system in user studies to investigate how the collaborative-learning environment we designed here would work for learning in science-museum exhibitions.

References

1. Dean, D.: Museum exhibition: theory and practice. Routledge Press, New York (1994)
2. Usaka, T., Yura, S., Fujimori, K., Mori, H., Sakamura, K.: A multimedia MUD system for the digital museum. In: Proc. 3rd Asia-Pacific Computer Human Interaction, pp. 32–37 (1998)
3. Umiker-sebeok, J.: Behavior in a museum: a semio-cognitive approach to museum consumption experiences. Signifying behavior. Journal of Research in Semiotics, Communication Theory, and Cognitive Science 1, 52–100 (1994)
4. Falk, J.H., Dierking, L.D.: The museum experience. Whalesback Press (1992)
5. Asai, K., Kondo, T., Kobayashi, H., Mizuki, A.: A geographical surface browsing tool using map-based augmented reality. In: International Conference Visualization (Vis 2008), pp. 93–98 (2008)
6. Kato, H., Billinghurst, M., Poupyrev, I., Imamoto, K., Tachibana, K.: Virtual object manipulation on a table-top AR environment. In: International Symposium on Augmented Reality, pp. 111–119 (2000)
7. Regenbrecht, H., Baratoff, G., Wagner, M.T.: A tangible AR desktop environment. Computer & Graphics 25, 755–763 (2001)
8. Duffy, T.M., Jonassen, D. (eds.): Constructivism and the technology of instruction: a conversation. Lawewnce Erlbaum Associates Inc., Hillsdale (1992)
9. Dede, C.: The evolution of constructivist learning environments: immersion in distributed virtual worlds. Educational Technology 35, 46–52 (1995)
10. Kafai, Y.B., Resnick, M. (eds.): Constructionism in practice: designing, thinking, and learning in a digital world. Lawrence Erlbaum Associates Inc, Mahwah (1996)
11. NASA: Apollo 17 Preliminary Science Report, SP-330 (1973)
12. Fraunhofer IGD: Virtual museum guide. Research News, Fraunhofer Press (Febuary 2010)
13. Digital Binocular Station: Edutainment using Augmented Reality - AR digital binocular station brings museum to life, http://www.digitalbinocularstation.com/
14. Louvre-DNP Museum Lab: The use of augmented reality (AR) technology to transmit information within a museum, http://www.museumlab.jp/english/tech/04tech.html
15. Woods, E., Billinghurst, M., Looser, J., Aldridge, G., Brown, D., Garrie, B., Nelles, C.: Augmenting the science centre and museum experience. In: International Conference on Computer Graphics and Interactive Techniques in Australasia and South East Asia (graphite 2004), pp. 230–236 (2004)

16. Qui, T.C.T., Nguyen, T.H.D., Mallawaarachchi, A., et al.: Magic Land: live 3D human capture mixed reality interactive system. In: Conference on Human Factors in Computing Systems (CHI 2005), pp. 1142–1143 (2005)
17. Bobrich, J., Otto, S.: Augmented maps. In: Symposium on Geospatial Theory, Processing, and Applications, in IAPRS, vol. 34 part 4 (2002)
18. Reitmayr, G., Eade, E., Drummond, T.: Localisation and interaction for augmented maps. In: International Symposium on Mixed and Augmented Reality, pp. 120–129 (2005)
19. Ishii, H., Ratti, C., Piper, B., Wang, Y., Biderman, A., Ben-Joseph, E.: Bringing clay and sand into digital design – continuous tangible user interfaces. BT Technology Journal 22, 287–299 (2004)
20. Kiyokawa, K., Billinghurst, M., Hayes, S.E., et al.: Communication behaviors of co-located users in collaborative AR interfaces. In: International Symposium on Mixed and Augmented Reality, pp. 139–148 (2002)
21. NASA: WorldWind and the Moon, `http://www.worldwind.arc.nasa.gov/moon.html`
22. Google: Google Earth, `http://earth.google.com/`
23. Hitlab, University of Wasington: ARToolkit, `http://www.hitl.washington.edu/artoolkit/`

Towards a Structural Model for Intention to Play a Digital Educational Game

Xu Sun and Effie Lai-Chong Law

Computer Science Department, Univeristy of Leicester
Leicester LE1 7RH, UK
{xs47,lcl9}@mcs.le.ac.uk

Abstract. Computer games are very popular among children. There is increasing interest to apply computer game in education to render learning more attractive and motivating. However, it remains to be explored what factors drive children to play educational games. In this paper, we investigate impacts of game narrative and player's characteristics (e.g. gender, personality) on children's intention to play a digital education game. The results revealed that effect of game narrative, gender, and sensation seeking personality are significant factors in predicting children's gameplay intention. Implications of this study are discussed from perspectives of academic research as well as educational game design.

1 Introduction

Computer games are today an important part of most children's leisure lives. A study funded by the Kaiser Family Foundation (KFF) finds that, 92% of children and adolescents age 2-17 play computer games daily in the United Sates [1]. On average, 2-7 year-olds spend 8 minutes a day, 8-13 year-olds spend 32 minutes a day, and 14-18 year-olds spend 20 minutes a day playing computer games. There is increasing interest in investigating how to use this powerful new medium to support children's learning. Research into the use of digital games in education is relatively novel, but has been growing rapidly in the last five years.

As computer gameplay has become a prominent form of entertainment and a potential way of learning, a comprehensive model of digital educational gameplay is needed for a better understanding of the process of gameplay and its impacts on users.

As a first step towards building such a comprehensive model, this research attempts to investigate impacts of game narrative and player's characteristics (such as gender, personality) on intention to play a digital educational game (DEG). The improved understandings of the driving factors behind such an intention can assist game developers in designing popular games or help teachers and instruction designer to create a powerful new resource to support learning in the information age.

Previous behavioral studies of people's DEG gameplay intention have mainly based on various latent variables pertaining to motivational theories, such as perceived ease of use, perceived usefulness, and perceived enjoyment [2, 3, 4, 5, 6]. However, not all researchers entirely agree on the source of such motivation. Some argue that game narrative attributes to the act of playing [7, 8]. Several theorists have suggested that narratives are the primary way that we assimilate knowledge of the

Z. Pan et al. (Eds.): Transactions on Edutainment IV, LNCS 6250, pp. 44–55, 2010.

world and make sense of it. Although game narrative is regarded fundamental to human understanding of intentional behavior [9], there have been a limited number of research studies investigating to what extent game narrative *per se*, without the support of a gameplay structure, can account for digital gameplay intention.

In contrast, a surge of research interest in the role of storytelling in DEGs has been witnessed in recent years [e.g. 10, 11]. It is relevant for us to address the nuanced distinction between narrative and storytelling, especially when the lay usage of the two terms misleadingly suggests they are synonyms. According to our interpretation, a narrative is a sequence of real or fictional events that can be presented in verbal (textual, audio without embellishing tones), non-verbal (i.e. graphical without animation) or mixed formats and different media (i.e. paper, computer screen); it is essentially *static*. In contrast, storytelling refers to the *dynamic* process of delivering a narrative with some enrichment techniques to enhance the impact of the narrative on audience. Formalizing such a process into a method with the support of a gameplay structure (i.e. injecting mini-games into a narrative) or model (i.e. adapting to player profile) is one of the goals that the current research on DEGs aims to achieve [e.g. 10, 11].

In the current study, we investigate the effect of a game narrative (prior to its being transformed into storytelling) on children's intention to play a DEG. It also examines the impacts of players' characteristics (such as gender and personality) on the gameplay intention.

2 Theoretical Background

2.1 Research of Educational Gameplay

Previous studies of educational game design are classified into three broad categories: 1) research on theoretical underpinning of game design; 2) studies of learning outcomes from gameplay and 3) research on effective game design.

The first strand of research – theoretical underpinning of game design - has focused on constructive and situated learning design principles. The key constructive principle refers to learning as an active process of creating, rather than acquiring, knowledge [7]. Situated learning suggests that learning should be situated in a specific context and embedded within a particular social and physical environment. Several researchers find that learning with well-designed educational games adheres to constructive principles [7, 12]. Some other researchers state that situated learning principle provides a meaningful framework for the study of games, given that games have an ability to situate learning in an authentic context and engage players in a community of practice [13].

The second strand of research – learning outcomes from gameplay - has focused on the effect of computer game on learning, including assessments of skills [e.g.14], abstract and conceptual knowledge [e.g. 15], visual capacity, motor activity, and spatial abilities [e.g. 16]. These assessments employed the mixed-method approach including pre- and post- domain specific tests, questionnaires, interviews, and video analysis. Findings of these efforts suggest that gameplay can enhance learning.

The third strand of research – effective game design - has mainly been concerned with how to develop games that are both engaging and educative. Several authors find that motivation to play (e.g. desire, pleasure, flow experience) is a significant characteristic of effective educational games design [2, 17, 18]. However, not all researchers entirely agree on the source of motivation. Some think narrative is an important element of effective educational game design [8, 12, 19]. Game based learning environments that incorporate a strong narrative can make the learning task to be meaningful [19].

This study is in the third strand of research, investigating which elements are the driving factors behind gameplay intention. It is expected to contribute to broadening our understanding of educational gameplay and its impacts on users. It can help designers to develop an educational game that will benefit and appeal to its users.

2.2 Game Narrative

Narrative provides a story and context for the game. The importance of narrative in computer game has been recognized by academics. Existing research studies have shown that narrative is a significant design element.

Fisch [8] discusses how to create effective interactive games based on his experiences in the creation of real life games and applications. He suggests placing the educational content within the game narrative in order to design effective educational game. By developing an educational game narrative, children can employ the targeted academic skills and knowledge as an integral part of playing the game.

In Malone‘s studies of what people like about computer games [20], he has interviewed elementary school students about their computer game experience addressing the question of what makes computer game fun to learn. As a result, he identifies that an appealing fantasy context through a narrative is an important reason why computer games are successful.

Similarly, Prensky [21] also investigates why people like to play computer game. In his study, narrative is found be a powerful motivational tool that can engender strong emotions in the player. He suggests using narrative in game environments to engage users into playing. Quinn describes that an important design parameter of an educational game is the need to entertain the learner [22]. The results show that computational graphics are not necessarily essential for an engaging game. The design and development of an appropriate narrative can be used to engage the learner instead.

Waraich analyzes the role of narrative for conceptualizing learning in an interactive learning environment [19]. The study reveals quantifiable knowledge gains in the interactive learning environment over traditional instruction. Waraich points out educational games should incorporate a strong narrative. He further suggests designing learning content couple with the narrative in order to make the learning meaningful to the learner.

2.3 Player’s Characteristics

Several individual characteristics may influence people’s intention to play game. A literature review reveals gender and personality as key factors influencing gameplay habit, gametype preference, gameplay behaviour and performance [23, 24, 26].

2.3.1 Gender

Over the years there has been an increasing amount of research on gender differences in playing games [24]. In the past decades, game players were predominately male.

Over the last ten years, digital games have moved into mainstream of people's entertainment, a growing number of females can be seen playing games. While games are no longer exclusively the realm of males, studies of computer game players reveal vast differences between males and females in both the quantity and quality of computer gameplay and thus characterize computer games as still men's game [25]. Till now, it is nevertheless the case that the intense game players are more likely to be male than female. Females may play the same games as males, but they may play them differently [26].

Several research studies explain that the gender gap in computer game involvement has to do with the design of game, such as unacceptable female role models (e.g. weak victim or exaggerated sexuality in female characters), much violence, high level of competition and insufficient amounts of social interaction [25, 26]. Although such game characteristics may attract male players, they may inhibit females from playing. Industry designers and academics point out that games need to emphasize different characters, contents and mechanics in order to appeal to females.

2.3.2 Sensation Seeking Personality Trait

Personality is a stable set of tendencies and characteristics that determine the commonalities and differences in people's psychological behavior that have continuity in time [27]. [28] Introduced an entertainment measure using the concept of 'incongruity' to scale the difficulty level of a game online to the capabilities of the human player. Personality is difficult to understand and difficult to test. Personality trait theories have served as the basis of several personality tests widely used commercially and in academic research. Sensation seeking is a personality trait defined by the need for varied, novel, and complex sensations and experience, and the willingness to take physical and social risks for the sake of such experience [29].

There are studies investigating how sensation seeking is linked to game player's behavior. Ravaja et al. [30] investigate the relationship between players' personality traits and their engagement with computer games. The authors conclude that the engagement with the games varies as a function of the sensation seeking trait of the player. Fang and Zhao [31] study the impacts of sensation seeking personality traits on gameplay. They suggest that sensation seeking personality trait is a significant predictor on enjoyment of computer gameplay through enhanced engagement during gameplay. Furthermore, significant gender differences in this trait have been documented: Males tend to possess higher achievement needs [29] and are keener sensation seekers than females [29]. Accordingly, males prefer games with confrontational and violent contents entailing fast responses [32]. In contrast, females appreciate storylines and personalities of game characters to be explored at a relaxing pace and value building relationships with game characters [33].

3 Research Framework and Hypotheses

Drawing upon the literature and prior empirical findings, we propose a research framework (Fig. 1) to investigate children's intention to play a DEG under development. The topic of the DEG is based on geography. The game story is about an alien

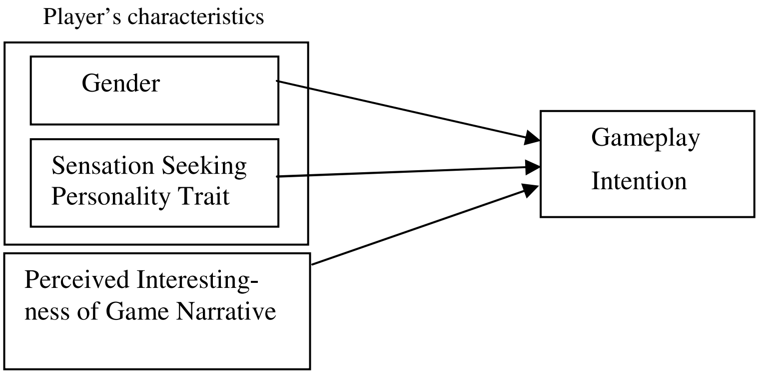


Fig. 1. Research framework

kidnapping a boy and their flying around the world to collect relevant geographical information. In this research framework, intention to play the game is a function of game narrative, gender and sensation seeking personality trait.

Narrative of game provides a story and context for the game. A strong narrative of a game makes the learning meaningful to the learner. It can be a powerful motivational tool that can engender strong emotions in the player [21]. It is likely that a player who is interested in the narrative of game will intend to play the game.

H1. Perceived interestingness in the game narrative is a significant factor influencing gameplay intention.

Many researchers have been interested in effect of gender on gameplay. Studies indicate there is gender gap in game use [25, 26]. It is likely that gender will influence the intention of gameplay.

H2. Gender is a significant factor influencing gameplay intention.

Some studies show sensation seeking personality trait plays a role in gameplay behavior [29, 30]. Computer games are designed to offer enjoyment and excitement. It is likely that high sensation seeking players may find a computer game more entertaining and will intend to play the game.

H3. Sensation seeking personality trait is a significant factor influencing gameplay intention.

4 Research Methods

A survey was conducted in five participating schools in UK to investigate the impacts of game narrative, gender, sensation seeking personality trait on gameplay intention. The target group of the game was children aged between 11 and 14. In total, 115 students responded to the survey; the distribution of age and gender is shown in Table 1. The teachers, who had agreed in response to an emailed invitation to get involved in the current study, administered the survey in their classrooms during regular lesson

Table 1. Gender and age of the survey respondents

	Girls	Boys	Total
Number	53	62	115
Mean age (St.Dev.)	12.85 (.53)	12.87 (.67)	12.86 (.61)

hours. The children completed the questionnaire individually and the average completion time was approximately 15 minutes. No tangible reward was granted to the teachers or the participating children.

The questionnaire consists of two major parts. Part A contains 5 close-end questions on the respondent's gender, age, gameplay habit, gametype preference, and affinity for geography. Part A also includes 8 items measuring sensation seeking personality trait. The 8 items are taken from the Form V of Sensation Seeking Scale (SSS-V) introduced by Zuckerman and his colleagues [29] (Table 3). Respondents were asked to rate each items with a 5-point Likert scale, with its leftmost and rightmost anchors being 'strongly disagree' and 'strongly agree', respectively.

Since its inception in 1978, different attempts have been undertaken to modify the response format from forced choice to Likert scale and to shorten the 40-item SSS-V to various lengths based on different rationales such as statistical validity and reliability, economical and motivational concerns in terms of response time and rate. Our 8-item version is an adapted version of Brief SSS (BSSS, [34] tailored for adolescents, including 2 items for each of the four subscales of the original SSS framework, i.e. experience seeking (ES), disinhibition (DI), thrill and adventure seeking (TAS), and boredom susceptibility (BS) (Table 3). The slight changes we have introduced are the order of the statements and replacement of one TAS item ('I like to do frightening things') with a more specific activity familiar to contemporary teenagers (i.e. movie/play viewing).

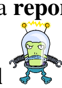
Table 2. Survey items

Constructs	Statements
Sensation seeking personality trait	▪ I would like to explore strange places (ES) ▪ I get very restless if I have to stay around home for any length of time [BS]. ▪ I like to have new and exciting experiences and sensations even if they are illegal [DI]. ▪ I like "wild" uninhibited parties [DI]. ▪ I like to take off on a trip with no pre-defined or definite routes or timetables [ES]. ▪ I prefer friends who are excitingly unpredictable [BS]. ▪ I would like to try parachute jumping [TAS]. ▪ I usually don't enjoy a movie or play where I can predict what will happen in advance [TAS].
Game narrative	▪ How interesting is the story for you? ▪ Would you like to play with the character in the story? ▪ What is the most/least like of the story? / How would you like to change it?
Gameplay Intention	Based on what you have learnt about the game, would you like to play it in the future?

Table 3. Narrative of the game

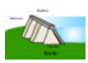
A 14-year old **boy**  is abducted by an **alien scout**  called Feon in a **UFO**  !

The boy is asked to help Feon explore the Earth by skillfully flying  the UFO to different cities in the world, and to write up a **report** about its geographical features.

In fact, Feon works for the **alien general**  whose plot is to conquer the Earth with a team of alien scouts and highly intelligent **alien scientist** advisor 

Only the first alien scout who hands in a complete report within the time limit is allowed to survive. The boy is in a dilemma. On the one hand, he befriends Feon and wants to help him. On the other hand, he is eager to save the Earth from the alien conquest.

One phenomenon on the Earth that Feon is particularly interested in is that of **Natural Disasters**. He and the boy fly over Europe searching for one taking place. The boy can communicate with an earthling to get some information.

They find a **dam**  that is about to burst – an event that will cause a  massive **flood**.

There is conflict between the boy and Feon; the boy is eager to save the local population. However, Feon does not bother to intervene so he can observe what happens. With some tricks, the boy is able to get Feon involved in the rescue. Together they modify the landscape using the UFO's technologies such as beaming to reduce the adverse effects of the flood.

Finally, the population of the local village is saved. The boy learns some important information about flood prevention. But his relationship with Feon is at stake ….

Part B first presents the game narrative (Table 2). Then two close-end questions are posed to investigate respondents' perceived interestingness of the game story and characters involved in the story. Next, three open-end questions on describing most and least liked aspects of the narrative and improvement suggestions are given. Finally a close-end question is presented to assess the respondent's intention to play the game in the future (Table 3).

5 Results and Discussion

The overall average of perceived interestingness of the game narrative was moderate (M=2.27, N=115, SD=0.95). The most interesting part of the narrative was 'the dam that is about to burst that will cause a massive flood'. Respondents considered this part exciting, unpredictable, and it involved lots of actions. Respondents did not like the part about 'writing a report on geographical features', because it was regarded boring, not playful, educational. They further made suggestions to improve the game story. The average intention to play was medium (M=3.03, N=115, SD=1.16). Boys (M=3.19, N=59, SD= 1.20) express higher intention to play the game than girls do (M=2.83, N=59, SD= 1.11).

Table 4. Correlation matrix of the research framework

		Sensation-seeking personality trait	Perceived interestingness of game narrative	Intention to play
Gender	Pearson Correlation	.264	-.152	.213
	Sig. (2-tailed)	.004*	.105	.034*
Sensation-seeking personality trait	Pearson Correlation		.056	.259
	Sig. (2-tailed)		.553	.005*
Perceived interestingness of game narrative	Pearson Correlation			.531
	Sig. (2-tailed)			.000**

A correlation matrix was computed and linear regression was conducted to explore the impacts of game narrative, player's characteristics (gender, personality) on the gameplay intention.

The correlation analysis suggests that gameplay intention is significantly correlated to gender, sensation seeking personality trait and perceived interestingness of game narrative. Besides, there were significant gender differences in sensation seeking trait with males being keener sensation seekers than females. This finding corroborates that of the earlier research (Section 2.3.2).

A further linear regression (r-square= 0.512) shows the significant effect of the predictor perceived interestingness of game narrative (beta coefficient= .551, t= 7.184, p<.001), significant effect of the predictor gender (beta coefficient= .193, t=2.43, p<.05) and significant effect of the predictor sensation seeking personality trait (beta coefficient= .177, t=2.25, p<.05).

Noteworthy is that the other demographic variables, including age, gameplay habit, gametype preference, and affinity for geography, had no significant correlation with the intention to play. However, gender was found to be significantly correlated with gameplay habit and with gameplay preference. Specifically, the boys tended to be

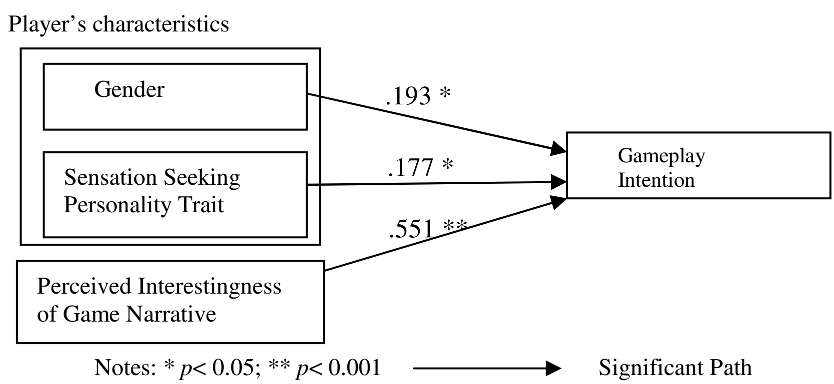


Fig. 2. Linear regression analysis of the research framework

more frequent computer game players than the girls; the most favourite gametype for the boys was shooting and that for the girls was active sports for fitness.

As shown in Table 4 and Fig. 2, the hypothesized paths from perceived interestingness of game narrative, gender and sensation seeking personality trait are *significant* in prediction of the gameplay intention. Therefore H1, H2, H3 are supported.

The present study shows that the effect of *perceived interestingness of game narrative* is significantly stronger than other factors in the framework. *Game narrative* is found to be a powerful tool to understand player's intention, consistent with the earlier findings [21]. This implies that game designers and instructional designer should put more emphasis on game narrative. The effect of *gender* also plays a significant influence on gameplay intention. The findings show that more boys tend to play the game than girls, as highlighted by [25]. This provides a need for game developer to encourage more female players. *Sensation seeking personality trait* has a significant influence on intention to play. This finding is consistent with the argument of [30, 31] who suggest that sensation seeking may lead to higher engagement in computer gameplay.

Table 5. Summary of hypothesis tests of research model

Hypotheses		ß	t-value	p-value	support
H1	Narrative→ Intention	.551	7.184	.000	Yes
H2	Gender→ Intention	.193	2.433	.017	Yes
H3	Sensation seeking personality trait → Intention	.177	2.255	.026	Yes

Notes: * p< 0.05; ** p< 0.001; the t-value is significant at p<0.05 while critical ratio values exceed 1.96.

6 Conclusion and Implications

Given the power of games to be an attractive and appealing source of entertainment and education, this study investigates the driving factors behind children's intention to play digital educational game. The results of this study provide support for the research framework presented in Figure1. Accordingly, game narrative, gender, sensation seeking personality trait are found to be important in predicting computer gameplay intention.

This study contributes to a theoretical understanding of the factors that influence entertainment-education oriented game usage. Previous behavior intention studies of computer game have focused on experiential motives such as, perceived ease of use, usefulness, and perceived enjoyment [2, 3, 4, 5, 6]. There have been a limited number of studies investigating to what extent game narrative is linked to intention to play DEGs. This present study argues that game narrative attributes to gameplay intention. The results of this study indicated this predictor is significant and can explain much of the variance in game usage.

This study also has several implications of game design: for *game developer*s, the findings imply that they should create narratives that act as a fictional interface to the game systems. The narratives mediate the player's interaction with the systems of rules via the game mechanics embedded in the narratives. Designers should focus on developing an appealing narrative for the game which will ultimately influence the gameplay intention.

Game narrative (without the support of a gameplay structure) can be applied as an extremely cost effective tool for designers to predict player's intention in the early game development. Game design ideas can be tested using narrative in the early game development process without any drain on technical resources and costs. Doing so enables game designers find problems on the design ideas before they waste money implementing something that doesn't work. Game narrative can also be used to understand users' motivations and elicit their requirements, thereby enabling the development of a realistic user model for a DEG. Using the game narrative allows users to explore as well as evaluate different aspects of the game (e.g. game context and characters). This process reveals at an early stage the areas that users like and dislike about a game narrative, providing game designer with useful data how to refine their initial ideas at a low cost.

One important concern found in this study is to consider player's characteristics such as gender and personality trait in order to produce an appealing narrative to its users. Game developers often use computer game design principles to inform the design but they often do not directly involve users in the design process [19]. This study suggests that it is important to include the users in the early stage of game design in order to develop a game that will appeal to them, for example, to encourage more female players by analyzing which game story and character female players prefer in the game. Moreover, player-centred design can be reached by using adaptive game technologies to tailor a game to individual players. By way of an example, a game can adapt to sensation seeking personality trait by aligning the game narrative with different levels of time pressure, difficulty and adventure.

From *instructional designer* perceptive of view, narrative-centred learning gaming environment can promote the deep, meaning-making activities that define constructive learning. Constructive theorists are particularly optimistic about the prospects of using narrative to learn for children [7]. Although narrative does find important in educational game development, the learning content must align with the narrative plotline for an educational game to be effective [8]. So far, there is still a need to investigate how to develop DEGs using an established pedagogy.

As game narrative and player's characteristic are important elements for effective educational game design, future research can investigate how to develop a theoretical framework for game narrative that is informed by the requirements of constructive learning and adaptive to different user characteristics (e.g. gender, personality). Besides, it will be intriguing to evaluate children's gameplay intention with different narrative presentation formats and media (e.g. audio, onscreen) to see whether the impacts observed in the current study remain.

References

1. Kaiser Family Foundation. Children and video games. Kaiser Family Foundation, Menlo Park (2002)
2. Dong, M.K.: The moderating role of locus of control on the links between experiential motives and intention to play online games. Computers in Human Behavior 25, 466–474 (2009)

3. Hsu, C.L., Lu, H.P.: Why do people play on-line games? an extended TAM with social influences and flow experience. Information and Management 41(7), 853–868 (2004)
4. Childers, T.L., Carr, C.L., Pecj, J., Carson, S.: Hedonic and utilitarian motivations for online retail shopping behavior. Journal of Retailing 77, 511–535 (2001)
5. Choi, D., Kim, J.: Why people continue to play online games: In search of critical design factors to increase customer loyalty to online contents. Cyber Psychology and Behavior 7(1), 11–24 (2004)
6. Venkatesh, V., Bala, H.: TAM 3: Advancing the Technology Acceptance Model with a Focus on Interventions (Manuscript in-preparation)
7. Dickey, M.D.: Ninja Looting for instructional design: The design challenges of creating a gamebased learning environment. In: ACM SIGGRAPH, Boston (2006)
8. Fisch, S.M.: Making educational computer games ”educational”. In: Interaction design and children, Boulder (2005)
9. Bruner, J.: The narrative construction of reality. Critical Inquiry 18(1), 1–21 (1991)
10. Champagnat, R., Delmas, G., Augeraud, M.: A Storytelling Model for Educational Games. In: The first International Workshop on Story-Telling and Educational Games, Maastricht, The Netherlands (2008)
11. Christou, I., Perdikaris, N., Tragazikis, P., Douros, A., Galani, E., Gouscos, D., Meimaris, M.: ”The Magic Potion” - An Adventure Game for Learning. In: The 2nd International Workshop on Story- Telling and Educational Games. Aachen, Germany (2009)
12. Dickey, M.D.: Three-dimensional virtual worlds and distance learning: Two case studies of Active Worlds as a medium for distance education. British Journal of Educational Technology 36(3), 439–451 (2005)
13. Halverson, R., Shaffer, D., Squire, K., Steinkuehler, C.: Theorizing games in/and education. In: The 7th international conference on Learning Sciences, Bloomington (2006)
14. Schrier, K.: Using augmented reality games to teach 21st century skills. In: The SIGGRAPH, Boston (2006)
15. Gee, J.P.: What video games have to teach us about learning and literacy. Palgrave MacMillan, New York (2003)
16. Burrow, A., More, G.: Architectural designers and the interactive audience. In: The second conference on Interactive entertainment, Sydney, Australia (2005)
17. Amory, A., Naicker, K., Vincent, J., Adams, C.: The use of computer games as an educational tool: Identification of appropriate game types and game elements. British Journal of Educational Technology 30(4), 311–321 (1999)
18. Jennings, M.: Best practices in corporate training and the role of aesthetics: Interviews with eight experts. In: The 2001 ACM SIGCPR Conference on Computer Personnel Research, SanDiego, CA (2001)
19. Waraich, A.: Using narrative as a motivating device to teach binary arithmetic and logic gates. In: The 9th annual SIGCSE Conference on Innovation and Technology in Computer Science Education, Leeds, United Kingdom (2004)
20. Malone, T.: Towards a theory of intrinsically motivating instruction. Cognitive Science 4, 333–369 (1981)
21. Prensky, M.: Digital Game Based Learning. McGraw-Hill, New York (2001)
22. Quinn, C.: Designing an instructional game: reflections on Quest for Independence. Education and Information Technologies 1, 251–269 (1996)
23. Fang, X.W., Susy, S., Chan, C.N.: An Online Survey System on Computer Game Enjoyment and Personality. HCI (4), 304–314 (2009)
24. Kafai, Y., Heeter, C., Denner, J., Sun, J. (eds.): Beyond Barbie and Mortal Kombat: New perspectives on gender and gaming. MIT Press, Cambridge (2008)

25. Ivory, J.D.: Still a man's game: Gender representation in online reviews of video games. Mass Communication & Society 9(1), 103–114 (2006)
26. Hartmann, T., Klimmt, C.: Gender and computer games: Exploring females' dislikes'. Journal of Computer-Mediated Communication 11(4) (2006)
27. Maddi, S.R.: Personality theories: a comparative analysis, 5th edn. Dorsey Press, Chicago (1989)
28. Zuckerman, M., Eysenck, S., Eysenck, H.J.: Sensation seeking in England and America: cross-cultural age and sex comparisons. J. Consult. Clin. Psychol. 46, 139–149 (1978)
29. Lankveld van, G., Spronck, P., Rauterberg, M.: Difficulty scaling through incongruity. In: Mateas, M., Darken, C. (eds.) The Fourth Artificial Intelligence and Interactive Digital Entertainment Conference, pp. 228–229 (2008)
30. López-Bonilla, J.M., López-Bonilla, L.M.: Sensation seeking and e-shoppers. Electronic Commerce Research 8, 143–154 (2008)
31. Ravaja, N., Salminen, M., Holopainen, J., Saari, T., Laarni, J., Järvinen, A.: Emotional Response Patterns and Sense of Presence during Video Games: Potential Criterion Variables for Game Design. In: The third Nordic conference on Humancomputer interaction, New York, pp. 339–347 (2004)
32. Fang, X.W., Zhao, F.: Sensation Seeking, Self Forgetfulness, and Computer Game Enjoyment. In: Smith, M.J., Salvendy, G. (eds.) Interface, HCII 2009, Part II. LNCS, vol. 5618, pp. 632–641. Springer, Heidelberg (2009)
33. Slater, M.D.: Alienation, aggression and sensation seeking as predictor of adoles-cent use of violent film, computer, and website content. Journal of Communication 53, 105–201 (2003)
34. Agosto, D.: Girls and gaming: A summary of the research with implications for practice. Teacher Librarian 31, 8–14 (2004)
35. Stephenson, M.T., Hoyle, R.H., Palmgreen, P., Slater, M.D.: Brief measures of sensation seeking for screening and large-scale surveys. Drug and Alcohol Dependence, 279–286 (2003)
36. Farrell, D., Kostkova, P., Lecky, D., McNulty, D.: Teaching Children Hygiene Using Prob lem Based Learning: The Story Telling Approach to Games Based. In: The 2nd International Workshop on Story-Telling and Educational Games, Aachen, Germany (2009)

Case Study of FISS: Digital Game Based Learning for a Broad Range of Ages

Johannes Bufe and Detlef Krömker

Institute Informatik
Goethe University Frankfurt am Main
Robert-Mayer-Str.10
60326 Frankfurt, Germany
{jbufe,kroemker}@gdv.cs.uni-frankfurt.de
http://www.gdv.cs.uni-frankfurt.de

Abstract. For several years serious games, e.g. videogames and simulations have been applied to support learning and to improve the learning process of the player. As Digital Game Based Learning (DGBL) is becoming more popular, the target audience is growing. But is serious game-supported learning a good way to address learners of all ages? And what are the characteristics of a possible game that is suitable and can be successfully used for this purpose? Prensky differentiates between two types of learners: the generation of learners as digital natives and that of digital immigrants depending on whether this generation grew up with new technologies like the Internet and videogames or not. The following article introduces and evaluates a serious game concept that should be suitable for learners of a broad range of ages, for digital immigrants as well as for digital natives.

Keywords: Game Based Learning, Serious Games, Case Study, Digital Natives, Digital Immigrants.

1 Introduction

Serious Game supported learning, known as Digital Game Based Learning, has become more popular over the last years. Edutainment games are not limited to classic learning contexts, they can also be used on video game platforms which were originally created for entertainment purposes only. Thus, for example, the Germany-based car manufacturer Daimler trains its engineers using simulations and games [1]. Language-training software based on/using games is also available and sold very well on the latest handheld entertainment systems. Thus DGBL is applied with high expectations regarding motivation and learning efforts and used by a growing audience opposed to traditional learning media [5]. The key expectation is to combine gaming and learning to create a more active experience, enhancing the learner's motivation, which induces greater learning effects. [7] With the success of Digital Game Based Learning (DGBL) the target audience has grown, including people of all ages. But every generation has its own learning

Z. Pan et al. (Eds.): Transactions on Edutainment IV, LNCS 6250, pp. 56–67, 2010.

habits and culture, as e.g. Prensky [14], [15] states. He describes the differences between people who have grown up with modern technologies, like videogames and the Internet, and those who had to "immigrate" into them. These differences have to be taken into consideration regarding serious games and their learning environments if they aim at being appealing to learners of all ages as well as using them successfully.

We propose a concept and the development of a serious game, which should be suitable for learners from both digital native and digital immigrant generations. Following this, an evaluation of the game is provided, analyzing the success depending on age and generation.

2 Digital Natives vs. Digital Immigrants

The design of a serious game is crucial to the success and learning outcomes of DGBL. It has to consider its learning contents and has to provide game mechanics that are suitable for its audience. So it is important to consider the learners characteristics, such as age or technical knowledge. This is more difficult if the audience is heterogeneous in terms of important characteristics. Thus it is essential that the design of learning media like serious games takes into consideration the difference between generations in regard to learning culture and determines whether there is an acceptance of games as an appropriate learning method [14], [15]. Many authors suggest that new technologies like computers, the Internet or mobiles changed the way a whole generation lives, thinks and learns so deeply, that traditional teaching and learning methods will not work for them. [21] Jon Tapscott described this "new generation" as the net generation, Horst Opaschowski [13] named them "generation @" and Marc Prensky referred them as digital natives. The authors state that people born after 1978 [13] - 1982 [14] belong to this new generation. Within this article we refer to them as digital natives and digital immigrants.

Prensky differentiates between a generation of learners as digital natives and as digital immigrants. He referes to people who have grown up with modern information technologies such as the Internet or videogames as digital natives. The fast progress of the development of information technology has influenced and formed the learning habits of this generation. Internet and mobiles allow a fast and direct communication which is independent of the location . New media are not sequential anymore, allowing the viewer to jump between various topics and to access more information at the same time. Especially videogames have become an important part of their culture. Prensky assumes that people born after 1980 belong to this generation. In comparison to the generations of digital natives, he defines digital immigrants. These are people who had to immigrate into newly developed technologies, while the digital natives have grown up with them. Their learning habits are often traditional, preferring classic sequentially structured media. Their teaching methods tend to ignore the needs and learning habits of the digital natives. Videogames are hardly accepted as a part of their culture and are regarded as inferior to classic learning media. The videogame industry has recently been trying to meet the needs of a broader audience by including games for digital immigrants where the created games are mostly not as complex as classic core games produced for digital natives. Is the conflict regarding learning habits between the different generations unsolvable or is it possible to create a game which is useful and suitable for both groups? In the next paragraph we present a game concept for DGBL and analyze key aspects in regard to digital natives/immigrants.

3 Serious Game: Gamedesign of the FIS-Simulation

Below we discuss the design of the serious game FISS ("Fertigungs- und Instandhaltungs- Strategie Simulation" - Production and Maintenance Strategy Simulation) especially under the conditions mentioned before. At first we analyze the initial situation, the educational objectives and the audience.

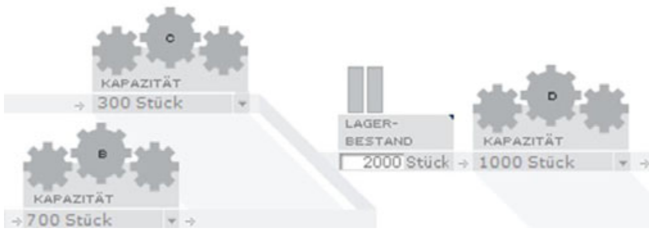


Fig. 1. Visualization of a plant

3.1 Initial Situation

In cooperation with the Goethe University Frankfurt Main, the car manufacturer Daimler updated its employee training program. Current courses were modernized by using the eLearning System Lernbar [11]. The training is now supported by interactive elements such as quizzes or videos. The content is also accessible outside the course via an online platform. The next step to improve the training is to implement the simulation of a factory, which allows the trainees to test the strategies they have learned. To work with the current Lernbar framework in on- and offline environments the simulation should be implemented using Adobe Flash. Also, the simulations should be available online after the training course has been completed. The primary objective of DGBL learning with FISS is to deepen the trainees' knowledge regarding maintenance strategies. Within FISS they should be able to test these strategies, without causing negative effects for a real factory. Therefore the trainees should manage a virtual factory and take care of different tasks, like hiring employees and assigning tasks to them, repairing and upgrading machines while trying to maximize the profits. The key to a successful learning process in DGBL scenarios is to create a state of intrinsic motivation. Thus serious games have to overcome the limitations of a simulation or of other learning software. When trying to create a suitable game it is vital to know the habits and preferences of the audience. Most of the trainees are male (ca. 95%) and have strong knowledge of computer technology. However the age of the trainees is very heterogeneous, reaching from 25 to 50 years. Thus different generations including digital natives and immigrants will play and learn with FISS. Most trainees lead a team of engineers, trying to maximize the factories' efficiency. The training takes three days and teaches advanced maintenance and employee management strategies.

3.2 Game Design Aspects

To create an immersive experience it is important to consider different aspects. We highlight Acceptance, Easy Access, Fun and Support of Learning Processes.

Acceptance: As mentioned in chapter 3, the main difference between digital natives and digital immigrants is that new technologies and videogames have become a natural part of the digital natives' life, whereas digital immigrants have had to adapt to these new technologies. Among digital natives videogames and edutainment software is accepted as a part of their culture and not seen only as a child's toy. In comparison, digital immigrants hardly accept games as cultural artifacts or appropriate learning tools [16]. But if a learning tool is not accepted it will hardly be successful. Thus it is important that FISS reaches a high acceptance among both generations of learners.

Serious Visual Design: The design of FISS has to visualize a lot of data, while keeping a less playful, more serious look. Most elements within the game are represented by abstract forms e.g. working machines as turning gears. Different states are indicated by using unique colors. Distracting elements were left out. Thus the look of FISS is tidy, similar to a schematic circuit diagram and adapted to the trainees' working environment.

Embedding FISS into a realistic context: Yet FISS could be an abstract simulation of any assembling factory. It is important to create a connection between the game and the trainee's job within the company. Thus FISS is introduced as a factory assembling axes for a specific car model with realistic parameters. All elements used within the game are labeled as their counterparts in the factory.

Realistic Plant Behavior: Most of the trainees have been working in a factory for several years and have developed a sense of how machines should behave under certain conditions. Although it is nearly impossible to match this sense, it is important that FISS fits their mental model as much as possible. Thus machine behavior is calculated by a mathematically correct model, trying to be most realistic. This should induce higher acceptance. [4]

Introduction to FISS: A serious game like FISS is based on a complex set of rules and variables. It must be ensured that the player understands the rules underlying the game and is able to handle the information offered. The training with FISS is limited to six hours allowing only one complete run of the simulation. With little time left for introduction and test runs, the player should still be able to judge game situations correctly. We based the introduction to FISS on three sets: Instruction while playing, enabling new features within the game and additional support through a haptic play board.

Instruction while playing: To create fast and easy access to the game the instruction was split into two parts. The first part consists of a short instruction at the beginning, introducing the basic elements of the game. Special features and the meaning of indexes are explained whenever needed within the game. This kind of introduction to a game should avoid overtaxing the player, while teaching the complete set of rules as mentioned in [6].

Enabling new features within the game: It is important not to ask too much of users that are new to FISS. To achieve this FISS starts with only a basic set of rules, allowing the game to work properly and offering an easy access to the game. To fully unfold the game's potential, new features are unlocked in periodic intervals. By the time the user is confronted with new options, he has practiced and has hopefully grown accustomed to the existing ones. These new features are introduced when they are unlocked (see "Instruction while playing") [20]. For example at the beginning of the game the employee management is simplified, allowing only the hiring of expensive temp workers. These can be discharged without restrictions and are qualified for every job. As the game advances the user will be able to recruit permanent employees, which are cheaper in the long term, but can only be released with a redundancy pay.

Support through haptic playboard: Most digital immigrant trainees have no experience in using games or simulations. So it can be an advantage to make use of something well known, before playing with the computer simulation only. Based on the fact that digital immigrants are used to board-games, we designed FISS as a computer supported board game in the first step. Employees are represented through pieces and allocated to a job by moving them to special sections. Custom events and actions are performed by picking "chance" cards. The design of the FISS playboard and that of computer simulation is quite similar. This parallelism of board game and computer simulation should support the adaptation to FISS for inexperienced trainees. Because playing FISS this way is time consuming, the playboard is only used in the first chapter, where the rules are limited to a basic set. Also it is not possible to play the FISS board game only, because complex calculations and visualizations are needed.

FISS as fun-to-play: Self-motivation is the key to the success of serious games [6]. Thus it is necessary to design FISS as fun-to-play in-order to create this kind of motivation. Garris [6] mentioned several design rules that are suitable in creating an entertaining serious game. We highlight challenge, fantasy as well as control and sensory stimuli.

Challenge: This game-design property means that the player is challenged within the process of playing, to reach his goals against difficulties. A challenge can happen when the game challenges the player, but also as a challenge between player's trying to achieve the best result. To increase the effect of the challenge, it is suitable to offer a reward within the game, like entering the players name in some sort of highscore table to show the achieved result in comparison to other players [12]. FISS is played within a team, while each team is responsible for its own virtual factory. Thus each team is in competition with the others, trying to maximize the factory's outcome. After a given time all teams have to present their current results to the audience of trainers and trainees. They have to judge their situation in relation to other teams and give estimates for the outcome at the end of the game. Additionally every player has a role with unique tasks within his team. The roles represent management jobs, such as factory chief or maintenance manager, and their success is measured through specific reference indices. Thus not only the teams are in active competition, but also the players are competing with each other. At the end of the training with FISS all players and teams present their score in front of the training audience. Thus every player is able to show his game efforts, which is rewarding in regards to the challenge design property.

Fantasy: Fantasy describes the ability of a game to show an imaginary world from different points of view. This allows the player to act and experiment, exploring different roles and rules within the virtual setting. Different studies suggest that DGBL using games with strong fantasy components is more appealing, inducing greater interest and learning success [12]. The player of FISS is assigned to the role of a member of the factory's leading staff. He has to act within this role and is bound to the rules of the virtual world. FISS is designed to offer a realistic experience, so all events and parts of the game are put into a realistic context with their own background. As presented in 4.2.2 the game becomes more complex in periodic intervals and every previous limitation is explained within the virtual world. Also additional events were implemented that require the player to adjust his strategy; e.g. security measures are temporarily affecting the maximum production rate.

Control & Sensory Stimuli: Effects enhancing the player's immersion and motivation can be supported through suitable sensory stimuli like animated graphics, videos or sound [18], [6]. Within a game these stimuli are often part of an interaction and not only consumed passively. The player takes control of the game causing sensory feedback. FISS uses a clear interface and animations to create a suitable and tidy look which was presented in 4.2.1. Sound effects are set aside because they are likely to disturb the learning scenario.

Supporting the Learning Process: Games have the ability to support active and critical learning [7]. The user has to reflect on his decisions and the content, while experimenting in a risk-free environment. In FISS, trainees should test the learned strategies and different options, while understanding possible risks and limits. The game design should offer enough time for decision making and reflection, allowing the trainees to choose and discuss the correct strategy. There are two options how to simulate time within a game: round based or real time. As discussed in [8], round-based game design has advantages in allowing a reflective process within the game. Thus FISS is designed as a round based game to support a reflective learning process. As mentioned before, the trainees have to present their efforts in specific intervals. The trainee has to rethink his strategies while explaining them to the audience and discussing the outcomes. Also, all trainees watch the other teams justifying their results and strategies. Thus after the presentation breaks, all teams should reflect on their strategy, trying to improve it actively while benefiting from other teams' experiences. The worst case thst happens when a team has chosen a wrong strategy and is suffering from the consequenc, is that they are unable to succeed without restarting the game. Although this is not intended to occur very often, the trainer has the option to interfere with the game. He is able to examine and undo decisions, load savegames and discuss possible mistakes with the team. Also FISS automatically saves the current game, so that the effects of possible system failures are minimized.

4 DGBL with FISS

FISS was played and evaluated within the employee training at Daimler in the 4th Quarter in 2008 and the 2nd Quarter in 2009. After a short introduction, the game was played for approximately 6 hours using the haptic playboard within the first rounds. To complete the game, the players had to pass 40 rounds. After half of the training session, after completing the 20th round, the players had to present the outcome of their virtual factory. Also they had to justify their strategies and predict the outcome of the game. After the trainees finished the game the final outcomes were presented and discussed with the audience. After the game was completed, trainers and trainees discussed the DGBL session with FISS regarding learning efforts, fun and gameplay.

4.1 Observations while Playing FISS

Most trainees started playing immediately, without having problems in understanding the games' rules and controls. All teams were keen to reach the best outcomes and highest score, by discussing different strategies and options. Most questions could be clarified by reading the digital lernbar manual, which explained new contents during the game. The haptic playboard was used extensively by most teams during the first quarter of the game, but only few teams used it for longer than 10 rounds. The presentation of the outcomes after the first half and at the end of the game were enjoyed by the audience and led to constructive discussions about possible strategies.

4.2 Feedback from the Players

After the presentation of the final results the trainees were asked to give feedback on playing and learning with FISS. All trainees agreed that playing FISS was useful and motivating, enhancing the training of maintenance strategies. The haptic playboard was criticized as time consuming and redundant when used after the first rounds. Also it was noticed that it was only possible to play the game once during the training. Thus wrong decisions at the beginning could not be undone and degraded the score at the end of the game. Most trainees wanted to play FISS again in order to aim for better scores and to try alternative strategies. Also some inconsistency between FISS and a real factory was noticed e.g. some teams fired most of their employees and stopped maintaining the machines at the end of the game. These actions would ruin a factory in the long run, but gave a short boost for the final score. There was also a debate about how realistically the machine behavior was simulated. Although most trainees judged the simulation as realistic, some trainees wanted an additional expert review on the machines' behavior. In the following we present an overview of the key design decisions and the corresponding feedback.

5 Evaluation: Is the Success of DGBL Dependent on whether the Players Are Digital Natives or Digital Immigrants?

Based on the question whether DGBL is suitable for an audience of all ages we described the concept and implementation of the serious game FISS. To initiate a successful process of learning we focused our game design on Fun to Play, Easy to Play, Acceptance, and Support of Learning Processes. In this chapter we present the evaluation of FISS under the aspects of the learner's generation and their in-game and test performance. We assume that if a serious game has a suitable game design, DGBL is able to engage learners of all ages - independent of which generation they are from.

5.1 Evaluation Design

In the afternoon before the DGBL session all trainees took a test to determine their knowledge of maintenance strategies. This pre-test was a multiple-choice test with ten items. The next day, after the training with FISS, the trainees took a second test to measure their improvement. Additionally all trainees were asked to answer evaluation questions about FISS and give personal information about

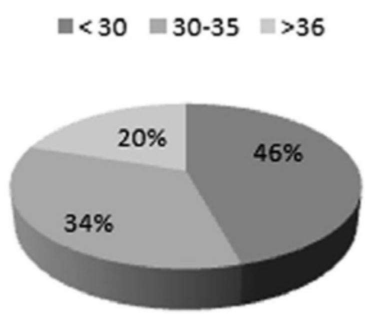


Fig. 2. Age distribution of the audience

their age or sex. The tests were linked through anonymous IDs. All trainees were assured that their data will remain anonymous and will not be forwarded to Daimler. We differentiate between three groups of trainees by age: Trainees under 30 years are part of the digital natives, while participants over 36 years are part of the digital immigrant generation. Trainees aged between 30 and 35 years are classified as part of a crossover generation between digital natives and digital immigrants.

5.2 Survey Results

The results of the survey are close to the feedback received in the discussion after the DGBL with FISS. The statement "I'm rating the training with FISS as very good overall" was rated at an average of 4.28 points on a scale from 0 to 5 points, where 5 points equals full approbation of the thesis (Sample size n = 35). Also the item "Learning with FISS was very motivating" was approved with 4.23 points. The acceptance of FISS as a suitable abstraction of a real factory was rated at an average of 3.5. Thus DGBL with FISS seems to be motivating and was accepted by the trainees as an appropriate way of learning. To determine differences between the different generations of learners, the results were related to the trainee's age as discussed above. To clarify the result we calculated the probability (r), that the rating given by a digital native trainee is below the rating from a digital immigrant by no more than 1 point. We presume that the samples are distributed normally.

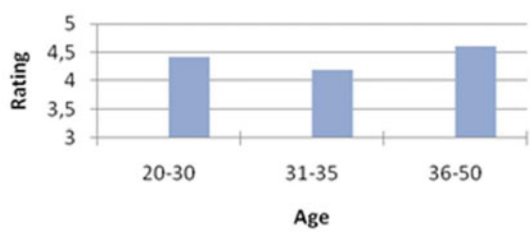


Fig. 3. I'm rating the training with FISS as very good overall [mean 4.28, r=0.79]

There is no significant correlation between the learner's generation and their survey results. All trainees, independent of their generation, appreciate learning using FISS. The r value states that the difference between the digital native and digital immigrant generation is low.

5.3 Performance Test

Next, we analyze the improvements made between the pre- and post-test. In a first step we examined the discriminatory power and difficulty level of the test items. Items with unsatisfying results were removed. An analysis of the test consistency with Cronbach Alpha results in 0.7; this is a good value in regard to the

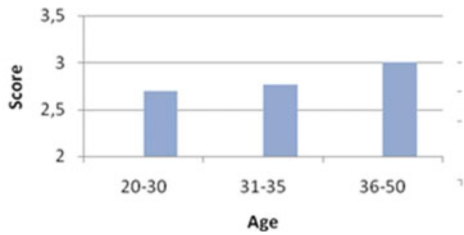


Fig. 4. Improvements between the pre- and post-test [mean 2.8]

reliability of the test [3] . Between the pre- and the post-test the participants improved by 2.8 points from 16.8 to 19.6. This is highly significant ($p<0.01$) and equates to an improvement of 16.6 percent. With a probability of 0.8 the difference between digital natives and immigrants is in a bandwidth of 2.8. There is no significant correlation between age (generation) and the improvements made during post- and pre-test - all participants improved their results.

5.4 In-Game Performance

In additionto the improvement between both tests, we measured the players' in-game performance. All trainees are rated by the fulfillment of their specified goals on a scale from 0 to 1. Though some goals proved to be more difficult, the ratings were scaled by the results of 82 game sessions. The ratings should be comparable between all players and different roles, when following this process. The result reflects the data above. There is no significant difference between digital natives, digital immigrants and the crossover generation. In this sample the digital natives achieved a slightly better rating, though the difference is within the sample's standard deviation. With a probability of 0.8 the difference between digital natives and digital immigrants is in a bandwidth of 0.2 points.

5.5 Summary

As seen in the survey and feedback evaluation, DGBL with FISS was accepted by learners of all ages. The participants state that the learning session was

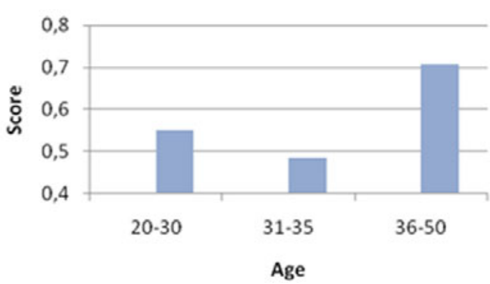


Fig. 5. In-game Performance [mean 5.3]

motivating and offered a suitable learning environment while putting the learning content into a realistic context. Also all trainees improved their skills significantly between the tests. We observed that in regard to test improvement and in-game performance no significant differences between digital immigrants and digital natives could be found. Thus learning with FISS - a suitable DGBL scenario - seems to be successful for all generations of learners in the same way.

6 Conclusion

Starting with the question of how to design a DGBL environment that is suitable for all ages, we examined the differences between digital natives and digital immigrants. Based on this analysis we developed a game-design for FISS - a serious game which should be used within the employee training program of Daimler. Though employees participating in this training are both digital natives and digital immigrants, FISS must be able to engage both groups of learners. As key aspects of a suitable game design we analyzed "Introduction to the game", "Acceptance", "Fun to play" and "Learn-process supporting elements" in particular. Following this we evaluated the training session with FISS in regard to trainee feedback, performance within the game and test improvements of the participants. FISS was rated as motivating and useful by the trainees, therefore enriching the training program of Daimler. We also observed that the trainees improved their skills by playing FISS. In the next step we compared the evaluation results with the participants' ages, where we distinguished between digital natives/immigrants and a crossover generation aged between 31 and 35 years. FISS was rated as "good" or "very good" by all participants independent of age or generation. Also, the performance within the game and test improvements differed only insignificantly between all groups of learners. Thus the game design of FISS proved successful in engaging learners of all ages and it can be assumed that DGBL supported learning is suitable for digital immigrants and digital natives in the same way. However videogames are a medium deeply connected with the generation of digital natives and foreign to most of the digital immigrant generation. Thus the question arises whether the learning habits between different generations are as fundamentally different as e.g. Prensky assumes. Though this question could not be answered within this study, it can be stated that serious game supported learning is able to enhance learning environments for different generations at the same time. To answer this question in a more general way it is necessary to investigate more DGBL environments which can be applied to generation-spanning audiences. Also it could be worth a further investigation if the game design aspects highlighted in this article are suitable for other learning contexts and types of serious games as well.

References

1. Bufe, J., Krömker, D.: FISS: Virtuelle Simulation für eine Mitarbeiterfortbildung. In: Proceedings Series of the German Informatics society (GI). Lecture Notes in Informatics (LNI), vol. P-153, pp. 67–78 (2009)

2. Callois, R.: Man, play and games. Free Press, New York (1961)
3. Cortina, J.M.: What is Coefficient Alpha? Journal of Applied Psychology 78(1) (1993)
4. van Eck, R.: Digital Game-Based Learning. It's Not Just the Digital Natives Who Are Restless: EDUCAUSE, Review, 41(2) (2006)
5. Freitas, S.: Learning in Immersive Worlds. A review of game-based learning. JISC e-Learning Programme, London (2006)
6. Garris, R., Ahlers, R., Driskell, J.: Games, Motivation, and Learning. A Research and Practice Model, Simulation Gaming 33(4), 441–467 (2002)
7. Gee, J.P.: What video games have to teach us about learning and literacy. Rev. and updated ed. Palgrave Macmillan, Basingstoke [u.a.] (2007)
8. Analysis: Turn-Based Versus Real-Time (11.06.2009), http://www.gamasutra.com/view/news/25920/Analysis_TurnBased_Versus_RealTime.php
9. Kaiser, A., Kaiser, R.: Metakognition. Denken und Problemlösen optimieren. Luchterhand, München (1999)
10. Kolb, D.A.: Experiential Learning: Experience as source of learning and development. Prentice Hall, Englewood Cliffs (1984)
11. Lernbar (2009), http://www.megadigitale.uni-frankfurt.de/et/LernBar/index.html
12. Malone, L.: Aptitude, learning, and instruction. Erlbaum, Mahwah (1987)
13. Opaschowski, H.W.: Generation @. Die Medienrevolution entläßt ihre Kinder: Leben im Informationszeitalter. Hamburg: British American Tobacco (1999)
14. Prensky, M.: Digital Natives, Digital Immigrants. NCB University Press, H. 9 (2001)
15. Prensky, M.: Digital Game-Based Learning. ACM Computers in Entertainment H.1 (2003)
16. Prensky, M., Raessens, J., Goldstein, J.H.: Handbook of computer game studies, pp. 108–121. MIT Press, Cambridge (2005)
17. Piaget, J.: Psychologie der Intelligenz. Klett-Cotta, Stuttgart (1980)
18. Rieber, L.P.: Seriously Considering Play. Designing Interactive Learning Environments Based on the Blending of Microwords Simulations, and Games, ETR&D, H. 44, 43–58 (1996)
19. Studiumdigitale (2009), http://www.megadigitale.uni-frankfurt.de/
20. Schell, J.: The art of game design. A book of lenses. A book of lenses. Elsevier/Morgan Kaufmann, Amsterdam (2008)
21. Schulmeister, R.: Is there a net generation in the house? Dispelling a mystification, vol. 5. E-learning and education, Köln (2009)
22. Vankus, P.: History and present of didactical games as a method mathematics teaching. Acta Didactica Universitatis Comenianae Mathematics (5) (2005)

Woodment: Web-Based Collaborative Multiplayer Serious Game

Viktor Wendel, Maxim Babarinow, Tobias Hörl, Sergej Kolmogorov, Stefan Göbel, and Ralf Steinmetz

Multimedia Communication Labs - KOM, TU Darmstadt,
Merckstr. 25, 64283 Darmstadt, Germany
{viktor.wendel,maxim.babarinow,tobias.hoerl,sergej.kolmogorov, stefan.goebel,ralf.steinmetz}@kom.tu-darmstadt.de
http://www.kom.tu-darmstadt.de/

Abstract. Recently, a lot of research has been investigated in the field of Collaborative Learning. On the other hand, Massively Multiplayer Online Games and browser games are a popular and commercially successful game genre. However, Browser Games commonly do not use the latest 3D web technology. Our simple idea is to combine learning concepts from the field of Serious Games with promising game technology for the development of browser-based Multi-User Learning Environments. Based on a brief introduction and state of the art analysis summarizing the latest trends in Collaborative Learning and Multiplayer (Online) Games in Chapter 1 and 2, in Chapter 3 we introduce methods and concepts of Woodment as a browser-based Serious Multiplayer Game to teach and explore a customizable learning content in a game-based and playful manner. Chapter 4 provides technical aspects and features of a first prototypical implementation of Woodment. Finally, in Chapter 5 the main achievements and first evaluation studies are summarized and further Research and Technology Development steps are pointed out.

Keywords: Serious Games, Digital Educational Games, Collaborative Learning, Multi-User Browser Game.

1 Motivation

Millions of people interact with each other every day using online virtual environments known as Massively-Multiplayer Online Role-Playing Games (MMORPGs). The average MMORPG gamer is 26 years old and spends about 22 hours each week in these online worlds [1]. A current trend in the video game industry is towards Browser Games, which are mostly Massively-Multiplayer Online Games (MMOGs) [2] running in a browser. With first Browser Games emerging, which make use of 3D-graphics technology, thus increasing the attractivity of Browser Games, it seems sound to utilize this technology to create platform-independent, attractive Serious Games.

Z. Pan et al. (Eds.): Transactions on Edutainment IV, LNCS 6250, pp. 68–78, 2010.

In this paper the authors introduce a new approach to Serious Games based on a 3D virtual world game running in a browser. The Woodment[1] project based on the Unity3D[2] game engine is such an environment including a collaborative Multiplayer Serious Game. It has been created in a practical course on game-based edutainment applications and is currently being further developed. We used design principles and patterns of commercial MMOGs [3,4,5] to increase the user's long-term motivation. As it runs in a browser, Woodment's architecture is platform-independent, thus being accessible for a wide target audience. The use of a Multi User Virtual Environment (MUVE) like Woodment has a lot of advantages, like the freely customizable learning content, an immersive experience, a strong social context, an interactive environment, and the possibility to learn with others in a collaborative way.

2 Related Work

A lot of research has been done in the field of Serious Games during the last years. In [6] Encarnação provides an encouraging overview over the chances and the future of Serious Games. Computer- and video games include a lot of fundamental gaming principles [7,8,9] which, combined with a scientific approach from the field of Computer-Supported Collaborative Learning (CSCL), can systematically increase the motivation to learn and teach knowledge of specific subjects. Among others, Prensky [10], Herz [3], and Mansour and El-Said [11] campaign for a merging of computer games and learning to Digital Educational Games (DEGs) and state first positive results. McFarlane et al. [12] warn of mismatches between game and curriculum content and provide interesting design issues concerning DEGs.

There are manifold examples of Serious Games and Digital Educational Games (DEGs), which contributed to an increasing public awareness for Serious Games. However, the majority of those games are single player games like 80Days [13] or the Eduventure II [14], based on the Oblivion[3] game engine, which is a good example of a Serious Game making use of a quite up-to-date 3D game engine. Among the multiplayer games, most approaches are based on modified game engines which originally were created for pure fun games, like the Neverwinter Nights 2[4] editor, the CryEngine[5] or the Unreal3[6] Game Engine. Apart from these, some DEGs have been placed in MUVEs like Second Life[7]. Especially concerning (Massive) Multiplayer Online Games a lot of research has been investigated. Achterbosch et al. provide an interesting overview concerning MMORPGs. Steinkühler [15], Delwiche [16], Childress and Braswell [17],

[1] www.woodment.com
[2] http://unity3D.com
[3] http://www.oblivion-game.de/
[4] www.atari.com/nwn2
[5] www.crytek.com
[6] unrealtechnology.com
[7] secondlife.com

and Voulgari and Komis [18] each provide very promising experimental examinations of MMO(RP)Gs for use as Serious Games. Another promising Serious Game settled in Second Life is Geoworlds [19].

In the field of collaboration, Johnson et al. [20] investigated cooperation and collaboration among students, whereas Dillenbourg [21] offers a promising overview over collaborative learning. Based on these thoughts, Hämäläinen et al. [22] investigated the effects of collaboration in a 3-D virtual game environment. Zagal et al. [23] adopted collaborative gaming aspects from board games and created useful design guidelines for the creation of collaborative multiplayer games like "Players must be able to trace payoffs back to their decisions" Furthermore, they warned of pitfalls in the creation of collaborative multiplayer games like "to avoid the game degenerating into one player making the decisions for the team, collaborative games have to provide a sufficient rationale for collaboration" or "players need to care about the outcome and that outcome should have a satisfying result". In Woodment we adopted these guidelines among others. First examples of collaborative Multiplayer Games [24,25,26] indicate the effectiveness of collaborative learning in MOGs.

3 Game Concept

When registering with Woodment via the web platform, users first select an avatar to represent themselves in the virtual world. This avatar is developed throughout the games played and visualizes a player's progress via a level and an Experience Points (XP) display similar to the mechanisms in ordinary RPGs.

Woodment represents an educational game, thus it contains educational as well as gaming elements. The gaming elements are designed to increase the players' motivation through fun. The learning elements are included in a way such that learning does not make a counterpart to the gaming elements but are included rather seamlessly. In Woodment, answering questions, which is a learning element, directly influences the gaming part of the game by providing bonuses for the one who solved the question and increasing his/her avatars's level which describes another additional motivation.

3.1 Mission

A Woodment match is played team versus team, each with three players. The players have to manage a virtual logging company. The human resources manager for example has to employ workers, who search for trees, cut them down and carry the wood to the team's main house. Workers are paid from the team's shared capital, as well as the occupations of the other character classes. Team behavior is supported by the fact, that each team owns a collective balance of the games resources (wood and gold). When starting a game, the players must select one of the following character classes: Risk Manager, Human Resources Manager or Procurement Manager. There can only be one player of each character class within a team. As the three different roles offer completely different tasks and challenges inside the team, they support a collaborative approach towards learning according to [27]. The main goal in Woodment is to collaborate as a team in order to lead a virtual enterprise, that tries to lumber the island as fast as possible. As soon as all trees are cut down, the game is over and the winning team is the one that gathered more wood.

The possibility to advance one's virtual character from game to game is an additional motivation to repeatedly play the game. Furthermore, players can unlock awards by reaching a specific level or other achievements like for example more than ten hours playing time. These awards are displayed within the web platform as well as ingame.

3.2 Gameplay

The player views his/her avatar in a third person perspective and leads it through the virtual world using the keyboard and the mouse. Woodment uses near-realistic 3D graphics. The gaming area is bounded in a natural way, as it takes place on an island. Woodment players can engage in the following activities:

- explore the island's 3D world (e.g. look for hidden learning content)
- manage the company
- react to unexpected events
- fight for the victory of the team (e.g. by disrupting the other team's workers)
- communicate with others via in-game chat
- level up (by answering questions collaboratively or alone
- create learning content using the web interface

3.3 Communication

As the most important tool for communication, the in-game chat also represents a crucial design element for collaboration. It should be easy to exchange ideas, strategies and knowledge. Woodment implements two different layers of real-time messaging: the Global Chat and the Team Chat, the latter only visible to members of one's own team. Text-based communication can lead to increasing reflection, because one can take a look at older messages any time. It is also a fairly slow kind of communication, which can be very useful to focus on certain aspects of a discussion and to understand every argument of a discussion as explained in [28].

3.4 Graphical User Interface

The Graphical User Interface is divided in five parts as shown in Figure 1. On the top left side, the statistics of the team members are displayed (name, team role, level and XP). On the top right a box containing information over recent game events is placed. In the top center, the global variables like available gold or the comparison of cut down timber is displayed. On the bottom left side, a chat window is placed with tabs for a global chat, a team chat or both. On the bottom right side, the players statistics provide information about his/her current level

Fig. 1. Woodment, prototypical implementation

and an experience bar visualizing the progress towards the next level. In the bottom center, various buttons for the role-specific actions are placed. All these elements are permanent GUI elements and are placed over the 3D-world. If a question is triggered or a player wants to make some role-specific adjustments by clicking one if the buttons on the bottom, an additional window is displayed containing the question or the adjustment possibilities.

When moving through the world, other players can be seen as well as information about their name and level. This way, other players can easily identify one's knowledge and learning strength by looking at the level. These information are also visible on the web-interface. This supports the formation of balanced teams, clans, guilds and other self-organized groups and has a greater emotional value for users, than a plain score-list [3].

3.5 Collaboration

Collaboration is one of the key elements of Woodment. The teams compete for victory but winning is only possible as a team. Therefore, every member of a team has to work with its team members. The players can choose between improving their own performance or the team's many times (like for example spending game money for a personal benefit or for the company). Concerning learning, the players also are never forced to collaborate, but can always choose to act competitively or on their own. Most players will realize though, that they fail when working alone and in most cases it will be easier to solve questions as a team. Woodment users should learn about the benefits from collaboration. For this reason, the game rewards collaborative behavior. For instance, solving (difficult) questions as a team is generally higher rewarded than solving (easy) questions alone.

3.6 Learning

Multiple-Choice Platforms are placed at certain points of the gaming area. Those platforms serve as spawn points for Multiple-Choice Questions (MCQs) which the players can trigger and solve on their own or collaboratively (see 1). Thereby, players can gain experience points and gold. The question window includes a "Learn"-Button, that can display background knowledge, and furthermore an information panel, showing the team members that take part in answering it. Every time a player triggers an MCQ, his/her team members can decide whether they want to collaborate with him. If a question is solved as a team, the amount of rewarded experience points is significantly increased. Considering the complexity and the amount of knowledge needed to learn the given content, the players have to decide for themselves, whether they want to solve MCQs collaboratively or alone. Players receive knowledge when playing the game in two ways. Firstly, when a question can not be solved, the players are provided with the necessary background information to learn the content via the "learn"-button. As solving a question is rewarded in the game, there is always the motivation to learn something in order to be able to answer the question correctly on the next try. Furthermore, when discussing a question in a group, players can acquire knowledge from their fellows and additionally they can practise communication skills like debating or negotiation.

3.7 Authoring

The virtual world of Woodment is a platform, that needs to be filled with learning content. We offer an intuitive web framework, that has hardly any restrictions for a tutor to insert User Generated Content (UGC). Tutors can insert content via the web portal[8] in form of scripts, PDFs, video- and audio files or images. Further, they can include their own questionnaires in form of MCQs. When starting a server, one or more sets of learning content and related questions can be selected to be available in the game, thus forming the learning focus of the game. This way learning content and questions can easily be re-used and, if wished so, even made accessible to others.

[8] http://woodment.com/

The creation of more sophisticated learning content for virtual worlds as web applications needs standardized data formats and modular course-parts. However, combined with an authoring tool for Serious Games (like StoryTec [29]), the creation of more profound learning content variations for special teaching purposes is conceivable and commercially promising.

4 Prototypical Implementation

Woodment was designed to be platform independent. For this reason we use the Unity3D Game Engine which provides a browser plugin allowing to display content of the Unity3D game engine inside the browser window. The minimal system specification for Woodment requires a computer with an average hardware of today (2010) with an average graphics card and a browser that supports JavaScript and the freely available Unity Web Player[9], like Internet Explorer, Firefox, Safari and almost every other Mozilla-based browser.

4.1 Networking

As mentioned before, we used the Unity3D Engine[10] to implement the interactive environment, which allows us to make use of an integrated networking API for MOGs. Moreover, a connection with an Apache[11] server hosting a MySQL[12] database containing learning content and MCQs is possible. Whenever a client triggers an MCQ, an HTTP POST request is sent to the Apache server. A CakePHP[13] framework handles that request, loads the needed question data from the education database and sends a resulting XML file back to the client. The XML file is then used by the client (see Figure 2) to display an MCQ window within the 3D environment. Woodment's web portal is programmed using the open-source CakePHP framework utilizing the MVC (Model View Controller) design pattern.

4.2 AI

All Non-Player Characters (NPCs) are navigated through the online world using a state-based game Artificial Intelligence (AI). We implemented Finite State Machines because they are easy to code, easy to understand and debugger-friendly as shown in [30]. Pathfinding is solved using an out-of-the-box A* pathfinding system[14]. This system is easy to set up, it automatically scans the environment and creates a path network.

---

9 http://unity3d.com/webplayer/
10 http://unity3d.com/unity/
11 www.apache.org
12 www.mysql.com
13 http://cakephp.org/
14 www.arongranberg.com/unity/a-pathfinding/

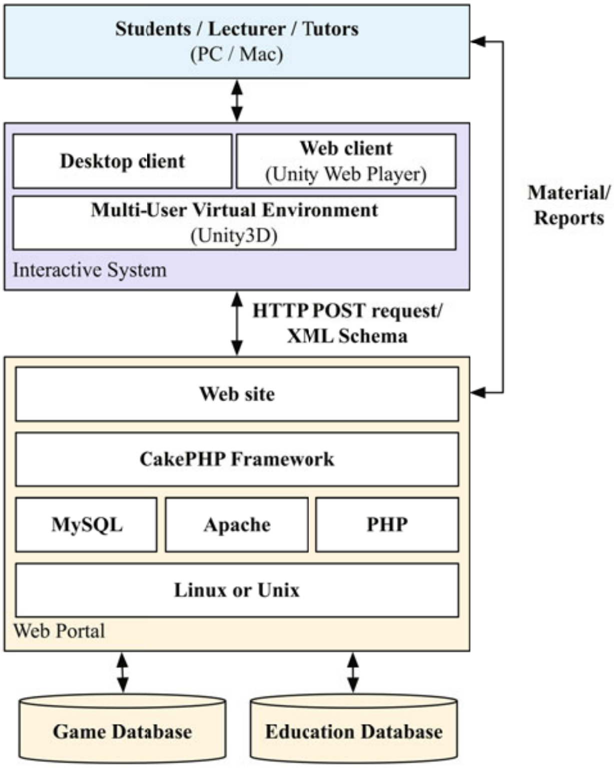


Fig. 2. Woodment architecture

4.3 Architecture

Architectures for MOGs have to be constructed in a way, that reduces latency and satisfies the player's expectations concerning this issue [31]. Traditionally, most MOGs implement a client-server architecture with only one authoritative server [32]. Woodment also uses a client-server architecture, but we use a non-authoritative approach to structure network gaming. In an authoritative approach, the server has to calculate the effect of every user input and send the updated game status to each client. Our network engine handles all input locally and updates only very important data via the server. The task of Woodment's server is mainly to transfer messages between the different clients, therefore reducing server load.

An ad-hoc game server can be created either by a tutor or a player, the creator of the server then acts as host. The server automatically gets registered with a central Unity3D server at which potential clients can lookup games with free slots. This way other players can find open game servers. Such a client-server architecture also creates anonymity for users, as clients are never connected directly and no IP addresses have to be revealed. From a developer's point of view, this approach is easier to implement than an authoritative server, because

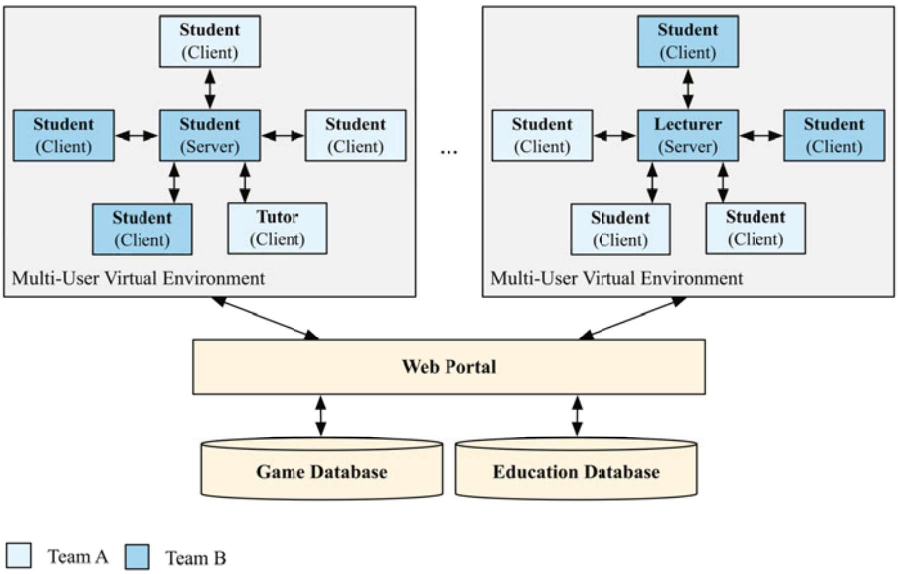


Fig. 3. Client-server network example

the clients calculate their own physics and events and send only the final result to the server which also reduces network traffic. The server-client architecture of Woodment can be seen in Figure 3.

5 Conclusion

The authors introduced Woodment as a browser-based Multiplayer Serious Game based on latest game technology including various collaborative learning concepts and a customizable learning content. A first prototype has technically proven that using 3D web technologies for collaborative Multiplayer Serious Games is a promising approach. First evaluation studies encourage to further development of the game. Especially the fact, the learning and assessment parts directly integrated within the 3D world as well as the possibility for online-collaboration are very well received. Based on that, the next steps will be to decorate the game world with more attractive features in order to increase fun and to create an authoring environment in order to provide teachers/trainers using Woodment with the ability to include customizable learning content. To make the game applicable for a lot of additional fields of knowledge we intend to provide potential authors with various types of tasks beyond the scope of multiple choice tests like a support of math related questions or gap texts. Moreover, further evaluation of the effectiveness of the implemented collaborative concepts is necessary. Therefore, an evaluation with a larger number of users will be performed for an assessment of playability and effectiveness of learning in Woodment.

References

1. Yee, N.: The demographics, motivations and derived experiences of users of massively-multiuser online graphical environments. PRESENCE: Teleoperators and Virtual Environments 15, 309–329 (2006)
2. Schultheiss, D.: Long-term motivations to play MMOGs: A longitudinal study on motivations, experience and behavior. In: DiGRA, pp. 344–348 (2007)
3. Herz, J.C.: Gaming the system: What higher education can learn from multiplayer online worlds. The Internet and the University: Forum, 169–291 (2001)
4. Bartle, R.: Designing virtual worlds. New Riders Games (2003)
5. Mulligan, J., Patrovsky, B., Koster, R.: Developing online games: An insider's guide. Pearson Education, London (2003)
6. Encarnação, M.: On the future of serious games in science and industry. In: Proceedings of CGames (2009)
7. Gee, J.P.: What video games have to teach us about learning and literacy. Comput. Entertain. 1, 20 (2003)
8. Greenfield, P.M.: Mind and media: The effects of television, video games, and computers. Harvard University Press, Cambridge (1984)
9. Squire, K.: Video Games in Education. International journal of intelligent simulations and gaming 2(1), 49–62 (2003)
10. Prensky, M.: Digital game-based learning. Comput. Entertain. 1(1), 21 (2003)
11. Mansour, S., El-Said, d.M.: Multi-players role- playing educational serious games: A link between fun and learning. The International Journal of Learning 15(11), 229–240 (2008)
12. McFarlane, A., Sparrowhawk, A., Heald, Y.: Report on the educational use of games: an exploration by teem of the contribution which games can make to the education process. Technical report, TEEM, St Ives, Cambridgeshire, UK (2002)
13. Göbel, S., Mehm, F., Radke, S., Steinmetz, R.: 80days: Adaptive digital storytelling for digital educational games. In: Proceedings of the 2nd International Workshop on Story-Telling and Educational Games (STEG 2009). CEUR Workshop Proceedings, vol. 498 (August 2009)
14. Wechselberger, U.: The eduventure ii. an approach to educational game design. In: CW 2008: Proceedings of the 2008 International Conference on Cyberworlds, Washington, DC, USA, pp. 397–404. IEEE Computer Society, Los Alamitos (2008)
15. Steinkuehler, C.A.: Learning in massively multiplayer online games. In: ICLS 2004: Proceedings of the 6th international conference on Learning sciences. International Society of the Learning Sciences, pp. 521–528 (2004)
16. Aaro: Massively multiplayer online games (mmos) in the new media classroom. Educational Technology & Society 9(3), 160–172 (2006)
17. Childress, M., Braswell, R.: Using massively multiplayer online roleplaying games for online learning. Distance Education 27(2), 187–196 (2006)
18. Voulgari, I., Komis, V.: Massively multi-user online games: The emergence of effective collaborative activities for learning. In: DIGITEL 2008: Proceedings of the 2008 Second IEEE International Conference on Digital Game and Intelligent Toy Enhanced Learning, pp. 132–134. IEEE Computer Society, Los Alamitos (2008)
19. Russell, D., Davies, M., Totten, I.: Geoworlds: Utilizing second life to develop advanced geosciences knowledge. In: 2008 Second IEEE International Conference on Digital Games and Intelligent Toys Based Education, November 2008, pp. 93–97 (2008)

20. Johnson, D.W., Johnson, R.T., Holubec, E.J.: Cooperation in the classroom. Interaction Book Co. (1988)
21. Dillenbourg, P.: What do you mean by collaborative learning? In: Dillenbourg, P. (ed.) Collaborative-learning: Cognitive and Computational Approaches, pp. 1–19. Elsevier, Oxford (1999)
22. Hämäläinen, R., Manninen, T., Järvelä, S., Häkkinen, P.: Learning to collaborate: Designing collaboration in a 3-d game environment. The Internet and Higher Education 9(1), 47–61 (2006)
23. Zagal, J.P., Rick, J., Hsi, I.: Collaborative games: lessons learned from board games. Simul. Gaming 37(1), 24–40 (2006)
24. Rodriguez, M.: A sustainable development simulation by a business school. In: Learning with Games (2007)
25. Zea, N.P., Sánchez, J.L.G., Gutiérrez, F.L., Cabrera, M.J., Paderewski, P.: Design of educational multiplayer videogames: A vision from collaborative learning. Adv. Eng. Softw. 40(12), 1251–1260 (2009)
26. Dempsey, K.B., Brunelle, J.F., Jackson, G.T., Boonthum, C., Levinstein, I.B., McNamara, D.S.: Miboard: Multiplayer interactive board game. In: 14th International Conference on Artificial Intelligence in Education (2009)
27. Fidas, C., Komis, V., Avouris, N.: Heterogeneity of learning material in synchronous computer-supported collaborative modelling. Comput. Educ. 44(2), 135–154 (2005)
28. Veerman, A.L.: Computer-supported collaborative learning through argumentation. In: Unpublished Doctoral Dissertation, University of Utrecht, The Netherlands (2000)
29. Mehm, F., Göbel, S., Radke, S.: Authoring story-based digital educational games. In: ICIDS 2009, 2nd International Conference on Interactive Digital Storytelling (2009)
30. Bourg, D.M., Seemann, G.: AI for game developers. O'Reilly Media, Inc., Sebastopol (2004)
31. Claypool, M., Claypool, K.: Latency and player actions in online games. Communications of the ACM 49(11), 45 (2006)
32. Smed, J., Kaukoranta, T., Hakonen, H.: Aspects of networking in multiplayer computer games. The Electronic Library 20(2), 87–97 (2002)

Learning with Virtual Reality: Its Effects on Students with Different Learning Styles

Elinda Ai-Lim Lee[1,2], Kok Wai Wong[1], and Chun Che Fung[1]

[1] School of Information Technology, Murdoch University, South Street Campus, Murdoch, Western Australia 6150

[2] FSKM, Mara University of Technology Sarawak Branch, Kota Samarahan Campus, 94300 East Malaysia

{elinda.lee,k.wong,l.fung}@murdoch.edu.au

Abstract. This study aims to investigate the effects of a virtual reality (VR)-based learning environment on learners with different learning styles. The learning outcomes were measured cognitively through academic performance, and affectively through perceived learning effectiveness and satisfaction. A pre-test-posttest design was employed for this study. A total of 232 students from four randomly selected co-education secondary schools in a city of East Malaysia participated in this study. There was no significant difference in the cognitive and affective learning outcomes for students with different learning styles in the VR-based learning environment. This shows that the VR-based learning environment offers promise in accommodating individual differences pertaining to learning styles.

1 Introduction

In recent years, VR has gained much attention as an alternative approach to traditional learning experiences in schools and colleges. This is mainly because of its ability to provide a highly interactive virtual environment that resembles the real world where user could experience a sense of "being there" when interacting with the virtual objects [1-5]. Research has shown an encouraging array of positive learning outcomes with VR such as better learning in geosciences [6], and better understanding in geometry [7]. In fact, VR has revolutionized how people of all ages learn and work. It has the potential to facilitate the acquisition of higher order thinking and problem solving skills [8]. However, there have been few studies that investigate individual differences in the use of VR [9]. Thus, much research is needed to investigate the impact of VR on learners with different aptitudes [8, 10]. Hence, this study aims to provide an answer to this question: Could VR accommodate learners with different learning styles?

2 Learning Style and VR

People learn in different ways according to their preferred learning style [11]. Kolb [11] defines learning style as one's preferred methods of perceiving and processing

Z. Pan et al. (Eds.): Transactions on Edutainment IV, LNCS 6250, pp. 79–90, 2010.

information based on experiential learning theory. A theory that propagates learning through experience and by experience [12]. It is through the transformation of experience that knowledge is created. Kolb [11] divides the learning process cycle into four learning modes: concrete experience (CE), reflective observation (RO), abstract conceptualization (AC), and active experimentation (AE). His experiential learning model articulates learning is an iterative process that generally begins with a concrete experience, which is followed by reflecting upon what have been observed in these experiences, then to assimilate and integrate conclusions into a theory by abstract conceptualization, and finally to test and apply new theories in new situations. Each individual is most likely to feel most comfortable in one of these four learning modes based on his or her preference along two primary dimensions: the concrete-abstract dimension and the active-reflective dimension [11]. In other words, in perceiving or taking in new information, people characteristically choose between a concrete (feeling) or an abstract (thinking) approach; in processing what they take in, they tend to choose between an active (doing) or a reflective (watching) approach[11]. By plotting the preference along these two primary learning dimension continua, Kolb identifies four types of learning styles: accommodator, assimilator, diverger and converger as shown in Fig. 1.

Accommodators prefer CE and AE. They like doing and experiencing things. They are risk-taker; tend to solve problems in an intuitive, trial-and-error manner; and rely on others for information. Assimilators, on the other hand, prefer RO and AC. They learn by watching and thinking. They are more concerned with abstract concepts and

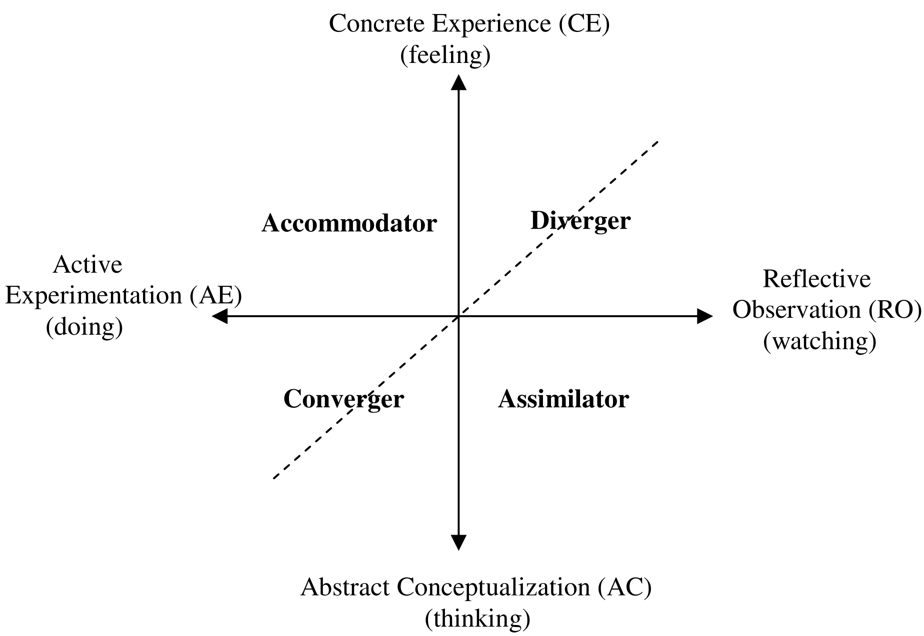


Fig. 1. Kolb's learning styles and learning modes (Adapter from Kolb, 1984)

ideas; good at putting information into a logical form; excel in inductive reasoning; less interested in people and prefer to work alone. Divergers prefer CE and RO. They learn by feeling and watching. They are imaginative; good at seeing things from different perspective; creative; emotional; sensitive to people; and have a broad cultural interest. Convergers, on the other hand, prefer AE and AC. They learn by doing and thinking. They tend to be a problem solver and decision maker; good at finding practical uses for ideas and theories; are generally deductive in their thinking; and relatively unemotional.

Bell and Foyler [13] posit that, experience – the main feature of VR is of great benefits to all learning styles, which means it could provide support to all four of Kolb's learning characteristics. Bricken [14] has also theorized about VR as a tool for experiential learning because VR supports active construction of knowledge through experiment, allows learner to experience the consequences and then to choose and apply the knowledge. Likewise, Fizpatrick [15] asserts that VR learning systems embrace the concepts and models of learning from experience by immersing learners in interactive 3D computer generated world which very closely emulates operations and equipment found in real world.

Chee's [16] - C-Visions (Collaborative Virtual Interactive Simulations), Müller and Ferreira's [17] MARVEL (Virtual Laboratory in Mechatronics) and Chen, Toh & Wan's [18] VR program for novice car driver have provided example applications of how VR can be designed to support Kolb's model of experiential learning. According to Chee [16], the first-person learning experience afforded by virtual environment allows learners to directly experience things that they seek to learn and they have the autonomy and control over their own learning experience. The virtual representations would help learners to concretize ideas, and would help what is otherwise unimaginable, imaginable and experienceable. Ultimately, learners could generalize their learning experiences to form appropriate rules and abstractions for the knowledge learned. VR is capable to support the experiential learning theory because it is able to provide "here-and-now" experience to test theories as well as giving instant feedback to change these theories [17]. Chen, Toh and Wan [18] have also found that their virtual environment that mimics the real world road scenarios provides concrete experience for learners to explore actively, and at the same time the text and image materials presented require reflective observation and abstract conceptualization.

Various learning style models have been developed by scholars such as VARK learning model, Honey and Mumford learning model and Felder-Silverman Learning model. As VR and Kolb learning style model are directly related, Kolb Learning Style Inventory that categorizes one's learning style based on the experiential learning theory was used in this study.

3 Research Objectives

The purpose of this research was to investigate the learning effects of VR on learners with different learning styles. The learning outcomes were measured cognitively through academic performance, and affectively through perceived learning effectiveness and satisfaction. Academic performance was measured through a summative assessment while perceived learning effectiveness and satisfaction were measured subjectively through a questionnaire. A VR software program, V-Frog™ was used as the VR learning material [19]. The specific objectives of this research are:

1. To determine the difference in the cognitive learning outcomes for students with different learning styles when learning with VR
2. To determine the difference in the affective learning outcomes for students with different learning styles when learning with VR

4 Research Hypotheses

In pursuance of the research purpose and objectives, the following four hypotheses were formulated for testing. A detailed description of how the learning styles were categorized is elaborated in section 5.3.4.

H1: There is no significant difference in the performance achievement for accommodator learners and assimilator learners.
H2: There is no significant difference in the overall improvement in performance for accommodator learners and assimilator learners.
H3: There is no significant difference in the perceived learning effectiveness for accommodator learners and assimilator learners.
H4: There is no significant difference in the perceived satisfaction for accommodator learners and assimilator learners.

5 Methodology

5.1 Research Design

A pretest-posttest design was used employed in this study. Participants were randomly assigned to take part in the experiment based on intact classes. Participants followed a lesson on frog anatomy with a desktop VR software program. All participants were required to sit for a pretest and a posttest, to answer the Kolb Learning Style Inventory and a set of questionnaire. Performance achievement was measured by the posttest scores, and the overall improvement in performance was measured by the gain scores (posttest scores minus pretest scores).

5.2 Population and Sample

The population was Form Four science students, aged between 15 and 17 years old of any co-education secondary schools that are well-equipped with multimedia computer laboratories in a city of East Malaysia. These students were chosen because they were within the targeted population as they have started to learn biology in Form Four. Four different co-education secondary schools were randomly selected. For each selected school, one to three intact classes were randomly chosen to participate in this study.

5.3 Instruments

5.3.1 Pretest and Posttest

Both pretest and posttest were similar in content but the order of the questions was different to avoid the set response effect. The tests include questions regarding frog anatomy for the modules covered in this study. The questions include sentence completion with the correct word(s); organ labeling and drawing; and multiple-choice questions. Content validity of these tests was determined by expert judgment. Three subject matter experts were requested to review the test questions and make a judgment about how well these items represent the intended content area. A pilot study was carried out in one co-education secondary school from the same city with forty seven randomly selected Form Four science students to obtain information that was useful to improve these tests. These included the item difficulty index, item discrimination index, and internal consistency measure. Item difficulty index is the proportion of students who answered an item correctly whereas item discrimination index measures how adequately an item discriminates between high scorer and low scorer on an entire test [20]. Six items were deleted in which five items were deleted due to poor discrimination and one was deleted as it had a low corrected item-total correlation ($r = 0.010$). As a result, the final version of the pretest and posttest contains 32 items with an alpha coefficient of 0.846. The item difficulty index was ranging from 0.27 – 0.85 which was of moderate difficulty [21].

5.3.2 Perceived Learning Effectiveness

Perceived learning effectiveness refers to the learning quality experienced by the participants. Based on the instruments of Benbunan-Fich & Hiltz [22], Mark, Sibley, & Arbaugh [23] and Martens, Bastiaens, & Kirschner [24], an eight-item instrument was developed to measure perceived learning effectiveness on the issue of identification, and integration and generalization of the lesson material. A five-point Likert scale ranging from (1) strongly disagree to (5) strongly agree was used to measure perceived learning effectiveness. Sample questions are: I learned a lot of factual information in the topics; I learned to identify the main and important issues of the topics; and I was able to summarize and conclude what I learned. The Cronbach's alpha coefficient for this instrument was 0.867.

5.3.3 Satisfaction

Students' perceived satisfaction in a desktop VR-based learning environment was measured using seven items adapted from Chou & Liu [25]. The original instruments have eight items with an alpha coefficient of 0.8625. This seven-item instrument with a five-point Likert scale with (1) strongly disagree and (5) strongly agree for measuring satisfaction has an alpha coefficient of 0.862. Sample questions are: I was satisfied with this type of computer-based learning experience; I was satisfied with the immediate information gained; and I was satisfied with the overall learning effectiveness.

Ideally, the Cronbach's alpha coefficient of a scale should be greater than 0.7 [26]. Thus, all instruments have a good level of internal validity as measured by the Cronbach's alpha.

5.3.4 Kolb Learning Style Inventory

The Kolb Learning Style Inventory Version 3.1 (KLSI 3.1) was used to categorize the learning style of each participate. The internal consistency for the scale scores of KLSI 3.1 is within the range of 0.52 – 0.84 [27].

Participants were required to complete 12 sentences that describe learning. Each item has four endings and the participants were required to rank the endings for each sentence according to how well he or she thinks each ending describes the way he or she learned.

Based on the Learning Style Inventory scoring, the scores for each learning phases: abstract conceptualization (AC), concrete experience (CE), active experimentation (AE) and reflective observation (RO) were first obtained for each participant. Then, CE score was subtracted from AC scores, and RO score was subtracted from AE scores to have two combination scores. These two combination scores were put on the Learning Style Type Grid to determine the participant's learning style, i.e., accommodator, assimilator, diverger or converger. (see Fig.1). On the Grid, accommodator falls on the top left quadrant; diverger falls on the top right quadrant; converger falls on the bottom left quadrant; and assimilator falls on the bottom right quadrant.

However, for this study, based on the method of Chen, Toh & Wan [18], instead of categorizing into four learning styles: accommodator, assimilator, diverger and converger, a dash diagonal line was introduced to equally separate the grid into two halves as shown in Fig. 1. Any diverger learner or converger learner with the two combination scores that fell below the diagonal line was classified as an assimilator learner. Likewise, if the two combination scores fell above the diagonal line, the participant was classified as an accommodator learner. Thus, assimilator included learners who fulfilled the Kolb's definition of assimilator, diverger learners with stronger Kolb's characteristics of RO than CE, and converger learners with stronger Kolb's characteristics of AC than AE [18]. On the other hand, accommodator included learners who fulfilled the Kolb's definition of accommodator, diverger learners with stronger Kolb's characteristic of CE than RO, and converger learners with stronger Kolb's characteristic of AE than AC [18].

5.4 Software

A desktop virtual reality program, V-Frog^{TM}, was used to provide the virtual learning environment to students. This software was developed and supplied by Tactus Technologies, Inc., New York. This virtual reality-based dissection simulator was developed using virtual surgery technology. Students can have hands-on learning experience with V-Frog^{TM}. They can cut, pull, probe, and examine a virtual specimen, as they would with a real frog. Thus, each dissection is different, reflecting the individual work of each student. Actions are repeatable and the content presentation is nonlinear. In each specimen window, there are viewpoint manipulation tools for students to rotate, slide and zoom the specimen. There is also a reset button to reset the position of the specimen. Additionally, in some specimen windows, dissection tools such as scalpel and tweezer for students to cut and peel the skin are provided. Moreover, there are also query tool that allows students to get information about a part of the specimen; magic wand tool that activates and brings parts of the specimen to life;

Fig. 2. Virtual frog dissection (Courtesy of Tactus Technologies)

and probe tool that examines an orifice in the specimen. Besides, a virtual endoscopy can be conducted with the endoscoping tool to explore the entire alimentary canal. There is also a V-Frog™ lab report to guide students through all the modules, highlighting key points and relationships. The existence of lab report icon on the screen indicates to students that information on the current screen can assist them to complete their lab report successfully.

5.5 Data Collection Procedures

Two weeks before the treatment, respondents were given a pretest regarding frog anatomy and the KLSI 3.1 to answer. Three modules of frog anatomy were selected for this study: Internal Anatomy, Digestive System and Circulatory System. Just before the treatment, respondents were given training on how to use the V-Frog™ software program. Immediately after the treatment, which took about 1.5 hours, the respondents were given the posttest and questionnaire to answer. The contents of the questionnaire involve background information, perceived learning effectiveness and satisfaction questions. A gap of two weeks between the pretest and the posttest was for the purpose of reducing the pretest sensitization threat.

6 Data Analysis

Frequency and proportion were used for descriptive statistics. Independent-samples t-test was used to determine the difference in the pretest, posttest, gain scores, perceived learning effectiveness and satisfaction for accommodators and assimilators.

7 Results

7.1 Distribution of Learners

Out of 232 students, 22 of them did not fully complete all instruments, that is, they were either absent on the day when the pretest or posttest was conducted, or did not return the questionnaire. Hence, only 210 participants were taken into consideration in the analysis. The sample was 42% (88) and 58% (122) in males and females, respectively. As for the learning styles, 57% (119) students were accommodator learners while 43% (91) were assimilator learners. The mean age of the participants was 16 years old.

Table 1. VR knowledge of the students

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	Know Nothing	96	45.7	46.2	46.2
	Some Knowledge	85	40.5	40.9	87.0
	Lots of Knowledge	8	3.8	3.8	90.9
	Some Experience	19	9.0	9.1	100.0
	Total	208	99.0	100.0	
Missing	System	2	1.0		
Total		210	100.0		

As shown in Table 1, based on those who had answered this question, almost half of participants, that is, 46.2% have no knowledge about VR, 40.9% have some knowledge, only 3.8% know a lot more about VR and 9.1% have some experiences in using VR.

7.2 Homogeneity of Pretest

Independent-samples t-test showed no statistically significant difference in the pretest scores for accommodator learners (M= 41.57, SD = 20.32 and assimilator learners (M = 45.19, SD = 19.45, t(208) = -1.301, p = 0.195), which inferred that all learners were homogeneous in the existing knowledge of the subject matter. Statistical tests were conducted at the alpha = 0.05 significance level.

7.3 Testing of Hypotheses

Independent-samples t-test was used to analyze the data. The assumptions for this test were performed before it was conducted, and this test was found to be appropriate for employment. Statistical tests were conducted at the alpha = 0.05 significance level. Levene's test was used to test the homogeneity of variances. There was no significant difference in the variance of both groups in the posttest (F = 0.029, p = 0.864), gain scores (F = 0.005, p = 0.946), perceived learning effectiveness (F = 0.000, p =0.991), and satisfaction (F = 0.044, p = 0.834). In other words, the basic assumption of homogeneity of variance was not violated.

7.3.1 Testing of H1

The statistical results did not reject the null hypothesis (p > 0.05). As shown in Table 2, there was no significant difference in the posttest score for accommodator learners [M = 65.05, SD = 15.72] and assimilator learners [M = 66.11, SD = 15.68, t(208) = -0.484, p = 0.629].

Table 2. Mean scores, standard deviation (SD) and t-test of posttest, gain scores, perceived learning effectiveness and satisfaction for accommodator learners (N = 119) and assimilator learners (N = 91)

Variables	Accommodator Mean (SD)	Assimilator Mean (SD)	t	df	*p*-value
Posttest	65.05 (15.72)	66.11 (15.68)	-0.484	208	0.629
Gain Scores	23.48 (19.27)	20.92 (19.26)	0.953	208	0.342
Perceived Learning Effectiveness	3.94 (0.54)	3.93 (0.53)	0.100	208	0.920
Satisfaction	3.97 (0.60)	4.09 (0.57)	-1.443	208	0.151

7.3.2 Testing of H2

The statistical results did not reject the null hypothesis ($p > 0.05$). Non-significant result was found in the gain scores for accommodator learners (M = 23.48, SD = 19.27) and assimilator learners [M = 20.92, SD = 19.26, t(208) = 0.953, $p = 0.342$]. (see Table 2.)

7.3.3 Testing of H3

The statistical results did not reject the null hypothesis ($p > 0.05$). As presented in Table 2, there was no significant difference in the perceived learning effectiveness for accommodator learners (M = 3.94, SD = 0.54) and assimilator learners [M = 3.93, SD = 0.53, t(208) = 0.100, $p = 0.920$].

7.3.4 Testing of H4

The statistical results did not reject the null hypothesis ($p > 0.05$). Non-significant results was found in the perceived satisfaction for accommodator learners (M = 3.97, SD = 0.60) and assimilator learners (M = 4.09, SD = 0.57, t(208) = -1.443, $p = 0.151$]. (see Table 2.)

8 Discussion

In terms of cognitive learning outcomes, there was no significant difference in the performance achievement and the overall improvement in performance between accommodator learners and assimilator learners. This implies that the effects of VR learning on both accommodator learners and assimilator learners were almost equivalent. Indeed, Bell and Foyler [13] have mentioned that, experience - the main feature of VR is of great benefits to all learning styles. In other words, VR could provide support to all elements of Kolb's model. A possible explanation to the findings is the virtual environments that mimic the real world provide concrete experience to the learners. Learners could actively construct knowledge through experiment [15]; experience the consequences as they could actively explore the virtual environment and have the autonomy and control over their own leaning [15, 16]. Moreover, the synthetic replica of objects by the virtual environment would help to concretize and reify ideas. Besides, the instructional material in text and images could provide reflective observation and abstract conceptualization to learners. Thus, the VR learning mode covers all four extreme ends of the information perception continuum and the information processing continuum of Kolb's model. In other words, the VR learning mode supports Kolb's model of experiential learning by providing concrete experience, abstract conceptualization, active experimentation and reflective observation. The findings are consistent with the study of Chen, Toh and Wan [18] in which their VR (guided exploration) mode benefited equally to both accommodator learners and assimilator learners.

Additionally, there was also no significant difference in perceived learning effectiveness and satisfaction between accommodator learners and assimilator learners. In other words, all learners regardless of their learning styles, had a similar perception of the learning quality in the VR-based learning environment and experienced the same level of satisfaction. Based on a 5-point scale with higher score indicating better perception in learning quality and satisfaction, both groups scored approximately four on perceived learning effectiveness and satisfaction. Thus, this indicates both groups have agreed that the VR-based learning environment allowed them to achieve the desirable learning effectiveness in identifying the issue learned, making generalization and conclusions. Likewise, both groups experienced a high level of satisfaction in the VR-based learning environment. The positive emotions generated when learning with VR could be the partial cause as to why both types of learners had a high perception on the affective learning outcomes. Research has shown that positive emotions experienced during learning could improve satisfaction and perception towards learning [28]. The esthetic elements such as colors, layout and graphic illustrations of the VR learning material could be the partial cause of the positive effects on perceived learning effectiveness and satisfaction for both types of learners [4]. In short, VR has appealed positively to the affective and emotional state of the learners irrespective of their learning styles. The high level of perceived learning effectiveness and satisfaction implies that learners irrespective of their learning styles are willing to adopt this technology and thus will have an impact on the success of this technology.

9 Conclusion

This paper has investigated the effects of VR on learners with different learning styles. The findings of this research have important educational implications of VR. Based on our findings, there was no significant difference in the performance achievement, overall improvement in performance, perceived learning effectiveness and satisfaction between accommodator learners and assimilator learners in the VR-based learning environment. This implies that VR provides equivalent cognitive and affective benefits to learners with different learning styles, and it could accommodate individual differences with regard to students' learning styles. These findings have provided evidence on the potential of VR to empower students' learning in secondary classrooms. It is suggested that future studies look into the effects of VR on other student characteristics in which the results would benefit individualized learning.

Acknowledgement

We thank Tactus Technologies, Inc., New York for providing us a license to use the V-Frog™ software for this research.

References

1. Chuah, K.M., Chen, C.J., Teh, C.S.: Incorporating kansei engineering in instructional design: Designing virtual reality based learning environments from a novel perspective. Themes in Sciences and Technology Education 1, 37–48 (2008)
2. Inoue, Y.: Concepts, applications, and research of virtual reality learning environments. International Journal of Social Sciences 2, 1–7 (2007)
3. Lee, E.A.-L., Wong, K.W.: A review of using virtual reality for learning. Transactions on Edutainment 1, 231–241 (2008)
4. Lee, E.A.-L., Wong, K.W., Fung, C.C.: Learning effectiveness in a desktop virtual reality-based learning environment. In: Kong, S.C., Ogata, H., Arnseth, H.C., Chan, C.K.K., Hiroshima, T., Klett, F., Lee, J.H.M., Liu, C.C., Looi, C.K., Milrad, M., Mitrovic, A., Nakabayashi, K., Wong, S.L., Yang, S.J.H. (eds.) Proceedings of the 17th International Conference on Computers in Education [CDROM], Asia-Pacific Society for Computers in Education, Hong Kong (2009)
5. Strangman, N., Hall, T.: Virtual reality/simulations. National Center on Accessing the General Curriculum, Wakefield (2003), `http://www.cast.org/publications/ncac/ncac_vr.html`
6. Li, F.C., Angelier, J., Deffontaines, B., et al.: A virtual reality application for distance learning of Taiwan stream erosion in Geosciences. In: Proceedings of the International Conference on Computers in Education (2002)
7. Song, K.S., Lee, W.Y.: A virtual reality application for geometry classes. Journal of Computer Assisted Learning 18, 149–156 (2002)
8. Rieber, L.: Seriously considering play: Designing interactive learning environments based on the blending of microworlds, simulations, and games. Educational Technology Research and Development 44, 43–58 (1996)
9. Crosier, J.K., Cobb, S.V.G., Wilson, J.R.: Experimental comparison of virtual reality with traditional teaching methods for teaching radioactivity. Education and Information Technologies 5, 329–343 (2000)
10. Gustafson, K.L.: The future of instructional design. In: Rieser, R.A., Dempsey, J.V. (eds.) Trends and Issues in Instructional Design and Technology, Prentice Hall, New Jersey (2002)
11. Kolb, D.A.: Experiential learning: Experience as a source of learning and development. Prentice-Hall, Englewood Cliffs (1984)
12. Miller, D., Ferreira, J.M.: Online labs and the MARVEL experience. International Journal on Online Engineering 1 (2005)
13. Bell, J.T., Fogler, H.S.: Ten steps to developing virtual reality applications for engineering education. Paper presented at American Society for Engineering Education Milwaukee, WI (1997)
14. Bricken, W.: Learning in virtual reality. University of Washington, Human Interface Technology Laboratory, Memorandum HITL-M-90-5, Seattle, WA (1990)
15. Fitzpatrick, J.J.: Interactive virtual reality learning systems: Are they a better way to ensure proficiency? (2008), `http://www.vrmariner.com/vr_learning.html`

16. Chee, Y.S.: Virtual reality in education: Rooting learning in experience. In: Proceedings of the International Symposium on Virtual Education 2001, Busan, South Korea (2001)
17. Müller, D., Ferreira, J.M.: Online labs and the MARVEL experience. International Journal on Online Engineering 1 (2005)
18. Chen, C.J., Toh, S.C., Wan, M.F.: Are learning styles relevant to virtual reality? Journal of Research on Technology in Education 38, 123–141 (2005)
19. Tactus Technologies: V-Frog: Virtual frog dissection (2007), `http://www.tactustech.com/products/vfrog/documents/V-Frog-White-Paper.pdf`
20. Cohen, R.J., Swerdlik, M.E., Philips, S.M.: Psychological Testing and Assessment: An Introduction to Tests and Measurement. Mayfield Publishing Company, California (1996)
21. Hopkins, K.D.: Educational and psychological measurement and evaluation. Allyn & Bacon, USA (1988)
22. Benbunan-Fich, R., Hiltz, S.R.: Mediators of the effectiveness of online courses. IEEE Transactions on Professional Communication 46, 298–312 (2003)
23. Marks, R.B., Sibley, S.D., Arbaugh, J.B.: A structural equation model of predictors for effective online learning. Journal of Management Education 29, 531–563 (2005)
24. Martens, R., Bastiaens, T., Kirschner, P.A.: New learning design in distance education: The impact on student perception and motivation. Distance Education 28, 81–93 (2007)
25. Chou, S.W., Liu, C.H.: Learning effectiveness in a web-based virtual learning environment: A learner control perspective. Journal of Computer Assisted Learning 21, 65–76 (2005)
26. Nunnally, J.C.: Psychometric Theory. McGraw Hill, New York (1978)
27. Kolb, A.Y., Kolb, D.A.: The kolb learning style inventory - version 3.1 2005 technical specifications (2005)
28. Um, E.R.: The effect of positive emotions on cognitive processes in multimedia-based learning. Dissertation, New York City University (2008)

Automatic Motion Generation Based on Path Editing from Motion Capture Data

Xiaoyue Guo, Shibiao Xu, Wujun Che, and Xiaopeng Zhang

LIAMA, CAS Institute of Automation, Beijing, China
NLPR, CAS Institute of Automation, Beijing, China
{xyguo,shbxu,xpzhang}@nlpr.ia.ac.cn, chewj@liama.ia.ac.cn

Abstract. In this paper, we present an integrated framework with easy and effective techniques to generate new motions of a character in real-time, using only a short segment of motion data. In this framework, a sequence of segment copies are strung together and then smoothed by a transition method at a junction of two consecutive segments. We use IK thought for the transition to solve the foot-slide problem at the junctions. Given a new motion path, we combine the IK constraints and the sequence of smoothed motion capture data to adjust the human-like model. Based on only a sample motion segment, our method can generate a sequence of motions following an arbitrary path specified by the user. Furthermore, it is easy to implement and can be applied to highly kinetic motions. The experimental results demonstrate our scheme.

Keywords: Animation, Human walk, Inverse kinematics, Motion capture data.

1 Introduction

Character animation is an important research topic in Computer Graphics community, which mainly focuses on human modeling and human motion simulating. Animation of human walking is a crucial problem. It is getting more and more popular and many synthetic scenes involve virtual human characters in both 3D films and 3D video games. However, realistic human walking generation is still a challenge because we are very familiar with human behaviors and highly sensitive to any factitious motion.

Motion capture technique is helpful to preserve the detail information of the performer's behavior for physical motion realness as much as possible, such as the personality or the mood of the performer. But what we obtain with motion capture equipment is just a set of motion data performed by an actor according to requirements, providing a motion toward a specific direction and lasting a certain time. If the requirement changes or a new one is proposed, it's not only inconvenient but also expensive to hire the actor in special clothes for motion capture to perform the required motion. It highly demands that new motion can be automatically generated from the existing motion data library.

In this paper, we make efforts to solve a problem of generating a realistic motion from capture data but in a different motion path. We propose easy and effective techniques for the free walking animation of human in real time regardless of the new

Z. Pan et al. (Eds.): Transactions on Edutainment IV, LNCS 6250, pp. 91–104, 2010.

path shape. In our method, only a short segment of walking data along a straight line is required. We transform it into a lasting motion with a new path. We also overcome the feet-sliding problem in path editing using constraint-based inverse kinematics method. Our method can reuse the existing data while preserving the physical reality. The experiments demonstrate that our method can be put into a variety of applications, e.g. the character detouring a road-block or approaching through a different route. It can also be applied to other periodical motions, e.g. running.

This paper is organized as follows. First we review some related work in Section 2. Then we introduce some basic knowledge on human skeleton model and regular pattern of human walking in Section 3. In Section 4, we consider each problem of our method in turn. The implementation and the experimental results are presented in Section 5. Finally, we conclude with a discussion of future work in Section 6.

2 Related Work

By far, the common methods for character animation can be mainly classified into three categories: kinematic animation, dynamics animation and animation based on motion capture data [1, 2].

- **Kinematic animation**

Kinematic methods treat the human body as a massless linkage. Kinematic animation includes forward and inverse kinematics. Forward kinematics specifies the state vector of an articulated character over time. This specification is usually done for a small set of key-frames, and the positions between them are usually obtained by interpolation techniques. The choice of appropriate interpolation functions has been well studied by A. Watt and M. Watt [3]. Designing key-frames is the task of an animator and the quality of the resultant motion deeply depends on his/her skills. Physical and biomechanical knowledge is helpful to design high-quality key-frames. The inverse kinematics (IK) is more effective than forward kinematics when the motion are limited by some constraints, such as that the support foot is not allowed to penetrate into the ground, to flow above the earth, or to slide slightly. Girard and Maciejewski [4] proposed a method to generate the motion of joints using IK. In this method, the animator specifies the positions of feet in the world coordinate system, and then calculates the rotation angles from the foot joint to the hip joint using inverse Jacobian matrix. It is also useful to solve the foot-slide problem. Boulic *et al.* [5] first use a standard forward kinematics approach to generate interpolated key-positions. Tolani *et al.* [6] developed a set of inverse kinematics algorithms to calculate the motion of human limbs.

- **Dynamics animation**

Dynamics methods also regard humans as an articulated joint system or a multi-rigid-body system, but each component has a mass. These methods can be classified as forward dynamics based method, the trajectory optimization method and controller based method. They may be used either for adding constraints that guarantee certain realistic details to a predefined motion, or for directly synthesizing the walk. Both of forward dynamics based method and trajectory optimization method are based on the dynamic equations of multi-rigid-body system. In order to avoid the singularity of Euler angles, Nikravesh [7] proposed a matrix form equation of dynamics by expressing the rotations of the rigid body as unit quaternions. Armstrong *et al.* [8] and Wilhelm *et al.* [9] proposed a method using forward dynamics integration to simulate the human behavior, but it is not yet intuitive since it is accomplished through setting the force and the torque indirectly. The trajectory optimization method can be understood as the inverse dynamics. As for the controller based method, the proportional-derivative (PD) controller is used in [10, 11, 12], and a finite state automation is built to describe the whole movement. In each joint, the PD controller generates the torque to drive human moving. Both kinematics and dynamics are often mixed to improve the generation quality. Van Overveld *et al.* [13] used the mixed method to stimulate human walk in curve path: At first, controlling the movement of the leg by kinematic methods, and then stimulating and optimizing movement of the other joints by dynamics methods.

Various kinematics and dynamics methods drive the model in character animation to behave as a human, and the utilization of physical and biomechanical knowledge also enhance reality. However, even the most advance motion model cannot generate motions as realistic as a living person does, needless to say the high computation cost. The development of motion capture techniques makes it feasible to preserve the physical realness.

- **Animation based on motion capture data**

Many researches have been carried out to edit motion capture data to make the character behave beyond the captured motion. Rose *et al.* [14] presented a framework of motion blending based on scattered data interpolation with radial basis functions. For a smooth motion transition, they used the sinusoidal function for smoothing and a verb graph (also called motion map) to describe the process. Park *et al.* [15, 16] proposed an on-line motion blending scheme to control a virtual character in real time from one behavior motion to another. To retarget the new motion to the character, an importance-based approach in [17] is used. But the postures of the other joints are computed algorithmically, not so realistically as recorded in capture data. Semancik *et al.* [18] transformed the individual movements to continuous motion by motion blending, in which the motion of various step lengths can be generated, but the motion orientation is limited. Li *et al.* [19] proposed a method based on functional principal component analysis. In this method they transformed the original motion with high dimension to the low dimension; then adjust parameters to blend new motions and control the character. But the analysis of data is linear, which cannot fully describe the relativity of motion sequences. M. Gleicher [20] provided a method allowing for path-based editing based on existing motion data, but it cannot deal with a sharp turning. It is similar to our method, but our method is easier to implement and fits to a sharp turning.

3 Preliminaries

We will give related knowledge on skeleton model and human walking in this section.

3.1 Skeleton Structure of Human Model

Many manufacturers produce motion capture equipments. They employ similar skeleton models except for slight differences in joints for captured motion information.

Here, we choose the skeleton model as shown in Fig. 1, but our method can be applied to the others.

In this model, 18 joints are selected in all besides the End-Sites. The root, also called hips, has 6 degrees of freedom (DOF), including 3 DOFs for translation and 3 DOFs for rotation, while the others have 3 only for rotation, respectively. For the rotational DOFs, we transform the Euler angle into quaternion form to get rid of the Gimbal Lock problem. Gimbal lock is the loss of one degree of freedom that occurs when the axes of two of the three gimbals are driven into the same place and cannot compensate for rotations around one axis in three dimensional spaces. The End-Sites have no DOFs, but only to present the model more vividly.

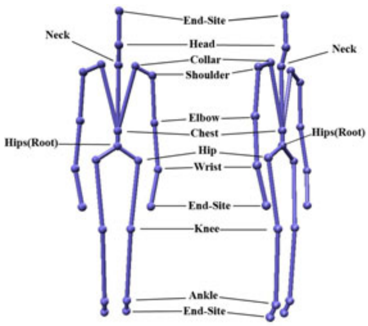


Fig. 1. Skeleton structure of human model

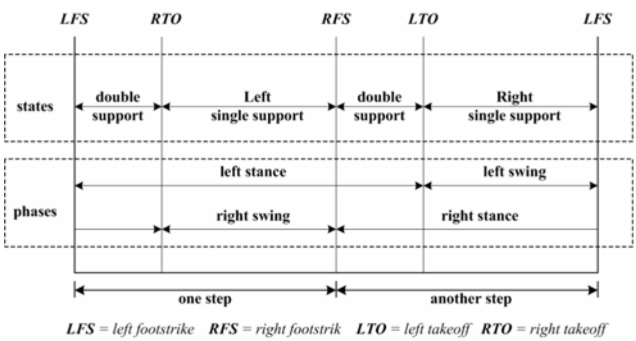


Fig. 2. Characteristic phases of a walking motion [1]

3.2 Circular Motion of Human Walking

Human walking is a kind of regularly circular behavior. An abstract walking model is a succession of separate phases in biomechanics [21, 22]. It characterizes human walking as four phases according to the time when foot-strikes *FS* (the foot contacts the ground) and takeoffs *TO* (the foot leaves the ground). In gait terminology, a stride is defined as a complete cycle from a left foot takeoff to another left foot takeoff, while the part of the cycle between the takeoffs of the two feet is called a step. Four foot-strike and takeoff events occur during a stride: left takeoff (*LTO*), left foot-strike (*LFS*), right takeoff (*RTO*), and right foot-strike (*RFS*). This leads to the characterization of two motion phases for each leg: the period of support is referred as the stance phase and the non-support is referred as the swing phase. As Fig. 2 shows, the continuous shifting between two phases is a walking cycle.

4 Motion Editing

In this section, we will explain in detail how to generate a new walking motion without limits of duration and original path. At first, we transform a segment of one walking cycle into a sequence of cycles in the direction of capture data. Foot-slide is then eliminated by our IK method, which is simple to implement. Next, a new path is designed as required to guide the character's walking. After that, the posture is modified to meet the new path. This step is the key to generate a pleasing motion with arbitrary walking path. We deal with the upper part of the body and the lower part separately. The general flowchart of the whole system is shown in Fig.3.

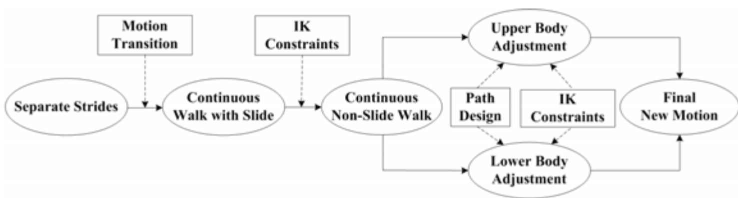


Fig. 3. General flowchart of the system

4.1 Generation of Continuous Walking

As we mentioned in Section 3.2, the continuous walk is formed by a series of separate strides linked end to end. So we regard a stride as a bead and string it repeatedly to produce a necklace. We get a continuous walk map, as shown in Fig. 4.

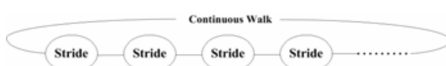
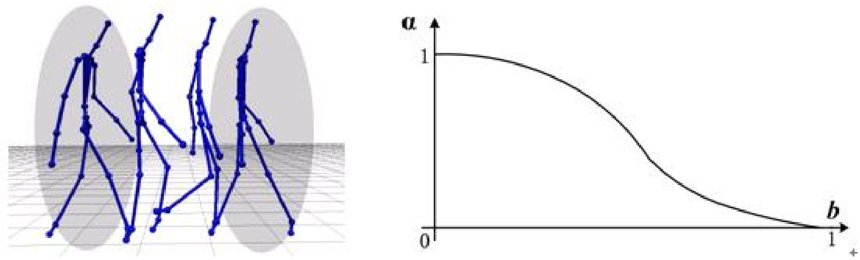


Fig. 4. Continuous walk map

It should be noted that the last posture of a stride is usually similar to the first one, possible with nuances. As shown in Fig. 5(a), four frames are extracted from one stride of walking motion. The two shaded frames are the initial and the last, respectively. Their related postures are quite similar. However, simple links as Fig. 4 usually cause visual shakes at the junction between the last frame of the preceding stride and the first of the succeeding because of such nuances. In order to solve this problem, we use motion transition techniques to smooth the shake. In [15], a motion blending method was proposed to generate a new motion between two different motions, in which the transition is realized by the sinusoidal function [14], as shown in Fig. 5 (b).



Sampled walking of one stride (Left),
Blending function (right)

Fig. 5. Samples of one stride and the blending function

We also use the transition function, but we use different transition period. In other work, the transition begins at a certain key frame in the preceding stride and ends at the first frame of the succeeding. In our work, the transition end at a certain key-frame of succeeding stride instead. This choice makes the transition perform more naturally.

The transition function provides a weight α to mix two source postures to target postures in transition period. Let $T = [t^{*,0}_I, t^{*,1}_I]$ be the period of the transition in Stride I as shown in Fig. 6., where $t^{*,0}_I$ and $t^{*,1}_I$ represent the first and last frame numbers of the period counting from the initial frame of Stride I; b means the beginning part of the current stride to be linked with the preceding stride and e means the ending part of the current stride to be linked with the succeeding one. Usually, the beginning part period is equal to the ending part period.

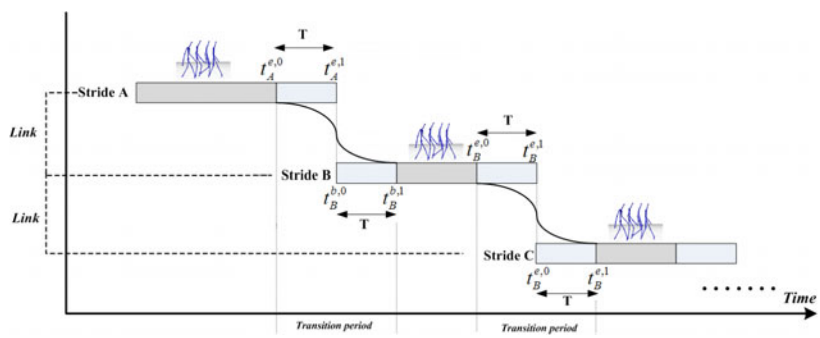


Fig. 6. Linking of strides using the transition function

For the transition between two consecutive Strides A and B, the root joint's position $\mathbf{p}_0(t^i_{A,B})$ of the transition motion at each frame $i \in T$ is computed as:

$$\mathbf{p}_0(t^i_{A,B}) = \alpha \cdot \mathbf{p}_0(t^{e,0}_A) + (1-\alpha) \cdot \mathbf{p}_0(t^{b,1}_B) \quad (1)$$

where $\mathbf{p}_0(t^{e,0}{}_A)$ is the root position of the first frame in the ending part of A and $\mathbf{p}_0(t^{b,1}{}_B)$ is the root position of the last frame in the beginning part of B. The sinusoidal function mentioned previously is presented as:

$$\alpha = 0.5\cos(b\pi) + 0.5 \tag{2}$$

where $b=(f-t_A{}^{e,0})/(t_B{}^{e,1}- t_A{}^{e,0})$, f is the number of the current frame. Equation (1) is suitable for translational DOFs.

For rotational DOFs, the angles are represented with quaternion, so the slurp (spherical linear interpolation) function [23, 24] is employed:

$$\mathbf{q}_n(t^i_{A,B}) = \mathbf{q}_n(t^{b,1}_B)\cdot(\mathbf{q}^{-1}_n(t^{b,1}_B)\cdot\mathbf{q}_n(t^{e,0}_A))^{(1-\alpha)} \tag{3}$$

where $\mathbf{q}_n(t^i{}_{A,B})$ is the nth rotation DOF of the transition motion at frame $i \in T$, $\mathbf{q}_n(t^{e,0}{}_A)$ is the nth rotational DOF at the first frame in the ending part of A and $\mathbf{q}_n(t^{b,1}{}_B)$ is the nth rotational DOF at the last frame in the beginning part of B. The total transition period equals to two times of T. This period should last appropriately, usually less than a step time. Finally, we use the synthesized motion at each frame $i \in [t_A{}^{e,0}, t_B{}^{e,1}]$ as the target motions in the transition period.

The transitions between other strides are dealt with in the same way. In our method, it is unnecessary to make time warping and to find the key times [15] before transition.

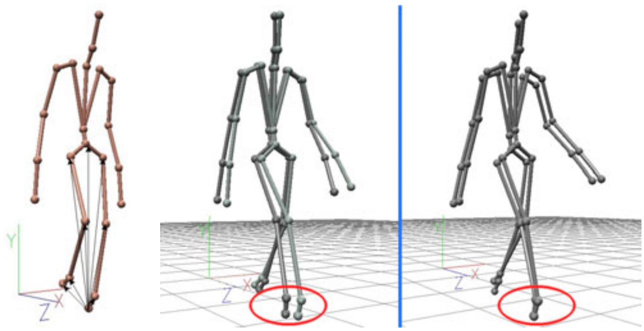
4.2 Elimination of Foot-Slide

In Section 4.1, we have not considered the foot-slide phenomenon in transition yet.

To eliminate the sliding, an IK constraint is subjected to the supporting foot on the ground. IK constraint is commonly used, but it is usually accompanied with complicated algorithms. In this work we propose a new method to do so, which is much easier to implement. We have pointed out in Section 2 that there are two phases of each leg during walking, and for most time, two legs stay in different phases. We define the key-time as the time point when one foot transforms from swing state to stance state. When the key-time begins, the constraint of one foot comes to work to fix the position of this foot on the ground, and hold on until another foot strikes the ground. When the constraint of one foot begins, the constraint of the other ends.

In BVH data from motion capturing, the translational DOFs of the root give the character's position in the world coordinate system. The positions of other joints are determined by the rotational DOFs with their father joints. The position of the supporting foot is recorded when the supporting foot strike the ground before the transition period according the BVH data. Since the position of the supporting foot is fixed on the ground, the positions of other joints can be determined based on this position.

Assume the position of the supporting foot is $\mathbf{p}$ in the world coordinate system. We transform all joints into this system. As shown in Fig. 7 (Left), suppose the left foot is the supporting foot and its position is $\mathbf{p}_{le}$ in a synthesized frame with a foot-slide. The positions of the other skeleton joints are obtained in much the same way according to the names defined in Fig. 1, e.g. the left knee is $\mathbf{p}_{lk}$. The arrows in this figure represent the difference vectors from $\mathbf{p}_{le}$ to the other joints and reflect the positions relative to $\mathbf{p}_{le}$. For example, the difference vector between $\mathbf{p}_{lk}$ and $\mathbf{p}_{le}$ is $\mathbf{v}_{lkle}= \mathbf{p}_{lk} - \mathbf{p}_{le}$. The positions of the other joints can be computed by translating their difference vectors to the position of the supporting foot, e.g. $\mathbf{v}_{lkle} + \mathbf{p}$. Finally, a realistic walking motion without foot-slide is generated. Fig. 7 (Middle and Right) shows the comparison between the walking with and without constraints. In red circle region of Fig. 7 (Middle), the supporting right foot is sliding between two frames. After processing with our method, the right foot is in the fixed position when it works as the supporting foot, as shown in Fig. 7 (Right).



The difference vector (Left), Strides' link without (Middle) and with (Right) constraints

Fig. 7. The difference vector and elimination of foot-slide

4.3 Design of Path

We now propose a scheme to design a new path.

The length of the path is related to the number of frames in motion. A walking curve can be thought of as many a set of line segments connecting head-to-tail. Assuming that the whole motion has N frames and the length of motion path is L, the path is sampled into $N-1$ segments by N points, each of which has the length of $L/(N-1)$. As shown in Fig. 8, the angle θ between two segments changes the walking direction along a curve path.

A path can be represented as a function to fix the coordinates of the junctions. Here we denote the generic path as c(t) and the path can be defined as:

$$\begin{cases} \mathrm{s}(n) = \mathrm{c}(n \cdot \Delta t) \\ \Delta t = T / N \qquad n = 0,1,2,\cdots,N-1 \\ L = \int_0^T \left| \mathrm{c}'(t) \right| dt \end{cases} \tag{4}$$

where s(n) is the n-th samples of the path; L is the length of the path; T is the lasting time; N is the number of frames. c(t) can be set as various curves , such as a circle, a sinusoidal curve, or a B-spline curve.

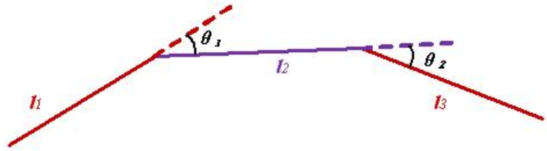


Fig. 8. The straight-line segments and angles between them

4.4 Adjustment of Skeleton for New Path

After the above operations, we are ready to adjust the whole skeleton to walk along the designed route. This a key stage to achieve a good performance.

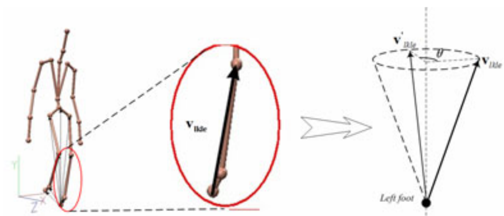


Fig. 9. The rotation of difference vector byθ

The angle θ between two consecutive segments indicates that the character needs to rotate by this angle in order to follow the designed path. In motion capture data, the skeleton rotates around the hips root. However, the intuition tells us that the whole body rotates around the supporting foot rather than the root joint when turning. Another fact we discovered is that, at a turning, people rotates the whole body with stepping forward, not making the rotation during the double support phase (See Figure 2). This discovery is helpful for us to solve the problem.

Although the rotation and forward stepping action proceed synchronously, we deal with them separately. Since the generation of postures at a frame has been completed in foot-slide elimination, the rotation of the skeleton is the last problem to meet the angleθ. The rotation around the supporting foot can be done by rotating other joints, respectively. It is factually the rotation of corresponding difference vectors around the vertical axis crossing the foot. We take $\mathbf{v}_{lkle}$ for example.

As in Fig. 9, the rotation of the left knee around supporting foot is complete by the rotation of $\mathbf{v}_{lkle}$ around the foot. We make $\mathbf{v}_{lkle}$ rotate byθ to the new position $\mathbf{v'}_{lkle}$ as:

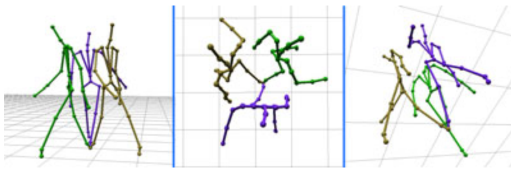
$$\mathbf{v}'_{lkle} = \mathbf{R}_{\theta}\mathbf{v}_{lkle} \quad (5)$$

where $\mathbf{R}_{\theta}$ is the rotation matrix byθ. The treatment can be applied not only to a generic path but also to a sharp turning, which is not considered [15]. But the character needs to rotate by a larger angle at a sharp turning.

Here, we divide the skeleton into the upper part and the lower part at the chest joint. The above operations are performed on joints of the lower half of the body. The upper part can be dealt with in the same manner as the lower half. But a better and easier way is first to translate the chest joint to the skeleton to be generated by using the difference vector between the supporting foot and the chest joint and then to rotate it byθ. Unlike the lower part, the joints of the upper body are all sub-nodes of the chest joint, so the rotation of the upper body can be controlled totally by the rotation of the chest joint. This method simplifies the processing and reserves quite as much the motion details.

5 Experimental Results

Experiments were performed on a Pentium PC (2.4GHz Core2 Quad Processor, 3GB memory).



(a) Side view (b) Top view (c) Bottom view

Fig. 10. Rotation around the supporting foot

Fig. 10 shows the rotation around the fixed supporting foot by 360°. In order to demonstrate the unchanged postures after rotating arbitrary angles, we display the result at three viewpoints: from the side view, from the overlooking view and from the bottom view. Every model rotates by 120°in turn.

Since the rotation of the skeleton by any angle is achieved, walking along arbitrary path can be achievable too. As shown in Fig. 11(a), the original motion is toward the positive z-axis. We choose four different paths for the character to walk: a straight line at 45°from the direction of the positive x-axis in (b), a circle in (c), an ellipse in (d) and a sinusoidal curve in (e). The original walk lasts 75 frames and the resultant walks last 8 times longer. The rate to play is 60fps. All results are satisfying. It should be noted that the path is only a guide direction for human walk in our method. The character doesn't walk exactly on it.

The experiments demonstrate that our method can be used to a variety of applications, such as making the character walk bypass the encountering objects or approach the goal location through different route. Besides human walk, it can also be applied to other periodic motions, e.g. running. Because of low computational cost, the new motion can be generated in real time, which is be beneficial to some real-time applications, such as 3D video games.

Fig. 12 shows the same experiments but for a female character. The original walk has 38 frames and the resultant walks last 8 times longer. The rate to play is 30fps. The performance is also good.

As pointed out in section 4.4, our method is also suitable to the sharp turning. To demonstrate it, we sharpen the turning of the path and put two different characters on this path for experiments.

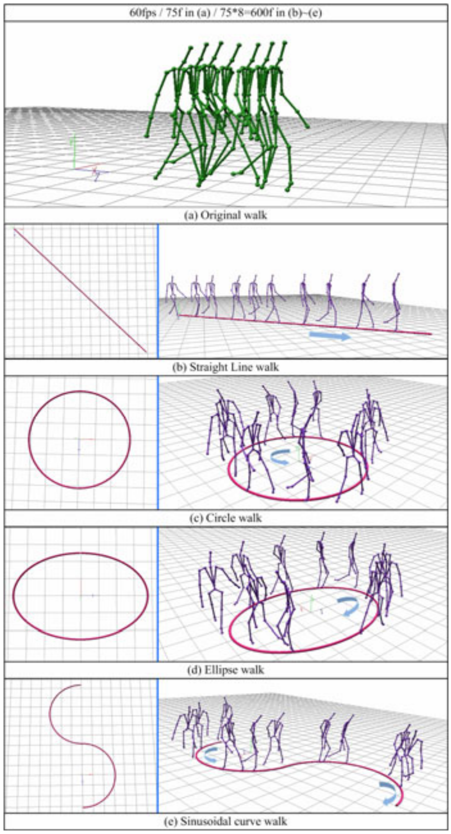


Fig. 11. Various paths of walk

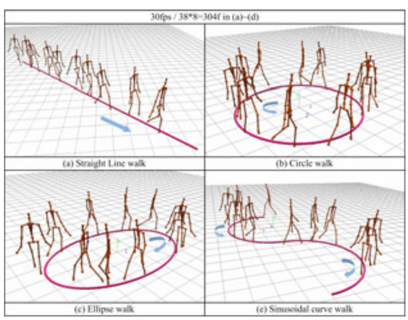
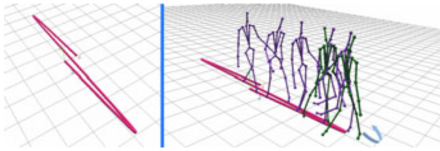
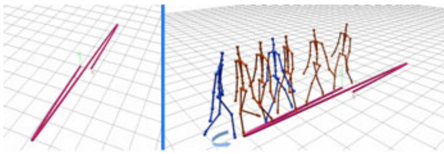


Fig. 12. Various paths of walk for female character



(a)Sharp turning (b) Walk on sharp turning

Fig. 13. The walk on the sharp turning (male character)



(a)Sharp turning (b) Walk on sharp turning

Fig. 14. The walk on the sharp turning (female character)

In Fig. 13, the purple character represents the motion before the turning, while the green represent the after. In Fig. 14, we use the character of a female one with different motion style. The orange character represents the motion before the turning, while the blue represent the after. We can see the performance at the turning is pleasing.

6 Conclusions

Motion capture can record physical information of human motion faithfully. Based on the captured data, highly realistic character animation can be regenerated in contrast to model based methods. In order to enhance the utility of motion libraries, it strongly demands that a single motion segment can be applied in a wide variety of settings. For this purpose, four issues need to be addressed effectively: continuous walking generation, foot-slide elimination, path design and skeleton adjustment for a new path. However, most current work focuses on some parts but not a whole one. In [20], all of these issues are considered, but its framework is just a simple combination of some existing methods, and it could fail for highly kinetic motions, e.g. a sharp turning.

In this paper, we have proposed an integrated framework for the purpose. Our method employs effective techniques for automatic generation of free walking in real-time with optional lasting duration and motion path, based on captured motion data. Our method is purely geometric and is easy to implement, in which there are no complicated mathematical algorithm and notion of physics. Furthermore, it can fit to a highly kinetic motion, e.g. a sharp turning. The experiments demonstrate the effects of our feasible scheme. It can be put into a variety of real-time applications. Besides human walk, our method can also be applied to other circular motions, e.g. running.

Our framework still has much room to be improved. For example, the resultant walk by our current method does not follow the desired path perfectly. We plan to address them to make it more robust and flexible in the future work.

Acknowledgements

The authors would like to thank Jaroslav Semancik for the software platform code and BVH data on his homepage. We would also thank Haiming Wang for the discussion with him during the work. This work is supported by National Natural Science Foundation of China Projects No. 60970093, No. 60872120, and No. 60902078; by National High-Tech Research and Development Plan, 863, of China under Grant No. 2008AA01Z301; by the Strategic Cooperation between Project Guangdong Province and Chinese Academy of Sciences No. 2009B091300044, and by Industry-University-Research Project of Department of Education in Guangdong Province No. 2008B090500207.

References

1. Multon, M., France, L., Cani-Gascuel, M.-P., Debunne, D.: Computer animation of human walking: a survey. Journal of Visualization and Computer Animation 10(1), 39–54 (1999)
2. Xia, S., Wang, Z.: Recent advances on virtual human synthesis. Science in China Series F: Information Sciences 52(5), 741–757 (2009)
3. Watt, A., Watt, M.: Advanced Animation and Rendering Techniques. Theory and Practice (1992)
4. Girard, M., Maciejewski, A.A.: Computational Modeling for the Computer Animation of Legged Figured. Computer Graphics 19(3), 263–270 (1985)
5. Boulic, N., Thalmann, M., Thalmann, D.: A Global Human Walking Model with Real-time Kinematic Personification. The Visual computer 6(6), 344–358 (1990)
6. Tolani, D., Goswami, A., Norman, I.: Real-Time Inverse Kinematics Techniques for Anthropomorphic Limbs. Graphical Models 62(5), 353–388 (2000)
7. Nikravesh, P.: Computer-aided Analysis of Mechanical Systems, pp. 1–370. Prentice Hall, EngleWood Cliffs (1988)
8. Armstrong, W.W., Green, M.: The dynamics of articulated rigid bodies for the purposes of animation. In: Proc. Graphics Interface 1985, pp. 407–415 (1985)
9. Wilhelms, J.: Using dynamic analysis for realistic animation of articulated bodies. IEEE Computer Graphics and Applications 7(6), 12–27 (1987)
10. Yin, K., Loken, K., van de Panne, M.: SIMBICON: Simple biped locomotion control. ACM Transactions on Graphics (Proc. SIGGRAPH) 26(3), article 105 (2007)
11. Wang, M., Fleet, J., Hertzmann, A.: Optimizing walking controllers. ACM Transactions on Graphics (ACM SIGGRAPH Asia) (2009)
12. Coros, S., Beaudoin, P., van de Panne, M.: Robust task-based control policies for physics-based characters. ACM Transactions on Graphics (SIGGRAPH Asia) (2009)
13. Van Overveld, C.W.A.M., Hyeongseok, K.: Small steps for mankind: towards a cinematically driven dynamic simulation of curved path walking. The Journal of Visualization and Computer Animation 5(3), 143–165 (1994)
14. Rose, C., Bodenheimer, B., Cohen, M.F.: Verbs and Adverbs: Multidimensional Motion Interpolation. IEEE Computer Graphics and Applications 18(5), 32–40 (1998)
15. Park, S.I., Shin, H.J., Shin, S.Y.: On-line locomotion generation based on motion blending. In: Proceedings of the 2002 ACM SIGGRAPH/Eurographics symposium on Computer animation (2002)

16. Park, S.I., Shin, H.J., Kim, T.H., Shin, S.Y.: On-line motion blending for real-time locomotion generation. Research Articles, Computer Animation and Virtual Worlds 15(3-4), 125–138 (2004)
17. Shin, H.J., Lee, J., Gleicher, M., Shin, S.Y.: Computer puppetry: an importance-based approach. ACM Transactions on Graphics 20(2), 67–94 (2001)
18. Semancik, J., Pelikan, J., Zara, J.: Interactive Synthesis of Constrained Motion from Example Movements. In: Visualization, Imaging, and Image Processing: Fourth IASTED International Conference Proceedings (2004)
19. Li, C., Wang, Z., Xia, S.: Motion synthesis for virtual human using functional data analysis. Journal of Software 20(6), 1664–1672 (2009)
20. Gleicher, M.: Motion path editing. In: Proceedings of the 2001 symposium on Interactive 3D graphics, pp. 195–202 (2001)
21. Enoka, R.M.: Neuromechanical Basis of Kinesiology, Edition II. Human Kinetics (1994)
22. Nilsson, J., Thorstensson, A., Halbertsam, J.: Changes in leg movements and muscle activity with speed of locomotion and mode of progression in humans. Acta Physiol. Scand., 457–475 (1985)
23. Kim, M.J., Kim, M.S., Shin, S.Y.: A general construction scheme for unit quaternion curves with simple high order derivatives. In: Proceedings of SIGGRAPH 1995, pp. 369–376 (1995)
24. Shoemake, K.: Animating rotations using quaternion curves. In: Proceedings of SIGGRAPH 1985, pp. 245–254 (1985)

Exploration of Metaphorical and Contextual Affect Sensing in a Virtual Improvisational Drama

Li Zhang

School of Computing, Teesside University, UK. TS1 3BA
l.zhang@tees.ac.uk

Abstract. Real-time affect detection from open-ended text-based dialogue is challenging but essential for the building of effective intelligent user interfaces. In this paper, we report updated developments of an affect detection model from text, including affect detection from one particular type of metaphorical affective expression (cooking metaphor) and affect detection based on context. The overall affect detection model has been embedded in an intelligent conversational AI agent interacting with human users under loose scenarios. Evaluation for the updated affect detection component is also provided. Our work contributes to the conference themes on engagement and emotion, interactions in games, storytelling and narrative in education, and virtual characters/agents development.

Keywords: Metaphor, affect detection, context and character profiles.

1 Introduction

Our original system has been designed for age 14-16 secondary school students to engage in role-play situations under emotionally charged loose scenarios (such as bullying and disease) in virtual social environments for drama performance training. A human director will monitor remotely on the virtual improvisation to ensure the improvisation of all the human characters is on the right track (e.g. staying in their characters). In order to reduce the human director's burden, first of all, we have developed a textual affect detection component, EMMA (emotion, metaphor and affect) on detecting simple and complex emotions, meta-emotions, value judgments etc [1]. Metaphorical expressions are another effective channel to convey emotions and feeling, which have drawn researchers' attention. Fainsilber and Ortony [2] commented that "an important function of metaphorical language is to permit the expression of that which is difficult to express using literal language alone". In our previous work [1], several metaphorical affective expressions (such as animal metaphor ("X is a rat") and affects as external entities metaphor ("joy ran through me") [3]) have been intensively studied and efforts have also been made to detect affect from such expressions.

The work presented here reports further developments on affect detection from one particular comparatively complex metaphorical expression with affect implication, i.e. cooking metaphor ("the lawyer grilled the witness on the stand", "I knew I was cooked when the teacher showed up at the door") (http://knowgramming.com/cooking

Z. Pan et al. (Eds.): Transactions on Edutainment IV, LNCS 6250, pp. 105–116, 2010.

_metaphors.htm). Since context plays an important role in the interpretation of the affect conveyed by the user during the interaction, we have used linguistic contextual analysis and cognitive emotional modeling based on Markov chain modeling and a dynamic algorithm to interpret affect from context in our application.

Our affect detection component has been embedded in an intelligent conversational agent, engaged in a drama improvisation with human users under loose scenarios (school bullying and Crohn's disease). The conversational AI agent also provides appropriate responses based on the detected affect from users' input in order to stimulate the improvisation. In both scenarios, the AI agent plays a minor role in drama improvisation. E.g. it plays a close friend of the bullied victim (the leading role) in school bullying scenario, who tries to stop the bullying.

We have also analyzed affect detection performance based on previously collected (other) transcripts from user testing by calculating agreements via Cohen's Kappa between two human judges and between human judges and the AI agent with and without the new development respectively in order to verify the efficiency of the metaphorical and contextual affect sensing.

The content is arranged as follows. We report relevant work in section 2 and the new developments on affect detection from cooking metaphor in section 3. Affect interpretation based on context is discussed in section 4. System evaluation and conclusion are presented respectively in section 5 & 6.

2 Relevant Work

Textual affect sensing is a rising research branch for natural language processing. ConceptNet [4] is a toolkit to provide practical textual reasoning for affect sensing for six basic emotions, text summarization and topic extraction. Shaikh et al. [5] provided sentence-level textual affect sensing to recognize evaluations (positive and negative). They adopted a rule-based domain-independent approach, but they haven't made attempts to recognize different affective states from open-ended text input.

Although Façade [6] included shallow natural language processing for characters' open-ended utterances, the detection of major emotions, rudeness and value judgements is not mentioned. Zhe and Boucouvalas [7] demonstrated an emotion extraction module embedded in an Internet chatting environment. It used a part-of-speech tagger and a syntactic chunker to detect the emotional words and to analyze emotion intensity for the first person (e.g. 'I' or 'we'). Unfortunately the emotion detection focused only on emotional adjectives, and did not address deep issues such as figurative expression of emotion (discussed below). Also, the concentration purely on first-person emotions is narrow. There has been relevant work on general linguistic cues that could be used in practice for affect detection (e.g. Craggs and Wood [8]).

There is also well-known research work on the development of emotional conversational agents. Elliott et al. [9] demonstrated tutoring systems that reason about users' emotions. They believe that motivation and emotion play very important roles in learning. Virtual tutors have been created in a way that not only having their own emotion appraisal and responsiveness, but also understanding users' emotional states according to their learning progress. Aylett et al. [10] also focused on the development of affective behaviour planning for the synthetic characters. Cavazza et al. [11] reported a conversational agent embodied in a wireless robot to provide suggestions for users on a healthy living life-style. Hierarchical Task Networks (HTN) planner and semantic interpretation have been used in this work. The cognitive planner plays an important role in assisting with dialogue management. The system normally works reasonably well within the pre-defined domain knowledge, but it will strike when open-ended user input going beyond the planner's knowledge has been used intensively during interaction. The system we present here intends to deal with such challenge.

Our work is distinctive in the following aspects: (1) affect detection from metaphorical expressions; (2) real-time affect sensing for basic and complex emotions, meta-emotions, value judgments etc in improvisational role-play situations from open-ended user input; (3) affect detection for second and third person cases ('you', 'she', 'they'); and (4) affect interpretation based on context profiles.

3 Further Development on Metaphorical Affect Detection

Without pre-defined constrained scripts, the human users could be creative in their role-play within the highly emotionally charged scenarios. The language used by the school students during their role-play is highly diverse with various online chatting features. Diverse metaphorical language phenomena have also been used to express emotions during the interaction, which challenge any natural language processing system greatly which intends to pursue affect/opinion mining and draw our attention for further study.

3.1 Affect Detection from Cooking Metaphor

In our previous study, we noticed that food has been used extensively as metaphor for social position, group identity, religion, etc. E.g. food could be used as a metaphor for national identity. British have been called 'roastbeefs' by the French, while French have been referred to as 'frogs' by the British. It has also been used to indicate social hierarchy. E.g. in certain Andean countries, potatoes have been used to represent poor rural farmers of native American descent and white flour and bread have been used mainly to refer to wealthy European descent. In our school bullying scenario, the big bully has called the bullied victim (Lisa) names, such as "u r a pizza", "Lisa has a pizza face" to exaggerate that fact that the victim has acne. Another most commonly used food metaphor is to use food to refer to a specific shape. E.g. body shape could be described as 'banana', 'pear' and 'apple' (http://jsgfood.blogspot.com/2008/02/food-metaphors.html). In our previous application, we have made useful attempts to detect affect from such metaphorical expressions.

Another type of metaphor that has great potential to carry affect implication is cooking metaphor. Very often, the agent himself/herself would become the victim of slow or intensive cooking (e.g. grilled, cooked, etc). Or one agent can perform cooking like actions towards another agent to realize punishment or torture. Examples are as follows, "he knew he was cooked when he saw his boss standing at the door", "he basted her with flattery to get the job", "she knew she was fried when the teacher handed back her paper" etc.

In these examples, the suffering agents have been regarded as food. They bear the results of intensive or slow cooking. Thus these agents who suffer from such cooking actions carried out by other agents tend to feel pain and sadness, while the ‘cooking performing’ agents may take advantage of such actions to achieve their intentions, such as persuasion, punishment or even enjoyment. The syntactic structures of some of the above examples also indicate the submissive of the suffering agents. E.g. in the instances, passive sentences (“he knew he was cooked when he saw his boss standing at the door”) have been very often used to imply unwillingness and victimised of the subject agents who are in fact the objects of the cooking actions described by the verb phrases (“X + copular form + passive cooking action”). In other examples, the cooking actions have been explicitly performed by the subject agents towards the object agents to imply the former’s potential willingness and enjoyment and the latter’s potential suffering and pain (“A + [cooking action] + B”).

Thus in our application, we focus on the above two particular types of expressions. We use Rasp [12], a statistical-based domain-independent robust parsing system for English, to recognize user input with such syntactic structures (‘A + copular form + VVN’, ‘A + VV0/VVD/VVZ (verb) + B’). Many sentences could possess such syntactic structures (e.g. “Lisa was bullied”, “he grills Lisa”, “I was hit by a car”, “Lisa was given the task to play the victim role”, “I steamed it” etc), but few of them are cooking metaphors. Therefore we need to resort to semantic profiles to recognize the metaphorical expressions. Rasp has also provided a syntactic label for each word in the user input. Thus the main verbs were identified by their corresponding syntactic labels (e.g. ‘given’ labeled as ‘past participle form of lexical verbs (VVN)’, ‘likes’ and ‘grills’ labeled as ‘-s form of lexical verbs (VVZ)’) and the semantic interpretation for their base forms is discovered from WordNet [13]. Since WordNet has provided hypernyms (Y is a hypernym of X if every X is a (kind of) Y) for the general noun and verb lexicon, ‘COOK’ has been derived as the hypernym of the verbs described cooking actions. E.g. ‘boil’, ‘grill’, ‘steam’, and ‘simmer’ are respectively interpreted as one way to ‘COOK’. ‘Toast’ is interpreted as one way to ‘HEAT UP’ while ‘cook’ is interpreted as one way to ‘CREAT’, or ‘CHEAT’ etc. One verb may recover several hypernyms and in our application, we collect all of them. Another evaluation profile [14] is resorted to in order to recover the evaluation values of all the hypernyms for a particular verb. If some hypernyms are negative (such as ‘CHEAT’) and the main object of the overall input refers to first/third person cases or singular proper noun (‘him’, ‘her’, or ‘Lisa’), then the user input (e.g. “he basted her with flattery to get the job”) conveys potential negative affect (e.g. pain and sadness) for the human objects and potential positive affect (e.g. persuasion or enjoyment) for the subjects. If the evaluation dictionary fails to provide any evaluation value for any hypernyms (such as ‘COOK’ and ‘HEAT UP’) of the main verbs, then we still assume that ‘verbs implying COOK/HEAT UP + human objects’ or ‘human subjects + copular form + VVN verbs implying COOK/HEAT UP’ may indicate negative emotions both for the human objects in the former and the human subjects in the latter. Two example implementations are provided in the following. E.g. for the input “I was fried by the head teacher”, the processing is as follows:

1. Rasp identifies the input has the following structure: ‘PPIS1 (I) + copular form (was) + VVN (fried)’;

2. 'Fry' (base form of the main verb) is sent to WordNet to obtain its hypernyms, which include 'COOK', 'HEAT' and 'KILL';

3. The input has the following syntactic semantic structure: 'PPIS1 (I) + copular form (was) + VVN (Hypernym: COOK)', thus it is recognized as a cooking metaphor;

4. The three hypernyms are sent to the evaluation profile to obtain their evaluation values. 'KILL' is labeled as negative while others can't obtain any evaluation values from the profile;

5. The input is transformed into: ''PPIS1 (I) + copular form (was) + VVN (KILL: negative)'

6. The subject is a first person case, thus the input indicates the user who is speaking suffered from a negative action and may have 'sad' emotional state.

Another example input is "she grilled him in front of the class". The hypernyms of 'grill' only contain 'COOK' (recognized as a cooking metaphor), but we cannot trace it from the evaluation profile to return an evaluation value. However, since the input has the following syntactic and semantic structure: "PPHS1 (she) + COOK + PPHO1 (him)', it implies that a third-person subject performed a cooking action towards another human object. The human object ('him') may experience 'sad' emotional state.

Sometimes, our processing may not be able to recognize the metaphorical phenomena (the cooking metaphor). However it is able to interpret the affect conveyed in the user input. E.g. if we have input "he basted her with flattery to get the job", the recognition process is as follows:

1. Rasp recognizes the input has the main syntactic structure of "PPHS1 (he) + VVD (basted) + PPHO1 (her)';

2. 'Baste' (base form of the main verb) is sent to WordNet to obtain its hypernyms, which include 'MOISTEN', 'BEAT', and 'SEW';

3. The hypernyms do not include 'COOK', thus it is not recognized as a cooking metaphor;

4. The hypernyms are sent to the evaluation profile to obtain their evaluation values. We obtain 'negative' for the hypernym 'BEAT'.

5. The input is interpreted as: 'PPHS1 (he) + VVD (BEAT: negative) + PPHO1 (her)'. It implies that a third-person subject performed a negative action towards another human object.

6. The human object ('her') may experience 'sad' emotional state.

Although our processing is indeed at a very initial stage, it has pointed out promising directions for figurative language processing. After our intention to improve the performance of affect sensing from individual turn-taking user input, we focus on improvement of the performance using context profiles. In future work, we also intend to use metaphor ontology to recognize and generate metaphors.

4 Affect Sensing from Context Profiles

Our previous affect detection has been performed solely based on individual turn-taking user input. Thus the context information has been ignored. However, the contextual and character profiles may influence the affect conveyed in the current input.

In this section, we are going to discuss linguistic contextual indicators, cognitive emotion simulation from communication context and our approach developed based on these features to interpret affect from context.

4.1 Linguistic Indicators for Contextual Communication

In our study, we noticed some linguistic indicators for contextual communication in the recorded transcripts. One useful indicator is (i) imperatives, which are often used to imply emotional responses to the previous speaking characters, such as "shut up", "go on then", "let's do it" and "bring it on". Other useful contextual indicators are (ii) prepositional phrases (e.g. "by who?"), semi-coordinating conjunctions (e.g. "so we are good then"), subordinating conjunctions ("because Lisa is a dog") and coordinating conjunctions ('and', 'or' and 'but'). These indicators are normally used by the current 'speaker' to express further opinions or gain further confirmation from the previous speakers. Moreover, (iii) short phrases for questions are also used frequently in the transcripts to gain further communication based on context, e.g. "where?", "who is Dave" or "what". (iv) Character names are also normally used in the user input to indicate that the current input is intended for particular characters, e.g. "Dave go away", "Mrs Parton, say something", etc. Finally there are also (i) some other well known contextual indicators in Internet relay chat such as 'yeah/yes followed by a sentence ("yeah, we will see", "yeah, we cool Lisa")', "I think so", 'no/nah followed by a sentence', "me too", "exactly", "thanks", "sorry", "grrrr", "hahahaha", etc.

Since natural language is ambiguous and there are cases that contextual information is required in order to appropriately interpret the affect conveyed in the input (e.g. "go on then"), our approach reported in the following integrates the above contextual linguistic indicators with cognitive contextual emotion prediction to uncover affect conveyed in emotionally ambiguous input.

4.2 Emotion Modeling in Communication Context

There are also other aspects which may influence the affect conveyed in the communication context. E.g. in our application, the affect conveyed by the speaking character himself/herself in the recent several turn-taking, the 'improvisational mood' that the speaking character is created, and emotions expressed by other characters, especially by the contradictory ones (e.g. the big bully), have great potential to influence the affect conveyed by the current speaking character (e.g. the bullied victim). Sometimes, the story themes or topics also have potential impact to emotions or feelings expressed by characters. For example, people tend to feel 'happy' when involved in discussions on positive topics such as harvest or raising salary, while people tend to feel 'sad' when engaged in the discussions on negative themes such as economy breakdown, tough examination or funeral.

In our application, although the hidden story sub-themes used in the scenarios are not that dramatic, they are still highly emotionally charged and used as the signals for potential changes of emotional context for each character. E.g. In the school bullying scenario (which is mainly about the bully, Mayid, picking on the new comer to the school, Lisa. Lisa's friends, Elise and Dave, are trying to stop the bullying. The school teacher, Mrs Parton, also tries to find out what is going on), the director mainly provided interventions based on several main sub-themes of the story to push the improvisation forward, i.e. “Mayid starts bullying Lisa”, “why Lisa is crying”, “why Mayid is so nasty/a bully”, “how Mayid feels when his uncle finds out his behavior” etc. From the inspection of the recorded transcripts, when discussing the topic of “why Lisa is crying”, we noticed that Mayid (the bully) tends to be really aggressive and rude, while Lisa (the bullied victim) tends to be upset and other characters (Lisa’s friends and the school teacher) are inclined to show anger at Mayid. For the improvisation of the hidden story sub-theme “why Mayid is so nasty/a bully”, the big bully changes from rude and aggressive to sad and embarrassed (e.g. because he is abused by his uncle), while Lisa and other characters become sympathetic (and sometimes caring) about Mayid. Usually all characters are trying to create the ‘improvisational mood’ according to the guidance of the hidden story sub-themes (provided via director’s intervention). Therefore, the story sub-themes could be used as the indicators for potential emotional context change. The emotion patterns expressed by each character within the improvisation of each story sub-theme could be very useful for the prediction of the affect shown in a similar topic context, although the improvisation of the characters is creative within the loose scenario. It will improve the performance of the emotional context prediction if we allow more emotional profiles for each story sub-theme to be added to the training data to reflect the creative improvisation (e.g. some improvisations went deeper for a particular topic).

Therefore, a Markov chain (http://en.wikipedia.org/wiki/Markov_chain) is used to learn from the emotional context shown in the recorded transcripts for each sub-theme and for each character, and generate other possible reasonable unseen emotional context similar to the training data for each character. Markov chains are usually used for word generation. In our application, they are used to record the frequencies of several emotions showing up after one particular emotion. A matrix has been constructed dynamically for the 12 most commonly used emotions in our application (caring, arguing, disapproving, approving, grateful, happy, sad, threatening, embarrassed, angry/rude, scared and sympathetic) with each row representing the previous emotion followed by the subsequent emotions in columns. The Markov chains employ roulette wheel selection to ensure to produce a greater probability to select emotions with higher frequencies than emotions with lower occurrences. This will allow the generation of emotional context to probabilistically follow the training data, which may reflect the creative nature of the improvisation.

Then a dynamic algorithm is used to find the most resembling emotional context for any given new situation from Markov chain’s training and generated emotional contexts. I.e. by using the algorithm, a particular series of emotions for a particular story sub-theme has been regarded as the most resembling context to the test emotional situation and an emotional state is recommended as the most possible emotion for the current user input. Since the most recent affect histories of other characters and relationships between characters may also have impact on the affect conveyed by the current speaking character, the recommended affect will be further evaluated (e.g. a most recent ‘insulting’ input from Mayid may cause Lisa and her friends to be ‘angry’).

At the training stage, first of all, the school bullying transcripts collected from previous user testing have been divided into several topic sections with each of them belonging to one of the story sub-themes. The classification of the sub-themes is mainly based on the human director’s intervention which was recorded in the transcripts. Then we used two human annotators to mark up the affect of every turn-taking input in the transcripts using context inference. Thus for each character, we have summarized a series of emotions expressed throughout the improvisation of a particular story sub-theme. We made attempts to gather emotional contexts as many as possible for each character for the improvisation of each sub-theme in order to enrich the training data. The following is a small portion of one recorded transcript used for the training of the Markov chain. The human annotators have marked up the affect expressed in each turn-taking input.

DIRECTOR: why is Lisa crying?
Elise Brown [caring]: lisa stop cryin
Lisa Murdoch [disagree]: lisa aint crying!!!!
Dave Simons [caring]: i dunno! y u cryin lisa?
Mayid Rahim [rude]: cuz she dnt realise she is lucky to b alive
Elise Brown [angry]: beat him up! itss onlii fat..he'll go down straight away
Mayid Rahim [insulting]: lisa, y u crying? u big baby!
Mrs Parton [caring]: lisa, r u ok?

E.g. the emotional context for Mayid from the above example is: 'rude' and 'insulting' (we use one letter to represent each emotional label, thus in this example, i.e. 'R I'), and in the similar way, the emotional contexts for other characters have been derived from the above example, which are used as the training data for the Markov chain for the topic "why Lisa is crying". We have summarized the emotional context for each story sub-theme for each character from 4 school bullying transcripts and used them for the training of the Markov chain. The topics considered at the training stage include: "Mayid starts bullying", "why is Lisa crying", "why is Mayid nasty/a bully" and "how does Mayid feel if his uncle knew about his behavior?"

At the test stage, our affect detection component, EMMA, is integrated with an AI agent and detects affect for each user input solely based on the analysis of individual turn-taking input itself as usual. The above algorithms for context-based affect sensing will be activated when the affect detection component recognizes 'neutral' from the current input during the emotionally charged proper improvisation after all the characters have known each other and went on the virtual drama stage. First of all, the linguistic indicators are used to identify if the input with 'neutral' implication is a contextual-based input. E.g. we mainly focus on the checking of the five contextual implications we mentioned previously, including imperatives, prepositional phrases, conjunctions, simplified question sentences, character names, and other commonly used contextual indicators (e.g. "yeah", "I think so"). If any of the above contextual indicators exists, then we further analyze the affect embedded in the input with contextual emotion modeling reported here.

E.g. we have collected the following transcript for testing. Normally the director intervened to suggest a topic change (e.g. "find out why Mayid is a bully"). Thus for a testing situation for a particular character, we use the emotion context attached with his/her user input starting right after the most recent director's intervention and ending at his/her last second input, since such a context may belong to one particular topic.

DIRECTOR: U R IN THE PLAYGROUND (indicating bullying starts)
1. Lisa Murdoch: leave me alone! [angry]
2. Mayid Rahim: WAT U GONNA DU? [neutral] -> [angry]

3. Mayid Rahim: SHUT UR FAT MOUTH [angry]
4. Elise Brown: grrrrr [angry]
5. Elise Brown: im telin da dinna lady! [threatening]
6. Mayid Rahim: go on den [neutral] -> [angry]
7. Elise Brown: missssssssssssssss [neutral]
8. Elise Brown: lol [happy]
9. Lisa Murdoch: mayid u gna gt banned [threatening]
10. Mayid Rahim: BY HU [neutral] -> [angry]

The affect detection component detected that Lisa was ‘angry’ by saying “leave me alone!”. It also sensed that Mayid was ‘neutral’ by saying “WAT U GONNA DU (what are you going to do)?” without consideration of context. From Rasp, we obtained that the input is a simplified question sentence (a linguistic contextual indicator). Thus, it implies that it could be an emotional situation caused by the previous context (e.g. previous input from Lisa) and the further processing for emotion prediction is activated. Since we don’t have an emotional context yet at this stage for Mayid (the very first input from Mayid after the director’s intervention), we cannot resort to the Markov chain and the dynamic algorithm currently to predict the affect. However, we could use the emotional context of other characters to predict the affect for Mayid’s current input since we believe that an emotional input from a character, esp. from an opponent character, has great potential to affect the emotions expressed by the current speaking character.

In the most recent chat history, there is only one input from Lisa after the director’s intervention, which implied ‘anger’. Since Lisa and Mayid have a negative relationship, then we predict Mayid currently experiences negative emotion. Since capitalizations have been used in Mayid’s input, we conclude that the affect implied in the input could be ‘angry’. However, EMMA could be fooled if the affect histories of other characters fail to provide any useful indication for prediction (e.g. if Lisa implied ‘neutral’ in the most recent input, the interpretation of the affect conveyed by Mayid would be still ‘neutral’).

EMMA also detected affect for the 3rd, 4th, and 5th user input in the above example (based on individual turn-taking) until it detected ‘neutral’ again from the 6th input “go on den (go on then)” from Mayid. Since it is an imperative mood sentence (a linguistic contextual indicator), the input may imply a potential (emotional) response to the previous speaking character. Since we couldn’t obtain the affect embedded in the imperative purely based on the analysis of the input itself, the contextual processing is required. Thus the emotional context profile for Mayid is retrieved, i.e. [angry (the 2nd input) and angry (the 3rd input)]. The Markov chain is used to produce the possible emotional context based on the training data for each sub-theme for Mayid.

The following are the generated example emotional profiles for the sub-theme “Mayid starts bullying” for the Mayid character:

1. T A A N A A [‘threatening, angry, angry, neutral, angry and angry’]
2. N A A A [‘neutral, angry, angry, and angry’]
3. D A I A A A N A [‘disapproval, angry, insulting, angry, angry, angry, neutral, and angry’]
4. I A A N [‘insulting, angry, angry and neutral’]

The dynamic algorithm is used to find the smallest edit distance between the test emotional context [angry and angry] and the training and generated emotional context for the Mayid character for each sub-theme. In the above example, the second and fourth emotional sequences have the smallest edit distance (distance = 2) to the test emotional context and the former suggests 'angry' as the affect conveyed in the current input ("go on den") while the latter implies 'neutral' expressed in the current input. Thus we need to resort to the emotional context of other characters to justify the recommended affects. From the chatting log, we find that Lisa was 'angry' in her most recent input (the 1st input) while Elise was 'threatening' in her most recent input (the 5th input). Since the bully, Mayid, has negative relationships with Lisa (being 'angry') and Elise (being 'threatening'), the imperative input ("go on den") may indicate 'angry' rather than 'neutral'. Therefore our processing adjusts the affect from 'neutral' to 'angry' for the 6th input.

In this way, by considering the linguistic contextual indicators, the potential emotional context one character was in, relationships with others and recent emotional profiles of other characters, our affect detection component has been able to inference emotion based on context to mark up the rest of the above test example transcript (e.g. Mayid being 'angry' for the 10th input). However our processing could be fooled easily by various diverse ways for affective expressions and creative improvisation (test emotional patterns not shown in the training and generated sets). We intend to adopt better emotion simulation tools, more linguistic hints, psychological (context-based) emotional theories etc for further improvements. Our processing currently only focused on school bullying scenario. We are on our way to extend the context-based affect sensing to other scenarios to further evaluate its efficiency.

5 Evaluation

We carried out user testing with 220 secondary school students from Birmingham and Darlington schools for the improvisation of school bullying and Crohn's disease scenarios. Generally, our previous statistical results based on the collected questionnaires indicate that the involvement of the AI character has not made any statistically significant difference to users' engagement and enjoyment with the emphasis of users' notice of the AI character's contribution throughout. Briefly, the methodology of the testing is that we had each testing subject had an experience of both scenarios, one including the AI minor character only and the other including the human-controlled minor character only. After the testing sessions, we obtained users' feedback via questionnaires and group debriefings. Improvisational transcripts were automatically recorded during the testing so that it allows further evaluation of the performance of affect detection component.

Therefore, we produce a new set of results for the evaluation of the updated affect detection component with metaphorical and context-based affect detection based on the analysis of some recorded transcripts of school bullying scenario. Generally two human judges (not engaged in any development stage) marked up the affect of 150 turn-taking user input from the recorded another 4 transcripts from school bullying scenario (different from those used for the training of Markov chains). In order to verify the efficiency of the new developments, we provide Cohen's Kappa inter-agreements for EMMA's performance with and without the new developments for the detection of the most commonly used 12 affective states. In the school scenario, EMMA played a minor bit-part character (Lisa's friend: Dave). The agreement for human judge A/B is 0.45. The inter-agreements between human judge A/B and EMMA with and without the new developments are presented in Table 1.

Table 1. Inter-agreements between human judges and EMMA with and without the new developments

	Human Judge A	Human Judge B
EMMA (previous version)	0.38	0.30
EMMA (new version)	0.40	0.32

Although further work is needed, the new developments on metaphorical and contextual affect sensing have improved EMMA's performance comparing with the previous version. Most of the improvement is obtained for negative affect detection based on the contextual inference. Other improvements have been made by the metaphorical processing. We also intend to compare our system's affect sensing performance with that of a well-known affect sensing tool, ConceptNet [4], as another way for future evaluation.

6 Conclusion

First of all, we have made a step towards automatic affect sensing from textual metaphorical figurative language. Also for the first time, context and character profiles have been employed in our study to interpret affect from user input. However, there is still a long way to go in order to successfully process the rich diverse variations of metaphorical language and other figurative expressions, such as humor, lies, irony etc. We also intend to adopt more linguistic contextual hints, psychological (context-based) emotional theories and better simulation tools (e.g. hidden Markov modeling) for further improvements on context-based affect detection.

Overall, our work provides automatic improvisational agents for virtual drama improvisation situations. It makes a contribution to the issue of what types of automation should be included in human-robots interaction, and as part of that the issue of what types of affect should be detected and how. It also provides an opportunity for the developers to explore how emotional issues embedded in the scenarios, characters and dialogue can be represented visually without detracting users from the learning situation.

References

1. Zhang, L., Barnden, J.A., Hendley, R.J., et al.: Affect Detection and Metaphor in E-drama. Int. J. Continuing Engineering Education and Life-Long Learning 18(2), 234–252 (2008)
2. Fainsilber, L., Ortony, A.: Metaphorical uses of language in the expression of emotions. Metaphor and Symbolic Activity 2(4), 239–250 (1987)

3. Kövecses, Z.: Are There Any Emotion-Specific Metaphors? In: Athanasiadou, A., Tabakowska, E. (eds.) Speaking of Emotions: Conceptualization and Expression, pp. 127–151. Mouton de Gruyter, Berlin (1998)
4. Liu, H., Singh, P.: ConceptNet: A practical commonsense reasoning toolkit. BT Technology Journal 22 (2004)
5. Shaikh, M.A.M., Prendinger, H., Mitsuru, I.: Assessing sentiment of text by semantic dependency and contextual valence analysis. In: Paiva, A.C.R., Prada, R., Picard, R.W. (eds.) ACII 2007. LNCS, vol. 4738, pp. 191–202. Springer, Heidelberg (2007)
6. Mateas, M.: Ph.D. Thesis. Interactive Drama, Art and Artificial Intelligence. School of Computer Science, Carnegie Mellon University (2002)
7. Zhe, X., Boucouvalas, A.C.: Text-to-Emotion Engine for Real Time Internet Communication. In: Proceedings of International Symposium on Communication Systems, Networks and DSPs, Staffordshire University, UK, pp. 164–168 (2002)
8. Craggs, R., Wood, M.: A Two Dimensional Annotation Scheme for Emotion in Dialogue. In: Proceedings of AAAI Spring Symposium: Exploring Attitude and Affect in Text (2004)
9. Elliott, C., Rickel, J., Lester, J.: Integrating Affective Computing into Animated Tutoring Agents. In: Proceedings of IJCAI 1997 Workshop on Intelligent Interface Agents (1997)
10. Aylett, R., Louchart, S., Dias, J., et al.: Unscripted Narrative for Affectively Driven Characters. IEEE Computer Graphics and Applications 26(3), 42–52 (2006)
11. Cavazza, M., Smith, C., Charlton, D., et al.: A 'Companion' ECA with Planning and Activity Modelling. In: Proceedings of the 7th International Conference on Autonomous Agents and Multi-Agent Systems, Portugal, pp. 1281–1284 (2008)
12. Briscoe, E., Carroll, J.: Robust Accurate Statistical Annotation of General Text. In: Proceedings of the 3rd International Conference on Language Resources and Evaluation, Las Palmas, Gran Canaria, pp. 1499–1504 (2002)
13. Fellbaum, C.: WordNet, an Electronic Lexical Database. The MIT press, Cambridge (1998)
14. Esuli, A., Sebastiani, F.: Determining Term Subjectivity and Term Orientation for Opinion Mining. In: Proceedings of EACL 2006, 11th Conference of the European Chapter of the Association for Computational Linguistics, Trento, IT, pp. 193–200 (2006)

Synchronizable Objects in Distributed Multimedia Applications

Jun Hu and Loe Feijs

Department of Industrial Design
Eindhoven University of Technology
5600MB Eindhoven, The Netherlands
{j.hu,l.m.g.feijs}@tue.nl

Abstract. In training and gaming systems, distributed multimedia are often used, in which the basic content elements must be conveyed or presented in a synchronized order at synchronized moments over multiple devices and in many cases over a network. These content elements are often presented or represented as "Synchronizable Objects" with which their control and management fall into a design pattern. This paper uses the pattern language to capture the common features of these "Synchronizable Objects", in combination of the formal Object-Z specification to treat the architectural construct. The proposed pattern can be applied for content elements with or without intrinsic timing in distributed multimedia applications. Examples are given to show how this pattern can be applied in distributed applications.

Keywords: multimedia, synchronizable object, design pattern, formal specification.

1 Introduction

Training and gaming systems often use distributed multimedia (text, audio, graphics, animation, video, and recently movements and behaviors of physical objects), to convey information in synchronized order and at synchronized moments over multiple devices and in many cases over a network. These content elements are often presented or represented as "Synchronizable Objects" with which their control and management fall into a design pattern. In this article we try to follow the pattern language [1] to describe its intent, context, forces, solution and consequences. Often in object-oriented design patterns, the solution is described using the Unified Modeling Language (UML) in combination of examples of source code. However UML as such is not sufficient to capture the details of the common features in pattern specification, and examples of source code are often full of too much of implementation details [2]. A formal and object-oriented specification language, Object-Z [3,4], is then used to compensate the insufficiency of UML. We then give two examples in which the design pattern is applied, one in gaming, the other in training.

Z. Pan et al. (Eds.): Transactions on Edutainment IV, LNCS 6250, pp. 117–129, 2010.

2 Synchronizable Object

2.1 Intent

The Synchronizable Object (SO) pattern extracts the common inter-media synchronization behavior from the media objects with intrinsic timing and the ones without, such that distributed media objects can be synchronized at distinct and possibly user definable synchronization points.

2.2 Context and Forces

Many media objects have intrinsic timing (audio, video, animation), and some don't but require the content to be presented in a given order (for example, speech [2,5,6] and movements [7]), where as the others are static (text, image and graphics). Some media appear active (video, audio, animation and TTS) and have a automatic and successive behavior, whereas some others are passive (presentation slides, linked web pages) and require external drive to move forward. How to synchronize all these different media is not trivial. Especially in an ambient intelligent environment, media are considered in a generic sense, including both digital and physical objects. Designers have the freedom to create new forms of media hence new media types should be able to be incorporated into the system easily.

In JMF[8], timed media objects or streaming media objects that have intrinsic timing, for example, audio and video objects, are treated differently from media objects that do not have intrinsic timing. While synchronization mechanisms (Libraries and APIs) are well defined for timed media objects, other media objects such as graphics and text are treated as different data types for which no synchronization mechanisms are defined. To synchronize with the timed media objects, the controlling interfaces of streaming media have to be implemented on top of these non-timed types, which is not only unnatural, but also sometimes hard to impose the semantics of timed media on the non-timed ones.

In PREMO [9] and STORYML[10,11,12], all the media objects are assumed to be active, that is, every object has its own process driven by a *Timer* (PREMO) or a *MediaClock* (STORYML), even if the object does not have intrinsic timing. This approach simplifies the synchronization by providing a common time-based mechanism for all the objects, however may increase the process resource overload caused by the unnecessary active process that need not to be actually active.

Moreover, when distribute these media objects over a network, it is necessary to clearly sperate the timing concerns and the synchronization control and to minimize the communication load [2,13].

2.3 Solution

The design of this pattern is very much inspired by the structure appeared in many multimedia standards and systems, especially JMF, PREMO amd IPML [2]. The formal specification of the synchronization objects is based on

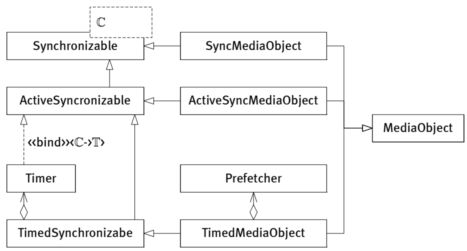


Fig. 1. Synchronizable Objects, where the *Synchronizable* can be distributed

the foundational objects described in [14], omitting exception handling mechanisms that are considered not to be important in a pattern specification, and applying a different event handling design. The design is then extended with the notion of time and the controlling structure inspired by JMF.

Structure. As the basis, the concept of the coordinate system is introduced. Every SO has a progressive behavior – to be started, step forward until the end if there is an end, and possibly repeat from the beginning. A step is modeled as a coordinate point and the sequence of the steps forms a coordinate system. An *ActiveSynchronizable* object automatically steps forward in its own process and a *TimedSynchronizable* automatically moves forward but also with timing constraints. The timing constraints are applied to the *TimedSynchronizable* by synchronizing with a *Timer*, which itself is an *ActiveSynchronizable* with a moving-forward or a moving-backward timeline at a specific speed. Synchronizable media objects are SO's at all the levels of the inheritance hierarchy, but also implement the presentation behavior as a *MediaObject*. In addition, a *TimedMediaObjects* needs a *Prefetcher* that keeps fetching enough amount of data ahead of the presentation process, so that immediate starting and continuous presentation are possible(Fig. 1).

The synchronization between SO's is event driven. A *Synchronizable* may trigger *StateSyncEvent*s when it changes its synchronization state (stopped, ready, started and paused), may also issue other events that are attached to a coordinate when the coordinate has been passed. Every SO dispatches the synchronization events to the event handlers that are interested in certain sets of events and sources. They an object of the type *EventDispatcher* that selects the registered *EventHandler* objects and dispatches the synchronization events of the interests. The *EventHandler* object in turn invokes concrete event handling operations in other objects. The *TimedSynchronizable* object synchronizes the internal state with a timer (reacting on the *StateSyncEvent*s), and present the data on the paces of that timer (reacting on the *TimedSyncEvent*s), which also demonstrates the use of this event-driven approach (fig. 2).

Synchronization states. A SO is a simple finite state machine with four states stopped, ready, started, paused. The initial state is stopped. A SO must get ready before it can be started. A ready state is necessary to model many

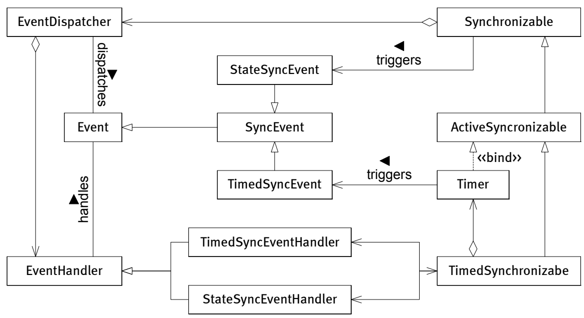


Fig. 2. Event based synchronization, in which the events can be be distributed

multimedia objects that requires computational resources to be allocated and certain amount of data to be prefetched before it can be started immediately. Once it has been started, it can be paused and be restarted again. A free type *SyncState* is introduced to identify these states:

$$SyncState ::= \mathsf{stopped} \mid \mathsf{ready} \mid \mathsf{started} \mid \mathsf{paused}$$

Coordinates. A SO controls the position and progress along its own coordinates. Different SO's may use different coordinates. For example, a video stream might use frame numbers as coordinates, and a clock might simply take $\mathbb{T}$ as its coordinates. A SO may have a function to map the coordinates to time and time constraints can be added to the operations – however the coordinates are not necessarily always time related, for example, a TTS engine may take white spaces or punctuation locations between words or sentences as coordinates. The concept of coordinates is more general than the time. To model the progression steps, the coordinates synchronization coordinates $\mathbb{C}$ is introduced and defined as a discrete set of discrete *points* that can be ordered with a relation $<$ that is irreflective, transitive, antisymmetric and total.

Repeating positions. A client of a SO might want to traverse the coordinates more than once, or possibly traverse in an infinite loop, for example, repeating a multimedia presentation several times or repeating it forever. This means a given point within the coordinates may be visited multiple times. A new type is defined to combine a point in the coordinates with a visit number, $Position == \mathbb{C} \times \mathbb{N}$, and a total order is defined over *Position*:

$$\_\ prec\ \_ : Position \leftrightarrow Position$$

$$\forall\, c_1, c_2;\ n_1, n_2 \bullet (c_1, n_1)\ prec\ (c_2, n_2) \Leftrightarrow n_1 < n_2 \lor (n_1 = n_2 \land c_1 < c_2)$$

Progression directions. A SO may advance along its coordinates, or backward in an inverted direction:

$$Direction ::= \mathsf{forward} \mid \mathsf{backward}$$

Synchronization Events. The coordinates can be attached with synchronization events. *Event* is defined as the super class for all the events in the system:

Event

$source : \mathbb{O}_{\circledcirc}$

$SetSource \mathrel{\widehat=} [\, \Delta(source)\ s? : \mathbb{O}_{\circledcirc} \mid source' = s? \,]$

which has an attribute *source* storing the reference to the originating object if desired[1]. Other details are left to subclasses. To distinguish synchronization events from other events (for example, user input events), *SyncEvent* is defined as a subclass of *Event*. Subtyping can be used to identify different types of *SyncEvent* objects, and the subclasses may also add other state variables to carry extra information. Instead of giving all the possible different types in advance, which is often not feasible, subtyping is a more flexible and extensible solution for identifying objects [15].

Synchronizable. Let's first give a visibility list that declares the visible state variables and operations:

Synchronizable

$\upharpoonright($ *eventDispatcher*, *repeatNumber*, *repeatCount*, *syncState*, *begin*, *end*, *current*, *syncElements*, *direction*, [Visible state variables]

INIT, *Start*, *TryStart*, *Stop*, *Pause*, *Visit*, *Step*, *setRepeatNumber*, *setDirection*, *setBegin*, *setEnd*, *setBeginEnd*, *attachSyncEvent*, *detachSyncEvent*, *addSyncElements*, *removeSyncElements* [Visible operations]

)

State variables. The state schema is then defined as follows:

Readable state variables:

$eventDispatcher : \downarrow EventDispatcher$

$repeatNumber : \mathbb{N}$ [How many times to repeat]

$repeatCount : \mathbb{N}$ [How many times repeating completed]

$syncState : SyncState$ [Current synchronization state]

$syncElements : \mathbb{C} \leftrightarrow \downarrow SyncEvent$ [Synchronization elements]

[1] The symbol $\mathbb{O}$ is defined as $\mathbb{O} == \downarrow Object$ where *Object* is the root superclass for all the classes in the system. $\downarrow Object$ is the set of all the objects of the class *Object* and its subclasses – the object universe of the system.

In this specification, a generic free type is also introduced to include the null value: $X_{\circledcirc} ::= X \mid \mathsf{null}$ for a given type X.

$begin : \mathbb{C}_{\circledcirc}$ [Where to begin]
$end : \mathbb{C}_{\circledcirc}$ [Where to end]
$current : \mathbb{C}_{\circledcirc}$ [Current location in the coordinates]
$syncSpan : \mathbb{P}\,\mathbb{C}$ [Synchronization span between $begin$ and end]
$direction : Direction$ [traversal direction]

Local state variables that are invisible to others:
$positions : \mathbb{P}\,Position$ [Positions to be traversed]
$curPosition : Position_{\circledcirc}$ [Current position]
$lastPosition : Position_{\circledcirc}$ [The position visited last time]
$\_ \prec \_ : Position \leftrightarrow Position$ [Order of traversal]

$$syncSpan = (\textbf{if } begin = \mathsf{null}\ \textbf{then}\ \mathbb{C}\ \textbf{else}\{c : \mathbb{C} \mid begin \leqslant c\}) \cap (\textbf{if } end = \mathsf{null}\ \textbf{then}\ \mathbb{C}\ \textbf{else}\{c : \mathbb{C} \mid c \leqslant end\})$$
[$begin$ and end define the synchronization span]

$$positions = syncSpan \times (\textbf{if } repeatNumber = 0\ \textbf{then}\ \mathbb{N}\ \textbf{else}\{n : \mathbb{N} \mid n < repeatNumber\})$$
[Positions are always defined by $syncSpan$ and $repeatNumber$]

$$\textbf{if } curPosition = \mathsf{null}\ \textbf{then}\ current = \mathsf{null} \wedge repeatCount = 0\ \textbf{else}\ curPosition = current \mapsto repeatCount$$
[Decompose current position to its coordinate and repeat count]

$$direction = \mathsf{forward} \Rightarrow (\_ \prec \_) = (\_\ prec\ \_)$$
$$direction = \mathsf{backward} \Rightarrow (\_ \prec \_) = (\_\ prec^{\sim}\ \_)$$
[Traversal direction defines the order of the positions]

The class aggregates an object of $\downarrow EventDispatcher$. Aggregation allows to separate them in independent processes and to dispatch the events asynchronously. The attributes *repeatNumber* keeps the total number of the traversal loops. The set *syncSpan* defines the span of the synchronization coordinates – The client of a SO may define the boundaries to limit the synchronization to a subset of coordinates. The relation *syncElements* attaches *SyncEvent* objects to certain coordinates. The attribute *positions* defines the positions (pairs of coordinates in *syncSpan* and the repeat counter) that will be passed during traversal. The attributes *curPosition* and *lastPosition* keeps track of the current traversal position, and the position that has just visited last time. The relation $\prec$ defines the order of the traversal according to the specified *direction*. The attributes *positions*, *curPosition* and *lastPosition* are for internal use and hence they are not visible to other objects.

Upon initialization, the repeat number by default is 1 and the repeat count starts from 0. The synchronization state is set to be stopped and there are no *syncEvent* attached to any coordinates. The *syncSpan* covers all the possible coordinates and the traverse direction is forward. The attribute *curPosition* and *lastPosition* are set to null, which means there is no last visited position, and the current position is not given yet.

$$
\begin{array}{l}
\textit{Init} \\
\hline
repeatNumber = 1 \land repeatCount = 0 \\
direction = \mathsf{forward} \land syncSpan = \mathbb{C} \\
syncState = \mathsf{stopped} \land syncElements = \varnothing \\
curPosition = \mathsf{null} \land lastPosition = \mathsf{null} \\
\hline
\end{array}
$$

Synchronization operations. Many operations may cause the synchronization state change and trigger corresponding *StateSyncEvent*. A schema *Transit* is defined to catch this common behavior:

$$
\begin{array}{l}
Transit_0 \\
\hline
\Delta(syncState, syncState') \\
e! : StateSyncEvent_{\circledcirc} \\
\hline
\textbf{if } syncState \neq syncState' \\
\quad \textbf{then } e! \neq \mathsf{null} \land e!.source = self \land \\
\qquad e!.oldState = syncState \land e!.newState = syncState' \\
\quad \textbf{else } e! = \mathsf{null} \\
\hline
\end{array}
$$

$$
\begin{array}{l}
Transit \mathrel{\widehat{=}} Transit_0 \mathbin{⨾} ([\, e? : StateSyncEvent_{\circledcirc} \,] \bullet \\
\qquad (\textbf{if } e? \neq \mathsf{null} \textbf{ then } eventDispatcher.Dispatch))
\end{array}
$$

The operations that manipulate the synchronization state are defined next. The operation *Stop* puts the SO into the stopped state from any other state. Stopping an object also causes its repeat counter, the current position and the last visited position to be reset to their initial state.

$$
\begin{array}{l}
Stop_0 \\
\hline
\Delta(syncState, repeatCount, curPosition, lastPosition) \\
\hline
syncState' = \mathsf{stopped} \land repeatCount' = 0 \\
curPosition' = \mathsf{null} \land lastPosition' = \mathsf{null} \\
\hline
\end{array}
$$

$$Stop \mathrel{\widehat{=}} Stop_0 \land Transit$$

The operation *Ready* gets the stopped SO ready for starting. This operation has no effects if the object is in any state other than stopped:

$$
\begin{array}{l}
Ready_0 \\
\hline
\Delta(syncState) \\
\hline
syncState = \mathsf{stopped} \Rightarrow syncState' = \mathsf{ready} \\
syncState \in \{\mathsf{started}, \mathsf{paused}, \mathsf{ready}\} \Rightarrow syncState' = syncState \\
\hline
\end{array}
$$

$$Ready \mathrel{\widehat{=}} Ready_0 \land Transit$$

A SO can be started when it is ready or paused:

$$
\begin{array}{|l}
\textit{Start}_0 \\
\hline
\Delta(syncState, curPosition) \\
\hline
syncState = \mathsf{ready} \Rightarrow \\
\quad curPosition' = minimum(\_ \prec \_, positions) \\
\quad \textbf{if } curPosition' = \mathsf{null} \textbf{ then } syncState' = \mathsf{stopped} \\
\qquad \textbf{else } syncState' = \mathsf{started} \\
syncState = \mathsf{paused} \Rightarrow syncState' = \mathsf{started} \\
syncState = \mathsf{started} \Rightarrow syncState' = syncState \\
\hline
\end{array}
$$

$$Start \mathrel{\widehat=} Start_0 \land Transit$$

Once the SO has been started, it can be put in to the paused by the operation *Pause* defined below:

$$
\begin{array}{|l}
\textit{Pause}_0 \\
\hline
\Delta(syncState) \\
\hline
syncState = \mathsf{started} \Rightarrow syncState' = \mathsf{paused} \\
syncState \in \{\mathsf{stopped}, \mathsf{ready}, \mathsf{paused}\} \Rightarrow syncState' = syncState \\
\hline
\end{array}
$$

$$Pause \mathrel{\widehat=} Pause_0 \land Transit$$

Other operations. "Setter" operations are needed to change the value of the state variables explicitly. These operations are straightforward, nonetheless it should be noticed that because the postcondition of every operation must yield to the predicates in the state schema, changing values of the state variables may implicitly change the value of the others. We omit the specifications (and hereafter) due to the limited space.

A set of event operations need to be defined. For example *SetEventDispatcher* that associates an event dispatcher to the SO, *Attach* and *Detach* operations to add or remove the synchronization events to the coordinate, and *DispatchEvents* takes a sequence of *SyncEvent* sets as input and sends the events simutaneously to the event handler for dispatching.

The progression of a SO is modeled by specifying the coordinates as milestones, unfolding repetition into positions, and defining the visiting order as per direction. The object has now the static basis for dynamic progression. A default "stepping" behavior bring in the dynamic aspects by assuming the next position to visit is the closest position after the current one, without knowing whether it is driven by its own process or by anything else.

A SO has now been defined as such so that no media specific semantics is directly attached to it. For example, there is no notion of time although it is crucial for time-based media objects. Subclasses, realizing specific media control should, through specification, attach concrete semantics to the object through their choice of the type of the internal coordinate system, through a proper specification of what “visiting a position” means, and through a proper specification of how the object should move from the current position to the next.

Other Classes. Other classes depicted in Fig. 1 and Fig. 2, such as *ActiveSynchronizable*, *SyncMediaObject*, *Timer*, *TimedSynchronizable*, *TimedMediaObject*, *EventDispatcher* and *EventHandler*, are omitted here due to the limited space. The complete specifications can be found in [2].

2.4 Consequences

Benefits

Separated timing concerns: Timing constraints are added only when they are necessary and the timer can be shared – which saves concurrent processing resources. Static media objects, for example, static pictures, can be easily implemented as passive SO’s with a simple coordinate system that only has one coordinate and it will stay static until there is an external drive to move it forward – which naturally stops it because there is no next coordinate. If there is a need for the picture to stay presenting itself for a certain amount of time, it can be then implemented as an *TimedSynchronizable* that uses a timer. For media objects that have a progressive presenting behavior and that are not constrained by the time, for example, a TTS object, the *ActiveMediaObject* class does what is needed. However, if the TTS object need to presented at a speed of a video stream, a timer can then be shared to make a *TimedMediaObject* so that the TTS object can speak faster or slower.

Flexible concurrency: The concurrency of the structure can be from none to fully concurrent.

Simple and unified synchronization interface: SO’s at different levels have the same synchronization controlling interface and the same state transition scheme, no matter whether the coordinate system is time based or not. This makes it easy to incorporate new types of SO’s without influencing existing types.

Open for extension: The actual presentation behaviors of media objects are left open, and the event-based synchronization mechanism can be extended for other purposes. For example, user interaction events can be easily added to the structure in addition to the state transition events and the timing events.

Liabilities

Hard to debug: When multiple processes are involved, *Syncronizable* can be hard to debug since the inverted flow of control oscillates between the framing infrastructure and the parallel callback operations on the synchronization controls and event handlers, especially when they are distributed over devices [16].

Event dispatching efficiency depends on the hosting platform and the performance of the network: This is a liability of all event based synchronization mechanisms. In the specification, events are sometimes required to be dispatched simultaneously using a conjunction operator. It is possible to emulate the semantics of the simultaneous operation using multiple threads, however this can only be efficient when the number of the included threads is small.

3 Examples

3.1 DeepSea

In the EU funded project ICE-CREAM we investigated how to make compelling experiences for end-users based on the possibilities of integrating technologies for interactive media, for example, DVB-MHP, MPEG-4, 3D graphics and Internet technologies [17,18]. Technological options that address different levels of interactivity for end-users were investigated and implemented in prototypes. In one of the applications a video is enhanced with fictional content, which uses 3D graphics and animations to enhance the users viewing experience. This application is a prototype of an interactive 3D movie about a deep-sea adventure in a submarine. Aside from the combination of video streams, 3D computer graphics, text, still pictures, soundtracks and multilingual interfaces, it also enhances the user¡¯s experience by distributing the interactive objects to multiple devices (lights, portable displays, robotic toys) to create ambience effects in the performance space - the end user¡¯s environment (fig. 3).

Fig. 3. Distributed media objects in DeapSea

In DeepSea, sound, video objects and 3D animations are *TimedMediaObjects* that have intrinsic timing. Some of the still pictures and text were modeled as static *Synchronizables* that are controlled only by other objects or user interaction, but some of them were *TimedSynchronizabes* because they appear and disappear based on its internal timing mechanism. Lighting effects and robotic behaviors were new forms of *TimedMediaObject* that act and react based on the timing and interaction events. All these objects are synchronized over a network based on an IPML script [13].

3.2 Simulation-Based Delivery Training

The numbers of high-risk pregnancies and premature births are increasing due to the steadily higher age of pregnancy. Medical simulators are used in training the doctors to deal with emergent perinatal situations. To enhance the training effect, sophisticated simulators are integrated into a realistic training environment that takes into account the medical instruments and team aspects. The training environment becomes increasingly complex and requires a clear structure for different training scenarios and flexible hardware configurations. Eindhoven

Fig. 4. Team training with delivery simulators at MMC

University of Technology is cooperating with Máxima Medical Center, aiming at the next generation simulation based training facilities. The result of this effort is the design and implementation of the MedSim system [19]. The system brings software and hardware devices and components into the training room than a single patient simulator, one of which is for example the device designed with audio-visual feedback for medical staff to perform Cardiopulmonary Resuscitation (CPR) more efficiently [20,21]. We aim at an open system architecture that is flexible and extensible enough for the industry to introduce further development and future technologies into simulation based team training. In the design of the delivery simulation system, distributed sensors and actuators are integrated into the mother manikin, the baby manikin, as well as the environment for the purpose of medical team training (fig. 4).

Here the behaviors of manikins are modeled as *TimedSynchronizables* and the projections of 3D graphics are *TimedMediaObjects*. Separated input from the sensors (for example, the positions of the team members detected by the camera) and output from the actuators (skin color and muscle tone of the baby for example) are modeled as *Synchronizables* so that they can be initiated, enabled and disabled at certain moments. The components of the system are coordinated by a scenario based script during a training session [22].

4 Conclusions

The Synchronized Object pattern can be applied for content elements with or without intrinsic timing in distributed multimedia applications. The pattern language is used here to capture the common features of these objects. Although the complete specifications are not included here, the Object-Z examples of a few classes in this pattern are enough to show how a formal method can be applied to treat the architectural construct and compensate the insufficiency of object-oriented language such as UML. The Object-Z specification provides a better amount of generality and abstraction than UML and concrete examples, with the power of both formalization and object orientation for elementary element building blocks – it has the properties that an architectural specification should have [23]. However, the Object-Z specification is not intended to replace natural language specifications. Instead, formal definitions are complementary to existing means (natural language, code samples, etc.). Two examples in gaming and training were briefly introduced without getting into the implementation details, but only showing that the described pattern is applicable for different situations in distributed multimedia applications.

References

1. Alexander, C., Ishikawa, S., Silverstein: A Pattern Language: Towns, Buildings, Construction. Oxford University Press, Oxford (1977)
2. Hu, J.: Design of a Distributed Architecture for Enriching Media Experience in Home Theaters. Technische Universiteit Eindhoven (2006)
3. Duke, R., Rose, G.: Formal Object-oriented Specification using Object-Z. Macmillan Press Limited, London (2000)
4. Smith, G.: The Object-Z spectification language. In: Advances in formal methods, Kluwer Academic Publishers, Dordrecht (2000)
5. Bartneck, C., Hu, J., Salem, B., Cristescu, R., Rauterberg, M.: Applying virtual and augmented reality in cultural computing. International Joural of Virtual Reality 7(2), 11–18 (2008)
6. Hu, J., Bartneck, C., Salem, B., Rauterberg, M.: Alice's adventures in cultural computing. International Journal of Arts and Technology 1(1), 102–118 (2008)
7. Hu, J.: Move, but right on time. In: Workshop on design and semantics of form and movement (DeSForM 2005), Newcastle upon Tyne, pp. 130–131 (2005)
8. Sullivan, S., Brown, D., Winzeler, L., Sullivan, S.: Programming With the Java Media Framework. John Wiley & Sons, Chichester (1998)
9. Duke, D.J., Herman, I., Marshall, M.S.: PREMO: A Framework for Multimedia Middleware: Specification, Rationale, and Java Binding. LNCS, vol. 1591, p. 1. Springer, Heidelberg (1999)
10. Hu, J.: Distributed Interfaces for a Time-based Media Application. Post-master thesis, Eindhoven University of Technology (2001)
11. Hu, J.: StoryML: Towards distributed interfaces for timed media. In: ten Kate, W. (ed.) Philips Conference InterWebT 2002, NatLab, Eindhoven (2002)
12. Hu, J.: StoryML: Enabling distributed interfaces for interactive media. In: The Twelfth International World Wide Web Conference, Budapest, Hungary (2003)
13. Hu, J., Feijs, L.: IPML: Structuring distributed multimedia presentations in ambient intelligent environments. International Journal of Cognitive Informatics & Natural Intelligence (IJCiNi) 3(2), 37–60 (2009)
14. Duke, D., Duce, D., Herman, I., Faconti, G.: Specifying the PREMO synchronization objects. Technical Report ERCIM-01/97-RD48, ERCIM (1997)
15. Simons, A.J.H.: The theory of classification, part 4: Object types and subtyping. Journal of Object Technology 1(5), 27–35 (2002)
16. Hu, J., Feijs, L.: An adaptive architecture for presenting interactive media onto distributed interfaces. In: Hamza, M. (ed.) AI 2003, pp. 899–904 (2003)
17. Hu, J., Janse, M., Kong, H.J.: User experience evaluation of a distributed interactive movie. In: HCI International 2005, Las Vegas (2005)
18. Feijs, L., Hu, J.: Component-wise mapping of media-needs to a distributed presentation environment. In: The 28th Annual International Computer Software and Applications Conference (COMPSAC 2004), Hong Kong, pp. 250–257 (2004)
19. Hu, J., Peters, P., Delbressine, F., Feijs, L.: Distributed architecture for delivery simulators. In: International Conference on e-Health Networking, Digital Ecosystems and Technologies (EDT 2010), Shenzhen, China, pp. 109–112 (2010)

20. Chen, W., Oetomo, S.B., Feijs, L.M.G., Andriessen, P., Geraets, F.K.M., Thielen, M.: Rhythm of Life Aid (ROLA) – an integrated sensor system for supporting medical staff during cardiopulmonary resuscitation (CPR) of newborn infants. Submitted to IEEE Transactions on Information Technology in Biomedicine (2009)
21. Oetomo, S.B., Feijs, L.M.G., Chen, W., Andriessen, P.: Efficacy of audio-promoted rate guidance for insufflation and chest compressions and feed-back signalling for the pressure of chest compressions during cardio-respiratory resuscitation (CPR) of newborn infants. In: The annual meeting of the Society for Pediatric Research (SPR 2009), Baltimore, US (2009)
22. Hu, J., Feijs, L.: A distributed multi-agent architecture in simulation based medical training. In: Chen, Q. (ed.) Transactions on Edutainment III. LNCS, vol. 5940, pp. 105–115. Springer, Heidelberg (2009)
23. McComb, T., Smith, G.: Architectural design in Object-Z. In: ASWEC 2004: Proceedings of the 2004 Australian Software Engineering Conference (ASWEC 2004), Washington, DC, USA, pp. 77–86. IEEE Computer Society, Los Alamitos (2004)

Anisotropic Cloth Modeling for Material Fabric*

Mingmin Zhang, Zhigeng Pan, and Qingfeng Mi

State Key Lab. of CAD&CG, Zhejiang University, Hangzhou, 310027, China
{zmm,zgpan}@cad.zju.edu.cn, mqfbreeze@126.com

Abstract. Physically based cloth simulation has been challenging the graphics community for more than three decades. With the developing of virtual reality and clothing CAD, it has become the key technique of virtual garment and try-on system. Although it has received considerable attention in computer graphics, due to its flexible property and realistic feeling that the textile engineers pay much attention to, there is not a successful methodology to simulate cloth both in visual realism and physical accuracy. We present a new anisotropic textile modeling method based on physical mass-spring system, which models the warps and wefts separately according to the different material fabrics. The simulation process includes two main steps: firstly the rigid object simulation and secondly the flexible mass simulation near to be equilibrium. A multiresolution modeling is applied to enhance the tradeoff fruit of the realistic presentation and computation cost. Finally, some examples and the analysis results show the efficiency of the proposed method.

Keywords: Cloth simulation, material fabric, anisotropic modelling.

1 Introduction

The wide range of textiles and clothing in daily-life has motivated much computer graphics research on cloth modeling and rendering. Cloth modeling plays the key role in textile simulation. Mori[1] introduced an interactive system called Plushie, which could design plush toy from 2D pattern to 3D model. Decaudin[2] put the virtual garment on the characters.

The modeling of flexible material, such as textile object, is different from that of rigid body and fluid object. The rigid object has its own internal structure that remains stationary, while the moving and displacement of fluid object totally rely on the repulse objects. However, the textile object has its own structure that deforms according to the repulse object moving. The deformation degree relies on the textile material (wool or nylon) and fabric (plain or twill). The physically based clothing animation takes consider of various physical features of real textile in real world, such as the quality, spring and damping. The dynamics equation can be solved by Newton Law which takes the positions and velocities as inputs.

According to the elasticity theory, the flexible object deformation is first described as the object response movement under the spring, external force and damping[3]. 2D garment clip is produced by Thalmann[4] and then combined to 3D clothing. This

* It is supported by the project of Nation Science Foundation of China (Grant 60970076).

Z. Pan et al. (Eds.): Transactions on Edutainment IV, LNCS 6250, pp. 130–144, 2010.

method is improved by Parks[5] and Baraff[6] to be more realistic looking. The textile animation is generated in Breen [7] by energy minimization, taking use of the famous spring-mass model. More enhanced methods[8,9,10] are presented to make simulation more accurate and useful, although they can hardly be adopted soon in real-time application.

The textile dynamics model of Ling[11] simulates the textile motion under the wind, in which the textile surface can be divided into m*n impenetrable fluid surfaces. The damping of each fluid surface can be received by Navier-Stokes equation, while the textile deformation is obtained by Terzopoulos flexible deformation model[3]. The flexible model is extended to triangle garment model in Volino[12] and the performance of different numerical solution of the previous studies is analyzed. The textile simulation progress is divided into several steps with multiresolution model in Zhang[13]: in the first step, modeling, force analyses, collision and animation are played on the coarse model till equilibrium. Next, taking the first equilibrium as input state, the finer model goes to balance.

Among the above methods with different solutions, the classical numerical solution includes explicit integration and implicit integration. The explicit numerical integration is more flexible, even though it has some disadvantages, such as the high computational cost by small time clip in order to avoid the collision penetration, while implicit integration can speed up the convergence with unrealistic feeling. The hybrid methods of explicit and implicit integration are developed by Boxeman[14] and Bridson[15]. Since the elastic parameters of spring are difficult to determine, several works[16][17] use video data of moving cloth to estimate the elastic and damping parameters.

The artificial damping presentation in the implicit method of Baraff model is mitigated by the carefully tailored bending semi-implicit model of Choi[18]. A new implicit approach is presented by Seungwoo[19] to improve the simulation falsity with artificial friction. The continuum medium force theory is put into mass-spring system to enhance the reality of garment animation[20].

Lots of current research work go to the key issue of collision detection and response, for example, Govindaraju[21] precomputes a chromatic decomposition of a mesh into non-adjacent primitives using graph coloring algorithms to achieve higher culling efficiency and reduce the number of false positives. Cutler[22] makes a layer, separated from cloth mesh, over which the wrinkles can be controlled and produced regardless of mesh shapes and fineness.

In current textile industry, the most widespread materials are wovens, which are made of two kinds of threads called wefts and warps. The wefts and warps interweave by means of special structures, and then give rise to special pattern and texture. This special structure is woven fabric. The ordinary fabric can be categorized as plain, twill and satin (as Fig. 1 shows)

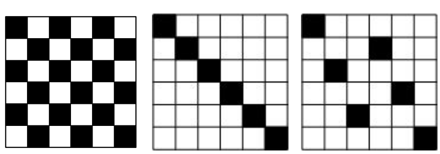


Fig. 1. The three prime textile fabrics (plain, twill and satin)

The traditional CAD system processes the rigid model, which is well developed now. Nevertheless it is unsuitable for textile objects. Textile clothing is a flexible material, which can be deformed greatly by stress. The deformation degree depends on the material's composing, the fabric structure, seam and clipping et al.

In the following section, an anisotropic textile modeling approach, which makes model along the warp direction and weft direction separately, is presented to speed up the process and represent the characteristic of anisotropic textile fabric.

2 Anisotropic Textile Modeling

The clothing modeling and dynamic force response are both anisotropic. The anisotropic cloth model includes:

- Hardly isotropic fabric, such as basket fabric and plain fabric
- Warp based woven: the rigid of weft is larger than that of warp
- Weft based woven: the rigid of warp is larger than that of weft

On the ground of the different yarn quality of weft and warp, the great numerical gap between weft and warp direction causes fault convergence vibration in isotropic system. Therefore, an anisotropic method, by means of separating the modeling of weft and warp direction, is proposed to normalize the parameters of weft and warp (such as Young modulus and Poisson's ratio). Some textile concepts are given first to ease the description.

(1) Interlacing

One interlacing means a woven cross between a weft and a warp. The interlacing numbers of wefts and warps are provided in Equation(1):

$$L_{warp} = \sum_{i=1}^{R_{wrp}} l_i^{warp} \ , \ L_{weft} = \sum_{i=1}^{R_{wft}} l_i^{weft} \tag{1}$$

where l_i^{warp} is the interlacing number of warp i, l_i^{weft} is the interlacing number of weft i. R_{wrp} and R_{wft} are the warp and weft number respectively.

(2) Float

The weft float and warp float are provided in Equation(2):

$$Fl_{warp} = \frac{\sum_{i=1}^{R_{wrp}} \frac{R_{wft}}{2 * l_i^{warp}}}{R_{wrp}} \ , \ Fl_{weft} = \frac{\sum_{i=1}^{R_{wft}} \frac{R_{wrp}}{2 * l_i^{weft}}}{R_{wft}} \tag{2}$$

The fabric number between two adjacent neighbor interlacings is:

$$\phi_{warp} = \frac{R_{wrp} R_{wft}}{2 L_{warp}} \ , \ \phi_{weft} = \frac{R_{wrp} R_{wft}}{2 L_{warp}} \tag{3}$$

To model the virtual textile, some parameters are very important. Taking the weft based woven as an example, in which $L_{warp} < L_{weft}$, one particle in warp direction is corresponding to k particles in weft direction, where $k = \left\lceil \frac{L_{weft}}{L_{warp}} \right\rceil$. The bigger ϕ is, the bigger the spring coefficient is, while the smaller Fl is, the bigger the spring coefficient is (described in section 2.2 in detail).

2.1 Separately Modeling for Weft and Warp

Comparing with the bending deformation, the stretch and condensation of yarn are much smaller. Therefore, the initial step of simulation assumes the length of weft and warp thread to be unchangeable. The wrinkle is produced by the bending of the intersected particles. The initial process can release some deformation energy into the middle balance state. This method can not only increase the convergence speed, but also solve the over-stretch problem in Feynman system [23].

Assuming the material is composed of $n*m$ grids, called particles, the particle in the i^{th} row and the j^{th} col is represented as (i,j), whose position is shown as $\vec{x}_{i,j} \in R^3$ in the world coordinate. The mass of particle (i,j) is $m_{i,j}$. The force acted to particle (i,j) is $\vec{f}_{i,j} \in R^3$. The distance of the two adjacent particles in warp direction is $l^v_{i,j} = v_{i,j} - v_{i-1,j}$, while in weft direction is $l^u_{i,j} = u_{i,j} - u_{i,j-1}$.

Taking the weft based woven as an example, the rigid of warp is larger than weft and the wrinkle effect in weft direction is greater than that of warp direction. In our multi-resolution model, one particle in warp direction is related to k particles in weft direction. Therefore, the position in weft direction that can lead to bending is more than that of warp (shown in Fig. 2 and Fig. 3).

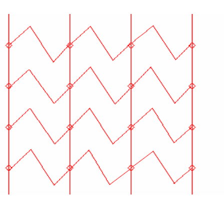


Fig. 2. The anisotropism of warp and weft

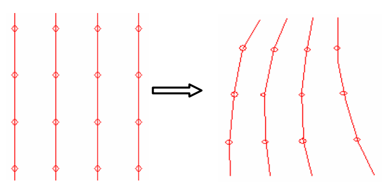


Fig. 3. Deformation of warp direction

(1) The constraints in warp direction

Assuming the rigid of warp is larger than that of weft, let one particle in warp direction be related to k particles in weft direction, which is connected by k-splines.

For particle (i,j), whose position is $\vec{x}_{i,j}$, the resultant force $\sum \vec{f}$ played to the particle can be separated as three components: $\vec{f}^v$ along with the warp direction, $\vec{f}^u$ vertical to $\vec{f}^v$, and $\vec{f}^z$ parallel to the z direction (shown in Fig. 4).

Then the position, velocity, acceleration can be calculated by Equation (4), where $\vec{x}^v_{i,j}$, $\vec{x}^u_{i,j}$ and $\vec{x}^z_{i,j}$ are the position components of particle (i, j) in warp, weft and z direction. $\omega_{i,j}$ is the angular velocity of particle (i, j). Fig. 5 shows the simulation of "initial equilibrium— attaching force— force cancelled – balance again".

$$\begin{aligned} \vec{f}^{\,v} &= m_{i,j} \cdot \omega_{i,j}^{\ 2} \cdot l^v_{i,j} \\ \omega_{i,j} \cdot l^v_{i,j} &= \frac{d\vec{x}^n_{i,j}}{dt} \\ \frac{d^2\vec{x}^t_{i,j}}{dt^2} &= \frac{\vec{f}^{\,t}}{m_{i,j}} \\ \frac{d^2\vec{x}^z_{i,j}}{dt^2} &= \frac{\vec{f}^{\,z}}{m_{i,j}} \end{aligned} \tag{4}$$

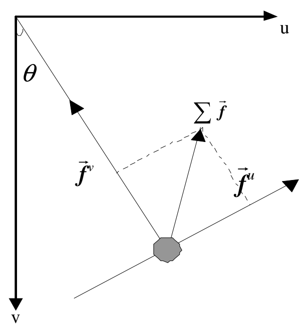


Fig. 4. The force acted on the particle

(2) The constraints in weft direction

After receiving the position of warp particles according to the warp constraint, the weft constraint is modified to keep the distance of the adjacent particles under the $l^u_{i,j}$ neighbor region of the balance position, otherwise, the spring system, described in next section, is applied to minimize the energy caused by the external forces.

As the example in Fig. 2, one line(two particles) in warp direction is related to 3 lines in weft direction. It looks like a ball gemel, which connected 3 bars. Each bar can rotate around the conjunctive point freely. Within a certain distance, the three rigid bars stretch and compress without starting the spring system. The material bending occurs at the intersection of particles. The wrinkles and tucks that reflect on the textile deformation can be produced with the help of enough particles.

It is the first step of our clothing simulation to separate the modeling of warp and weft direction and make the cloth bending by means of the rigid gemel movement. Then with the help of external and internal spring forces, the clothing makes bending, stretching, shearing, and then to final equilibrium.

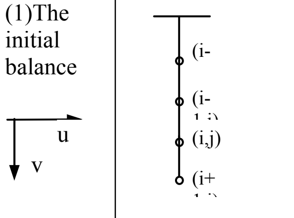
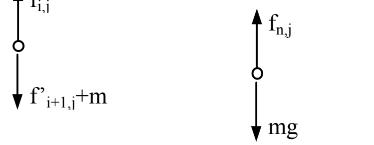
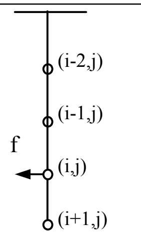
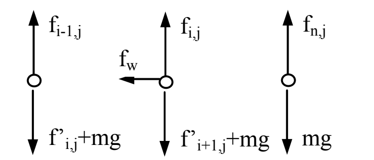
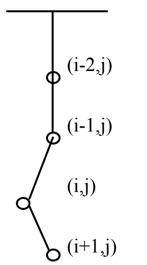
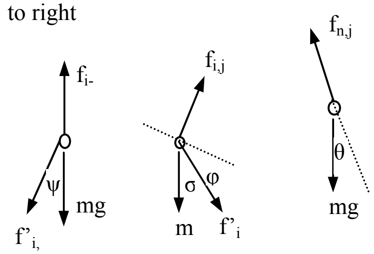
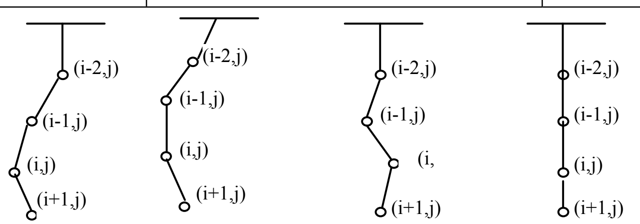
State	Sketch	The force played on the particle	Equation
(1)The initial balance		Except the lowermost The lowermost: 	Particle*(i,j)*: $f_{i,j} = f'_{i+1,j} + mg$ The lowermost*(n,j)*: $f_{n,j} = mg$
(2) External force acted on particle*(i,j)*		The particle*(i-1,j),(i,j),(n,j)* from left to right 	Particle*(i,j)* at *V* direction: $f_{i,j} = f'_{i+1,j} + mg$ Particle*(i,j)* at *U* direction: $f_w = ma_{i,j}$
(3) External force cancelled.		The particle *(i-1,j),(i,j),(n,j)* from left to right 	particle*(i-1,j)*: $\begin{cases} f_{i-1,j} = f'_{i,j}\cos\psi + mg \\ f'_{i,j}\sin\psi = ma_{i-1,j} \end{cases}$ particle *(i,j)*: $\begin{cases} f_{i,j} = f'_{i+1,j}\cos\varphi + mg\cos \\ mg\sin\sigma + f'_{i+1,j}\sin\varphi = ma \end{cases}$ particle *(n,j)*: $\begin{cases} f_{n,j} = mg\cos\theta \\ mg\sin\theta = ma_{n,j} \end{cases}$
(4) Balance Again			

Fig. 5. The simulation of “initial balance— attaching force— force cancelled – balance again”

2.2 The Anisotropic Spring System

The clothing appearance depends mainly on its shearing and bending deformation. The shearing effect comes from the in-plane movement, and the bending effect is from the out-plane stress. The force can be calculated from the related position of neighbor particles. To maintain the characteristic of clothing, there are three kinds of springs:

- Structure spring: between particle *(i,j)* to particle *(i+1,j)*, particle*(i,j)* to particle*(i,j+1)*;
- Shearing spring: between particle *(i,j)* to particle *(i+1,j+1)*, particle *(i+1,j)* to particle *(i,j+1)*;
- Bending spring: between particle *(i,j)* to particle *(i+2,j)*, particle *(i,j)* to particle *(i,j+2)*.

The springs between particle *(i,j)* to its neighbors are shown in Fig. 6. According to the anisotropic feature of clothing, the spring played to the clothing can be separated along the warp direction and the weft direction. The clothing appearance, responding to the springs, depends on the three kinds of the spring coefficients. Therefore, the three kinds of spring can be described as Equation (5),(6) and (7).

$$F_{struct}(r_{ij}) = \begin{cases} 0, & 0 < r_{ij} < \xi_1 \\ k_1(r_{ij} - \xi_1), & \xi_1 < r_{ij} < \xi_2 \\ c_0(r_{ij} - \xi_2)^5, & r_{ij} > \xi_2 \end{cases} \tag{5}$$

where $r_{ij} = | l_1^u - l_0^u |$, l_0^u is the distance of adjacent particles in warp direction while in initial balance state, l_1^u is that under the condition that the gemel opens or closes to some extent, with $r_{ij} < \xi_1$, which means the line between the adjacent particles moves as a rigid bar. When $\xi_1 < r_{ij} < \xi_2$, the strain is linear to the stress according to the Newton Law. When $r_{ij} > \xi_2$, the stress is very large that should be processed specially (shown in Fig. 7).

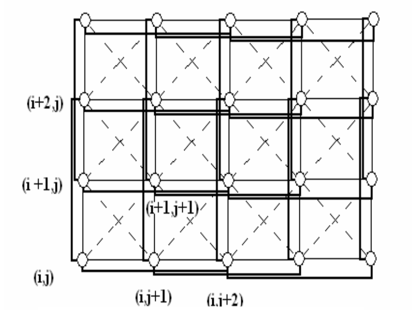


Fig. 6. Spring system of cloth

$$F_{bend} = k_2 \tan(\pi - \theta_{ij}) \tag{6}$$

where, the bending spring reaches to 0 when $\theta_{ij} = \pi$.

$$F_{shear} = k_3 \tan(\Phi_{ij}) \tag{7}$$

where, the shearing spring reaches to 0 when Φ_{ij} =0.

For the 2D textile material, rather than the 1D fiber, the warp strain will influence the weft stress deformation. Let the warp direction be *i*, the weft direction be *j*, the relationship of strain and stress is shown in Equation (8).

$$\begin{bmatrix} \sigma_{ii} \\ \sigma_{jj} \\ \sigma_{ij} \end{bmatrix} = \frac{1}{1 - V_i - V_j} \begin{bmatrix} (1-V_j)E_i & V_j E_i & 0 \\ V_i E_j & (1-V_i)E_j & 0 \\ 0 & 0 & G(1-V_i-V_j) \end{bmatrix} \begin{bmatrix} \varepsilon_{ii} \\ \varepsilon_{jj} \\ \varepsilon_{ij} \end{bmatrix} \quad \frac{E}{3} \le G \le \frac{E}{2}, \quad 0 \le V_i, V_j \le 0.5 \tag{8}$$

where ε is the stress, σ is the strain, *E* is the Young modulus, *V* is the Poisson's ratio and *G* is the stiffness weighting.

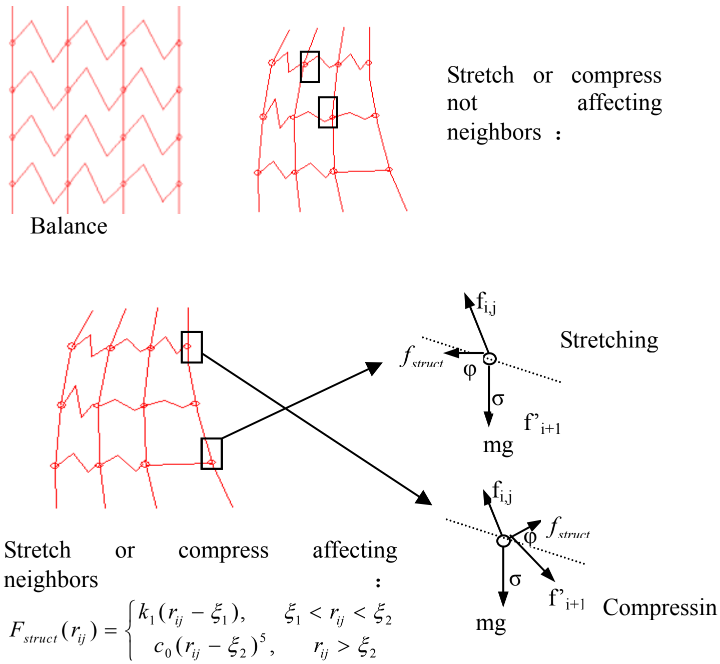


Fig. 7. The spring acted on the cloth

2.3 Force Model and Its Field for Cloth Animation

The physically based cloth animation relies on Equation (9), where x is the position of particle, the matrix M expresses the mass distribution. E is the scale function of x, which represents the internal energy of clothing. $\frac{\partial E}{\partial x}$ is the internal spring, such as structure, shearing and bending force, which has been described at above section. F, the function of x and $\dot{x}$, gives the external force acted on the clothing (such as wind, gravity and damping).

$$\ddot{x} = M^{-1}\left(-\frac{\partial E}{\partial x} + F\right) \tag{9}$$

The typical external forces are the wind and its resistance. The wind, something different from the other external forces, is an indirect force, affecting the cloth motion through the relative motion of air flow over the cloth surface.

Based on low-speed unsteady aerodynamic theory, a force distribution model[11] is described to estimate the unsteady force distributions on cloth surfaces. Instead of pre-defined concentrated force models, this wind model is based on physical model, which can be repeated without modifying the parameters interactively. However, its disadvantages lie on: (1) The wind speed should meet up to the clothing deformation. If the wind speed changing is much faster than the cloth deformation, many vortexes result in excessive computation cost. If the wind speed changing is much slower than the transformation of cloth movement, the wind plays as a fixing force; (2) It is very intricate to analyze the relative movement of the wind and the particles.

Under most conditions, the aerodynamics of air flow around a moving object is unsteady, nonlinear, complicated and highly turbulent. Fortunately, when air flows at a low speed (<= 72KM/h), it can be reasonably simplified to be incompressible, irrigational, and inviscid. With this in mind, we apply hydrodynamics theories to explain the relationship of wind model and the cloth movement. Replacing the aerodynamics theory with hydrodynamics depends on the assumption that, under some condition (for example, free and unclosed environment), the air is something like fluid that can not be compressed.

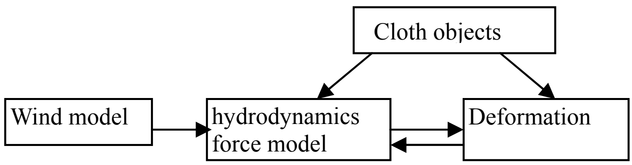


Fig. 8. The wind model and the cloth deformation

The process of generating dynamic cloth animation is based on the physical wind model, hydrodynamics force model, and deformation(shown in Fig. 8). First, a free stream wind model is simulated and then used as the input for the hydrodynamics field force model to produce the force distribution over the cloth surface. Finally, the movement of the surface under these forces can result in the deformation of the cloth.

The wind model is shown in Fig. 9. The force played to the particles is expressed in Equation (10). For each particle on the cloth surface, $\vec{f}_w$ is the wind force in the center of wind field, $\vec{d}$ is the vector that links the cloth particle and the wind center, $\|\vec{d}\|$ is the distance from wind center to the cloth particle, φ is the angle between the vector d and the mean normal of the particles, c_μ and c_d determine the influence of the surface density and the distance to the power of wind.

$$\vec{f} = c_\mu \cos\varphi \frac{\vec{f}_w \bullet \vec{d}}{c_d \cdot (1 + \|\vec{d}\|^2)} \vec{f}_w \tag{10}$$

The wind affects the movement of the cloth, meanwhile is influenced by the cloth after collision. Assuming the wind to be a fluid, some definitions are given in the first place: the stream line is tangent with the velocity vector, the stream field is the space and time function of a certain fluid. The fluid particle of wind moves along the stream line, colliding with the cloth, then releasing the force. As shown in Fig. 10, the stream line is a 3D curve, defined as $f(x, y, z) = 0$. According to the Newton Law, the kinetics equation is $\frac{\partial v_x}{\partial x}+\frac{\partial v_y}{\partial y}+\frac{\partial v_z}{\partial z}=0$. That means the volume of the fluid remains unchangeable. The velocity of fluid particle at *x, y, z* direction v_x, v_y, v_z constrain with each other.

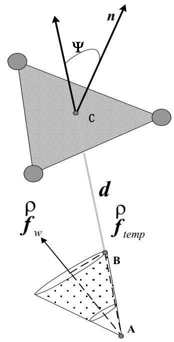


Fig. 9. The wind model

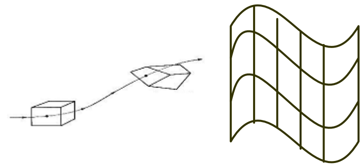


Fig. 10. The wind particle acted on the cloth

As Fig.11 shows, Force *G* at *X,Y,Z* direction are dG_x, dG_y, dG_z, and X_G、Y_G、Z_G are the unit forces on X,Y,Z direction:

$$dG_x=\rho dxdydzX_G,\ dG_y=\rho dxdydzY_G,\ dG_z=\rho dxdydzZ_G \tag{11}$$

The presses at face 1-2 and 3-4 are:

$$p_A=p-\frac{1}{2}\frac{\partial p}{\partial x}dx \tag{12}$$

$$p_B=p+\frac{1}{2}\frac{\partial p}{\partial x}dx$$

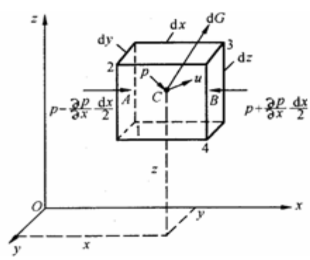


Fig. 11. The movement of stream particle

If the fluid particle is at its balance state, then**:**

$$dG_x+(p-\frac{1}{2}\frac{\partial p}{\partial x}dx)dydz-(p+\frac{1}{2}\frac{\partial p}{\partial x}dx)dydz=ma_x \tag{13}$$

$$\bar{a}=\frac{d\bar{v}}{dt},a_x=\frac{dv_x}{dt},a_y=\frac{dv_y}{dt},a_z=\frac{dv_z}{dt}$$

The movement equations of the fluid particle along X,Y,Z direction are as follows**:**

$$x_G-\frac{1}{\rho}\frac{\partial p}{\partial x}=\frac{dv_x}{dt} \tag{14}$$

$$y_G-\frac{1}{\rho}\frac{\partial p}{\partial y}=\frac{dv_y}{dt}$$

$$z_G - \frac{1}{\rho}\frac{\partial p}{\partial z} = \frac{dv_z}{dt}$$

According to the above analyses, the wind is treated as particles with its own power and direction. After colliding with the cloth surface, the wind moves according to Equation (14), and is damped by the Newton Law.

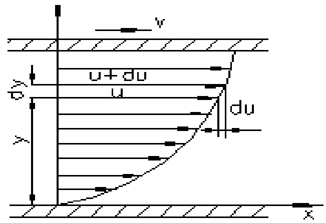


Fig. 12. Wind attenuation

Since the wind is produced by the flowing air, when the air flows along the cloth surface, the shearing force, called damping, is produced. Assuming the wind is Newton fluid, the damping f_d is direct proportion of the interface area A, and the relative velocity difference dV is inverse proportion of the perpendicular distance dy, shown in Fig. 12:

$$f_d = -\mu A \frac{dV}{dy} \tag{15}$$

where μ is a factor related to the fluid feature, which called friction parameter.

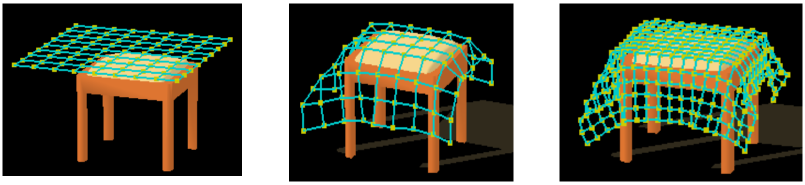
Unfortunately, the above friction Equation (15) is not able to capture dynamic aspects of cloth in our experiments. Therefore it is modified to a simple nonlinear alternative. The average velocity on the surface is decomposed into two components, one normal to the surface (V_n), and the other tangential (V_t). Total damping is the linear function of tangential velocity and the quadratic function of normal velocity as Equation (16).

$$f_d = -\mu A\left(\frac{k_n |V_n|^2}{1 + k_f |V_n|^2}\frac{V_n}{|V_n|} + k_t V_t\right) \tag{16}$$

where k_f controls the degree of nonlinearity. If the damping is too small, the dynamic energy can not be convergent to make the particles be balance. If the damping is too large, vibration occurs. Here, k_t comes from Stokes' Law, k_n matches the low Reynold's number flow. The V_n^2 term in the denominator makes the damping asymptotic as $V_n \to \infty$, which is partially motivated by the drag crisis.

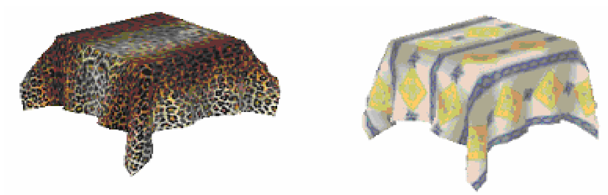
3 Our Implementation and Experimental Results

The multiresolution model, shown in Fig 13, is applied to improve the tradeoff of the realistic presentation and expensive computation[24]. To enhance the performance, the process of modeling, force analyses, collision and animation are played on the coarse model (shown in Fig.13a and 13b). Then the simulation goes to equilibrium (shown in Fig. 13b). Next, the first equilibrium is taken as input state. The finer model goes to balance (shown in Fig. 13c). The processing can progress as time permits.



(a) Initial state (b) Coarse grid balance (c) Fine grid equilibrium

Fig. 13. Multiresolution model



(a)Woolen (b) Nylon

(c)Different repulse objects

(d) Curtains

Fig. 14. The drapes of different fabric cloth

The algorithm steps are as follows:

Step 1: For each particle, initiate the force accumulator.
Step 2: For each particle, calculate the internal spring (structure, bending, shearing) and external force (wind) to force accumulator.
Step 3: Collision detection, which produces new force to force accumulator.
Step 4: Put damping to force accumulator.
Step 5: Detect the surface penetration under a certain time clip. If penetration, go to next step. If not, go to Step 7.

Step 6: If penetration, modify the time clip according to the Courant-Friedrich-Levy (CFL) constraint. Then go to Step 5.
Step 7: According to the force accumulator, modify the position, velocity and acceleration by Runge-Kutta integration.
Step 8: According to the multiresolution processing, take the balance state of coarse model as input state. The finer model goes to balance. The processing can progress as time permits.

Some experiments are performed to present the clothing simulation, yet demonstrate on the complex dynamics of cloth motion. Two types of fabrics, woolen and nylon with different repulse objects (shown in Fig. 14 and Fig. 15), are taken as examples. By modifying the parameters, it can exhibit a wide range of static and dynamic behavior.

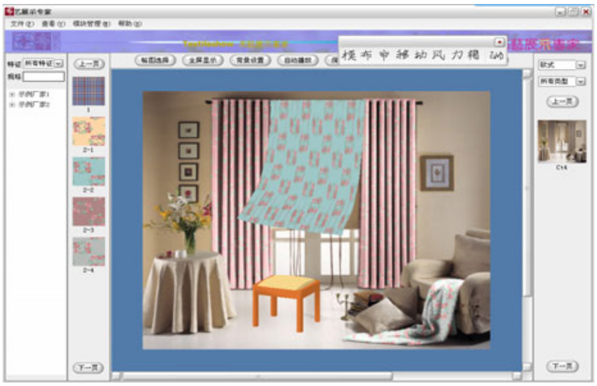


Fig. 15. The drapes of different fabric cloth

4 Conclusion

This paper presents an anisotropic modeling method by means of separating the weft and warp, which is the first step of simulation, and then taking the first state as input to get the next balance state by mass-spring method. Comparing with the traditional mass-spring system [7], the proposed approach has the following advantages:

(1) The mass-spring model focuses on the local appearance, which makes wrinkle from one particle, and extends the energy to other place through the spring structure. Under most condition, the local feature is more concerned than the global one. However, in the proposed approach, no matter the modeling structure or the forces and damping are considered on global surface.

(2) In traditional mass-spring system, one particle can be deformed greatly (more than 10%). It is unrealistic. Although modifying the spring parameters can avoid this phenomenon, the additional computation has to be paid to ensure convergence. We use separate modeling to avoid the great deformation. The computation cost can be controlled by multiresolution model.

The hydrodynamics theory, which assumes the air flow to be incompressible, irrotational, and inviscid, is applied to explain the relationship of wind model and the cloth movement in the proposed approach. A wind model, acted on the cloth surface through its stream field, is created in which the wind force is a point function on the field. The cloth deformation and movement are caused by the general field force, which can include not only the wind but also others (such as electromagnetic force) in the future.

References

1. Mori, Y., Igarashi, T.: Plushie: an interactive design system for plush toys. ACM Tran. on Graphics 26(3) (2007)
2. Decaudin, P., Julius, D., et al.: Virtual garments: A fully geometric approach for clothing design. Computer Graphics Forum (Eurographics 25(3), 625–634 (2006)
3. Terzopoulos, D., Platt, J., Barr, A., Fleischer, K.: Elastically deformable models. Computer Graphics(SIGGRAPH) 21, 205–214 (1987)
4. Thalmann, N.M., Yang, Y.: Techniques for cloth animation. In: New Trends Animation and Visualization, pp. 243–256 (1991)
5. Parks, D., Forsyth, D.: Improved integration for cloth simulation. In: Proc. of Eurographics, Computer Graphics Forum (2002)
6. Baraff, D., Witkin, A., Kass, M.: Untangling cloth. Proc. of ACM SIGGRAPH 22(3), 862–870 (2003)
7. Breen, E., House, H., Wozny, J.: Predicting the drape of woven cloth using interacting particles. Computer Graphics Proceedings 28(5), 365–372 (1994)
8. Chen, C., et al.: The physically based curtain animation in real time. The Journal of Software 11(9), 1228–1236 (2000)
9. Goldenthal, R., Harmon, D., Fattal, R., Bercovier, M., Grinspun, E.: Efficient simulation of inextensible cloth. In: Proceedings of SIGGRAPH 2007, vol. 26(3), p. 49 (2007)
10. Kaldor, J.M., James, D.L., Marschner, S.: Simulating knitted cloth at the yarn level. In: Proceedings of SIGGAPH 2008, pp. 1–9 (2008)
11. Ling, L.: Animation of stochastic motion of 3D cloth objects. Computers & Graphics 21(6), 769–776 (1997)
12. Volino, P., Magnenat-Thalmann, N.: Comparing Efficiency of Integration Methods for Cloth Simulation. Computer Graphics International 2001, 265–274 (2001)
13. Zhang, D., Matthew, M.F.: Cloth simulation using multilevel meshes. Computers & Graphics 25(3), 383–389 (2001)
14. Boxerman, E.: Speeding up cloth simulation. Master's thesis, University of British Columbia (2003)
15. Bridson, R., Marino, S., Fedkiw, R.: Simulation of clothing with folds and wrinkles. In: Eurographics/SIGGRAPH Symposium on Computer Animation 2003, pp. 28–36 (2003)
16. White, R., Crane, K., Forsyth, D.: Capturing and animating occluded cloth. ACM Trans. on Graphics, SIGGRAPH 26(3) (2007)
17. Scholz, V., Stich, T., et al.: Garment motion capture using color-coded pattern. In: Eurographics (2005)
18. Choi, K., Ko, H.: Stable but Responsive Cloth. In: Computer Graphics Proceedings, Annual Conference Series, ACM SIGGRAPH, San Antonio, Texas, pp. 604–611 (2002)
19. Oh, S., Ahn, J., Wohn, K.: Low damped cloth simulation. The Visual Computer 22(2), 70–79 (2006)

20. Etzmuss, O., Gross, J., Strasser, W.: Deriving a Particle System from Continuum Mechanics for the Animation of Deformable Objects. IEEE Transactions on Visualization and Computer Graphics 9(4), 538–550 (2003)
21. Govindaraju, K., Knott, D., Jain, N., et al.: Interactive collision detection between deformable models using chromatic decomposition. In: SIGGRAPH Proc. 2005, pp. 991–999 (2005)
22. Cutler, L.D., Gershbein, R., Wang, X.C., Curtis, C., Maigret, E., Prasso, L., Farson, P.: An artdirected wrinkle system for cg character clothing. In: Proceedings of ACM SIGGRAPH/Eurographic Symposium on Computer Animation, pp. 117–125 (2005)
23. Feynman, C.R.: Modeling the appearance of cloth, MSc Thesis, Department of Electrical Engineering and Computer Science. MIT, Cambridge (1986)
24. Zhang, M.: Fabric-based cloth simulation and realistic presentation. Doctoral Dissertation, Department of Computer Science, Zhejiang Univ. (2008)

A Virtual Reality Simulator Prototype for Learning and Assessing Phaco-sculpting Skills

Kup-Sze Choi

School of Nursing, The Hong Kong Polytechnic University, Hong Kong
kschoi@ieee.org

Abstract. This paper presents a virtual reality based simulator prototype for learning phacoemulsification in cataract surgery, with focus on the skills required for making a cross-shape trench in cataractous lens by an ultrasound probe during the phaco-sculpting procedure. An immersive virtual environment is created with 3D models of the lens and surgical tools. Haptic device is also used as 3D user interface. Phaco-sculpting is simulated by interactively deleting the constituting tetrahedrons of the lens model. Collisions between the virtual probe and the lens are effectively identified by partitioning the space containing the lens hierarchically with an octree. The simulator can be programmed to collect real-time quantitative user data for reviewing and assessing trainee's performance in an objective manner. A game-based learning environment can be created on top of the simulator by incorporating gaming elements based on the quantifiable performance metrics.

Keywords: Cataract Surgery; Phacoemulsification; Surgical Simulation; Virtual Reality; Medical Education.

1 Introduction

Surgery is an acquired skill requiring master-apprentice teaching, repeated practice, and broad exposure to medical cases. Learning the skills in computer-simulated virtual environments is a promising adjunct to conventional surgical education since the flexibility and reusability it offers can reinforce experiential learning by allowing trainees to practise over and over again at any time in various simulated scenarios. Also, making errors in virtual environments poses no risk on real patients. The performance can be recorded quantitatively for progress tracking. For these benefits, various virtual reality (VR) based simulators have been developed for surgical training, from minimally invasive surgery, endoscopic surgery to open surgery.

In eye surgery, dexterous and microscopic manipulations with high accuracy are required. Surgical errors may damage the delicate ocular tissues and lead to blindness. However, like other operations, the skills are acquired through real ophthalmic surgery and repeated practice is needed to gain experience. A risk-free and efficient methodology is thus in demand in ophthalmic surgery education. This paper presents

Z. Pan et al. (Eds.): Transactions on Edutainment IV, LNCS 6250, pp. 145–156, 2010.

a VR-based system for learning cataract surgery with the technique of phacoemulsification. The system is particularly developed for simulating phaco-sculpting, i.e. the removal of the cataractous lens by ultrasound phacoemulsification probe. An immersive virtual environment is created using 3D models of the lens and the surgical tools. Haptic device is used as the 3D user interface. Phaco-sculpting is simulated by interactive deletion of the constituting tetrahedrons of the volumetric lens model. Contacts between the virtual probe and the lens are effectively identified by using hierarchical octree. User performance can be recorded with quantifiable metrics.

The rest of the paper is organized as follows. A brief summary of virtual surgical simulators developed for eye surgery is first given in Section 2. The procedures for cataract surgery with phacoemulsification are then introduced in Section 3. The simulation techniques involved in the development of the proposed virtual phaco-sculpting simulator prototype are presented in Section 4. The quantitative measures that can be provided by the simulator prototype for performance tracking and assessment are discussed in Section 5. The work is concluded in Section 6.

2 Virtual Eye Surgery

Application of virtual reality to eye surgery dates back to early work on the simulation of retinal laser photocoagulation [1], lens removal by virtual phacoemulsification in cataract surgery [2] and cornea modification by simulated scalpel in radial keratotomy [3]. Attention was then drawn to the simulation of vitrectomy, the removal of the transparent gel (vitreous) filling the eyeball [4-7]. As a result of the research effort, commercial eye surgery simulators such as EyeSi [7] has become available in the market. On the other hand, the effectiveness of using virtual simulators for learning ophthalmic surgery and in the transfer of skills are also investigated, which are evaluated with quantitative measures of proficiency [8]. Promising results have been obtained from clinical trials conducted to evaluate and compare the skill level between experts and novices [9-11].

3 Phacoemulsification

Phacoemulsification, also known as "phaco", is a common surgery technique to remove cataractous lens from the eye. It is a preferred technique among others since tiny incisions and thus minimal discomfort are only involved, which enables fast recovery of vision [12]. In phacoemulsification, a small incision is first made at the edge of the cornea to allow for the insertion of surgical tools. This step is known as corneal incision. The membrane at the anterior part of the capsule containing the cataractous lens is then torn in a circular fashion to expose the lens, which is referred to as circular continuous capsulorrhexis, or CCC. Next, an ultrasound probe is inserted through the corneal incision to reach the cataractous lens. The vibrating tip of the probe (namely, phaco tip) is then manoeuvred carefully to break, i.e. emulsify, the clouded lens in to small fragments, which are simultaneously sucked away through a duct mounted at the phaco tip. The procedure of cataract removal by the ultrasound probe is referred to as phaco-sculpting. Phaco-sculpting can be performed with the "divide and conquer" approach where the cataractous lens is first sculpted to produce a cross-shape trench, which provides mechanical advantage to facilitate the cracking of the lens nucleus [13]. The lens is further emulsified to become four separated quadrants and removed one by one. An artificial intraocular lens (IOL) is then implanted in the cornea to replace the cataractous lens. This study focuses on the simulation of the making of cross-shape trench during the phaco-sculpting procedure.

4 Virtual Phaco-sculpting Simulation

4.1 Volumetric Lens Model

To simulate phaco-sculpting in virtual environments, 3D virtual models of the lens and the ultrasound probe are required, as shown in Fig. 1. The virtual probe is built directly with computer-aided design tools. The geometry of the lens is modelled with a tetrahedral volumetric mesh and is created by applying Delaunay tetrahedralization to the hollow triangulated surface model built for the lens. In the tetrahedral lens model, the centre of mass of each tetrahedron is inside the tetrahedron. As the cataractous lens is sculpted away bit by bit using the phacoemulsification probe, this process is simulated by interactively deleting the tetrahedrons in contact with the tip of the virtual probe, followed by updating the tetrahedral volumetric mesh accordingly. This

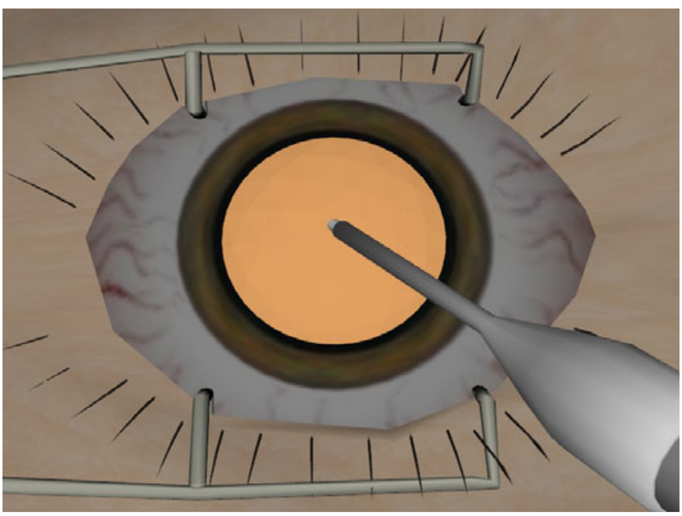


Fig. 1. The virtual environment developed for simulating phaco-sculpting

approach is suitable for simulating the bit-by-bit sculpting process and avoids tedious topology update usually encountered in interactive 3D volume cutting [14]. The downside is that a large number of tetrahedral elements are necessary to render high-quality visual effect of lens sculpting. In the current implementation, more than 10,000 tetrahedrons are required to model the lens. To ensure the simulation to run at interactive rates, a time-efficient algorithm is required to detect the collisions between the probe and the tetrahedrons. Brute-force search for the collided tetrahedrons is certainly prohibitive.

4.2 Data Structures

In the proposed simulator, collision detection is achieved by using bounding boxes arranged with a hierarchical octree to partition the 3D space containing the lens, as shown in Fig. 2. At initialization, the tetrahedrons inside the subspace bounded by a leaf box of the octree are collected with a list in the data structure of that leaf box. In the process of sculpting, the outer tetrahedrons are always removed first before the probe can reach the inner ones. The tetrahedrons inside each leaf box are therefore depth-sorted so that the ones closer to the viewpoint (i.e. the surface facing the user) can be visited first to reduce the time required for searching the collided tetrahedrons. Since a trench is created in the lens by hollowing it out progressively, the tetrahedrons to be removed are usually next to those tetrahedrons being removed. To this end, the data structure of tetrahedron registers its four neighbours, if any, so that when the phaco tip collides with a tetrahedron, the neighbours can be identified directly by assessing the data structure to detect subsequent collisions. The data structures used in the system is depicted in Fig. 3.

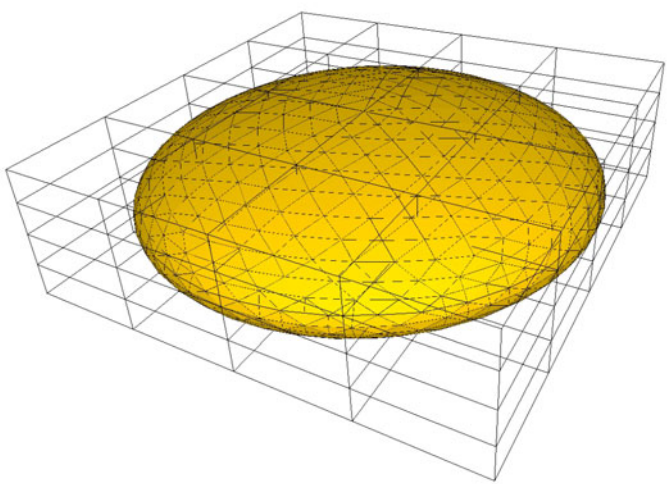


Fig. 2. Using bounding boxes to partition the tetrahedrons in the virtual lens model

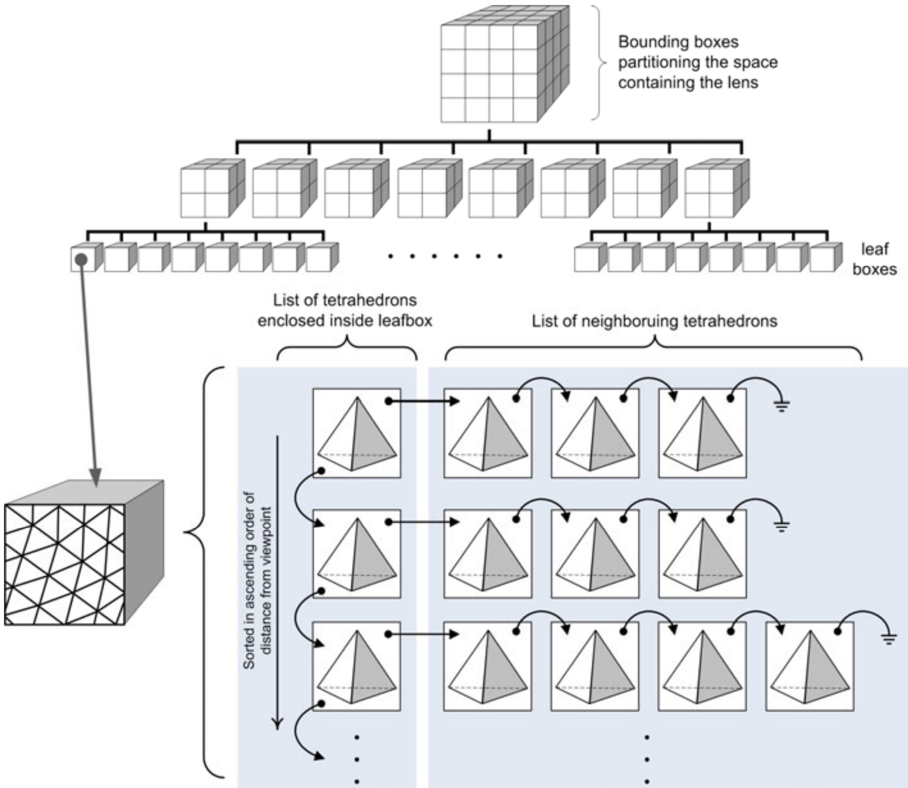


Fig. 3. Data structures used to manage the access of bounding boxes and tetrahedrons in the system during collision detection. The arrangement of bounding boxes by octree is illustrated with an octree of three levels.

4.3 Collision Detection

To simulate the sculpting of cataractous lens, it is necessary to identify the tetrahedrons in contact with the phaco tip and then remove them from the simulation process. This is achieved by first searching the octree downward from the root to the children boxes, by breadth-first traversal, for the leaf box intersected by the phaco tip. If probe-lens collision has occurred, the tetrahedron intersecting with the phaco tip will be identified by further searching the tetrahedrons enclosed within that leaf box. Details of the collision detection algorithm are described in Fig. 4.

In the simulator, the phaco tip is modelled as a bounding sphere and the test for collisions is carried out simply by evaluating the distance between the centre of tetrahedron and the tip. Collision is identified when the distance is less than the radius of the bounding sphere. This is depicted schematically in Fig. 5. The tetrahedron list associated to each leaf box is updated dynamically by deleting the collided tetrahedrons during the sculpting process. The neighbours stored in the data structure of the tetrahedron are also updated accordingly. The entire collision detection process repeats whenever the probe re-enters the space bounded by the octree.

Besides, by controlling the radius and depending on the resolution of the tetrahedral mesh, the bounding sphere at the phaco tip may enclose multiple tetrahedrons which are removed simultaneously to simulate sculpting. While precise edge-face intersections between the probe and the tetrahedrons can be computed to detect for collisions, it is more time-consuming and the level of accuracy provided is not necessary.

In the proposed simulator, only the tetrahedrons on the surface of the virtual lens are rendered in order to improve real-time performance. A list of surface tetrahedrons is maintained and updated interactively in response to the process of sculpting, where surface tetrahedrons are removed and inner tetrahedrons are exposed to become surface tetrahedrons.

1. Define $\boldsymbol{T}=\{t_i \mid i=1, \ldots, n\}$ as the set of n tetrahedrons comprising the virtual lens.
2. For each tetrahedron $t_i \in \boldsymbol{T}$, initialize the data structure of t_i with its centre position, vertex position and neighbouring tetrahedron.
3. Define $\boldsymbol{O}=\{b_j \mid j=1, \ldots, m\}$ as the octree that arranges m bounding boxes, from b_1 to b_m, in the space containing the lens.
4. Initialize $\boldsymbol{O}$ by partitioning the space containing the lens with different levels of b_j's.
5. For each leaf bounding box b_j in $\boldsymbol{O}$, initialize the data structure of b_j with a list of tetrahedrons enclosed inside b_j, denoted by $\boldsymbol{T}(b_j)$. The tetrahedrons in list $\boldsymbol{T}(b_j)$ is depth-sorted, where $\boldsymbol{T}(b_j) \subset \boldsymbol{T}$.
6. Perform breadth first search from the root of $\boldsymbol{O}$ downwards to identify, if any, the leaf bounding box b^* intersecting the phaco tip.
7. If b^* exists, search from $\boldsymbol{T}(b^*)$ to identify a tetrahedron t_0 enclosed inside the bounding sphere $\boldsymbol{S}$ centered at the phaco tip. Recursively search the neighbors of t_0 for all tetrahedrons inside $\boldsymbol{S}$, denoted by $\boldsymbol{T}(\boldsymbol{S})$.
8. Remove tetrahedrons $\boldsymbol{T}(\boldsymbol{S})$ from the rendering list.
9. Update $\boldsymbol{T}$ with $\boldsymbol{T}(\boldsymbol{S})$. Update $\boldsymbol{T}(b_j)$ in the data structure of b_j with $\boldsymbol{T}(\boldsymbol{S})$.

Fig. 4. Detection of collided tetrahedrons

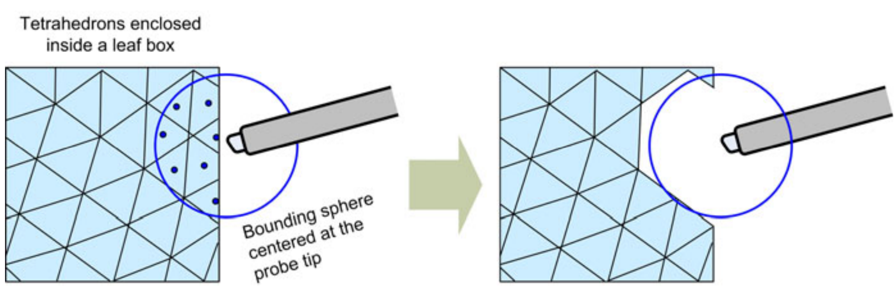


Fig. 5. In the collided leaf box, tetrahedrons with centre of mass falling inside the bounding sphere of the phaco tip are removed

4.4 Nucleus Rotation

During the process of phaco-sculpting, the cataractous lens is required to be appropriately rotated in order to facilitate the making of cross-shape trench. This manoeuvre can be modelled in the simulator by allowing the virtual lens to rotate about the optical axis. Here, it is assumed that there exists a constant damping torque D about the optical axis. To rotate the lens, the applied torque T has to be large enough to counteract D before the residual torque can induce an angular acceleration α. That is, $T - D = I\alpha$, where I is the moment of inertial of the lens. Furthermore, quasi-static equilibrium is also assumed such that the initial angular velocity of the lens at the beginning of each simulation time step h is small and negligible. Hence, the angular displacement $\theta = \frac{1}{2}\alpha h^2$. The modelling of nucleus rotation is depicted schematically in Fig. 6. The force F applied to the lens is modelled by using the distance between previous and current position of the lens-probe contact points, where F is considered to be linearly dependent of the distance. The moment of inertia I of the lens is obtained by considering the constituting tetrahedrons as mass points of equal weight. While the virtual lens is being sculpted, I is changed on the fly to reflect the reduction in inertia due to interactive removal of tetrahedrons.

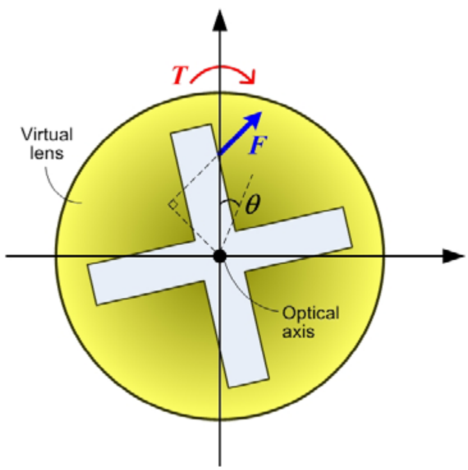


Fig. 6. Modelling of lens rotation

4.5 Real-Time Interactive Simulation

The proposed simulation system is implemented on a personal computer with an Intel Core 2 Duo 2.66 GHz CPU and 2GB RAM. A six degrees-of-freedom haptic device (PHANToM® Omni™, SensAble Technologies, Inc.) is employed as 3D user interfaces of the system. It is used to manoeuvre the virtual ultrasound probe and sculpt away the virtual cataractous lens in the 3D space. The real-time position and orientation of the probe are read continuously from the haptic device and fed into the simulation system. When collisions occur, the visual and haptic responses are computed on the fly to render the objects graphically on the screen and to drive the haptic devices. User is able to feel the interactions between the virtual probe and the lens through the stylus of the haptic device. To render the feedback forces, viscous drag force is generated to act against the movement the phaco tip when lens-probe collision occurs during phaco-sculpting. The drag force is proportional to the number of tetrahedrons enclosed inside the bounding sphere attached to the phaco tip, and the velocity of the phaco tip. In phaco-sculpting, the lens is emulsified and sculpted bit by bit to produce a cross-shape trench. Snapshots of the making of cross-shape trench on the virtual cataractous lens are shown Fig. 7. The simulation proceeds at interactive frame rates.

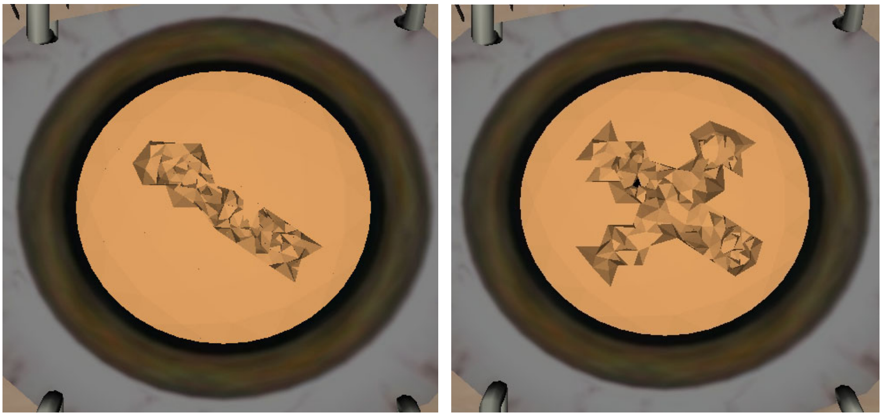


Fig. 7. Snapshots of the making of cross-shape trench on the virtual cataractous lens

In the simulator prototype, phaco-sculpting of clouded lens in cataract surgery is simulated by interactive deletion of the constituting tetrahedrons in the 3D lens model. Since a large number of tetrahedrons are required to ensure the visual quality, direct probe-tetrahedron collision detection could be too slow for interactive performance. The algorithm adopted in the prototype partitions the space containing the tetrahedral lens model using bounding boxes arranged with hierarchical octree. Instead of performing exhaustive search on all the tetrahedrons (say, N_{tetra}>10,000), collision detection is carried out systematically through the steps below:

(i) Search the leaf boxes (e.g. N_{bbox}=64 for an octree with 3 levels) for the one that encompasses the phaco tip,
(ii) Search the collided tetrahedrons inside that leaf box, and
(iii) Visit the neighbours of those collided tetrahedrons to detect further collisions.

On average, there are N_{tetra}/N_{bbox} tetrahedrons inside a leaf box, which is about two orders of magnitude less than N_{tetra} when N_{bbox}=64, for example. As a result, this greatly reduces the time required for collision detection. In particular, if the number of tetrahedrons inside the leaf boxes is moderate, the collided tetrahedrons can be identified directly and quickly with step (ii), making it unnecessary to go through step (iii) where additional overhead and implementation effort are needed for bookkeeping and updating the neighbours in the tetrahedron's data structure during the interactive sculpting process.

5 Quantitative Measures

The purpose of creating a trench in the shape of a cross in cataractous lens is to facilitate the cracking of the lens nucleus. The trenches should be wide enough to allow the access of both the phaco tip and another tool to the two sides of a trench. Forces can then be applied by using the tools to each side of a trench in opposite direction to crack the nucleus. Besides, the trenches should also be deep enough to make the separation of the sculpted nucleus possible and easier when forces are applied for cracking. Since the posterior lens surface is curved, to deepen the trenches, the phaco tip should go downwards, along the bottom and then upwards when a forward-moving stroke is being performed. Ideally, the trenches should be at right angles to each other so that they would crack towards the centre of the nucleus and the resulting quadrants will be about the same size [13]. To objectively evaluate trainees' performance in the process of virtual phaco-sculpting, quantifiable metrics can be used to measure their skill level. The simulator can be programmed to collect real-time user data, including completion time of creating the cross-shape trench, the amount of tissues removed from the lens, the width and depth of the trenches and the crossing angle, as well as the path of the phaco tip and the forces applied during sculpting. These quantitative data can be used to monitor learning progress and assess skills competency.

In the proposed simulator, the tetrahedrons removed during virtual training can be recorded, from which the amount of tissues removed from the cataractous lens can be

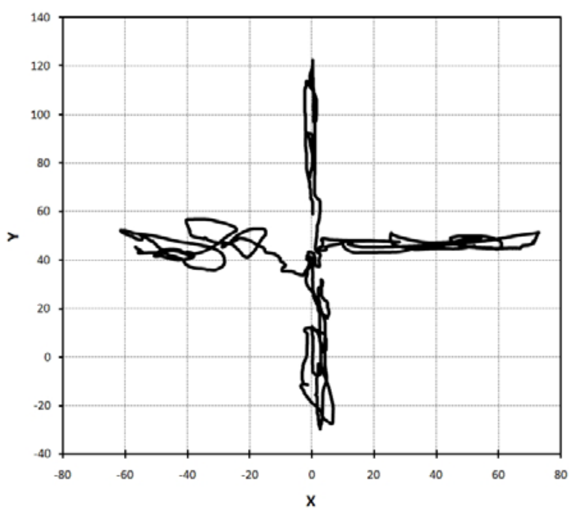


Fig. 8. The trajectory of the phaco tip in the process of making a cross-shape trench is recorded by logging the position of the haptic device's stylus endpoint

estimated by counting the number of tetrahedrons removed, and the volume of tissues emulsified for creating the cross-shape trench can then be estimated. The width and depth of the trenches and the crossing angle can also be estimated with the position of the removed tetrahedrons. On the other hand, the trajectory of the phaco tip during phaco-sculpting is readily available by reading the position data of the stylus endpoint of the haptic device, which can be visualized conveniently as shown in Fig. 8. Similarly, the force data can be obtained from the haptic device, from which the interactions can be plotted as a function of time for analysis after virtual training. An example of the force-time graph is given in Fig. 9. With these quantitative measures, skills competency can be evaluated objectively and automatically by using the virtual surgical simulator.

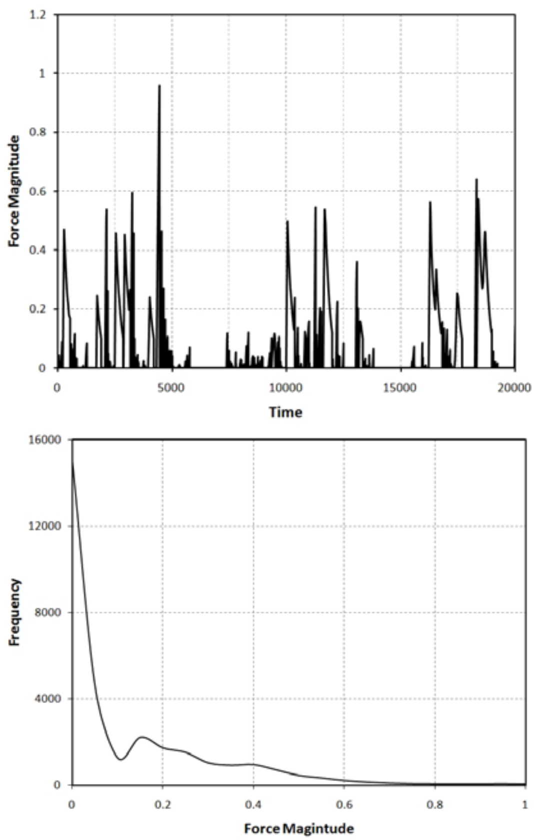


Fig. 9. With the virtual surgical simulator, the magnitude of forces over time (top) and their distribution (bottom) during the process of phaco-sculpting can be visualized and analyzed after the training

6 Conclusion

In this paper, the design and implementation of a virtual reality simulator for phacoemulsification training is reported. Currently, the virtual training simulator is being implemented on a passive stereoscopic LCD display with polarizing glasses to create the illusion of depth and enhance the level of immersion of the system.

One common error in real phaco-sculpting is that trainees happen to manoeuvre the phaco tip to sculpt the lens nucleus but the phaco probe is indeed moved in the air without touching the nucleus. The cataract is eventually not removed or just sculpted very slightly, wasting the ultrasound power [13]. The virtual phaco-sculpting simulator will be further developed to include the capability for detecting this error. Future work will also be performed, based on top of the virtual training simulator, to realize a serious game for learning phaco-sculpting in cataract surgery through gameplay. This can be achieved by incorporating gaming elements such as time-attack scenarios, bonus and game levels, based on the quantifiable performance metrics, to provide a game-based learning environment. This is beneficial in that the process of game playing is relevant to the pedagogical strategies of active learning and experiential learning. Game-based learning has also proven to be effective for knowledge transfer.

Acknowledgments. This work was supported in part by the Research Grants Council of the Hong Kong Special Administrative Region (No. PolyU 5145/05E and PolyU 5147/06E). The author is grateful to the support of Dr. S. Karpf and Dr. A.W. Siu to this project. He would also like to thank Dr. V. Woo who demonstrated the ocular procedures in real cataract surgery, which is very helpful to the research.

References

1. Meseure, P., Karpf, S., Chaillou, C., Dubois, P., et al.: Low-Cost Medical Simulation: A Retinal Laser Photocoagulation Simulator. In: Graphics Interface, pp. 155–162 (1995)
2. Sinclair, M., Peifer, J., Haleblian, R., Luxenberg, M., et al.: Computer-Simulated Eye Surgery. Ophthalmology 102(3), 517–521 (1995)
3. Hunter, I., Jones, L., Sagar, M., Lafontaine, S., et al.: Ophthalmic Microsurgical Robot and Associated Virtual Environment. Computers in Biology and Medicine 25(2), 173–182 (1995)
4. Hikichi, T., Yoshida, A., Igarashi, S., Mukai, N., et al.: Vitreous Surgery Simulator. Arch. Ophthalmology 118, 1679–1681 (2000)
5. Neumann, P., Salder, L., Gieser, J.: Virtual Reality Vitrectomy Simulator. In: Goos, G., Hartmanis, J., van Leeuwen, J. (eds.) MICCAI 1998. LNCS, vol. 1496, pp. 910–917. Springer, Heidelberg (1998)
6. Verma, D., Wills, D., Verma, M.: Virtual Reality Simulator for Vitreoretinal Surgery. Eye 17, 71–73 (2003)
7. Wagner, C., Schill, M., Männer, R.: Intraocular Surgery on a Virtual Eye. Communications of the ACM 45(7), 45–49 (2002)
8. Soderberg, P., Laurell, C.-G., Simawi, W., Skarman, E., et al.: Performance Index for Virtual Reality Phacoemulsification Surgery. In: Ophthalmic Technologies XVII, pp. 64261B-9. SPIE, San Jose (2007)

9. Feldman, B.H., Ake, J.M., Geist, C.E.: Virtual Reality Simulation. Ophthalmology 114(4), 828.e1–828.e4 (2007)
10. Soderberg, P., Laurell, C.-G., Simawi, W., Nordqvist, P., et al.: Virtual Reality Phacoemulsification: A Comparison between Skilled Surgeons and Students Naive to Cataract Surgery. In: Ophthalmic Technologies XV, pp. 164–169. SPIE, San Jose (2005)
11. Rossi, J.V., Verma, D., Fujii, G.Y., Lakhanpal, R.R., et al.: Virtual Vitreoretinal Surgical Simulator as a Training Tool. Retina 24(2), 231–236 (2004)
12. Benjamin, L.: Cataract Surgery. In: Benjamin, L. (ed.) Surgical Techniques in Ophthalmology. Elsevier Saunders, Edinburgh (2007)
13. Caesar, R.L.: Benjamin. Phacoemulsification: Step by Step. Butterworth-Heinemann, Edinburgh (2003)
14. Bielser, D., Glardon, P., Teschner, M., Gross, M.: A State Machine for Real-Time Cutting of Tetrahedral Meshes. Graph. Models. 66(6), 398–417 (2004)

An Augmented Reality Nanomanipulator for Learning Nanophysics: The "NanoLearner" Platform

Florence Marchi[1,2], Sylvain Marliere[3], Jean Loup Florens[3], Annie Luciani[3], and Joel Chevrier[1,2]

[1] Institut Néel, CNRS UPR 2940, 25 Avenue des Martyrs, 38042 Grenoble Cedex 9, France
[2] Université Joseph Fourier - BP 53 - 38041 Grenoble Cedex 9, France
[3] ICA-ACROE, 46 Av. Felix Viallet, 38031 Grenoble Cedex, France
florence.marchi@grenoble.cnrs.fr

Abstract. The work focuses on the description and evaluation of an augmented reality nanomanipulator, called "NanoLearner" platform used as educational tool in practical works of nanophysics. Through virtual reality associated to multisensory renderings, students are immersed in the nanoworld where they can interact in real time with a sample surface or an object, using their senses as hearing, seeing and touching. The role of each sensorial rendering in the understanding and control of the "approach-retract" interaction has been determined thanks to statistical studies obtained during the practical works. Finally, we present two extensions of the use of this innovative tool for investigating nano effects in living organisms and for allowing grand public to have access to a natural understanding of nanophenomena.

Keywords: Nanomanipulator, nanophysics education, virtual reality, haptic interaction, multisensorial interfaces.

1 Introduction

With the recent and very promising development of the nanotechnology industry, most of the scientific universities have opened master level formations on nano sciences that include mainly nanobiology, nanochemistery, nanophysics and nano-engineering courses. In this frame, innovative and specific practical works (PW) have to be created. In parallel, as this new technology should have a great impact on our life, it is crucial to educate the 'grand' public to nanoworld specificities. In this frame, the development of interactive and smart systems dealing with this inaccessible world and dedicated to 'grand' public should play a major role to convey the main differences between the macro and the nanoscale.

Nowadays, atomic force microscopy (AFM) is the most widespread techniques to measure force down to atom bonding and to characterize the topography of surfaces from micro to atomic scale, AFM is also the most powerful tool to manipulate nano-objects on a surface [1]. The AFM principle is based on the scan of a surface with a tip that is ideally terminated by a single atom at its apex. The tip is attached to a free end of a cantilever. When the tip apex interacts with the surface for tip-sample distances in the nanometers range, the local attractive or repulsive forces act on the tip

Z. Pan et al. (Eds.): Transactions on Edutainment IV, LNCS 6250, pp. 157–175, 2010.

inducing the cantilever bending up or down in order to keep the tip in equilibrium. By recording the bending variations of the cantilever during the scan, the topography surface can be reconstructed and displayed as a 3D image. It has to be pointed out that AFM image of surface landscape result of a serial process of pixel by pixel force measurement, it means that from local tactile sensations a visual representation is produced where the fully artificial colors code the height variation of the topography surface. Indeed, visual representation freezes interesting information, for this reason some groups have coupled a haptic system to an AFM [2] in order to match with all the richness of the measured data. Nevertheless this approach is not usual and one of the main limitations in using the conventional interface of the AFM microscopes in teaching and learning nanophenomena is its lack of interactive and real-time information. The only possibility for students to understand what happens in the nanoworld is only through formal equations. In addition, the collected data are only visually represented as topographic images or curves; their interpretation is not an intuitive task and requires time and solid theoretical background.

One can consider that it is a pity considering that the nanoworld is at physical scales just under the human sensorial accessible scale. It is situated at a frontier point between macroscopic physics and atomic physics where the macroscopic laws start to be non valid. In the nanoworld, adhesion and friction phenomena [3] are much more relevant than the gravitational phenomenon. When manipulating nano-objects, various forces, which can have different origins as capillarity, Van der Waals or electrostatic attraction, are playing an important role and have to be taken into account prior classical macroscopic collisions as usually processed in Virtual Environments. Indeed, the strategies and tools used to interact with a nano object cannot be simply transferred from the macroscopic scale.

Consequently the efficient learning of what happens at this frontier point may help students to understand in depth the scale transformations. New experimental methods of teaching and experimenting are needed to make understandable the nanoworld phenomena to a large number of students. Two main questions are raised to tackle new challenge: (1) How to adapt teaching of nanoscience and to reconcile human perception with mathematical and scientific views of reality; (2) How to extend or redesign usual virtual realities platform, including real time architecture and models of virtual objects, to serve the objective of learning nanophysics by using multisensory-haptic real time exploration. A similar situation has been observed in Astronomy teaching, the solar system can be considering as an inaccessible world where the distance and time constants are very different to human world. To overcome this situation, Y. Yair et al developed a 3D virtual environment where the learner could journey in the solar system [4].

In order to render tangible the principle of nanoscale interactions we have linked an Atomic Force Microscope (AFM) to a physically-based multisensorial haptic Virtual Reality platform, the whole computer-assisted teleoperation chain being called the NanoLearner Augmented Reality Platform (NanoLearner AR Platform). The NanoLearner configuration without the AFM connection is called the NanoLearner Virtual Reality Platform (NanoLearner VR platform). Using the NanoLearner VR platform, a student is immersed in a virtual nanoscene through a physically-based real-time simulation workstation equipped with a haptic device (i.e. a force feedback device) and generating visual and auditory representations of the nanoscene in accordance to the dynamics of the phenomena. He/she belongs to the dynamics of the chain in interaction since he/she is acting on scene and perceiving the nano-scale effects through the main sensorial channels: haptic, visual and auditory.

The paper is divided in several sections. Firstly, a review of the state of the art in nanosciences practical works implying interactive and multimodal interfaces as well as in nanophysics practical works using AFM; secondly a description of the NanoLearner Platform and its use and impacts in educational aims. The last part presents two extensions of the use of this innovative tool: (1) the NanoLearner AR Platform to investigate nano effects in living organisms and a specific NanoLearner VR system for allowing grand public to have access to a natural understanding of nanophenomena.

2 State of the Art in Nanoscience Interactive PW and in Nanophysics AFM PW

Some PW experiences have been carried out using VR and multimodal interfaces or using real Atomic Force Microscopes for nanosciences education.

2.1 Virtual Reality and Multisensory Feedbacks in Nanoscience

Some previous educational approaches have used virtual reality applications with haptic, visual, and/or auditory feedback in different ways to improve scientific learning in molecular biology [5] or in nano-chemistry [6]. For example, in the aim of teaching atomic bonding, a research group from Baltimore has coupled a haptic system to their own-developed molecular modeling software [7]. Their work has shown a significant improvement in the time required to build a simple molecule compared to other practical methods, the authors concluded that the use of a haptic device improves the understanding process of the bonding rules.

The University of North California [8] conducted a very promising educational project directed to middle and high school students. This project is based on the use of haptic system located in the classroom, coupled through the Web to an AFM located in a biophysics laboratory in order to interact with biological samples. The haptic device, a Phantom commercial device, is plugged to a computer where the AFM software is launched and it drives the AFM head through the web. Thanks to this set-up, middle and high school students touched different virus as adenoviruses and tobacco mosaic ones. The main result of this work is the great enthusiasm induced by this experience, which emerges from the students comments, published also in the school newspaper. These kinds of novel interactive PW had a significant impact on the interest of all students even the ones who were not attracted by science. This successful experiment opens the route of the use of the nanomanipulator system to scientific collaboratories in order to provide students access to specialized scientific instruments, data and experts [9,10]. For research use, some groups combined an Atomic Force Microscope (AFM) or Scanning Tunneling Microscope (STM) with haptic devices and virtual reality platforms to facilitate nanomanipulation [11,12]. On the same purpose, Li and colleagues [13] have improved the on-line observation of the changes in nano-environment by developing the real-time visual display combined with the real-time force feedback. All these very promising works have also a fruitful echo to the emerging nano-robotics field [14,15].

2.2 AFM Practical Works for Nanophysics

At the university level undergraduate and master students can follow scanning probe microscopy practical works linked to theoretical courses on nanosciences and nanotechnology. Indeed, Scanning Probe Microscopes (SPM) is an easy and well adapted gate to explore the material properties at the nanometer scale. For example, Murphy et al. [16] implemented a method of using atomic and magnetic force microscopy to led students in differentiating ranges and origin of forces involved in imaging. Ringlein et al. [17] have developed simulations to describe and illustrate the atomic origins of friction observed by AFM. This kind of PW will expand in a near future thanks to the development of cheap AFM instruments dedicated to teaching purpose[1]. Nevertheless any of these PW has addressed the deep understanding of these various nano-interactions that represent the key to explain the organization of the material at this scale and also to manipulate efficiently nano-object. The main limitation to develop this kind of practical works is due to the use of a real AFM to probe the nano-interaction and also to manipulate an object, which is time consuming and delicate (high risk of tip, object or surface damages). In addition, as the AFM interface has been developed for researchers, the feedback representations are restricted to image or graphic curves that required a non-obvious background in nanosciences to be interpreted.

At high school level recently, a teaching model of the AFM for pupils that combines Lego blocks, strip cut from used CD, a laser pointer and magnets, has been build and tested [18]. It is a cheap and easy way to increase the high school pupil awareness of the instrumentation dedicated to nanoscale. Nevertheless as only the magnetic force acts on the macroscopic AFM probe, it does not underline the complexity and variety of the interactions (Van der Waals, capillarity, electrostatic...) at the nanometer range. In fact this experience does not explore the notion of ‘contact’ at the nanoscale.

In order to improve the learning and teaching of the key physics phenomena of the nanoworld and the associated instrumentation (AFM) to graduate students but also to high school pupils and ‘grand public’, we have develop the “NanoLearner” platform.

3 Design of the “NanoLearner” Educational Platform

The developed NanoLearner Augmented Reality platform (Figure 1) brings together (1) an Atomic Force Microscope (AFM), (2) the Real-Time multisensory physically-based simulation workstation with a high quality haptic device, (3) multisensory representations based on physical modeling by which they are coherently correlated.

The real-time simulation workstation placed between the AFM and the haptic device insures a high quality synchronized data flow between the experimentalist and the nanoworld. It is in charge of a simulation of a physical based model of an

---

[1] NT-MDT manufacturer offers a nanoeducator SPM, Veeco manufacturer offers **Caliber Atomic Force Microscope.**

intermediate mean (tool) between the real nano scene and the operator ; this tool acts as an augmented environment where the operator can use his/her natural senses (haptic, sound and vision) to control the interactions at the nanoscale.

The simulation of this multisensory physically-based model runs in real–time and is synchronized to the inputs and outputs of the teleoperation main chain at the frequency of 3 kHz. This condition leads to a stable delay of maximum 1/3 ms between the inputs and outputs of on one side the haptic device and on the other side the AFM. The simulation is also in charge of producing visual and auditory feedbacks. Thus, haptic, visual and acoustical feedbacks constitute a complete real-time multisensory rendering of the physical phenomenon.

In order to offer a suitable environment matching any kind of nanomanipulation, the user should be able to manipulate the real and virtual nano-scene through a haptic device with a variable number of degrees of freedom (DoF). The haptic device used [19] is a modular interface. According to the specific manipulation, the experimentalist can use either a key of a piano-like keyboard (figure 1) that offers the same up-and-down movements as the AFM tip during an approach-retract interaction or a three-dimensional stick that allows 3DoF movements needed for a nano-object manipulation.

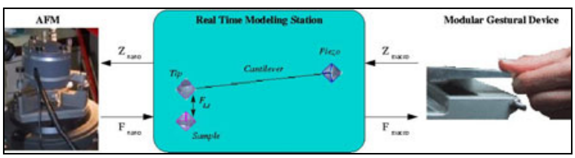


Fig. 1. The whole augmented NanoLearner Educational Platform

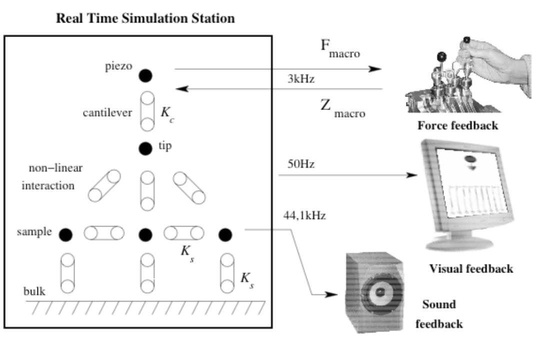


Fig. 2. Implementation of the virtual reality NanoLearner platform

The hand movements of the experimenter are the inputs of the physically-based model. This last computes the virtual scene displacements and deformations that are simultaneously used to produce the visual and auditory feedbacks. The forces produced by the real tip-sample interaction are returned through the real-time simulation workstation to the haptic device. Thus, the experimenter acts on the nanoworld, actively seeing, hearing and feeling the nanophenomenon.

The virtual part of the NanoLearner platform (Figure 2) consists in simulating by a virtual physically-based model, the whole real system composed of the AFM probe and the surface at the atomic scale. The piezoelectric tube, the tip, the surface atoms and the bulk of the sample are modeled by punctual masses that are permanently in interaction with the external environment, as well as among them (Figure 2).

The cantilever is modeled by a spring with zero rest length and with a stiffness K_c driving the tip on the vertical axis. The elasticity of the sample surface at the atomic level is modeled by springs (with a stiffness K_S) linking the atoms on the horizontal plane as well as in the vertical direction with the bulk. The tip-sample interaction (force, F_{LJ}) is modeled by a non-linear atom-atom interaction derived from the Lennard-Jones potential U_{LJ} (equation 1).

$$\left|\vec{F}_{LJ}\right| = \left|-\overrightarrow{grad}\cdot U_{LJ}\right| = \left| \frac{12B}{z^{13}} - \frac{6A}{z^{7}} \right| \tag{1}$$

Where A and B are the constants depending on the materials and the working environment, z is the tip-surface distance. The force variation with z is highly non linear.

In this virtual nanoscene, the modeling AFM probe is working in contact mode where the probe is in quasi-equilibrium at any time, this behavior is easily described by equation (2):

$$\vec{F}_{T/S}(z) + \vec{F}_{T/C}(z) = \vec{0} \tag{2}$$

Where $F_{T/S}$ is the tip-sample force and $F_{T/C}$ is the tip-cantilever force ($F_{T/C} = K_C * \Delta z$)

During all the interaction, the spring elongation Δz, corresponding to the cantilever deformation, is proportional to the tip-sample force. It should be noticed that as equation (2) is non linear, it leads to a hysteresis behavior according to the value of the cantilever spring constant (a low cantilever spring constant is required to measure weak interaction but it enhances the hysteresis behavior).

3.1 The Haptic Rendering

The haptic interface presents the morphology of three degrees of freedom (DoF) stick (fig.3). Since the virtual scene takes place in a 2D space, only 2 Dof are necessary to control the tip movements. We used the 3rd Dof to constrain the stick to remain in the 2D (quasi)plane of the manipulation, acting as a virtual guiding. The use of such virtual guiding presents two mains advantages. First, the level of parasitic effects that are usually generated in mechanical guiding such as friction or mechanism backlash can be significantly reduced. Secondly, the geometrical parameters of the 2Dof guiding like the curvature radius or the guiding surface shape are tunable. In the present case, we use a plane guiding.

The back-front axis Y of the stick point corresponds to the z axis of the tip in the nano-scene, while the X joystick axis corresponds to the lateral x axis in the 2D nano-scene and the motion along the users vertical axis Z are constrained to follow the XY horizontal directions in which the 2D virtual space of the represented nano-scene is mapped.

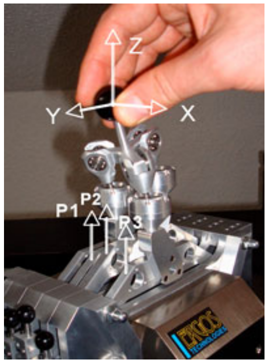


Fig. 3. The Ergos haptic device. The actuator/ sensor multi-axis base is equipped of a 3D Joystick. The 3(R-S) kinematic transformer provides the 3D natural motions (*XYZ*) of the manipulated point from the three parallel translational movements of the base axis (*p1,p2,p3*).

The haptic device belongs to a family of devices called the Ergos systems [11] which has been specially developed for high performance demanding applications. Two main parts compose the system (1) a special actuator-sensor base provides a three translational parallel motions; (2) a parallel isostatic cinematic structure transforms the basis movements into the three natural (*XYZ*) motion axes when the user manipulates the point of the device (black sphere on figure 3).

The data flows which are associated with this haptic interface configuration are shown in the diagram figure 4. The orientations of theses force and motion flows correspond to the basic impedance mode. It globally consists in sensing the motions and actuating the forces.

Following this chain, the haptic interface positions and forces flows are considered at 3 different levels:

- At the *sensing / actuating level*: the (*p1, p2, p3*) and (f_1, f_2, f_3) are the position sensor outputs and the force actuators inputs.

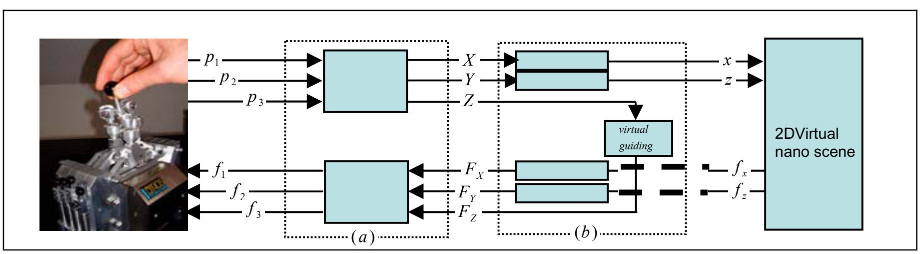


Fig. 4. The flow diagram related to the haptic interface. (a) Kinematics stage and (b) the real/virtual transformation stage that includes the virtual guiding.

- At the *virtual space level:* the (x,z) coordinates of the virtual "tool point" which is controlled by the haptic interface, and the related forces (f_x,f_y), consisting in the reaction forces of the virtual environment to this virtual tool.
- At the *User space level* : the manipulated point position and related feedback force at this point are represented by a (X, Y, Z) and (F_x, F_y, F_z) variables.

Then the global processing consists in two main stages.

- The kinematics stage performs a $(p1, p2, p3)$-> (XYZ) and a (F_x, F_y, F_z)-> (f_1,f_2,f_3) transformations. These transformations are determined from the haptic interface kinematics configuration knowledge. Since the displacement amplitude remains small and the device kinematics structure does not lead to hard non-linearity, the calculation consists in the linear approximation of the kinematics model. This linearization presents two main interests : (1) A reduced calculation costs, (2) In the case of a linear kinematics, the inverse transformation is no longer dependant on the position variables as it would be in the case of a non-linear transformation.

Then the kinematics stage transformations are defined by the following equations:

$$(X,Y,Z) = T(p_1,p_2,p_3) \quad (3)$$

$$(f_1,f_2,f_3) = {}^tT(F_X,F_Y,F_Z) \quad (4)$$

- Where T is the tangent matrix at the medium position point, to the device kinematics transformation.
- The real to virtual transformation stage performs the operator's space to virtual space positions and force forward and backward transformations. As indicated the Y and X user space motions are directly mapped on the 2D virtual space motions (the z and x axis) through simple scaling coefficients. The transduction stage includes the virtual guiding process. This last consists in simulating a uni-dimensional visco-elastic link along the Z axis which results in a maintaining the haptic image of the 2D virtual space in the horizontal $Z=0$ plane. This constrain is felt as a damped spring force whose the user can tune the hardness.
- These stage transformations are completely defined by the following three equations :

$$(x,z) = \alpha(X,Z) \quad (5)$$

$$(F_X,F_Y) = \beta(f_x,f_z) \quad (6)$$

$$F_Z = kZ + b\dot{Z} \quad (7)$$

where (α,β) are the scaling coefficients and (k,b) the stiffness and damping coefficients of the virtual guiding.

3.2 Visual Rendering

Figure 3 shows the visual feedback of a virtual nanoscene manipulation in 1D corresponding to a force curve experiment [20,21]. The black sphere represents the piezoelectric actuator allowing the displacement of the AFM probe in 1D, the user is moving it through the haptic device. The triangle represents the tip apex; the straight line between the two represents the cantilever. Its length varies according to the intensity and the nature (attractive or repulsive) of the tip-surface force acting on the tip apex (triangle); it is the equivalent of the cantilever deformation in the real nanoscene. The surface is represented at atomic level, each small sphere corresponds to an atom, and they are linking between them and to the bulk by atomic bound representing as straight lines. This basic representation matches with the well-known approach used by the chemists to teach molecules and crystal formation from atoms to undergraduate level. This mere graphical representation fits with the recommendation made by J. M. Ottino[22] about the scientific illustration : "A sensible first rule would be that pictures must not divorce from science and scientific plausibility". In addition, this non-continuous representation of the surface conveys the fact that at this nanoscale the concept of permanent spatial shape is not relevant because the shape is not independent of the probe used to characterize it. Indeed, a characterization performs with a scanning tunneling microscope (STM measures the tunneling current between the tip and the surface revealing the conductive properties of the sample) will not provide the same resulting shape as the one performed with an AFM. Thanks to this operational representation, the main difference in the concept of shape between macro and nanoscopic systems is restituted. The main contribution of the virtual reality to this operational visual representation relies on the dynamic aspect: the spatial deformation of the atom bonds according to their interactions with the environment is possible. It is a crucial point because it points out the un-static behavior of the atomic bonds.

This virtual nanoscene simulates working conditions in the elastic regime, which means that no plastic and irreversible deformations are taking into account, tip or sample cannot be destroy. These conditions are the classical ones when force spectroscopy is carrying on a conventional sample with the standard procedure. Nevertheless, the length variations of the cantilever spring are huge in the virtual nanoscene compared the reality. Indeed in the real scene, the vertical cantilever deformations are very weak compared to its horizontal length (a ratio of 10^4). In fact, the dynamical virtual nanoscene acts as a zoom on the relevant key parameters.

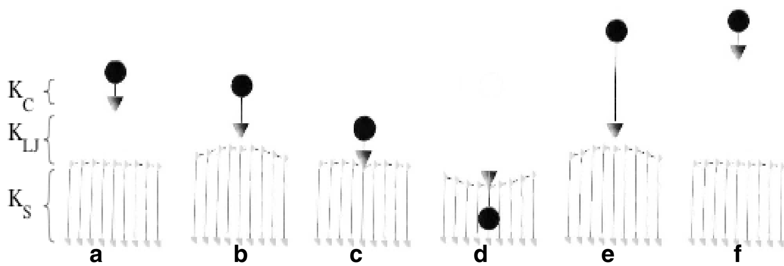


Fig. 5. Successive visual aspects of approach-retract process: a) the stiffness K_{LJ} is lower that the cantilever one K_C, b) jump to contact, K_{LJ} is higher than K_C, c) equilibrium position, the attractive force equal repulsive force, d) the tip pushes onto the surface, the repulsive force dominates the interaction leading to the spring compression e) retract phase where attractive force dominates, this step underlines the non-linear behavior of the approach-retract phenomena and f) snap off, the tip is not anymore in interaction with the surface

It should be underlined that in the case presented on figure 5, the surface stiffness is not very high compared to the cantilever stiffness, this explains why the surface is deformed at the jump to contact (fig 5.b) and just before the snap off (fig 5.e). To avoid the surface deformation K_s has to be higher than K_c and the force gradient of F_{LJ}, which corresponds to the classical case of probing a hard surface. Thanks to the virtual reality, most of the parameters can be easily and quickly changed offering to the experimentalist the possibility of exploring a wide range of different situations.

3.3 Sound Rendering

In order to stress a continuous sound to represent consistently the evolution of the physical states of the virtual nanoscene, particularly the current running point of Lennard-Jones interaction between the AFM probe and the nano surface, we exploit the intrinsic white noise of the haptic device. This acoustical white noise is modulated by the slope, i.e. the elasticity, of the Lennard-Jones interaction function defined by the equation (1), causing physically coherent sound modulations (Figure 6). Consequently, the sound modulations represent the evolution of physical state of the nanoscene along the approach-retract evolution. This sonification process is different than usual sonification processes in which the relation between acoustical parameters, such as the pitch of the sound for example, are controlled directly and one by one independently by input data [23], for example by the distance or by the force. In the physically-based method presented here, the physical system (here the tip interacting with the nanosurface) plays as a physical filter of the wide spectrum acoustical inputs. Consequently, the evolutions of all the parameters of the sound (pitch, timbre, amplitude) are not independent. It is similar to the vocal system in which the vocal cords signals are modulated by the vocal track and vocal cavity to produce vowels in speech or singing. Thanks to this correlation between the evolutions of the sound parameters, the evolution of the sound is physically plausible, enhancing the "naturalness" of the relation between the action of the experimenter, the physical states of the virtual nanoscene and the sound.

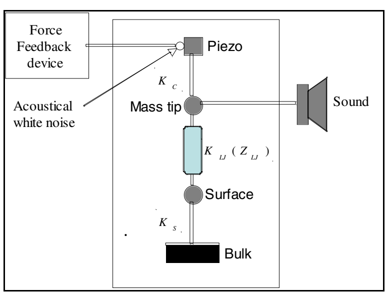


Fig. 6. The sound production process

We could think that the use of a white noise in the model seems not correspond, at a first glance, to a real situation. First of all, this noise exists naturally in haptic device. But more, this type of noise exists also in a real nanoscene. Indeed, in a real nanoscene, the Brownian motion (molecules movement due to the thermal energy E

with $E=k_BT$ where k_B is the Boltzmann constant and T the temperature) excites the AFM cantilever inducing very small oscillating amplitude around its resonance frequency f_0 ($f_0= [(1/2\pi).(k/m^*)^{1/2}]$, where m^* is the effective mass and k the cantilever spring constant). The Brownian motion behaves as a white noise, this external and uncontrollable excitation fixes the smallest force that is possible to detect with the AFM instrument [24].

Finally, thanks to the association between the functional visual feedback based on atomic representation, the real-time sound feedback modulated according to the variation of the interaction and the force feedback based on the cantilever deflection, the implemented model allows a complete metaphor of the approach-retract phenomenon at the macroscopic scale that has the property to be "physically consistent" with the studied nano-phenomenon. A student does not only see how the interaction is propagated but hears the evolution of the tip-surface interaction and feels the forces involved in, all these perceptions being consistently correlated.

4 Methodology for the Evaluation of the Educational Properties of the Nanolearner Platform

Students of nanotechnology master followed 3-hours practical work focused on the approach-retract phenomenon using the NanoLearner platform. All students had a theoretical background in physics but no experience in multisensory renderings and Virtual Reality. In practice, sixty students have been separated in thirty couples randomly and each couple followed this PW. The PW starts with device description and familiarization the VR NanoLearner Platform.

The educational aim of this PW is to teach the specificity of the interactions between objects at the nanoscale through the complex approach-retract effect that is the main one use in AFM to characterize the contact. In parallel, we have conducted during these PW, a study to determine the role of each sensorial renderings (haptic, visual and acoustic) and each combination of them in the identification of the key phases and points in the approach-retract phenomenon.

To perform this study on the role of each modality, a protocol has been applied. It consists of using the same virtual model (a hard surface, $K_S>>K_C$) where the student controls the vertical piezoelectric tube position z_{PS}, via the haptic device but with different feedbacks "ON" as summarized in table 1. During all this experiment, the students cannot change any parameter values. After each AR performance, students have to answer to the following set of questions by YES, NO or Undefined.

Table 1. Steps of the rendering evaluation

N°	1	2	3	4	5	6	7
Force	ON	*OFF*	*OFF*	ON	*OFF*	ON	ON
Sound	*OFF*	ON	*OFF*	ON	ON	*OFF*	ON
Vision	*OFF*	*OFF*	ON	*OFF*	ON	ON	ON

In order to evaluate the impact of the NanoLearner Platform as educational tool, the chosen questions are oriented to the physics of the phenomenon.

This method is different than others [25] in which the questions asked as more like that "do you think that haptic rendering helped you to understand the phenomenon ?", addressing more the feeling and the perception of the students, and evaluating more the user-friendliness of the platform rather than its efficiency in the understanding of the physical phenomena.

Four questions are asked, directly related to the physical phenomenon, i.e. the approach-retract phenomenon when observed with an AFM:

Q1a: Do you detect a fast variation in force intensity during the approach phase?

Q2a: Do you detect several kinds of forces during the approach phase?

These questions are also asked for the retract phase (Q1b & Q2b) and the right answer is YES.

Q3: Does the snap-on occur at the same position of the piezoelectric tube than the snap-off? The right answer is NO.

Q4: Is there any difference in the maximum of force intensity between the approach and the retract phases? The right answer is YES since the force intensity is higher during the retract phase than the approach one.

5 Results and Analysis

The student answers to the above set of questions (Q1 – Q4) are presented in figures 7, 8, 9 and 10. The statistical chart of the figure 7, which is related to the first question Q1a – Q1b on the way of variation of force intensity, reveals that:

- Around 50% of the students could answer correctly using only one rendering and there are small differences between haptic, visual and sound rendering.
- Some associations of two renderings can improve the percentage of correct answer up to 80%. Nevertheless, all the possible associations do not produce similar results. It seems that force rendering combined with sound rendering, and, in a second position, force and visual renderings combination offer more useful information to answer correctly. The sound and visual association does not improve the degree of correctness when compare to the use of only one rendering.
- The association of the three renderings provides the same result as the best renderings coupling (force with sound).
- The percentage of "UNDEFINED" answers is higher (up to 30%) when the force rendering is *not* used.

These observations reveal that the small variation in the force intensity during *the approach* phase (Figure 7 - Q1a) is detected more accurate thanks to the sound rendering (55%) rather than the force or visual rendering (below 50%). On the contrary, the force intensity variation during the *retract* phase (Figure 7 - Q1b) is well restituted both by force and visual renderings. This difference can be explained by the fact that, during the retract phase, the snap-off induces a stronger variation of the force intensity than in the approach case (around ten times stronger). This augmentation improves the efficiency of force and visual feedbacks in the phenomenon detection. Indeed, the statistical results confirm it, the force and visual renderings furnishing the higher percentages among the trials with only one sensorial feedback. From this observation, it can be interpreted that the modulation of the sound is the most suitable candidate to translate *small* force variations.

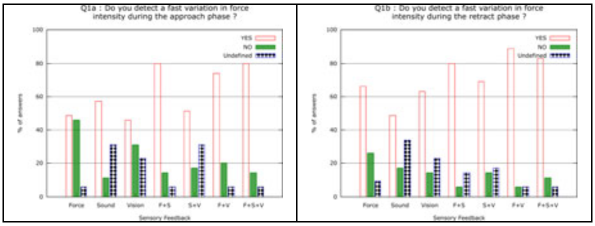


Fig. 7. Statistical charts related to question Q1

As the three associations of two renderings are not equivalent, it seems that force rendering supports complementary and specific information to the sound or the vision. Indeed the coupling of the force feedback with the sound or the visual one allowed a better detection of the force intensity variation illustrated by the higher percentage of correct answers. The high level of "UNDEFINED" answer for the sound-visual configuration confirms this interpretation.

The answers of the second question Q2 on the detection of several kinds of forces are summarized in the statistical chart of the figure 8, and reveal that:

• Thanks to the force feedback, more than 80% of the students answered correctly.

• The sound and visual renderings confer poor results, respectively 25% and 45% of correct answer.

• The association of two or three renderings when force feedback is included improves slightly the percentage of correct answer: 90% instead of 85% (force rendering alone).

The observation of this chart shows clearly that the determination of the *nature* of the force (attractive, repulsive or null) is mainly done thanks to the force rendering. This is true for the approach and the retract phases.

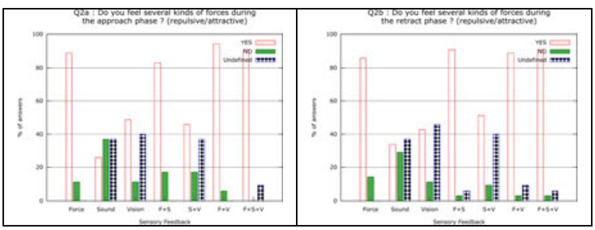


Fig. 8. Statistical charts related to question Q2

The chart of the figure 9-left, related to the question Q3 (the difference in the snap-on and snap-off position thresholds), reveals that:

• The use of one rendering provides a percentage of correct answer between 55% and 65% (visual on the first position, followed by sound and finally force feedback).

• With the association of two renderings the percentage of correct answers reaches 90%.

• The association of all the three renderings confers a result equivalent to the association of two.

From this statistical chart, it can be deduced that each rendering supports specific information about the spatial position. Indeed, any association of two renderings improves significantly the level of correct answer.

The answers to the last question on the difference in force intensities between approach and retract phases in the *attractive regime*, are grouped in the chart of Figure 9-right. It shows that:

• Among one sensorial renderings, the visual and force ones are more efficient than the sound.

• The association of force and visual renderings provides the best percentage of correct answer (85%). Among the possible combinations of two renderings, one specific combination (force and vision) improves significantly the correct answers.

This chart outlines the predominance of the force and visual renderings combination. It seems that force and visual renderings support complementary representations. In addition to the information furnished by the haptic device, the visual scene underlines the cantilever elongation (more accentuated in the retract phase, hardly visible in the approach phase), fact that matches the student's physical knowledge on the relation between the spring elongations with the force increasing.

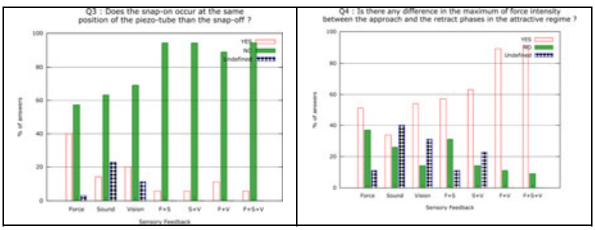


Fig. 9. Statistical charts related to question Q3 (Left) and to question Q4 (Right)

In general, these statistical data show a percentage higher than 80% when the three feedbacks are "on". However, the association of two well-chosen renderings could be enough to detect the approach-retract phenomenon. Each sensorial rendering supports specific and complementary information thus it is difficult to determine in advance the best two renderings association that makes understandable one type of interaction. A more efficient and more general method is the use of all three sensorial renderings, which allows the major part of the students to describe accurately this complex phenomenon, no matter the type of materials in interaction.

6 Extending the Use of NanoLearner Platform

This section deals with two extensions of the NanoLearner concept: one applied on the interaction with the insect world and the second one for the "grand public", in the frame of a European exhibition on Nanotechnologies, focusing on the notion of contact between nano and macroscale.

6.1 Interaction with a Drosophilae Leg

The interest of interactive multisensory augmented reality system can be also demonstrated in other experiments related to micro-manipulation of active structures, as biological organisms. During an educational project on MEMS (Micro Electro Mechanical System) [26], a group of four master students has to find a way to characterize Drosophilae leg reaction to an external stimulation. Drosophilae leg can be considered as a biological microsystem very close to MEMS, as its diameter and size is around 100µm and it can react to a mechanical or electrical stimulation. To succeed in this goal, two main challenges have to be solved: (1) to fabricate a specific AFM probe with a tip area of the same scale than the drosophilae foot (Figure 10) to apply a homogenous force on all the foot surface, (2) to adapt and control tip-foot interaction according to the leg reaction in real time in order to avoid partial or total destruction of the foot during the experience.

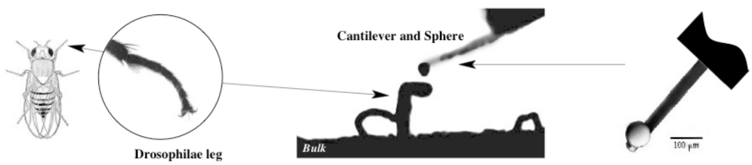


Fig. 10. Setup for interacting with *Drosophilae* foot: on the left, a picture of a drosophilae leg, on the center the leg glued on the substrate and the specific probe is above it, on the right a picture of the specific probe

The specific AFM tip has been fabricated from a classical AFM cantilever tipples - 300µm length, 35µm width, 1.5µm thickness- where a 50µm diameter latex sphere has been glued (Figure 14). The use of a modified AFM tip was required for two reasons: (1) the surface interaction had to take place with all the surface of the foot, (2) the cantilever stiffness has to be small (~0.01N/m versus 1N/m for classical stiffness of contact cantilever) to minimize the risk of exerting a strong force intensity which could lead to a total or partial destruction of the *Drosophilae* leg/foot.

The second challenge is to adapt in real time the intensity of the applied force on the foot in function of the leg behavior. The AFM is usually used to image and characterize samples, which cannot be deformed in vertical direction on such long distances as a leg. Indeed the fly leg can retract or stretch by itself on long distance - from few hundred of nanometers to few micrometers - according to the intensity of the pressure exerted on the foot.

Thanks to the nanomanipulator, the user controls through the force feedback system, the tip-foot interaction by adapting in real-time his (her) hand impedance as illustrated in Figure 11. During this experience, the student has been able to describe his (her) feeling in the same time he (she) was interacting with the leg. The students produced a movie in which they explain what they are feeling when they were interacting with the fly leg in association with the physical phenomenon.

This experience underlines the fact that the force feedback device linked to the AFM probe constitutes an adapted interface for the human hand to reach the submicrometer scale. It opens the door to the real time interaction with living insect through a micro/nano-hand, where humans could use their feelings, knowledge and own previous experiences to discover and act on this extraordinary under-microscopic living world.

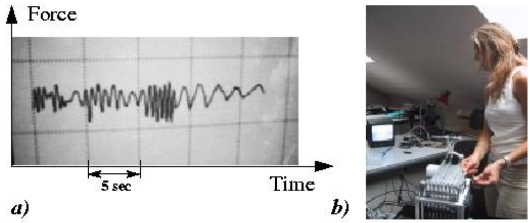


Fig. 11. (a)Temporal variation of tip-foot force for different speed manipulations - (b) Human interaction with the Drosophilae leg using the ERGOS nanomanipulator

6.2 A Dedicated NanoLearner Platform for the "Grand Public"

An integrated workstation has been implemented for large grand public scientific exhibitions at Cité de la Science in France and CERN in Switzerland (fig.12). This NanoLearner explores the notion of contact at the human and nanoscale. The most obvious and simple way to experience interaction with a flat rigid object using an instrument is to touch its surface through a stick. In this simulation, the rigid object, a piece of wood, is represented by a continuous and uniform rectangular shape and the stick is representing as a drumstick. In this context, no need to explain in detail to the user what is this situation: the user holds the stick in his (her) hand and uses it to touch the table. The user identifies straightforward that the force feedback device is the system enabling to move the virtual stick and to feel the virtual contact. In this reference situation, the user experiences a daily well-known situation of contact with a macroscopic object by seeing, feeling and hearing it. Conversely, in the virtual nanoscene, the surface is represented at atomic level with a series of atoms to underline the discontinuous aspect of the matter at the nanoscale; the stick is represented as a cantilever with a tip at its free end like for a real AFM probe, this choice makes the user sensitive to the real instrument employed by the surface scientist. In this virtual nanoscene the user experiences the complex notion of contact at the nanoscale via the

Fig. 12. "Put you hand in the nanoworld" grand public exhibition: how contacts are different at the macroscopic and nanoscopic scales

multi-sensorial interface. In fact, all the game for the user is to come closed as possible to the surface but without getting stuck on it; in that way, the user experience the specific long-range contact acting at nanoscale and can realize that at this scale the "walls" are sticking and that it possible to feel an object without enter in mechanical contact with it. By exploring the two virtual scenes, the user can compare the two situations and understand the specificity of the contact at the nanoscale.

At the present time, more than 10 000 people have had access to this natural understanding of physical nano-phenomena by their senses through direct actions and sensory feedbacks [27].

7 Conclusion

The NanoLearner VR Platform has been successfully used as an educational nanotool to teach the specific notion of contact in the nanoworld to university master students but also to the grand public. As the experimenter handles directly the nanoscene, a change in the physical interaction is transferred in a variation perceived by human sensorial channels.

At master level the NanoLearner VR Platform was used in PW to teach the complex approach-retract phenomenon. During these practical works, a study was carried on sixty students, in order to determine the role of each sensorial rendering in the identification of the key phases and points in the approach-retract phenomenon. This study was based on four questions related to the physical phenomenon and reveals several main results. Firstly, the *force feedback* plays a major role in the detection of the *intensity* and *nature* of tip-surface interaction. Secondly it emphasizes the surprising predominant role of the sound to detect *small* force variations. Thirdly, the visual rendering plays a major role on the identification of the piezoelectric actuator position. Finally, the combination of two well-chosen renderings enhances significantly the identification of the approach-retract phenomenon key phases. This optimal renderings combination is not the same for the different key phases and points. All these results demonstrate the potential of such platform to reconcile human perception with non-intuitive physical phenomenon. In the future, we will develop virtual nanoscenes where nano-objects on a surface could be moved, grasp and transport via AFM probes. In this way, the students could realize that the manipulation strategies at the nanoscale obey on effects as adhesion and friction and not on the gravity.

A dedicated version of the NanoLearner platform for the grand public has been developed where two virtual models were implemented in order to compare the notion of contact between the macroscopic and the nanoworld. In the virtual nanoscene the AFM probe was designed closed to a real one to grow public awareness to the surface instrumentation at the nanoscale. These experiments in grand public exhibitions, with very young people encourage us to continue to explore such physical based multi-sensory metaphors in order to offer a direct sensory access to physics of all citizens.

The augmented reality NanoLearner platform was a tool of choice to interact in real time with the world of insect via a specific AFM probe where a micro-sphere replaced the tip. In this experiment the students have shake the "hand" of a Drosophilae and probed qualitatively the elasticity of its leg. This first attempt opens the way to very promising and exciting experiments on reactive samples.

Acknowledgements

This work has been supported by the French National Center of Research (CNRS), the French Ministry of Culture and by the FP6 Network of Excellence IST-2002-002114 - Enactive Interfaces.

References

1. Bhusban, B.: Handbook of Nanotechnology. Springer, Heidelberg (2004)
2. Tourney, C.: Cubism at Nanoscale. Nature Nanotechnology 2, 587–589 (2007)
3. Israelachvili, J.: Molecular and surface forces, 2nd edn. Academic Press, NewYork (1991)
4. Yair, Y., Mintz, R., Litvak, S.: 3D virtual reality in science education: an implication for Astronomy Teaching. Journal of Computers in Mathematics and Science Teaching 20, 293–305 (2001)
5. Sankaranarayanan, G., Weghorst, S., Sanner, M., Gillet, A., Olson, A.: Role of haptics in teaching structural molecular biology. In: Proc. of the 11th Symposium on Haptics Interfaces for Virtual Environments and Teleoperator Systems, Los Angeles, CA, pp. 363–366 (2003)
6. Trindade, J., Fiolhais, C., Gil, V., Teixera, J.C.: Virtual Environment of water molecules for learning and teaching science. In: Proc. of the Eurographics Workshop, GVE 1999-Computer Graphics and Visualization Education 1999, Portugal, pp. 153–157 (1999)
7. Sauer, C., Hastings, W.A., Okamura, A.M.: Virtual Environment for Exploring Atomic Bonding. In: Proc. of EuroHaptics 2004, Munich, pp. 232–239 (2004)
8. Jones, M.G., Bokinsky, A., Andre, T., et al.: NanoManipulator applications in education: The impact of haptic experiences on students' attitudes and concepts. In: Proceedings of the IEEE Computer Science Haptics 2002 Symposium, pp. 295–298. IEEE Computer Society, Orlando (2002)
9. Sonnenwald, D.H., Li, B.: Scientific collaboratories in higher education: exploring learning style preferences and perceptions of technology. British Journal of Education Technology 34, 419–431 (2003)
10. Jeschke, S., Thomsen, C., Sinha, U.: Collaborative working environment for virtual and remote experiments in nanoscience and nanotechnologies. In: Proceedings of World Conference on E-Learning in Corporate, Government, Healthcare and Higher Education, pp. 2055–2060 (2006)
11. Taylor, R.M., Robinett, V., Chi, L., et al.: The Nanomanipulator: A Virtual-Reality Interface for a Scanning Tunneling Microscope. In: Proc. SIGGRAPH 1993, Computer Graphics, Anaheim, CA, pp. 127–134 (1993)
12. Vogl, W., Sitti, M., Ehrenstrasser, M., Zäh, M.: Augmented Reality User Interface for Nanomanipulation using Atomic Force Microscopes. In: Proc. of Eurohaptics 2004, Munich, Germany, pp. 413–416 (2004)
13. Li, G., Xi, N., Yu, M., Fung, W.K.: Development of augmented reality system for AFM-based nanomanipulation. IEEE/ASME Trans. on Mechatron. 9, 358–365 (2004)
14. Ferreira, A., Mavroidis, C.: Virtual reality and Haptics for Nanorobotics. IEEE Robotics&Automation Magazine, 78–92 (2006)
15. Dong, L., Nelson, J.B.: Robotics in the Small. IEEE Robotics&Automation Magazine, 111–121 (2007)
16. Murphy, W.L., Spalding, G.C.: Range of interactions, an experiment in atomic and magnetic force microscopy. Am. J. Phys. 67(10), 905–908 (1999)

17. Ringlein, J., Robbins, M.O.: Understanding and illustrating the atomic origins of friction. Am. J. Phys. 72(7), 884–891 (2004)
18. Planinsic, G., Kovac: When nano goes to school. J. Physics Education 43, 37–45 (2008)
19. Florens, J.L., Luciani, A., Cadoz, C., Castagné, N.: ERGOS: Multi-degrees of Freedom and Versatile Force-Feedback Panoply. In: Proc. of EuroHaptics 2004, Munich, Germany, June 5-7, pp. 356–360 (2004)
20. Capella, B., Dietler, G.: Force-distance curves by atomic force microscopy. Surface Science Reports 34, 1–104 (1999)
21. Marliere, S.: Multisensorial interaction with a nanoscale phenomenon: the force curve. In: Proc. of Euro-Haptics 2004, pp. 246–252 (2004)
22. Ottino, J.M.: Is picture worth 1,000 words? Nature 421, 474–476 (2003)
23. Wang, D., Zhang, Y., Sun, Z.: Multi-modal Virtual Reality Dental Training System with Integrated Haptic-Visual-Audio Display. In: Robotic Welding, Intelligence and Automation. Lecture Notes in Control and Information Sciences, vol. 362, pp. 453–462. Springer, Heidelberg (2007)
24. Stark, R.W., Drobek, T., Heckl, W.M.: Thermomechanical noise of a free v-shaped cantile-ver for atomic-force microscopy. Ultramicroscopy 86, 207–215 (2001)
25. Williams, R.L., He, X., Franklin, T., Wang, S.: Haptics-Augmented Undergraduate Engineering Education. In: International Conference on Engineering Education (2004)
26. Maboulian, R.: Surface processes in MEMS technology. Surface Science Reports 60, 209–268 (1998)
27. Marlière, S., Florens, J.L., Marchi, F., et al.: Implementation of perception and action at nanoscale. In: Proceedings of the 4th International Conference on Enactive Interfaces, pp. 181–184 (2007)

Fast Prototyping of Virtual Reality Based Surgical Simulators with PhysX-enabled GPU

Wai-Man Pang[1], Jing Qin[2], Yim-Pan Chui[2], and Pheng-Ann Heng[2]

[1] Spatial Media Group, Computer Arts Lab, University of Aizu, Japan
[2] Dept. of Computer Science and Engineering, The Chinese University of Hong Kong
wmpang@u-aizu.ac.jp, {jqin,ypchui,pheng}@cse.cuhk.edu.hk

Abstract. We present our experience in fast prototyping of a series of important but computation-intensive functionalities in surgical simulators based on newly released PhysX-enabled GPU. We focus on soft tissue deformation and bleeding simulation, as they are essential but have previously been difficult to be rapidly prototyped. A multilayered soft tissue deformation model is implemented by extending the hardware-accelerated mass-spring system (MSS) in PhysX engine. To ensure accuracy, we configure spring parameters in an analytic way and integrate a fast volume preservation method to overcome the volume loss problem in MSS. Fast bleeding simulation with consideration of both patient behavior and mechanical dynamics is introduced. By making use of the PhysX built-in SPH-based fluid solver with careful assignment of parameters, realistic yet efficient bleeding effects can be achieved. Experimental results demonstrate that our approaches can achieve both interactive frame rates and convincing visual effects even when complex models are involved.

1 Introduction

Recent years have witnessed significant progress in virtual reality-based surgical simulation. Many simulators have been developed to enable doctors and students to learn human anatomy and conduct surgical planning, training, and rehearsal in a virtual environment [1]. The repeatability and reusability offered in these simulators can reinforce experiential learning by allowing trainees to practise repeatedly at any time in various scenarios or even rare cases, and inaccurate operations have no risk for patients. The skill level and performance can also be recorded and assessed quantitatively.

However, development of an interactive and robust simulator is still a difficult task. One of the main challenges is the involvement of many labour-and-computation-intensive but indispensable modules, such as simulation of soft tissue deformation, simulation of bleeding, simulation of physiological phenomena, force feedback and real-time visualization for large medical dataset. Furthermore, just like other engineering design domains, it is usually necessary to do many trials before finding correct models for special surgery. In this regard,

Z. Pan et al. (Eds.): Transactions on Edutainment IV, LNCS 6250, pp. 176–188, 2010.

prototyping tools and methods are required to support the rapid development of new simulators.

Among these challenges, realistic soft tissue deformation and bleeding simulation are commonly involved in most surgical simulators, especially in simulators for open surgery. Fast algorithms for these two modules have been researched these years. With the rapid improvement in the graphics hardware, researchers attempt to exploit the processing power of Graphics Processing Units (GPUs) for these physically-based simulations. However, designing GPU-friendly data structures usually requires a lot of labour.

In newly released GPUs, a hardware accelerated physics engine, called PhysX Engine, is integrated to support physically-based computation. Combined with the parallel processing capability of the GPU, the PhysX Engine provides an API to support fast physics processing and calculation. The idea of using the PhysX Engine to accelerate development of surgical simulators is motivated by the recent concept of "serious game," aimed at utilizing game technologies to provide cost-effective solutions for non-entertainment applications [2,3]. The PhysX Engine was originally designed to provide computing power for dynamic motion and interaction in video games and has found a lot of successful applications. It is naturally thought that PhysX Engine can be applied to the development of surgical simulators, as the physics processes involved in simulators share the same principles as the processes in games [4].

However, directly employing the functions provided in the PhysX Engine is far from fulfilling the requirements of a surgical simulator. As opposed to games, where eye-catching visual effects is enough for most cases, surgical simulators also require high fidelity of tissue behaviors corresponding to a specific operation. For examples, the parameters of the deformation models should be configured carefully, and the features of physiological phenomena should also be fully considered. These requirements need to be seamlessly fulfilled with the built-in functions of PhysX engine to improve the accuracy of the simulation.

We present our experience in implementing fast prototyping methods for surgical simulators based on newly released PhysX-enabled GPU. Both accuracy and interactivity are considered. Preliminary results indicate that it is a novel technique with high potential for rapid development of similar kinds of healthcare related simulators. The rest of this paper is organized as follows. Section 2 reviews previous work on rapid prototyping tools and packages for surgical simulation and hardware-accelerated modeling and simulation approaches. Section 3 gives the implementation details and results of a multilayered soft tissue deformation model based on PhysX engine. Section 4 presents fast bleeding simulation with consideration of both patient physiology and mechanical dynamics based on PhysX engine. Section 5 provides a short discussion and draws conclusions.

2 Related Work

Some projects seamlessly integrated many heterogeneous models or processes for rapid development of surgical simulators. K. Montgomery et al. proposed a real-time surgical simulation system development framework with soft-tissue modeling and multi-user, multi-instrument, networked haptics [5]. In addition, the framework is cross-platform and can run on both Unix and Windows platforms. M. C. Cavusoglu et al. implemented a framework, "Gipsi", for open source/open architecture for organ-level surgical simulation [6]. J. Allard et al. presented an open source framework, "SOFA", for medical simulation research [7], aiming at enabling component sharing/exchange, reducing development time and validating new algorithms. However, these toolkits for rapid development mainly focused on the integration of existing algorithms and paid little attention regarding how to accelerate these algorithms by newly released graphics hardware.

Hardware acceleration for rendering has long been used in many graphics applications including surgical simulation. Similar to rendering, physically-based deformation is another essential and computationally intensive component in surgical simulation. Many researchers [8,9,10] have tried to exploit the power of programmable GPU for physics computation, but they usually suffer from certain difficulties or limitations. For example the model's topology has to be carefully designed, and considerable amount of computation still needs to be done on CPU ahead. In addition, various data structures are usually required to be designed for the purpose of parallelization. Recently, a specialized hardware accelerator for physics called Physical Processing Unit (PPU) is released for dynamic motion, interaction, and physics computation in games [11]. Pang et al. [12] was the first to employ PPU to accelerate simulations in an orthopedics surgical trainer. However, some limitations such as the bottleneck of a standard PCI bus prohibit its application to more complex models. Half a year ago when this paper was written, the SDK for physical computation previously employed in PPU was ported to new GPUs and named as PhysX engine. This provides a novel platform for constructing surgical simulation systems. Recently, Maciel et al. have proposed solutions for constructing multimodal surgical simulation environments based on PhysX Engine [13]. However, their work focuses on how to integrate haptic devices with the PhysX engine to achieve high update rate of both haptic and visual rendering, while our work emphasizes providing solutions for real-time and realistic soft tissue deformation and bleeding simulation.

3 Multilayered Soft Tissue Deformation

Modeling biomechanical behavior in virtual surgical simulators is indispensable, as realistic deformation is important for training of surgeons. However, it is a challenging task due to the complicated mechanical interaction among tissues in an organ. In this section, we present a fast prototyping approach for multilayered soft tissue deformation based on the PhysX engine.

The simulation of biomechanical behavior of single specific tissue has been intensively investigated. However, to the best of our knowledge, there are few reports related to simulation of comprehensive behavior concerning multiple tissues. Therefore, we propose a multilayered mass-spring deformable model, in which the layers of the model are consistent with the anatomical tissue structure in real human organs. In this work, we focus on open orthopedic surgery where heterogeneous tissues such as skin, fat, muscle, and bone are involved. The layered model can effectively simulate the interaction between tissues and facilitate assignment of different properties across different tissue layers.

3.1 Geometry Modeling

Manual editing of a multilayered model can be done with help of knowledge of thicknesses of each layer in an organ from anatomy references, although such data may not be accurate enough. In order to generate a multilayered model with anatomical details, one of the possibilities is to adopt a real dataset, e.g. the Chinese Visible Human (CVH) dataset. The CVH dataset is segmented to the tissue level. Tissues such as skin, adipose tissues, muscle, and bone are labeled. We extract the interfaces between all tissue layers and represent the interface as polygonal surfaces. The extracted surface meshes are stored as a list of triangular facets in Stereolithography (STL) format.

Having obtained all the surface meshes of interfaces between tissues, they are ready for use in the construction of multilayered mass-spring model. The edges of the mesh are converted to springs in the multilayered model, and correspondingly the apexes turn into masses. Each tissue interface is first reparameterized into a regularized quad-based mesh like the one shown in Fig. 1(a). This is important as "unstructured" facets may introduce difficulties regarding generating computationally efficient deformable models. Then, the construction process is performed from upper to lower layers, and tetrahedral structures are formed between layers, as shown in Figs. 1(b) & (c).

Fig. 2 shows the results of a complete construction process for a virtual upper limb extracted from the CVH. In this example, all layers are first reparameterized to a quad-based mesh based on user specified resolution, e.g. 39×10 in our case (see Fig. 2(d)). Then, the construction of tetrahedral structures which connect adjacent layers is performed. However, in case of a upper limb model, the masses in lower layers are crowded together because of the reduced radius of the inner tissue interfaces. Therefore, we try to heuristically reduce the density

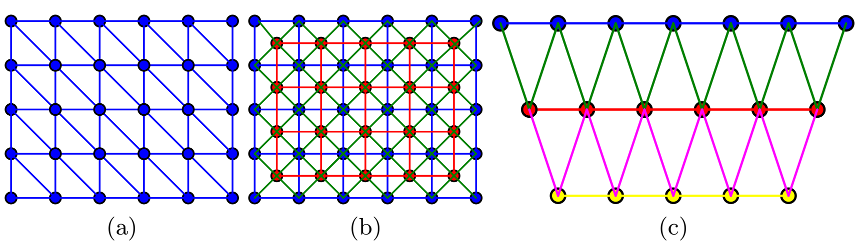
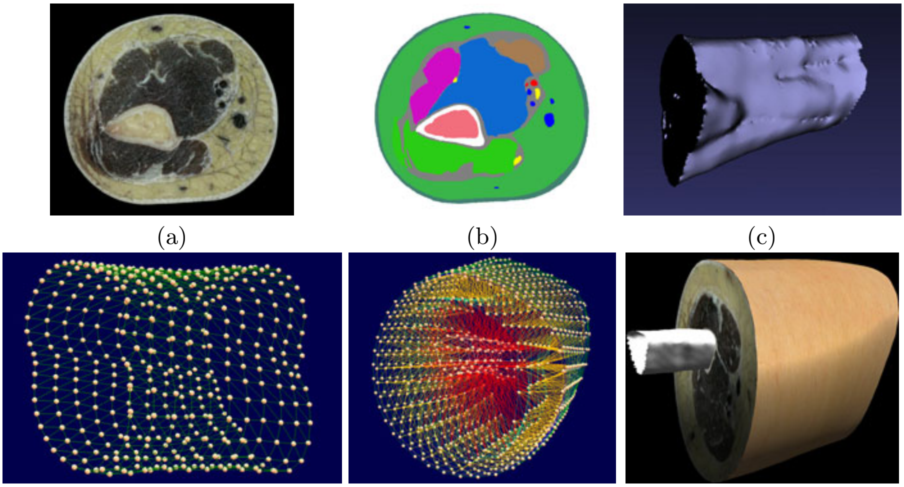


Fig. 1. An example of the reconstruction process. (a) the original regularized grid mesh, (b) tetrahedral structures formed between layers, and (c) the front view of the multilayered volumetric mesh.



(a) (b) (c) (d) (e) (f)

Fig. 2. The construction process of virtual multilayered deformable model for human upper limb. (a) one of the original CVH images (b) segmentation results of the CVH image (c) the muscle surface extracted from the segmented data (d) the muscle surface after reparameterization (e) the multilayered model (different colors denote different group of springs) (f) the textured virtual upper limb.

of mass by merging several adjacent apexes in the lower interfaces based on the curvature of tissue interface. After the merging of apexes, the connection springs are also merged in order to keep each pair of masses connected by only a single spring. Fig. 2(e) shows the generated multilayered mass-spring model for human upper limb, and Fig. 2(f) shows the volume rendering result by 3D texture mapping. After construction of a multilayered model, an XML file describing the geometrical model is generated for deformation and visualization modules.

3.2 Stiffness Configuration in MSS

A key issue with mass-spring systems (MSS) is the assignment of proper parameters, e.g. stiffness of springs, so that the model manifests the true mechanical behavior of the material that it emulates. However, due to the discretized nature of MSS, it is not obvious how to configure its parameters for physically consistent real tissue deformation.

The tuning of parameters may be determined by numerous manual trial-and-errors until the simulated response is seemingly realistic, but this process is tedious. Heuristic approaches have been proposed to identify model parameters automatically by using a more accurate simulated deformation as the benchmarking reference. Optimization techniques such as genetic algorithms [14], simulated annealing [15], and evolutionary algorithms [16,17] can be used to obtain the parameters. However, the setup of optimization procedures requires specific knowledge and careful design to perform the parameter configurations. A better approach, employed in this paper, is based on mathematical derivation from mechanical equations. Lloyd et. al [18] provided a formal derivation for such purposes. In their work, they are trying to equate the stiffness matrices from MSS and FEM, then solve for the stiffness coefficient k of the springs in MSS, as expressed by,

$$k_{i,j} = \sum_e Et\frac{3}{4}. \tag{1}$$

Here we sum over all adjacent element e to nodes i and j. E is Young's modulus and t is thickness of the plane stress elastic model, whereas the Poisson's ratio v is assumed to be $\frac{1}{3}$. This equation works for an equilateral triangle, so we have all the springs equal. To extent it to arbitrary triangles, the authors suggest an approximation by introducing the term $\frac{A_e}{A_0}$ and have the equation,

$$k_{i,j} = \sum_e Et\frac{\sqrt{3}}{4}\frac{A_e}{A_0}. \tag{2}$$

A_e and A_0 are the areas of adjacent triangles and the triangle containing the spring (i, j) respectively.

This equation is simple and easy to be implemented in a 2D MSS. For the 3D case, with a tetrahedral element, the following formula is obtained after minimizing the error analytically:

$$k_{i,j} = \sum_e El\frac{2\sqrt{2}}{25}, \tag{3}$$

where l represents the edge length.

Based on Equation 3 and proper mechanical properties from the literature [19,20], we can configurate our MSS spring parameters properly. In our case, we are using a human upper limb model that consists of 3 tissue layers: dermis, hypodermis, and muscle tissues. The following table lists the parameters involved:

As in Equation 3, we require the edge length l of the tetrahedron, which corresponds to the spring rest length in our MSS. Since springs within the same layer should behave the same, they usually have similar rest length. We try to estimate and assign the same stiffness k to all springs within the same layer. This layer stiffness is computed from the average edge length from springs within a layer. Table 1 shows the stiffness values estimated for different layers of springs.

Table 1. Mechanical properties and assigned stiffness for different skin layers

Layer	Young's modulus (kPa)	Average spring length (cm)	Stiffness (Nm^{-1})
Dermis	35.0	0.92	145.7
Hypodermis	2.0	1.03	9.3
Muscle	80.0	2.47	893.9

3.3 Volume Preservation

Combined with the analytic spring stiffness, the multilayered deformable model is fed into PhysX-enabled GPU. However, in the current implementation of MSS in PhysX, volume preservation forces are not considered. Based on biomechanics literature, most human tissues, including muscle and skin, are approximately incompressible, and the overall volume is well maintained even during a large deformation.

To overcome the inherent drawback of volume loss of MSS, we integrate a fast volume preservation method [21] into our system. Instead of preserving the local volume of every element (typically a tetrahedron), this volume-preservation method maintains the entire volume of deformable objects using a surface mesh.

For a MSS with n masses, the dynamics of the masses can be formulated based on the Newton's law of motion:

$$\boldsymbol{x}(t + \Delta t) = \boldsymbol{x}(t) + \Delta t \dot{\boldsymbol{x}}(t + \Delta t) \tag{4}$$

$$\dot{\boldsymbol{x}}(t + \Delta t) = \dot{\boldsymbol{x}}(t) + \Delta t \boldsymbol{M}^{-1} \boldsymbol{F}(\boldsymbol{x}, t) \tag{5}$$

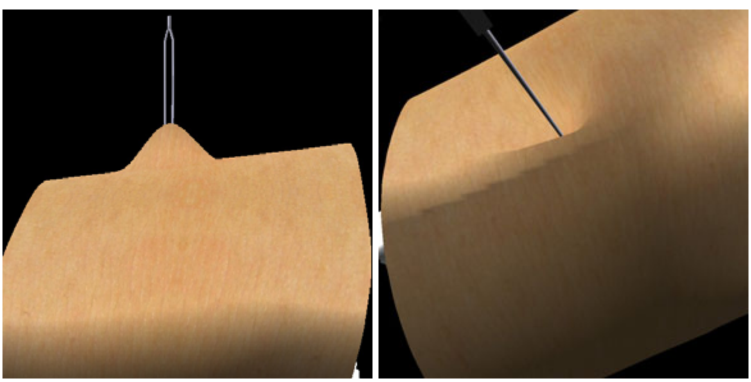
where $\boldsymbol{x}$ is the position vector of the masses; $\dot{\boldsymbol{x}}$ is their velocity vector; $\boldsymbol{M}$ is a $3n \times 3n$ diagonal matrix containing discrete nodal masses, and $\boldsymbol{F}$ is the force vector applied on the masses, including gravitational and spring forces.

To maintain the volume during the deformation, we apply the implicit constraint method [22] to add the constraint force into Eq. 5,

$$\dot{\boldsymbol{x}}(t + \Delta t) = \dot{\boldsymbol{x}}(t) - \Delta t \boldsymbol{M}^{-1} \Upsilon_{\boldsymbol{x}}(\boldsymbol{x}, t)\lambda + \Delta t \boldsymbol{M}^{-1} \boldsymbol{F}(\boldsymbol{x}, t), \tag{6}$$

where λ is a vector containing the Lagrange multipliers and $\Upsilon(\boldsymbol{x}, t)$ is the constraint vector. λ is a scalar in our case where only one constraint, volume preservation, is involved. $\Upsilon(\boldsymbol{x}, t)$ can be written as

$$\Upsilon(\boldsymbol{x}, t) = V - V_0 = 0. \tag{7}$$



(a) (b)

Fig. 3. Visualization results of a virtual upper limb deformations based on our modeling method. (a) deformation caused by a pulling force, (b) deformation caused by a pushing force.

$\Upsilon_{\boldsymbol{x}}(\boldsymbol{x}, t)$ in Eq. 6 equals $\partial V/\partial \boldsymbol{x}$ and can be calculated using information of triangle surfaces based on the divergence theorem.

Thus, in every time step during the simulation, we add the acceleration caused by volume preservation forces to the velocity calculated by PhysX. As only the surface of the volumetric mesh is involved, this method is fast and will not degrade the performance significantly [21]. Experiments show that our system can maintain a frame rate of $25F/S$ even when the number of the spring exceeds $20k$. In our system, we extend this method to support multilayered surfaces. Figs. 3 (a) and (b) show the visualization results of a virtual limb deformation based on our method.

4 Bleeding Simulation

In the development of bleeding simulation in virtual surgical systems, both the patient behavior and mechanical dynamics are of the greatest importance. Most traditional games use a life bar (or Hit Point bar) to represent the status of player: life will be deducted when the player does something wrong, gets hurt or is bleeding. Moreover, the change of player's behavior will be ignored even if they were bleeding over a long time, except the player loses all his life leads to end of the game. This lack of realism should be fixed in health care or surgical serious games: the modeling of patient behavior when bleeding, which is also referred as hemorrhage modeling, must be incorporated into the system. On the other hand, the simulation of bleeding dynamics also needs to be fast enough and realistic to meet the purpose of training. We could not consider a slow motioned training environment to be successful, except for demonstration purposes at the beginning of the training. Hardware acceleration, which has been developed with great success in gaming industry, will be employed to solve the simulation of dynamics quickly. In the following section, we separately introduce details on the fast prototyping of these two aspects related to bleeding simulation.

4.1 Hemorrhage Modeling

Patient behavior with blood loss is common in most surgery. As defined by the World Health Organization, there are five classes of hemorrhage. They are no bleeding (Class 0), petechial bleeding (Class I), mild bleeding (Class II), gross bleeding (Class III) and debilitating blood loss (Class IV). The corresponding percentage of whole blood loss for each stage are listed in Table 2.

Obvious changes in blood pressure (bp) and heart rate (hr) start to appear from Class II bleeding. For example, the difference between the systolic and diastolic blood pressure narrows, and the heart rate increases when blood loss becomes more significant. We can formulate the relationships similar to [23] as

$$bp(t) = bp_0[1 - (\frac{bv_0 - bv(t)}{bv_{max}})^{2.4}]^{1.9}, \tag{8}$$

$$br(t) = \frac{br_0 \times bp(t)}{bp_0} \tag{9}$$

Table 2. Status of hemorrhage model

Class	Status	% of blood loss
0	no bleeding	0
I	petechial bleeding	0 - 15
II	mild bleeding	15 - 30
III	gross bleeding	30 - 40
IV	debilitating blood loss	> 40

where t represents the elapsed time, and bp_0 and $bp(t)$ are the initial and current blood pressure in mmHg respectively. Similarly, br_0 and $br(t)$ are the initial and current bleeding rate in ml/min respectively. Blood pressure also depends on the whole volume of circulating blood. bv_0 and $bv(t)$ are the total and current circulating whole blood volume in ml respectively. bv_{max} denotes the maximum blood volume that can be lost before death (ml), typically equal to $0.8bv_0$.

4.2 Bleeding Dynamics

Human blood is a mixture of blood plasma and blood cells. Normally, red blood cells occupy about 45% of the blood volume while plasma comprises 55% [24]. As plasma is mostly water, so its dynamics, though not exactly the same as water, can acceptably be simplified to be similar to water, which is a newtonian fluid. By assuming constant temperature, the governing rules of dynamics become the well-known Navier-Stokes equations, which formulate conservation of momentum of fluid:

$$\rho \frac{\partial \boldsymbol{V}}{\partial t} = -\rho \boldsymbol{V} \cdot \nabla \boldsymbol{V} - \nabla p + \mu \nabla^2 \boldsymbol{V} + \rho \boldsymbol{g}. \tag{10}$$

Here ρ is the density; $\boldsymbol{V}$ is the vector velocity field; the vector operator ∇ is defined as $\nabla \equiv \boldsymbol{i}\frac{\partial}{\partial x} + \boldsymbol{j}\frac{\partial}{\partial y} + \boldsymbol{k}\frac{\partial}{\partial z}$; p is the pressure; $\boldsymbol{g}$ means the gravitational forces; and μ is viscosity.

To solve the Navier-Stokes equations, we employed a Lagrangian-based method called Smoothed Particle Hydrodynamics (SPH). Compared to other fluid simulation approaches, SPH is relatively simple in implementation and fast in computation [25] [26]. Moreover, it is provided and can be accelerated with PhysX engine, so that sophisticated blood effects can run at interactive frame rates. The basic mechanism of SPH is based on the interaction between other neighboring particles, as illustrated by the following equation.

$$Q_S(\boldsymbol{x}) = \sum_j m_j \frac{Q_j}{\rho_j} W(\boldsymbol{r}, h) \tag{11}$$

where

$$\int W(\boldsymbol{r}, h) d\boldsymbol{r} = 1. \tag{12}$$

A scalar quantity Q at position $\boldsymbol{x}$ is a weighted sum of contributions from neighboring particles. m_j and ρ_j are the mass and density of particle j, respectively, and $\boldsymbol{r} = \boldsymbol{x} - \boldsymbol{x_j}$. A_j is the field quantity at $\boldsymbol{x}_j$. The function $W(\boldsymbol{r}, h)$ is usually referred as the smoothing kernel, in which $lim_{h \to 0} W(\boldsymbol{r}, h) = \delta(\boldsymbol{x})$, where h is the support radius and δ is the Dirac-function.

In equation 10, we have three major quantities including mass, position, and velocity. So we can substitute the scalar quantity Q with these quantities. At every time step of the simulation, SPH is used to evaluate the density (ρ_i) and the force density (f_i^p, f_i^v) caused by pressure and viscosity at particle i based on the following equations, derived in [27]:

$$\rho_i = \sum_j m_j W(\boldsymbol{r}, h) \tag{13}$$

$$f_i^p = -\nabla p(r_i) = -\sum_j m_j (\frac{p_j + p_i}{2\rho_j}) \nabla W(\boldsymbol{r}, h) \tag{14}$$

and

$$f_i^v = \mu \nabla^2 v(r_a) = \mu \sum_j m_j \frac{\boldsymbol{V}_j - \boldsymbol{V}_i}{\rho_j} \nabla^2 W(\boldsymbol{r}, h) \tag{15}$$

where $\boldsymbol{r} = \boldsymbol{x_i} - \boldsymbol{x_j}$. Adding the gravitational force, we can get the total force densities. Thus based on Newton's second law, we can compute acceleration, velocity, and position of every particle. In the computation, we use built-in smoothing kernels ($W(\boldsymbol{r}, h)$) in the PhysX engine to ensure the stability of the simulation.

We can better approximate similar behavior and visual effect by carefully setting and adjusting the properties. Based on research of blood biomechanics, typical values of the density of blood plasma and blood cells are 1025 kg/m^3 and 1125 kg/m^3 respectively, while the whole blood has a specific density between 1.056 kg/m^3 and 1.066 kg/m^3.

In PhysX, Continuous Collision Detection(CCD) mechanism is provided for fast moving objects, which is suitable for simulating interactions between blood and soft tissues such as skin. Two built-in key parameters in PhysX, Coefficient of Restitution and Adhesion Factor, are parameterizable and dynamically adjusted to get different effects. The former controls how much the particles bounce when they collide with soft bodies, while the latter determines how easily particles slide along a surface.

Visualization of blood surface can be realized using a built-in module within the PhysX engine. The provided method employed splatting techniques to render the isosurface of the particles. Another possible method is the Marching Cubes8 [28]. Utilizing the new Geometry Shader features in modern GPUs supporting shader model 4.0, the marching cubes method can be performed entirely inside the GPU in a single pass. From our experiments, interactive or even real-time frame-rates can be achieved by acceleration with GPU. Finally, the isosurface is set to semi-transparent and illuminated in the fragment shader for display, as shown in Fig. 4.

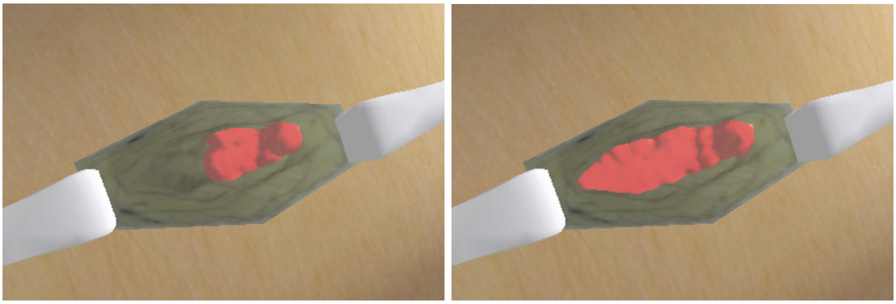


Fig. 4. Bleeding rendering using GPU-accelerated Marching Cube

The average frame rate of PhysX-based system for simulating fluid dynamics with SPH is about 49 fps with 5000 particles on a Pentium 4 Dual Core 3.2 GHz CPU, 4 GB memory and Geforce 8800 graphics card. For rendering marching cubes on GPU using geometry shader, we can achieve 72.5 fps for a volume of $64 \times 64 \times 64$ in size, or 17.48 fps for a volume of $128 \times 128 \times 128$ in size. All these figures ensure an interactive frame rate for bleeding simulation.

5 Discussion and Conclusion

We report a set of tools for fast prototyping of soft tissue deformation and bleeding simulation with PhysX-enabled GPUs. With the functionality provided by the PhysX engine and accompanying with hardware acceleration, short development time yet realistic real-time simulations can be easily achieved. We believe it is important for the future development of healthcare-related simulation system and especially edutainment or serious games. This not only saves time required for debugging, but more importantly, allows developers to concentrate on work flow or training materials and scenarios of related serious games. For example, we have recently successfully applied the related technologies to a serious game for blood management [29].

However, at the time this paper is written, the PhysX engine though is free, but not open source. This limits the customizability and optimizability of certain algorithms. For example, it is simply impossible to extend the governing equation of SPH to support non-newtonian fluids. We think that if the engine can allow more flexibility for customized changes, it can better serve the community of developers of serious simulators and games.

Acknowledgment

This project is supported by the Research Grants Council of the Hong Kong Special Administrative Region, under General Research Fund (Project No. CUHK4453/06M). This work is also affiliated with the Virtual Reality, Visualization and

Imaging Research Centre at CUHK as well as the CUHK MoE-Microsoft Key Laboratory of Human-Centric Computing and Interface Technologies. We thank Professor Michael Cohen at University of Aizu for his kindly help on advising the use of English in the paper.

References

1. Liu, A., Tendick, F., Cleary, K., Kaufmann, C.: A survey of surgical simulation: applications, technology, and education. Presence: Teleoperators and Virtual Environments 12(6), 599–614 (2003)
2. Michel, D., Chen, S.: Serious Games: Games that educate, train and inform. In: Thomson Course Technology, Boston, MA, USA (2006)
3. Narayanasamy, V., Wong, K.W., Fung, C.C., Rai, S.: Distinguishing games and simulation games from simulators. Comput. Entertain. 4(2), 9 (2006)
4. Rosser, J.C., Lynch, P.J., Cuddihy, L., Gentile, D.A., Klonsky, J., Merrell, R.: The impact of video games on training surgeons in the 21st century. Archives of Surgery 142(2), 181–186 (2007)
5. Montgomery, K., Bruyns, C., Brown, J., Sorkin, S., Mazzela, F., Thonier, G., Tellier, A., Lerman, B., Menon, A.: Spring: A general framework for collaborative, realtime surgical simulation. In: Medicine Meets Virtual Reality 2002, pp. 296–303. IOS Press, Amsterdam (2002)
6. Cavusoglu, M.C., Goktekin, T., Tendick, F.: Gipsi: A framework for open source/open architecture software development for organ-level surgical simulation. IEEE Transactions on Information Technology in Biomedicine 10(2), 312–322 (2006)
7. Allard, J., Cotin, S., Faure, F., Bensoussan, P., Poyer, F., Duriez, C., Delingette, H., Grisoni, L.: Sofa - an open source framework for medical simulation. In: Medicine Meets Virtual Reality. IOS Press, Amsterdam (2007)
8. Joachim, G., Rüdiger, W.: Mass-spring systems on the gpu. Simulation Modelling Practice and Theory 13, 693–702 (2005)
9. Mosegaard, J., Sorensen, T.: Gpu accelerated surgical simulators for complex morphology. In: Proceedings of the 2005 IEEE Conference on Virtual Reality, Washington, DC, USA, pp. 147–153. IEEE Computer Society, Los Alamitos (2005)
10. Taylor, Z., Cheng, M., Ourselin, S.: High-speed nonlinear finite element analysis for surgical simulation using graphics processing units. IEEE Transactions on Medical Imaging 27(5), 650–663 (2008)
11. Nealen, A., Muller, M., Keiser, R., Boxerman, E., Carlson, M.: Physically based deformable models in computer graphics. In: Eurographics State of Art. (2005)
12. Pang, W.M., Qin, J., Chui, Y.P., Wong, T.T., Leung, K.S., Heng, P.A.: Orthopedics surgery trainer with ppu-accelerated blood and tissue simulation. In: Ayache, N., Ourselin, S., Maeder, A. (eds.) MICCAI 2007, Part II. LNCS, vol. 4792, pp. 842–849. Springer, Heidelberg (2007)
13. Maciel, A., Halic, T., Lu, Z., Nedel, L.P., De, S.: Using the physx engine for physics-based virtual surgery with force feedback. The International Journal of Medical Robotics and Computer Assisted Surgery 5(3), 341–353 (2009)
14. Louchen, J., Provot, X., Crochemore, D.: Evolutionary identification of cloth animation models. In: Proceedings of Eurographics Workshop on Computer Animation and Simulation, pp. 44–54 (1995)

15. Deussen, O., Kobbelt, L., Tucke, P.: Using simulated annealing to obtain good nodal approximations of deformable bodies. In: Proceedings of Eurographics Workshop on Computer Animation and Simulation, pp. 30–43 (1995)
16. Bianchi, G., Harders, M., Székely, G.: Mesh topology identification for mass-spring models. In: Ellis, R.E., Peters, T.M. (eds.) MICCAI 2003. LNCS, vol. 2878, pp. 50–58. Springer, Heidelberg (2003)
17. Bianchi, G., Solenthaler, B., Szekely, G., Harders, M.: Simultaneous topology and stiffness identification for mass-spring models based on fem reference deformations. In: Barillot, C., Haynor, D.R., Hellier, P. (eds.) MICCAI 2004. LNCS, vol. 3217, pp. 293–301. Springer, Heidelberg (2004)
18. Lloyd, B., Székely, G., Harders, M.: Identification of spring parameters for deformable object simulation. IEEE Transactions on Visualization and Computer Graphics 13(5), 1081–1094 (2007)
19. Pailler-Mattei, C., Beca, S., Zahouani, H.: In vivo measurements of the elastic mechanical properties of human skin by indentation tests. Medical Engineering and Physics 30(5), 599–606 (2008)
20. Levinson, S.F., Shinagawa, M., Sato, T.: Sonoelastic determination of human skeletal muscle elasticity. Journal of Biomechanics 28(10), 1145–1154 (1995)
21. Hong, M., Jung, S., Choi, M.H., Welch, S.W.J.: Fast volume preservation for a mass-spring system. IEEE Comput. Graph. Appl. 26(5), 83–91 (2006)
22. Hong, M., Choi, M.-H., Jung, S., Welch, S., Trapp, J.: Effective constrained dynamic simulation using implicit constraint enforcement. In: International Conference on Robotics and Automation (April 2005)
23. Simpson, S.H., Menezes, G., Mardel, S.N., Kelly, S., White, R., Beattie, T.: A computer model of major haemorrhage and resuscitation. Med. Eng. Phys. 18(4), 339–343 (1996)
24. Fung, Y.C.: Biomechanics: Mechanical Properties of Living Tissues, 2nd edn. Springer, Berlin (1993)
25. Libersky, L.D., Petschek, A.G., Carney, T.C., Hipp, J.R., Allahdadi, F.A.: High strain lagrangian hydrodynamics: a three-dimensional sph code for dynamic material response. Journal of Computational Physics 109, 67–75 (1993)
26. Monaghan, J.J.: Simulating free surface flows with sph. Journal of Computational Physics 110, 399–406 (1994)
27. Müller, M., Charypar, D., Gross, M.: Particle-based fluid simulation for interactive applications. In: Proceedings of ACM SIGGRAPH Symposium on Computer Animation (SCA), pp. 154–159 (2003)
28. Lorensen, W.E., Cline, H.E.: Marching cubes: A high resolution 3d surface construction algorithm. In: SIGGRAPH 1987 Proceedings, vol. 21(4) (July 1987)
29. Qin, J., Chui, Y.P., Pang, W.M., Choi, K.S., Heng, P.A.: Learning blood management in orthopedic surgery through gameplay. IEEE Computer Graphics and Applications 30, 45–57 (2010)

Dance-Based ExerGaming: User Experience Design Implications for Maximizing Health Benefits Based on Exercise Intensity and Perceived Enjoyment

Alasdair G. Thin and Nicola Poole

School of Life Sciences, Heriot-Watt University, Edinburgh, EH14 4AS, United Kingdom
a.g.thin@hw.ac.uk

Abstract. Dance is a form of exercise that is considered to have widespread popular appeal and in particular to adolescent females. Dance-based body-movement controlled video games are a popular form of ExerGaming that is being adopted for use in school-based physical activity health promotion programs. The results of this study indicate that the game play mechanics and skill demands of the dance-based ExerGames would appear to have limited the subjects' level of physical exertion over the period of study. After training there was an increase in enjoyment rating for the Step Aerobics game which appears related to a perceptible improvement in game performance. It is therefore recommended that ExerGames should be designed with very low initial skill demands in order to maximize the user's level of exertion and to realize and reward progress, thereby helping to promote an enjoyable exercise experience and counterbalance any sense of exertional discomfort. Keywords: exercise; health promotion; exergaming; user experience; design; video game; enjoyment.

1 Introduction

ExerGames are now a significant social phenomenon [1] due to recent advances in video games technology, including the development of low cost body-movement sensitive controllers, coupled with the desire of video games companies to appeal to new markets. There is a wide spectrum of ways to interact with the games from motion capture camera systems, motion sensitive hand-held controllers, pressure sensitive boards and foot pad switches, through to sports boards with tilt sensors and adapted arm and leg ergometers. Historically video games have attracted a negative press due to their apparent encouragement of sedentary behavior. However, research findings do not single out any one form of "screen time" (e.g. television viewing, computer use, video game play) as being directly related to a positive energy balance (consuming more calories than expended), but rather it is an accumulation of sedentary behaviors that is detrimental [2]. Quite understandably, both the video games industry and the media are keen for something positive to say about video games [1,3].

Health agencies have begun to consider the potential of ExerGames because of both their popularity and the limited success of traditional health promotion approaches. Dance is a form of exercise that has widespread popular appeal and in particular to adolescent females. Dance-based body-movement controlled video games

Z. Pan et al. (Eds.): Transactions on Edutainment IV, LNCS 6250, pp. 189–199, 2010.

are a popular form of ExerGaming due to the success of the Dance Dance Revolution (DDR) series of games (Konami Inc.) and it has been adopted for use in several school-based physical activity health promotion programs [4].

For a large part of human history, being physically active has been a necessity in order to survive [5]. In contrast, the reality of 21st century life in the developed world is one of comprehensive mechanization to the extent that physical exertion has become almost exclusively a leisure time pursuit. Yet the term “working out” is loaded with negative connotations [6]. Exercise is generally seen as a chore and it is not something that is done for fun, but rather it is often uncomfortable and unpleasant and hardly what one would choose as a so called “leisure” (discretionary) time pursuit. It is not surprising therefore that the effects of a modern sedentary lifestyle are becoming evermore apparent with rising rates of diseases that have significant lifestyle factors (e.g. diabetes, heart disease, cancer) and the prevalence of overweight and obese individuals in the developed world continuing to rise. Assuming current trends, by 2015 in the United States, 5 in every 20 children and 8 in every 20 adults and will be obese and there will be 1.5 billion obese people worldwide [7]. On any scale, this is a big problem and the capacity to treat patients who are overweight or obese is already limited within health services and the projected increases in demand will only stretch the systems further. Effective preventative efforts are therefore urgently needed and on a grand scale.

Health promotion interventions have for a long time been focused on educating the population (or specific target groups within) with a view to changing attitudes and beliefs towards a particular health behavior. The underlying expectation is that when a person is given insight into the need to change their ways, they are sufficiently motivated to adopt a new set of behaviors with ensuing positive health benefits. Unfortunately such attempts at mass education of the population are undermined on an almost daily basis by a barrage of conflicting media reports as to what now happens to be “good” for peoples’ health and what is now “bad” for it [8]. Despite widespread appreciation of the health benefits of regular exercise, there is a lack of translation into significant behavioral change. This gap, and it is a large one, between peoples’ health intentions and their actions has been termed the “Intentional Gap” [9]. What are desperately needed are new health promotion intervention strategies that can overcome this chasm.

The Transtheoretical model of behavior change is the most widely used theoretical underpinning to interventions designed to promote physical activity [10] and much of the research effort has focused on the factors that affect a person’s motivation to start exercising e.g. the benefits of regular exercise in terms of future avoidance of various chronic diseases such as heart disease, diabetes and cancer. As a consequence health promotion interventions designed to increase levels of physical activity in the population have therefore placed a significant emphasis approaches and techniques to maximize a given person’s motivation to become active. However, rather disappointingly, efforts to encourage people to become physically active have to date been largely unsuccessful due to high rates of attrition [11]. It has been suggested that reason for this may be focus on rather remote and abstract outcomes (i.e. health benefits) rather than more immediate rewards e.g. enjoyment [12].

Rather surprisingly, the nature of the experience that people have whilst undertaking exercise and whether or not they find it an enjoyable experience is under-researched. Only recently has it been argued that they a major reason for the very limited success of physical activity interventions with that they have failed to consider the experiential aspects [10]. Thus whilst the health benefits of regular exercise are widely documented, it would be fair to say that exercise is often perceived as something to be endured rather than enjoyed [6]. This is likely to be at least in part due to the origins of the fitness industry in the physical culture movement, whereby exercise was approached in a work-like manner as a worthwhile investment in an individual's "physical capital" [6]. If the nature of the experience of performing the exercise is such a negative one and so lacking in enjoyment, it is perhaps little wonder that attempts by health promotion authorities to get people more physically active have so far been largely unsuccessful. It is only recently that the findings of an exercise intervention focusing on enjoyment, which sought to make physical education (PE) for high school girls more enjoyable and resulted in an increase in levels of physical activity, have been reported [12].

Interestingly, there is an analogous situation with regard to video games, whereby gameplay is normally described in terms of the various game components and how these motivate the user to play, rather than the motivation arising from user finding the nature of the experience of playing the game to be enjoyable [13]. When due consideration is given to the experiential aspects of gameplay, then from the player's perspective, the gameplay is more akin to an artistic performance, albeit one that is structured to a greater or lesser degree by the game rules and mechanics [13,14]. It has been argued that such a change in emphasis to focus on the experience provides a helpful and insightful route to innovative game design [13]. Furthermore, it has been proposed that when gameplay is viewed as a performance it encapsulates a number of different aspects (vectors) of how the event relates to the physical capacity of the player, their skill at playing the game, their performing to an audience, personal expression and it can also act as a vehicle for the expression of artistic talent and drama [15]. The relative focus on each of these vectors will be influenced by many factors including the nature of the game being played, the intentions of the player and the social context.

ExerGaming is a concatenation of "Exercise" and (Video) "Gaming" and draws on the rich heritage and experience of the video game industry in creating engaging interactive experiences. Simply viewing ExerGames as technology enhanced-exercise fails to do the genre justice [16]. In the same way that there is a large range of different types and genres of video games designed to appeal to different groups of gamers, there is also the possibility of different ExerGames appealing to differing individual interests of exercisers. By combining different game play scenarios, rules and mechanics there is the potential to create a large number of different highly enjoyable ExerGaming experiences which would be adjustable to the skill level and fitness of the player. Such flexibility to accommodate a wide range of interests is much more likely to result in a good match between user and game and therefore they would be expected to have an enjoyable exercise experience.

There is a real opportunity to capitalize on the wave of enthusiasm for ExerGaming that has turned it into a social phenomenon and to realize positive public health benefits. However, if excessive claims are made, or if ExerGaming is shoehorned into traditional health interventions or game design is not informed by exercise physiology and psychology, then this opportunity may well be lost. Given that the current generation of ExerGames have to a large extent been designed with entertainment rather than physical activity in mind, it is important therefore to properly evaluate the games with regard to their potential health benefits. There a need to accurately quantify the degree of physical exertion required to play a specific ExerGame (which in turn determines the level of energy expenditure). It is also essential to research the ability of ExerGames to provide an enjoyable and engaging experience and provide a gateway to an active lifestyle. The results of such evaluations will provide insight into future ExerGame design and development with a view to creating and enjoyable exercise experience and at the same time maximizing the health benefits. The aim of this study was therefore to investigate the level of physical exertion required to play three different dance-based ExerGames and also assess the interaction between game play factors and exercise intensity on the level of enjoyment that players experience whilst playing the ExerGames.

2 Methods

Subjects were recruited from the student population at Heriot-Watt University. The study was subject to local ethical committee approval and all subjects gave written, informed consent and underwent health screening. The study comprised of five sessions in total, an initial exercise testing and familiarization session, a pre-training session, two training sessions and a final post-training testing session.

2.1 Initial Exercise Testing and Familiarization Session

The first session comprised an incremental exercise test a cycle ergometer (Monark Ergomedic 874E, Monark Exercise AB, Varberg, Sweden) in order to assess each subject's fitness level in terms of their maximal aerobic exercise capacity (VO_2max). This test lasted for approximately 10 minutes duration and oxygen consumption (VO_2) was measured throughout using an ambulatory telemetric expired gas analysis system (Oxycon Mobile, Cardinal Health, Basingstoke, UK). It comprises a breathing mask worn over the mouth and nose and incorporates an impropeller flow turbine and a gas sample line which is connected to the micro-analyzers worn on the subject's back. The flow sensor and gas analyzers are calibrated before use using a pump with known rates of air flow and a precision gas mixture respectively. Flow and gas concentration are cross-product integrated on a breath-by-breath basis to determine the rates of oxygen consumption and carbon dioxide production. The device also incorporates a battery pack and a radio transmitter to relay the data to a laptop base station. In total the system has a mass of 1.2 kg and is held in place with a harness incorporating adjustable straps and did not hinder the subjects' movements whilst playing the ExerGames being studied. Subjects' height and weight were also measured during the familiarization session (portable stadiometer (model 225, Seca Ltd, Birmingham, UK) and weighing scales (model 770, Seca Ltd)). After a short rest period to recover from the fitness test, the subjects had a period of familiarization with each of the three dance-based ExerGames used in the study.

Dancing Stage (*Dance*) (Konami Digital Entertainment GmbH, Frankfurt/Main, Germany) is a version of the popular arcade game Dance Dance Revolution adapted for the Nintendo Wii games consoles. It comes with its own pressure sensitive mat and also makes use of the Wii's hand-held motion sensitive controllers to incorporate hand movements into the game. The objective of the game is to step on the various defined areas of the mat and shack your hands in time to choreographed sequence shown on the screen, the timing of which corresponds to the beats of the music track (Fig. 1A). As this game involves a large amount of muscle mass due to its involvement of both arm and leg movements, it might be expected to place high exertional demands on players.

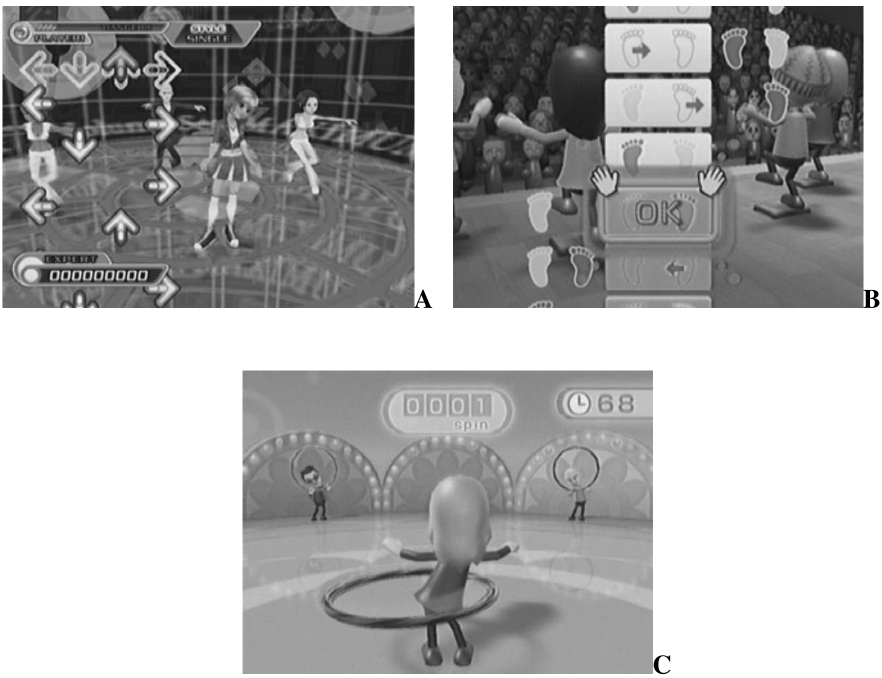


Fig. 1. Screenshots of Dance-based ExerGames used in study. A = Dancing Stage (*Dance*), B = Step Aerobics (*Step*), C = Hula Hoop (*Hula*).

The second ExerGame investigated in this study was Step Aerobics (*Step*) which is one of the mini-games that form part of the Wii Fit package that utilizes the balance board peripheral (Nintendo of Europe GmbH, Grossostheim, Germany). It involves performing a variety of stepping moves in time to a choreographed sequence (Fig. 1B) and also includes additional kicking actions. The balance board peripheral detects a player's actions by changes in the pressure exerted on different parts of the board. Stepping actions to raise an individual's boy mass require work to be performed against gravity, however the board is only 50 mm high and therefore only a moderate level of exertion is likely to be required.

The third and final game studied was Hula Hoop (*Hula*) and is another of the mini-games that form part of the Wii Fit package and is based on the popular children's toy of the same name. It requires the player to rotate their hips in a coordinated manner as fast as possible to accumulate as many revolutions as possible in the time limit (Fig. 1C). Additional hoops can also be "caught" by leaning to catch hoops that are thrown towards the player. Since it involves a large number of skeletal muscles in maintaining the hip rotation, this game is expected to be more intensive than it might otherwise appear.

2.2 Pre-Training Testing Session (*Pre-Tr*)

On the second occasion subjects performed each game for around five to six minutes in a randomized order. Subjects wore the ambulatory gas analysis system throughout the game play periods in order to determine the exertional demands of the game play in terms of the measured oxygen consumption (VO_2). After each game bout the Physical Activity Enjoyment Scale (PACES) questionnaire was administered [17]. It comprises 18 seven-point bipolar statements with a number of the items reverse scored. It instructs subjects to respond by rating "How they feel at the moment:" with regard to the physical activity they have just undertaken.

2.3 Training Sessions

Subjects undertook two training sessions on separate occasions where each game was played in a randomized order for approximately five minutes of actual game play, each in a similar manner to the *Pre-Tr* session except without any measurements being made.

2.4 Post-Training Testing Session (*Post-Tr*)

In the final session (*Post-Tr*) the subjects again played each game for around five to six minutes in a randomized order and the expired gas analysis (VO_2) and physical activity enjoyment scale (PACES) measurements were repeated as per the *Pre-Tr* session.

2.5 Game Scores

For each ExerGame played by the subjects during the *Pre-Tr* and *Post-Tr* testing sessions an experimenter noted down the player's game scores but did not give them any feedback on their performance. The time for each period of game play was sufficient to play only one game of Hula Hoop (*Hula*) and Step Aerobics (*Step*). The scoring for the former is depends on the number of hoop revolutions performed in the time period of the game and for the latter it is the number of correctly timed steps performed. The duration of the Dancing Stage (*Dance*) games were of a much shorter time period such that up to five games could be played in each session. Whilst the scoring was again based on the number of correctly timed moves, the score for each game is given as a letter grade (A-E) at the end of the game. In order to convert these to an overall score an "E" grade was allocated 1 point and so on up to 5 points for an "A" grade, giving a possible maximum score of 25. However subjects were not told their composite scores but rather only saw the on-screen letter ratings at the end of each game.

2.6 Statistical Analysis

In order to determine the level of physical exertion each subject was exercising at whilst playing the ExerGames, their measured oxygen consumption (VO_2) was expressed as a percentage of their oxygen uptake reserve (%VO_2reserve) (i.e. as a percentage of the difference between their maximum (VO_2max) and their resting VO_2 (assumed to be 3.5 ml/kg/min)) [18]. All data are reported as mean ± standard error of the mean (SEM) with the exception of the demographic data. Pre- and post-training data were compared using repeated measures ANOVA (SPSS 14.0 for Windows, SPSS Inc., Surrey, UK) with the Bonferroni adjustment for multiple comparisons.

3 Results

Twelve recruited and 11 (5 male, 6 female) completed the study. The subjects had a mean (±SD) age 20 ± 2 years, height 1.70 ± 0.10 m, weight 67.7 ± 10.6 kg, body mass index (BMI) 23.3 ± 2.7 kg/m^2 and a maximal aerobic fitness (VO_2max) 43.7 ± 4.4 ml/kg/min. There were no effects of game play order on the results (data not shown). There were no differences in mean level of physical exertion (%VO_2reserve) between the *Pre-Tr* and *Post-Tr* sessions for any of the games (Fig. 2).

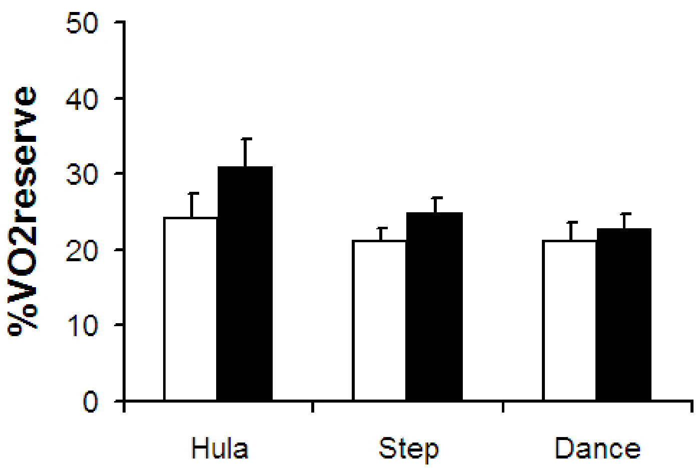


Fig. 2. Mean physical exertion (%VO_2reserve) for the pre-training (*Pre-Tr*: open bars) and post-training (*Post-Tr*: solid bars) sessions for each of the three Dance-based ExerGames. Hula = Hula Hoop, Step = Step Aerobics, Dance = Dancing Stage. Error bars indicate SEM.

After training (*Post-Tr*) there was a significant increase in mean PACES rating for the Step Aerobics (*Step*) game (Fig. 3).

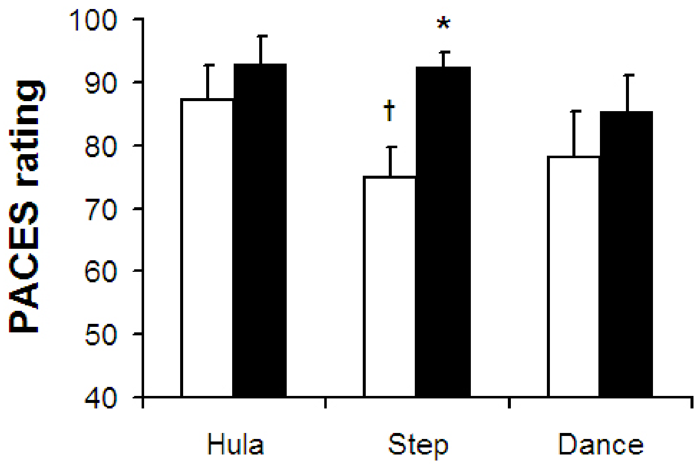


Fig. 3. Mean PACES ratings for the pre-training (*Pre-Tr*: open bars) and post-training (*Post-Tr*: solid bars) conditions for each of the three Dance-based ExerGames. * indicates significantly different from the *Pre-Tr* value (p<0.05). Hula = Hula Hoop, Step = Step Aerobics, Dance = Dancing Stage. Error bars indicate SEM. † indicates significantly different from Hula *Pre-Tr* (p<0.05).

The increase in mean PACES rating for the Step Aerobics (*Step*) game corresponded with an increase in the mean *Post-Tr* game score (Fig. 4). There was also an increase in the composite score for the Dancing Stage (*Dance*) game after training (Fig. 4).

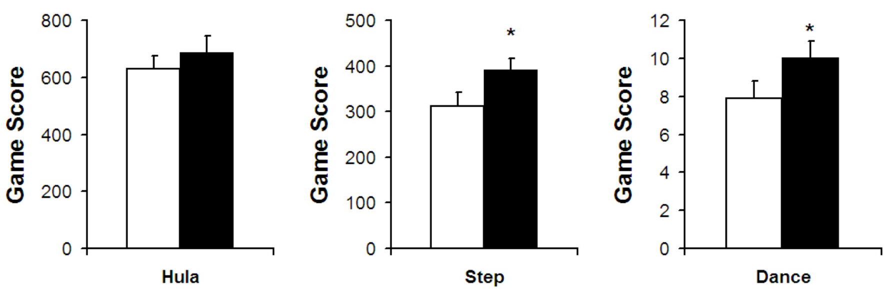


Fig. 4. Mean Game Scores for the pre-training (*Pre-Tr*: open bars) and post-training (*Post-Tr*: solid bars) conditions for each of the three Dance-based ExerGames. Hula = Hula Hoop, Step = Step Aerobics, Dance = Dancing Stage. Error bars indicate SEM. * indicates significantly different from the pre- training value (p<0.05).

4 Discussion

In terms of the *physical exertion* required to play the dance-based ExerGames, the results of this study indicate that, even after short period of training, it was at a level below the range required to improve aerobic fitness [18]. In contrast, an earlier study using the arcade version of the game (Dance Dance Revolution) reported a level of exertion that was just inside the range appropriate for individuals with a low level of fitness [19]. However, in terms of *energy expenditure*, the periods of ExerGame play in the current study equated to 3.5-4.5 times resting metabolic rate and would therefore make a useful contribution to maintaining energy balance and therefore preventing weight gain.

There was a significant increase in the physical activity enjoyment rating (PACES) for the Step Aerobics (*Step*) game after the training period (Fig. 3) and this corresponded with an increase in mean game score (Fig. 4). It is likely therefore that a clear sense of achievement was behind this ratings increase. Mean (±SD) *Post-Tr* PACES ratings for the *Hula* and *Step* games (93 ± 15 and 92 ± 8 respectively) were somewhat higher than the 85 ± 19 after 15 minutes of exercise on a cycle ergometer at a self-selected pace as reported in the validation study for the PACES questionnaire [17]. This positive difference corresponds to an effect size of +0.4 and which is statistically significantly higher for the *Step* game but only approaching significance (p=0.55) for *Hula* game due to a wider distribution of ratings.

The small improvement in the composite Dancing Stage (*Dance*) score (Fig. 4) equates to going from an "E^{+}" to a "D" grade. The composite scores were not computed until after the subjects had completed all the testing and they were not therefore aware of them at the time of completing the PACES questionnaire. It is therefore highly unlikely that the subjects would have perceived this small overall increase in performance on the *Dance* game. Furthermore, the nature of the game's scoring system would have tended to reinforce what were rated to be relatively poor performances by the subjects and would therefore if anything, be expected to have been somewhat discouraging to the subjects.

One of the most attractive and exciting aspect of the whole ExerGaming phenomenon is that it has the potential to enable physical educators to break free from various "one size fits all" constraints and adopt an approach that is customizable to the specific needs of each individual in a class. Not only is there scope for a wide range of genres of games to appeal to different players, but more fundamentally, if appropriately programmed, an ExerGame could adapt and respond to an individual player's levels of skill and fitness. This would help ensure that the game presents a challenge at an appropriate level for each individual. Furthermore, as the player's skill and fitness improve, the game would be able to upwardly adjust the difficulty level and exertional demands of the ExerGame.

There are a whole host of features which could be incorporated into an ExerGame in order to help make it a more enjoyable exercise experience including (but not limited to) offering helpful tips and advice, motor skill learning and development and the coaching of particular moves and combinations of movements, monitoring game play time in terms of the amount of physical activity performed and providing some form of progress chart and timely positive feedback. It is not necessary and indeed may not be desirable that some of these aforementioned features are incorporated into

ExerGames by directly mimicking the corresponding counterparts from traditional forms of exercise (e.g. some of the roles erstwhile performed by a personal trainer). Rather than having an explicit virtual representation of a human trainer these features could be subtlety incorporated into the ExerGame play [20]. Finally, given the aim of facilitating a habitually physically active lifestyle, ExerGames should be designed with the long-term in mind. Designers should therefore consider carefully how a player will progress through an ExerGame and how they can maintain interest and continue to present a real challenge to the player by increasing the exertional demands as their skill and fitness improve.

5 Conclusion

The results of this study indicate that the game mechanics and skill demands of the dance-based ExerGames would appear to have limited the subjects' intensity of physical exertion over the period of study. The observed increase in enjoyment rating for the Step Aerobics game after a period of training was most likely related to a perceptible improvement in game performance and there is evidence to suggest that dance-based ExerGames can provide an exercise experience which is more enjoyable than using traditional equipment such as an exercise bike. However, in order to realize significant health benefits, the design of future ExerGames must be informed by exercise physiology and psychology principles. It is therefore recommended that ExerGames should be designed with very low initial skill demands in order to maximize a player's level of exertion and to realize and reward progress, thereby helping to promote enjoyment and counterbalance any sense of exertional discomfort.

References

1. Sayre, C.: Video Games That Keep Kids Fit. Time Magazine (September 13, 2007)
2. Marshall, S.J., Biddle, S.J., Gorely, T., Cameron, N., Murdey, I.: Relationships between media use, body fatness and physical activity in children and youth: a meta-analysis. Int. J. Obes. Relat. Metab. Disord. 28, 1238–1246 (2004)
3. Hetherington, J.: The Wii Workout: Getting Healthy Through Gameplay. Animation World Magazine (March 31, 2008)
4. Schiesel, S.: P.E. Classes Turn to Video Game That Works Legs. The New York Times (April 30, 2007)
5. Booth, F.W., Chakravarthy, M.V., Gordon, S.E., Spangenburg, E.E.: Waging war on physical inactivity: using modern molecular ammunition against an ancient enemy. J. Appl. Physiol. 93, 3–30 (2002)
6. Maguire, J.S.: Fit for consumption: sociology and the business of fitness. Routledge, New York (2007)
7. Kumanyika, S.K., Obarzanek, E., Stettler, N., et al.: Population-based prevention of obesity: the need for comprehensive promotion of healthful eating, physical activity, and energy balance: a scientific statement from American Heart Association Council on Epidemiology and Prevention, Interdisciplinary Committee for Prevention (formerly the expert panel on population and prevention science). Circulation 118, 428–464 (2008)
8. Eaton, S.B., Strassman, B.I., Nesse, R.M., et al.: Evolutionary health promotion. Prev. Med. 34, 109–118 (2002)

9. Amireault, S., Godin, G., Vohl, M.C., Perusse, L.: Moderators of the intention-behaviour and perceived behavioural control-behaviour relationships for leisure-time physical activity. Int. J. Behav. Nutr. Phys. Act. 5, 7 (2008)
10. Marshall, S.J., Biddle, S.J.: The transtheoretical model of behavior change: a meta-analysis of applications to physical activity and exercise. Ann. Behav. Med. 23, 229–246 (2001)
11. Thurston, M., Green, K.: Adherence to exercise in later life: how can exercise on prescription programmes be made more effective? Health Promot. Int. 19, 379–387 (2004)
12. Dishman, R.K., Motl, R.W., Saunders, R., Felton, G., Ward, D.S., Dowda, M., Pate, R.R.: Enjoyment mediates effects of a school-based physical-activity intervention. Med. Sci. Sports Exerc. 37, 478–487 (2005)
13. Kristiansen, E.: Designing innovative video games. In: Sundbo, J., Darmer, P. (eds.) Creating experiences in the experience economy. Edward Elgar Publishing Ltd., Cheltenham (2008)
14. Behrenshausen, B.G.: Toward a (Kin) Aesthetic of Video Gaming: The Case of Dance Dance Revolution. Games and Culture 2, 335–354 (2007)
15. McGonigal, J.: Gamers Love Performance, Perform or Else: A Workshop on Extroverted Gameplay. DIGRA, Vancouver (2005)
16. Thin, A.G., Hansen, L., McEachen, D.: Flow Experience and Mood States whilst Playing Body-Movement Controlled Video Games (2009) (submitted for publication)
17. Kendzierski, D., DeCarlo, K.J.: Physical Activity Enjoyment Scale: Two Validation Studies. J. Sport Exer. Psy. 13, 50–64 (1991)
18. American College of Sports Medicine: ACSM's Guidelines for exercise testing and prescription. Lippincott Williams & Wilkins, Philadelphia (2000)
19. Tan, B., Aziz, A.R., Chua, K., Teh, K.C.: Aerobic demands of the dance simulation game. Int. J. Sports Med. 23, 125–129 (2002)
20. Bogost, I.: Persuasive Games: The Expressive Power of Videogames. The MIT Press, Cambridge (2007)

Learning Ultrasound-Guided Needle Insertion Skills through an Edutainment Game

Wing-Yin Chan[1], Dong Ni[1], Wai-Man Pang[1], Jing Qin[1], Yim-Pan Chui[1], Simon Chun-Ho Yu[2], and Pheng-Ann Heng[1,3]

[1] Department of Computer Science and Engineering, The Chinese University of Hong Kong, Hong Kong SAR, China
{wychan,dni,wmpang,jqin,ypchui,pheng}@cse.cuhk.edu.hk
[2] Department of Diagnostic Radiology and Organ Imaging, The Chinese University of Hong Kong, Hong Kong SAR, China
simonyu@cuhk.edu.hk
[3] Shenzhen Institute of Advanced Integration Technology, Chinese Academy of Sciences/The Chinese University of Hong Kong, China

Abstract. Ultrasound-guided needle insertion is essential in many of minimally invasive surgeries or procedures, such as biopsy, drug delivery, spinal anaesthesia, etc. Accurate and safe needle insertion is a difficult task due to the high requirement of hand-eye coordination skills. Many proposed virtual reality (VR) based training systems put their emphasis on realistic simulation instead of pedagogical efficiency. The lack of schematic training scenario leads to boredom of repetitive operations. To solve this, we present our novel training system with the integration of game elements in order to retain the trainees' enthusiasm. Task-oriented scenarios, time attack scenarios and performance evaluation are introduced. Besides, some state-of-art technologies are also presented, including ultrasound simulation, needle haptic rendering as well as a mass-spring-based needle-tissue interaction simulation. These works are shown to be effective to keep the trainees up with learning.

1 Introduction

With the fast advancement in technologies of virtual reality based surgical simulation in the last decade, many virtual simulators providing impressive visualization and haptic feedback are proposed, letting trainees to practice various surgical procedures. While from the pedagogical point of view, these are far from enough to be an effective learning platform. A better design in training scenarios, learning flow, riveting interactions between the users and system, as well as appropriate evaluation mechanisms is definitely our next step to amplify the efficiency of a training system. The aim of this work is to let trainees learn the surgical skills more actively, independently and efficiently.

The concept of "serious game (or edutainment game)" [1,2] has been proposed recently, referring to the emerging research in using video game elements for purposes apart from entertaining. Some of the examples include flight and driving

Z. Pan et al. (Eds.): Transactions on Edutainment IV, LNCS 6250, pp. 200–214, 2010.

training, physical exercise [3], culture and social learning [4], and healthcare education [5]. The process of game playing is believed to be important for the cognitive development of a student. In a more relaxing game-like environment, players usually exert more efforts and achieve better results. It is relevant to educational strategies of active and experiential learning. In healthcare, many game-based rehabilitation training systems have been developed to make the related process more interesting, so as to keep the patients concentrate on it and willing to repeat the training [6,7]. To this end, it is natural to consider about the integration of edutainment game elements into surgical training simulators in order to provide fun interaction and improve the learning efficiency and enjoyment.

In this paper, we report our experience in designing and implementing an edutainment game for ultrasound-guided needle insertion training with the state-of-the-art simulation techniques and a series of game elements. To our knowledge, this is the first work to introduce edutainment designs for the training of ultrasound-guided needle insertion procedures. Needle placement into a patient body on a specific target under the guidance of ultrasound is a common but difficult skill in interventional radiology. Safe procedure requires a high level of hand-eye coordination, which must be developed through intensive practice, especially when the tissue is under deformation or various target positions and needle shapes are considered.

To provide a more realistic environment compared to previous work, we propose and introduce some novel techniques like a new ultrasound image simulation approach and a mass-spring-based needle-tissue interaction simulation. To engage the trainees for intensive training, game elements including task-oriented scenarios, time attack scenarios and performance evaluation tools are introduced. The trainee is challenged by tasks with various difficulties defined according to the target size and position, the time limit and the insertion path. A pilot study demonstrates that our game-based simulation system is effective for the trainees in acquiring the necessary skills in the needle insertion process.

The following of this paper is organized as follows. In section 2, we briefly review related work on both serious game and virtual-reality based ultrasound-guided needle insertion training systems. Section 3 presents the design and implementation details of the proposed edutainment game. Section 4 provides the pilot evaluation results of our system and gives some discussions on several future directions.

2 Related Work

The primary purpose of most computer or video games is entertainment. Players enjoy the process of achieving goals in the gaming environment by making decisions and formulating strategies. The nature of game playing is relevant to active and experiential learning where trainees take initiatives to learn and solve problems, thereby gaining deep understanding in an active way. Furthermore, the reusability and repeatability offered by computer games facilitate the reinforcement of acquired knowledge. In this regard, computer games have been applied to serve “serious” non-entertaining purposes, e.g. military, education, government, corporation and healthcare [2]. In healthcare, serious games have been widely introduced for distraction therapy, physical fitness, self-management and health education, rehabilitation, mental health, and surgical training [2,6,7]. Unlike generic games or simulation games, the major design objective is not “fun” but skill development. To facilitate the transfer of skills, serious games should also provide testing and progress tracking to evaluate how much a trainee has learned from playing [8].

There were a few ultrasound-guided needle insertion simulation systems developed previously. For example, Franck et al. [9] had presented a training system using actual patient data with force feedback device. However, this system provides only CT or X-Ray images where physicians usually prefer to use ultrasound image as guidance. Simulator presented by Forest et al. [10,11] tries to improve the visualization of the simulated ultrasound images based on CT or MRI data. However, the acquisition mechanisms of these images are essentially different from those by ultrasound. Magee et al. [12,13] introduced an augmented-reality system for the training of ultrasound-guided needle insertion procedure with improved techniques for realistic simulated ultrasound images. The key limitation of this system lies on its lack of force feedback, as the virtual transducer and the needle are both simulated by a pair of magnetic based 3D position sensors. Apart from the mentioned technical shortcomings, none of the above systems consider the advantage provided by serious games, like a systematic guidance and edutainment elements for keeping trainee’s learning passion.

Small tissue deformations introduced by needle insertion take place during most needle insertion related surgical procedures. Simulation of such an effect would greatly enhance the realism of our biopsy training simulator. Traditionally, finite element method (FEM) has been applied to solve various deformation simulation. Though FEM is generally accepted as a rigorous approach, its computation demanding property doesn’t not guarantee a real-time response which prohibited its application on interactive systems. Besides, some other physics-based approaches have been suggested to simulate deformable objects. Among them, mass-spring model is considered as a computationally efficient model that suits a time-critical application. Choi et al. [14] introduced a force propagation model to simulate deformations. The proposed approach regards deformation as a result of force propagation among mass nodes through connected springs and is capable to deal with both single and multiple forces. This approach makes a good balance between realism and interactivity.

Our goal is to design system which makes use of state-of-the-art techniques while designed to be a serious game specific for ultrasound-guided needle insertion training.

3 System Implementation

An overview of our system architecture is shown in Fig. 1. The four major modules are **game scenario management**, **ultrasound imaging simulation**,

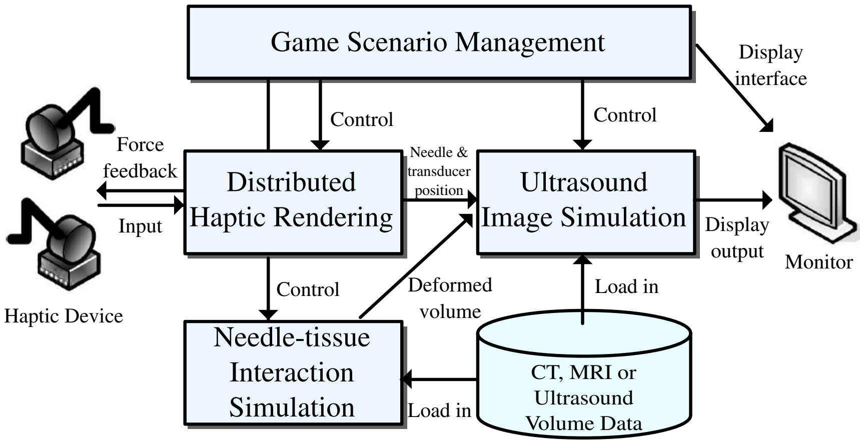


Fig. 1. System overview

distributed haptic rendering and **needle-tissue interaction simulation** modules. As mentioned before, apart from realistic simulation of surgical environment, another important issue for an attractive and effective training is the integration of game elements. These edutainment design thoughts and elements are implemented in the game scenario management module. In the section follows, we first introduce details in the game scenario management module, including game flow design, game rules design, game levels design, interface design and evaluation design. Then, many core technical issues in the other three modules will be described.

3.1 Game Design

There are two essential skills in most ultrasound-guided needle insertion procedures: (1) using a transducer to identify and locate the lesion sites (2) inserting and removing a specific needle under ultrasound guidance. Usually, these two tasks are operated by two different hands one after another. Surely, some adjustment of the transducer may be required after the needle is inserted. In the first task, a transducer is pressed against the skin and directs small pulses of high-frequency sound waves into the body to visualize internal organs and tissues in real-time. Unfortunately, the transducer can only visualize a planar region with a field of view around 60 degree. As a result, identifying and locating the targets is not an easy task, which may require sufficient anatomical knowledge and practical experience, especially when the lesions are small, overlap with other tissues or even deformed by respiration. In the second task, the trainees is required to insert a fine gauge needle into patient's body and target for specific target tumor or tissue with the help of ultrasound guidance. Usually the tactile sensation between the needle and tissues can provide helpful hints to the trainees in judging

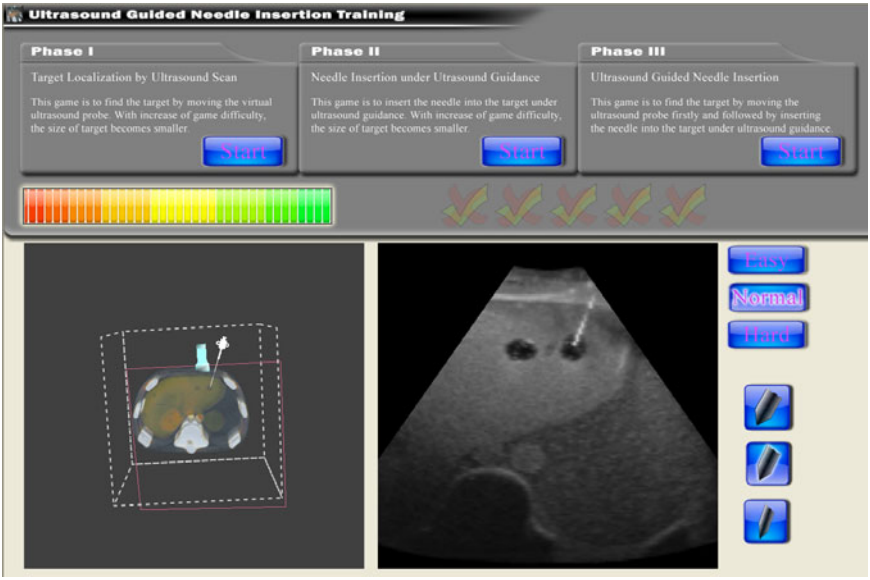


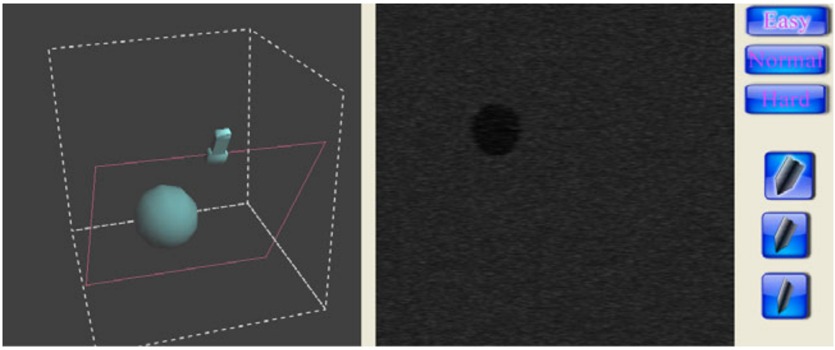
Fig. 2. The main interface of our system

the insertion depth of the needle. Intensive practices are therefore necessary to improve the hand-eye coordination for this task.

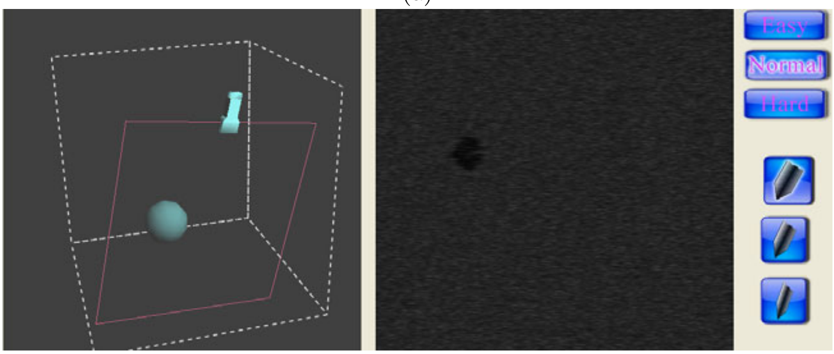
In this regard, the game management module of our system is divided into three phases, so that trainees can learn the skills progressively. The first and second phases provide game-like scenarios of seeking targets with ultrasound transducer and needle insertion, respectively. In the last phase, the trainees are asked to finish the whole ultrasound-guided needle insertion procedures with real data sets. The benefit for such a progressive learning scenario can be revealed from experimental results reported in section 4. We will demonstrate that this design helps the students learning the skills more solidly and efficiently than the traditional manner that the procedure is learnt in a whole at the beginning.

In all these three parts, a *time-attack* mode is adopted. This resembles most of the games to evaluate a player's performance to finish a specific task. A screen capture of our game user interface is shown in Fig. 2. To avoid success or failure by chance, our system requires the trainees to repeat five times for every task. We have three kinds of game difficulties, namely *easy*, *normal* and *hard*. The difficulty is defined based on the size of the target, the position of the target and the stipulated time. The virtual target is synthesized in order to make the adjustment of target size and position easily. Fig. 3 shows the ultrasound scan in different game levels of phase I. Fig. 4 displays the needle insertion under ultrasound guidance in phase II.

To provide a more intuitive user experience, two haptic devices are employed as the 3D user interface to control the transducer and the needle. The force feedback modeling of transducer is relatively simple, it will be blocked when touching



(a)



(b)

Fig. 3. Ultrasound scan in phase I, (a) target with large size in easy level, (b) target with smaller size in normal level

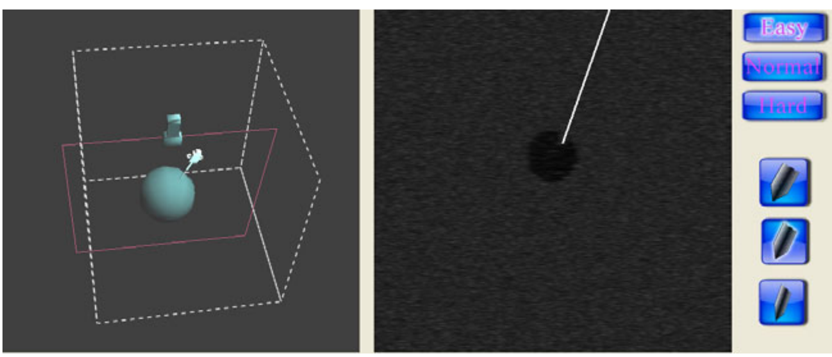


Fig. 4. Needle insertion in phase II

any object or skin. In contrast, as virtual needles with different gauges are provided, it will expert different touch sensation after inserted into the body or object. An integrated solution is implemented to finely model the pre-puncture forces, friction forces and cutting forces based on the gauges when the needle penetrate into skin, adipose tissues, muscle and internal organs [15]. A client-server based distributed visual-haptic framework is developed in order to facilitate effective load-balancing of visual and haptic rendering processes on different machines. More detail can be found in section 3.3.

At the end of each part, score obtained will be listed. The score comes from a weighted average of a series of metrics designed to evaluate the performance of the users, which will be detailed in section 4.

3.2 Ultrasound Image Simulation

Two different methods are deployed to generate 2D ultrasound images in our system in order to suit different pedagogical purposes. In the first method, ultrasound images are directly simulated from raw CT or MRI data in run-time, which enables the use of dynamic source data. However, the general appearance of these simulated ultrasound images may differ from real ultrasound images. In the second method, the ultrasound images are generated by slicing a pre-created compounded 3D ultrasound volume. While this method can guarantee the realism of simulated ultrasound image, the 3D ultrasound volume must be pre-processed. In this paper, we apply the first method in phase I and II in order to meet the requirement of changing the size and position of the targets based on different game levels in run-time, while adopting the second method in phase III for the purpose of providing realistic ultrasound images.

Fig. 5. Reconstructed ultrasound panorama

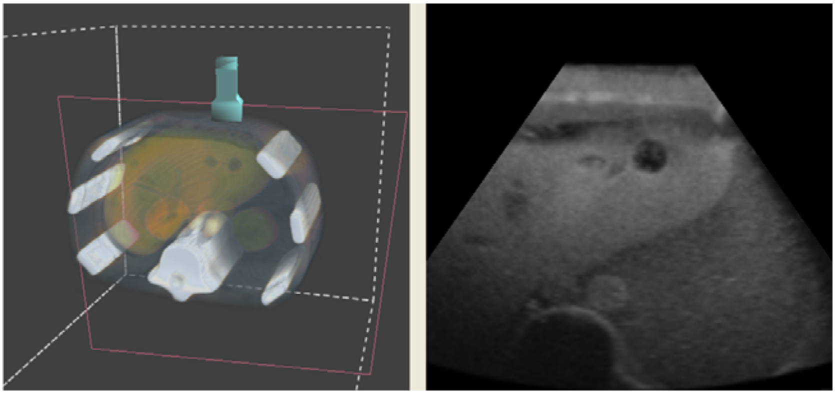


Fig. 6. Ultrasound image generated by correlating CT and ultrasound volume

In phase I and II, virtual targets are randomly generated in run-time. Corresponding artificial ultrasound imageries are generated in respond to users' interaction. Nevertheless, it is difficult to fully model all physical properties of ultrasonic wave propagations and reflections in run-time. In this regard, we adopted a fast image synthesis procedure to provide better interactivity. Recently, some studies have demonstrated that speckle pattern in ultrasound can be described as different probability distribution function (PDFs). Johan [16] further pointed out that the speckle has some similarity with a point spread function (PSF). In our system, we render the clipping plane of transducer using Gaussian envelope PSF to simulate speckle patterns. The simulated ultrasound image are shown in Fig. 3 and Fig. 4.

As for the second method, we adopt 3D SIFT algorithm [17] to register multiple ultrasound volumes and form a volumetric ultrasound panorama. Compared with traditional intensity based registration algorithms, the 3D SIFT algorithm permits the generation of panorama with high matching stability, because the 3D SIFT feature has merit of invariance to rotation and provides robust feature matching across a substantial range of addition of noise and change in illumination. Fig. 5 shows one slice of the volumetric ultrasound panorama. Then CT volume and ultrasound panorama are registered. Landmark pairs are first manually selected by experienced radiologists from the CT and US volume and they are used as the initialization of the registration process; mutual information is then used to optimize the result. After correlating the CT and Ultrasound volume, the ultrasound image is generated in run-time by slicing this pre-created 3D ultrasound panorama based on the pose of a virtual probe in the virtual environment (Fig. 6).

3.3 Distributed Haptic Architecture

In traditional visual-haptic applications, haptic rendering and graphic rendering are usually carried out in different threads to fulfill the respective refresh

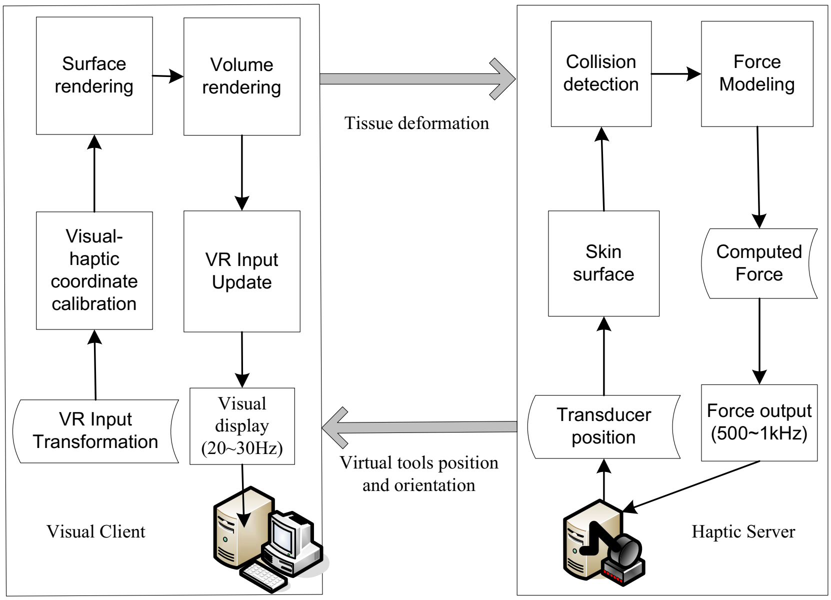


Fig. 7. Distributed visual-haptic framework of our system

requirement. Although this architecture has achieved some success in developing VR based surgical simulation system, we argue that it has several insurmountable difficulties in supporting more complex environment where multiple haptic devices are involved.

First, the great disparity in update rate between the graphic rendering and haptic rendering makes it easy to display inconsistent state. That is, more computational resources are assigned to haptic thread to fulfill its requirement of high update rate, e.g. 1000Hz. Meanwhile, the graphic thread cannot get enough resources to maintain necessary update rate for continuous display, especially when (1) the graphic models are complicated and many objects are simultaneously moving or deforming in the virtual environment or (2) the computational resources is limited. Although we can utilize high performance computers to deal with the second problem; cost efficiency, one of the most important advantages of VR based surgical simulation, will be lost. In addition, as the field of biomechanical modeling matures, the models employed in surgical simulation would be more and more complex. More computational resources are needed to render them in a realistic and fast manner. The problem should be further involved when multiple haptic devices are adopted in the applications.

In this work, we propose a distributed visual-haptic framework (Fig. 7). A client-server architecture is adopted to parallelize visual and haptic rendering in order to minimize performance overhead occurring in concurrently running haptic and visual rendering processes on a same machine. The visualization client is responsible for (1) collecting transformation from the haptic devices to get the current position and orientation of the virtual tools, (2) rendering 3D graphic models, including surface rendering and volume rendering. Simultaneously, the haptic servers can perform collision detection and haptic rendering at a relatively higher update rate. The other advantages of the proposed framework are it enabling the use of specialized computing architectures for both interfaces and achieve modularity during development.

3.4 Needle-Tissue Interaction Simulation

Besides realistic visual and haptic rendering, deformation of tissues introduced by needle advancement would help the physicians to judge the insertion path of the needle, thus affect the success of needle insertion procedures. Computer simulations of needle insertion procedures enable physicians to be exposed to both common and rare patient cases without risks to patient safety [18]. Therefore, needle placement exercise using computational simulators would greatly assist trainees acquiring surgical skills beforehand.

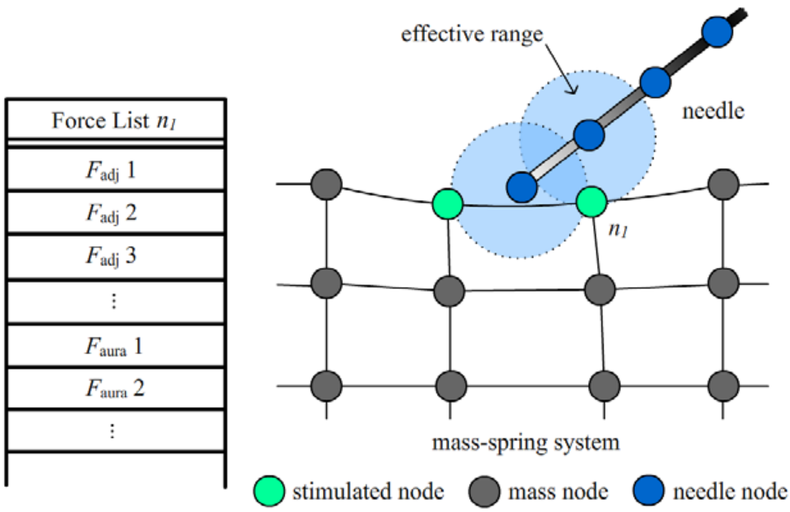


Fig. 8. Illustration of force propagation model in our mass spring system and the force list of node n_1 which records all the forces exerted by other nodes

We have adopted the force propagation model proposed by Choi et al. [14] to simulate tissue deformations. The major advantage of this approach is the low requirement of computational time which enables instant response to user actions. In our force propagation model, force exerted on *stimulated nodes* propagates locally to directly adjacent nodes. The effects experienced by these *stimulated nodes* are then further propagated as secondary forces to their adjacent nodes

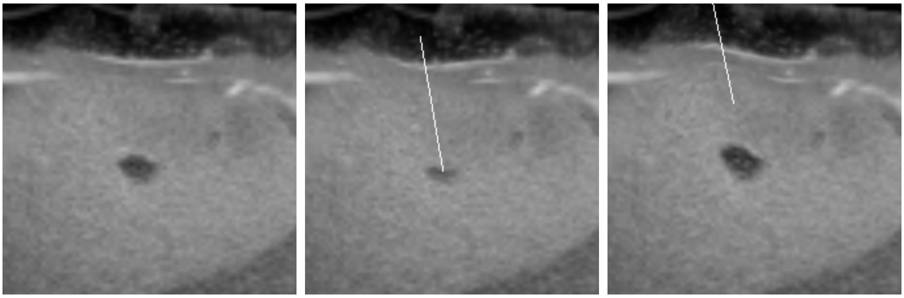


Fig. 9. Results from needle tissue interaction simulation, (a) cross section of original ultrasound volume, (b) compression occurs during needle insertion, (c) stretching occurs during needle removal

where summing up of all these effects would result in a global deformation. Each node in the system is considered as a *mass node* connected by elastic strings obeying the Voigt's model of visco-elasticity (every link comprises of a spring and a damper connected in parallel) and the dynamics of these nodes is formulated by the Newton's law of motion [14]. The mass-spring model in this work implements deformable objects as a homogeneous lattices, while a detail heterogeneous deformable model can be found in Lin et al. [19].

In our system, the needle is considered as a line strip comprising a number of *needle nodes* (See Fig. 8). Each of these *needle nodes* contains a spherical effective aura which can exert forces on *mass nodes* in the mass-spring system and labelling *mass nodes* as *stimulated nodes*. An aura would affect all *mass nodes* within its effective range. Every *mass node* in the system contains a list to store all the forces exerted by both adjacent nodes (F_{adj}) and *needle nodes* (F_{aura}). The force list is a collection of all forces acting on one particular node and would be processed after all the force list are built up. The purpose of this special design is to ensure the forces received from multiple closely located external force sources (or needle nodes in this system) are synchronized with same priority. After processing each force list, a node would somehow be perturbed and create ripples to its neighbor. This spreading effect would be terminated until a predefined penetration depth is reached. Figure 9 shows some cross section images of an ultrasound volume under needle motion.

4 Results and Discussion

A pilot user study of our system is conducted, where 10 users from medicine school and engineering school are invited to join this study. Participants from different educational backgrounds make our study more comprehensive since we can acquire comments not only on medical realism but also on pedagogical efficiency. Our system is implemented on machines with a configuration of 2.66GHz

Intel P4 Dual Core CPU, 4GB of RAM, and an Nvidia FX8800 GPU. The transducer and the needle are simulated using a SensAble PHANTOM Omni and a PHANTOM Premium 1.5 High Force/6DOF, respectively. The latter device can provide 6DOF force feedback, which is necessary to maintain the needle trajectory after puncturing the skin.

The 10 participants are equally divided into two groups (group A and group B). Group A is asked to finish all the three phrases while group B only conduct phrase III. Same evaluation matrices including completion time, success rate and position accuracy are automatically recorded for both groups. The completion time means the average time that the trainees finish their task successfully. The success rate means the portion of target is found by ultrasound scanning in phase I and the needle is inserted into the target successfully in phase II and III. The position accuracy is defined as the distance from the needle tip to the target center. To make the comparison more objective, we let the users of group B try several times of phrase III as a common surgical simulator without any game elements. Table 1 recorded the results of group A in phase I and II. The results of two groups in phase III are shown in Table 2. It is observed that the performance of group A is much better than that of group B in all matrices.This demonstrates that the proposed edutainment game can improve the training efficiency compared to traditional simulators.

Table 1. Performance of group A in phase I and II

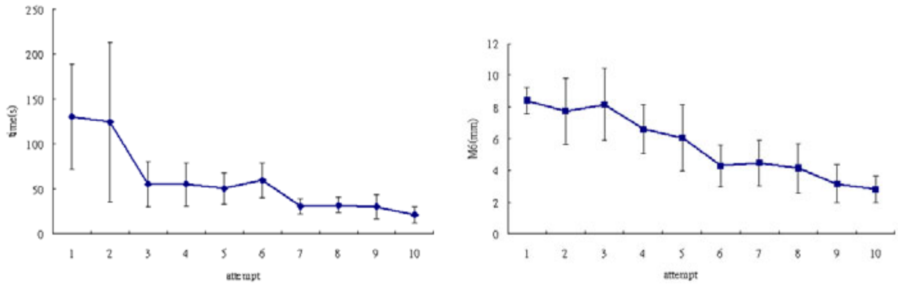
	Completion time(s)	Success rate
Phase I	11	72%
Phase II	82	61%

Table 2. Performance comparison of group A and B in phase III

	Completion time(s)	Success rate	Insertion accuracy(mm)
Group A	105	69%	7.1
Group B	130	35%	10.4

Another statical measurement, Friedman's trend test, has been carried out to examine the difference between the values from consecutive attempts. The learning curve patterns shown in Fig. 10(a) shows that the curve regarding completion time reaches plateau after the fourth attempt (Friedman's test, p=0.03). The minimal distance from needle tip to target (M6 in Fig. 10(b)) decreases gradually and reaches plateau after the sixth attempt (Friedman's test, $p = 0.02$). These results suggest that the trainees could gain significant improvement of related skills by using our virtual simulator.

Comments collected from the users show that our systematic design, friendly user interface and game-enabled interaction can increase their interest in learning ultrasound-guided needle insertion procedure. The different game levels and



(a) (b)

Fig. 10. (a) Learning curve regarding completion time of the biopsy procedure, and (b) M6 : Learning curve regarding the minimal distance from needle tip to target. In (a) and (b), the error bars indicate the standard deviation.

automatical performance reports provided are useful measurement that indicates their skill level. They are also helpful to keep track of the learning progress and to improve their performance gradually throughout the training. The new type of surgical training system integrating state-of-the-art techniques from both virtual reality and video game domains show great promise in medical education. Some users suggest that we should ensure trainees to establish a serious and correct attitude before letting them use this kind of system, since it would cause very bad consequences if they consider real surgery as a game. Based on this suggestion, we will integrate some medical code of practice into our system to enhance the responsibility of trainees.

5 Conclusion

An edutainment game for ultrasound-guided needle insertion training is proposed. Compared to traditional virtual-reality based surgical training system, our system not only provides a series of techniques to ensure the realism of the training environment, but integrates a lot of game design thoughts to enhance its usability in medical education. A pilot study demonstrates that our game-based simulation system is more effective for the trainees in acquiring necessary skills in the needle insertion process. Future investigations include modeling different needle shapes, elaborating the ultrasound image simulation algorithm, and conducting more evaluations for the effectiveness of the game elements.

Acknowledgment

This project is supported by the Research Grants Council of the Hong Kong Special Administrative Region, under General Research Fund (Project No.

CUHK4461/05M). This work is also affiliated with the Virtual Reality, Visualization and Imaging Research Centre at CUHK as well as the CUHK MoE-Microsoft Key Laboratory of Human-Centric Computing and Interface Technologies.

References

1. Blackman, S.: Serious games...and less! SIGGRAPH Comput. Graph. 39, 12–16 (2005)
2. Michel, D., Chen, S.: Serious Games: Games that educate, train and inform. In: Thomson Course Technology, Boston, MA, USA (2006)
3. Sinclair, J., Hingston, P., Masek, M., Nosaka, K.K.: Using a virtual body to aid in exergaming system development. IEEE Computer Graphics and Applications 29(2), 39–48 (2009)
4. Zielke, M.A., Evans, M.J., Dufour, F., Christopher, T.V., Donahue, J.K., Johnson, P., Jennings, E.B., Friedman, B.S., Ounekeo, P.L., Flores, R.: Serious games for immersive cultural training: Creating a living world. IEEE Comput. Graph. Appl. 29(2), 49–60 (2009)
5. Moreno-Ger, P., Blesius, C., Currier, P., Sierra, J.-L., Fernández-Manjón, B.: Online learning and clinical procedures: Rapid development and effective deployment of game-like interactive simulations. In: Pan, Z., Cheok, D.A.D., Müller, W., El Rhalibi, A. (eds.) Transactions on Edutainment I. LNCS, vol. 5080, pp. 288–304. Springer, Heidelberg (2008)
6. Jung, G.S., Kim, S.Y., Jung, S.K., Byun, S.D., Lee, Y.S.: Timed automata-based rehabilitation training game design for the affected lower extremity of hemiparetic patient. Transactions on Edutainment I, 17–27 (2008)
7. Kizony, R., Weiss, P.L., Shahar, M., Rand, D.: Theragame c a home based virtual reality rehabilitation system. In: Proc. of 6th Intl. Conf. Disability, Virtual Reality and Assoc. Tech., pp. 265–269 (2006)
8. Narayanasamy, V., Wong, K.W., Fung, C.C., Rai, S.: Distinguishing games and simulation games from simulators. Comput. Entertain 4(2), 9 (2006)
9. Vidal, F.P., Chalmers, N., Gould, D.A., Healey, A.E., John, N.W.: Developing a needle guidance virtual environment with patient specific data and force feedback. In: Proc. of Computer Assisted Radiology and Surgery, pp. 418–423 (2005)
10. Forest, C., Comas, O., Vaysière, C., Soler, L., Marescaux, J.: Ultrasound and needle insertion simulators built on real patient-based data. Stud. Health Technol. Inform. 125, 136–139 (2007)
11. Hostettler, A., Forest, C., Forgione, A., Soler, L., Marescaux, J.: Real-time ultrasonography simulator based on 3d ct-scan images. Stud. Health Technol. Inform. 111, 191–193 (2005)
12. Magee, D., Zhu, Y., Ratnalingam, R., Gardner, P., Kessel, D.: An augmented reality simulator for ultrasound guided needle placement training. J. of Med. Bio. Eng. and Comput. 45(10), 957–967 (2007)
13. Zhu, Y., Magee, D., Ratnalingam, R., Kessel, D.: A training system for ultrasound-guided needle insertion procedures. In: Ayache, N., Ourselin, S., Maeder, A. (eds.) MICCAI 2007, Part I. LNCS, vol. 4791, pp. 566–574. Springer, Heidelberg (2007)
14. Choi, K.S., Sun, H., Heng, P.A.: Interactive deformation of soft tissues with haptic feedback for medical learning. IEEE Transactions on Information Technology in Biomedicine 7(4), 358–363 (2003)

15. Chan, W.Y., Qin, J., Chui, Y.P., Yu, S.C.H., Ho, S.S.M., Heng, P.A.: A six degree-of-freedom (dof) haptic model for percutaneous needle insertion training. In: The 94th Scientific Assembly and Annual Meeting of the RSNA 2008 (2008)
16. Johan, M.T.: Speckle Formation, Analysis and Processing Applied to Ultrasound Tissue Characterization. Physics for Medical Imaging Applications 240(3), 151–176 (2007)
17. Ni, D., Qu, Y., Yang, X., Chui, Y.P., Wong, T.T., Ho, S.S.M., Heng, P.A.: Volumetric ultrasound panorama based on 3d sift. In: Metaxas, D., Axel, L., Fichtinger, G., Székely, G. (eds.) MICCAI 2008, Part II. LNCS, vol. 5242, pp. 52–60. Springer, Heidelberg (2008)
18. Chentanez, N., Alterovitz, R., Ritchie, D., Cho, L., Hauser, K.K., Goldberg, K., Shewchuk, J.R., O'Brien, J.F.: Interactive simulation of surgical needle insertion and steering. In: Proceedings of ACM SIGGRAPH 2009 (August 2009)
19. Lin, S., Lee, Y.S., Narayan, R.J.: Heterogeneous material modelling and virtual prototyping with 5-dof haptic force feedback for product development. International Journal of Mechatronics and Manufacturing Systems 1(1), 43–67 (2008)

Sketch-Based 3D Face Modeling for Virtual Character

Wei Jiang and Zhengxing Sun

State Key Lab for Novel Software Technology, Nanjing University,
Nanjing 210093, P.R. China
szx@nju.edu.cn

Abstract. It is very important for virtual character designers to generate 3D face models interactively and creatively in real-time. This paper investigates the use of sketch-based interface in 3D face modeling and focuses on how to map 2D sketchy features onto a 3D model. A mapping mechanism on the basis of contour lines is proposed. Representative points of the input are extracted, which retain the global and local shape originalities. Then the outer and inner contours are extracted automatically from a given 3D face model (as a template model), and a set of feature points are determined according to the 2D representative points. The displacements between the 2D representatives and the 3D feature points are calculated and used as the control parameters of model deformation. A novel face model is finally created by means of deforming the given model. The experiments show the efficiency and effectiveness of our approaches.

1 Introduction

To generate novel face models is an important task of 3D virtual character designers. Since the pioneering work of Parke in 1972 [1], 3D face modeling and animation has been a hot topic of computer graphics with many applications [2]. Existed facial shape modeling techniques can be categorized into two classes according to their data sources. One is to reconstruct face models from 3D data which are captured from 3D devices such as 3D laser scanners [3]. The other is to construct models from 2D data, such as face images [4] and video clips [5]. However, the facial data used in these techniques are usually captured from real faces. Therefore, only standard (normal) face models can be created while virtual characters usually have novel faces. Moreover, the user interfaces of some existed modeling tools such as Maya [6], which are on the basis of WIMP (Window, Icon, Menu and Pointing device) style, are not fit for creative design because users need to switch from one command mode to another frequently and their thinking processes will be interfered. Therefore, the existed techniques and tools are not fit for novel 3D face models, especially virtual characters.

Pen/paper is a traditional tool of 3D model designers. Nowadays, HCI (human computer interaction) researchers not only focus on sketch-based interface in 2D applications such as diagramming, but also pay more and more attention to

Z. Pan et al. (Eds.): Transactions on Edutainment IV, LNCS 6250, pp. 215–228, 2010.

3D applications such as model creation and modification. Sketch-based interface for modeling (SBIM) becomes a new research highlight in computer graphics [7]. Using sketch-based interfaces in modeling processes, artists can express their ideas more freely. Furthermore, if the tools can construct 3D models directly and automatically from sketch documents, they will free the artists from the burden of manual transformation of sketch documents into 3D models. However, while we try to use sketch-based interface in 3D face modeling for virtual character, two challenges arise. First, sketch-based interaction is discrete, non-accurate and ambiguous. The draft need to be filtered and parsed in order to interpret the author's intention. Second, a sketchy face is a kind of planar representation of human faces and does not contain enough depth information. How to map the 2D input onto the final 3D output and retain the shape originalities is the most import issue to be considered.

This paper works on these issues and proposes a novel shape mapping mechanism on the basis of face contours. Given a 2D facial sketch and a 3D face model (as template model), both 2D and 3D contours (outer and inner) are extracted at first. We have defined a set of points from the 2D contours as representative points, which retain both the global and local shape originalities of the sketchy face. Then we calculate the same amount of points from the 3D contours and obtain a 1-to-1 mapping between the 2D representatives and the 3D feature points. Finally, a deformation method is used to generate the 3D result. We have implemented a prototype and performed a few experiments. The generated models can retain the shape originalities from the input. Compared to the existed face modeling methods, our approach can create novel 3D face models quickly and fulfill the requirements of 3D character face design.

The rest of the paper is organized as follows. In section 2, we will address some related works. Section 3 will give an overview of our approaches. Section 4 will discuss how to process the original facial sketches and construct sketchy faces. The detailed shape mapping mechanism is proposed in Section 5, including calculation of 3D contours, extraction of 2D representatives, shape mapping, and model deformation. Section 6 will introduce the experiments and discusses the results. Finally, we will make a conclusion in Section 7.

2 Related Works

Freehand sketch is also a kind of 2D data input. Current image (or video clip) based 3D face modeling aims at 3D face reconstruction, i.e., the final face model is on the basis of a real person's face. Most of the existed modeling methods usually try to locate a few feature points in face images, which will be mapped to some fixed feature points from a 3D template model. Generally, these 3D feature points are specific landmarks of human faces (such as mouth corners). Because the image (or video) input is captured from a real person, only a few feature points are sufficient for reconstructing the 3D result. However, artists draw sketches to design novel faces that may be exaggerated and even abnormal, for example, a square face (global shape) or a huge and round nose (local shape).

Only a few fixed feature points are not capable of representing all the shape originalities of a sketchy face. A few works try to use standard face libraries to guide the modeling process [8]. However, it is impossible to construct a shape library for all the 3D characters that usually are abnormal in shape.

There have been a few researches which involve sketch-based interfaces in 3D face modeling and 3D character design. The modeling algorithms can be categorized into two classes. The first recognizes the inputs according to some pre-defined templates and then generated 3D models. Nataneli [9] has proposed a sketch-based expression generation system which recognizes the semantic category of the input and then generates expressions in a 3D model according to some templates. It is not fit for 3D face shape design. Gunnarsson [10] has introduced a sketch-based arbitrary 3D face modeling system. It uses a stand face library from which a statistical face model is obtained. The system interprets the input and find out a linear combination of faces from the library. However, library-based methods are not fit for 3D character design, because the designer's creative intention can not be predicted and described by the linear combination of normal faces. The second does not use templates and generates the model directly. Teddy [11] is a typical example, which considers the sketchy input as contours of 3D objects. The system extracts the inner skeletons of the contours and uses an inflation algorithm to construct plausible 3D models. However, it is not fit for modeling objects with complex surface such as character faces.

Furthermore, Chang [12] uses freehand strokes to design 3D expressions. The user selects an area to be modified in the 3D model, and then draws a stroke to indicate the final position of the selected area. The system uses an optimization algorithm to deform the model, which is very time consuming. Moreover, the input stroke should be line segment and only the inner of the face model can be deformed.

To sum up, the existed methods do not fulfill all the requirements of face modeling for 3D character design. Considered that it is very hard to construct a face model directly from a 2D input, we decide to use a deformation-based approach to construct the final 3D model. We use triangle mesh as the representation of 3D models, because it is supported by most of current 3D modeling tools. During a 3D face design session, the designer selects a face mesh according to his design intention at first, and he can also export the generated face mesh which can be opened and modified in other 3D modeling tools.

3 Approach Overview

The key of sketch-based modeling is to explore and transfer the shape originalities of the sketch input into the generated 3D model. However, because the sketch is in 2D space and the model is in 3D space, there is not a direct geometric relationship between them. We have to find out a proper data level in which they can be analyzed and mapped onto each other. A sketch document, which is drawn to describe a face, is composed of strokes which represent different contours of the outer face and the inner organs. If we find out the corresponding contour lines in the 3D model, we can easily build up the geometric relationship. The mapping of two lines can be achieved by point-based method, i.e., extracting and mapping points from both the 2D stroke contour and the 3D model contour. Then we can use a control point based deformation method to generate the final face model.

We use *facial sketch* and *sketchy face* to distinguish the different states of a sketch document in a modeling process. A facial sketch means an original sketch document (without processing) drawn by a face designer, while a sketchy face means a structured sketch document of which the strokes of different facial organs are separated from the original document and filtered. As shown in Fig. 1, the detailed workflow contains three stages as follows:

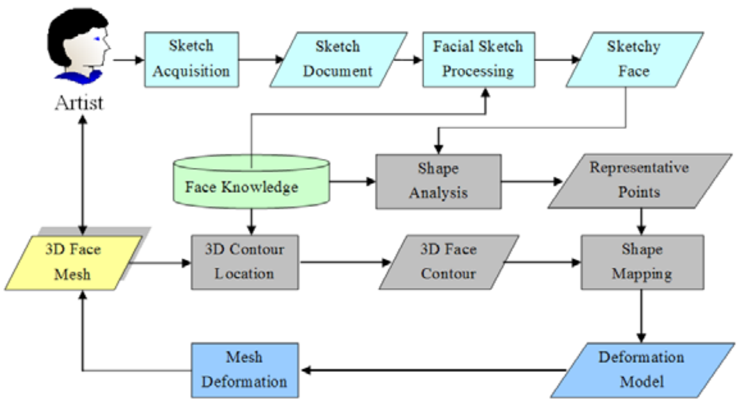


Fig. 1. The workflow of 3D face modeling from freehand sketch contains three stages

(a) Facial sketch processing: This stage focuses on the capture of facial sketch documents and the generation of sketchy faces. The sketch acquisition module provides users a sketch-based interface (with freehand input, gesture command, etc.), and captures the input as a facial sketch document. Then the face sketch processing module will filter the strokes and interpret the input as a sketchy face. Some prior knowledge of human faces, such as the spatial relationships of face organs will be used in the facial sketch processing module.

(b) Shape mapping: This stage aims at exploring the best mapping between a sketchy face and the final 3D face model. Strokes-to-organ correspondence has been calculated in a sketchy face in the facial sketch processing stage. The shape analysis module will analyze the contour strokes and extract a set of representative points. Inner and outer contours of the given 3D face model are also calculated. Then the shape mapping module calculate the corresponding feature points in the 3D face contour model and obtains the mesh deformation model (deformation parameters).

(c) Model deformation: This stage deals with the model generation in which the mesh deformation module deforms the given model according to the deformation parameters.

4 Sketchy Face Construction

When a person walks towards us, our visual system will perceive the outer contour of that person's face at first. As the person gets closer to us, we will perceive his face organs (approximate shape and spatial relationships). Then the organ contours become more clear, and finally we will discern the details of his face. While sketching a human face, we usually use a similar process, i.e., a global to local and coarse to fine drawing style.A typical facial sketch document is composed of a set of strokes, which are drawn to represent the outer contour of the face and the inner contours of face organs. Then a sketchy face can be seen a domain-dependent sketch document which is filtered and parsed from the input (facial sketch).

A sketchy face is composed of an Outer-Contour Component (corresponds to the outer face contour) and several Inner-Contour Components (correspond to different face organs). The Outer-Contour Component and the Inner-Contour Component are both composed of strokes. A Stroke is composed of a set of sample points captured by the input device between pen-down and pen-up events. A Sample Point is parameterized by values of Coordinates, Time Stamp, Pressure and Lean of Pen.

The original sketch document usually contains some noisy or even erroneous points. Sezgin et al. [13] indicate that there exist two causes: one is user error such as unproficient drawing skills; the other is device error which arises in the digitization of the user's input. As a result, the sketch input may not be the perfect representation of the user's intention and needs to be filtered. In our approaches, sketch filtering deals not only with the redundant sample points, but also with some domain-dependent circumstances. For example, Fig. 2 shows two strokes that represent the outer contour of a human face respectively. In general, the outer contour is an occluded shape which is similar to an eclipse. However, the stroke in Fig. 2(a) is not occluded and the one in Fig. 2(b) has an intersection which should be processed. We utilize the sketch processing algorithms from MagicSketch [14], which is a platform and algorithm library for implementing sketch-based systems and has been developed in our previous work.

The system also needs to determine which part of the face the stroke (input) belongs to. We have made an assumption that all the spatial relationships of facial organs are normal, i.e., the facial organs are displaced from top to bottom according to human anthropometry. We also assume that artists usually sketch the outer contours at first. Then we can group other strokes into different inner-contour components according to the spatial relationships between strokes and the facial prior knowledge. Finally, we will obtain a parsed sketchy face.

Fig. 3 demonstrates an example of a parsed sketchy face: B_1 is the outer contour; B_2s are the eyebrow contours; B_3s are the eye contours; B_4 is the nose contour; B_5 is the mouth contour; B_6s are the ear contours.

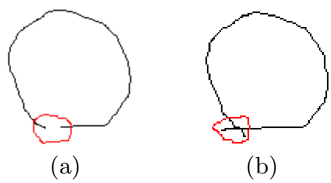


Fig. 2. Sketchy data filtering. (a) The stroke should be occluded. (b) There should not be an intersection.

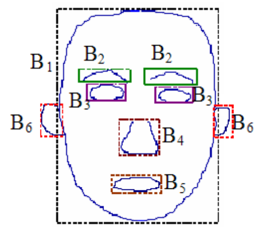


Fig. 3. An example of a parsed sketchy face

5 Shape Mapping and Model Generation

As discussed in the previous sections, we have refined and converted the problem of 2D/3D face shape mapping into the mapping of 2D/3D contours. However, two new problems arise as follows.

One is how to find out the contours from a given 3D face model (mesh). Existed face modeling approaches which are on the basis of model deformation usually use a kind of manual (or semi-manual) procedure to locate a few landmarks on the given 3D face model. It is a very time-consuming process, which is also very hard for non-professional users. Moreover, the tedious procedure will be repeated while new face models are used. Therefore, an automatic method is highly needed. We have integrated the knowledge of facial anthropometry, 3D model viewing and differential geometry, and proposed two algorithms to locate the face contours automatically.

The other problem is how to transfer the shape information obtained from the input into the generated model. Dislike other 2D-based face modeling methods, we have proposed a reverse process that the system extracts the feature points from the input at first and then locates a set of corresponding feature points in the 3D model.

5.1 3D Face Contours

An intuitive method to extract face contours from a 3D face mesh is to use range images. Then we can use image processing techniques such as edge detection to locate organs in the face model. However, the performance depends highly on the mesh quality, and the result is hard to be mapped to the discrete vertices.

Consider a face model and a view point in a 3D viewing environment. If the normal of a vertex is perpendicular to the current view direction (the dot product of the normal and the view vector is zero), the vertex may be a point in a contour line. Then, we can find the outer contour of the face mesh.

Assume that the 3D face mesh is adjusted in advance, i.e., the front faces the positive Z direction and the top faces the positive Y direction in an OpenGL environment. The 3D bounding box is calculated and the center is moved to the origin. Then we map Y-coordinates to interval [-1, 1], and recalculate X-coordinates and Z-coordinates according to the same scaling. A more detailed description is given below.

Algorithm 1. Outer contour extraction from a 3D face mesh
Input: A 3D face mesh
Output: The outer contour line
Description:

1 Compute each vertex's normal vector (using the mean vector of facet normals).
2 Set up the view point in the positive Z-axis.
3 At each vertex, compute the dot production of the normal vector and the view vector.
4 For each facet of the mesh:
  4.1 If the facet is not visible in the OpenGL environment, continue the iteration and process the next facet;
  4.2 If the dot products of the three vertices have different signs (i.e., two positives and one negative, or two negatives and one positive), there will be two zero crossings in the two edges of which the vertices have different signs. A linear interpolation method is used to calculate the zero crossings that will be added to the candidate point set.
5 Find out the outer contour from the candidate point set:
  5.1 Connect two candidate points if they are in a same facet;
  5.2 There will be several lines. Find out two lines which have the largest point amounts. One is the upper part and the other is the lower part of the outer contour.
6 The end.

If there are ears in the given face mesh, the obtained outer contour may contain the points of the ears. These points can be processed after the ear contours have been extracted. Fig. 4 demonstrates an example of outer-contour extraction. The coordinates of the view point are $(0, 0, 10)$. Fig. 4(a) shows the given 3D face mesh model; the face mesh is rendered with lighting in Fig. 4(b); Fig. 4(c) shows the extraction results.

(a) (b) (c)

Fig. 4. Outer-contour of a 3D face mesh. (a)The original face mesh. (b) The original face model rendered with lighting. (c) The outer-contour is extracted.

Then we will extract the inner contours. Gaussian curvature is an effective tool to analyze and describe shape information of surfaces. The value shows how much a surface bends and in which direction at a single point on the surface. A positive Gaussian curvature indicates that the local area is convex, and a negative Gaussian curvature indicates that the local area is concave. Note that each organ contour contains a lot of convex points (with high positive Gaussian curvatures), such as the mouth tips, we can locate the face organs by clustering the points of which the curvature values are higher than a threshold (determined by mesh size and can be set interactively). The detailed description of the algorithm is as follows.

Algorithm 2. Inner contours extraction from a 3D face mesh
Input: A 3D face mesh
Output: Inner contours
Description:

1 Compute each vertex's Gaussian curvature (product of two principle curvatures).
2 Set the threshold interactively and display each vertex whose Gaussian curvature is higher than the threshold.
3 According to facial anthropometry and the nose tip position (vertex with highest Z-coordinate), estimate the center points (X and Y coordinates) of the eyes, the ears and the mouth.
4 Separate the points into six clusters (two eye clusters, two ear clusters, a nose cluster and a mouth cluster):
   4.1 For each vertex, compute the distances between itself and the estimated centers (in XOY plane), and add it to the nearest cluster;
   4.2 For each cluster, calculate its centroid as new center point, and remove a few exceptional points (too far from the center);
   4.3 Re-cluster the remaining points according to step 4.1 and remove the exceptional points according to facial anthropometry (such as ratios of mouth size to the outer contour size).
5 Compute each point cluster's convex hull.
6 The end.

Fig. 5 shows an example of inner-contour extraction: Fig. 5(a) displays the points with high Gaussian curvature value; Fig. 5(b) shows the result convex hulls. Note

(a) (b)

Fig. 5. An example of inner-contour extraction. (a) Points with high Gaussian curvatures. (b) The results of inner-contour extraction.

that the convex hull of a point cluster is composed of discrete points; we need to connect neighbor points or uses a parabolic interpolation method to construct the final contour lines.

5.2 2D Representative Points

The original sample points of a contour stroke usually contain too many points. They are usually non-uniform and probably bring complexities and difficulties in a modeling process. We propose a method to extract a set of representative points from the strokes. There are two criteria in extracting the representatives: 1) the local detail of the shape can be retained; 2) the points are uniform in distribution.

Fig. 6 demonstrates how to extract the 2D representatives. An intuitive extraction is to divide the contour stroke into multiple segments with equal curve length. This method can achieve a high performance if the stroke is smooth, but it will lead to a disappointing result while be performed on a non-smooth sketch such as the one given in Fig. 6(a). Fig. 6(b) shows the result which is obtained by the above method. The red points are extracted points on the outer contour. As we can see in the figure, these points can not fulfill the first criterion, however. For example, point p (on the left) is very special and should be included in the point set.

We consider a point (like point p) which contains more shape information as a feature point of a stroke. The representative points must include all the feature points. In fact, a stroke segment between two neighbor feature points is very smooth and can be seen as a line segment. Therefore, we can use the polygonal approximation algorithms to find out the feature points. We apply the algorithm proposed by Sklansky et al. [15] to achieve the polygonal approximation. The algorithm is controlled by a threshold ε. As ε gets larger, there will be fewer vertices, and the shape accuracy is reduced. The threshold is also can be modified interactively in our approaches. Fig. 6(c) shows the extraction result with threshold $\varepsilon = 1$. Fig. 6(d) shows the extraction result with threshold $\varepsilon = 3$, of which the amount of points is smaller and the accuracy is reduced.

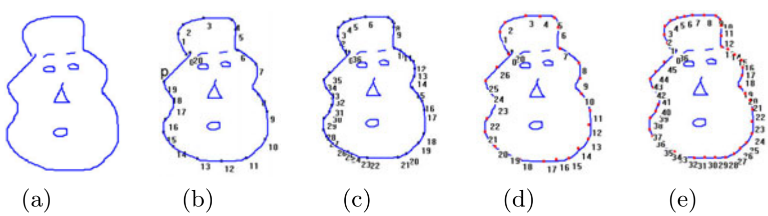


Fig. 6. Extraction of 2D representatives. (a)The original face sketch. (b)Points extracted with equal length. (c)Set $\varepsilon = 1$ and extract the representatives. (d)Set $\varepsilon = 3$ and extract the representatives. (e)Extract the representatives with $\varepsilon = 1$ and interpolate with new points.

However, only the extracted feature points are not sufficient. For example, considering the stroke segment between feature point 21 and 22 (in the bottom of the face), there is no other feature points on it. Therefore, the result points do not fulfill the second criterion. We need to add more points to the result. Let $MeanLength$ be the mean length of all the stroke segments, and $CurrentLength$ be the length of current segment. For each stroke segment which is longer than a threshold $1.5 * MeanLength$, we divided it into N pieces and add N-1 new points ($N = [CurrentLength/MeanLength]+1$). Fig. 6(e) shows the final result of representative points.

5.3 Shape Mapping

After we have obtained the 2D representative points, we can extract the same amount of feature points from the 3D contours. The 3D feature points are interpolated according to the ratios of each stroke segment's length to the entire 2D contour's length.

The top and bottom points which are located from both 2D and 3D outer contours will divide a face into two parts: one is the left part and the other is the right part. For mouth and eye contours, we try to find out the most left and the most right points to separate each contour into two parts: the upper part and the lower part. Because the nose contour is not occluded, we separate it into three parts: the left part, the right part and the lower part. Then we calculate the 3D feature points by interpolation.

Let $FP_{2D} = \{fp_{2d1}, fp_{2d2}, \cdots, fp_{2dN}\}$ be the set of representative points of the right 2D outer contour. Point fp_{2d1} is the top and point fp_{2dN} is the bottom. If the feature point set of the right 3D outer contour is $FP_{3D} = \{fp_{3d1}, fp_{3d2}, \cdots, fp_{3dN}\}$, in which point fp_{3d1} is the top and point fp_{3dN} is the bottom. Let $Length_{2d}$ be the length of the right 2D outer contour and $Length_{3d}$ be the length of the right 3D outer contour, then we can obtain the following equation:

$$\frac{Dis_{2d}(fp_{2d(i-1)}, fp_{2d_i})}{Length_{2d}} = \frac{Dis_{3d}(fp_{3d(i-1)}, fp_{3d_i})}{Length_{3d}} \qquad i = 1, \cdots, N \qquad (1)$$

where $Dis_{2d}(p_1, p_2)$ and $Dis_{3d}(p_1, p_2)$ are functions for calculating the Euclidean distance between two given points in 2D and 3D space.

Then we can use the same approach to process other contour parts. Finally, we can locate all the 3D feature points that construct the deformation model (including feature points and a displacement vector).

5.4 Model Deformation

We assume that users always draw the outer contour at first while designing a face. After the outer contour is drawn, our approach will deform the mesh according to the aspect ratio at first (we call it global deformation). As soon as a new stroke is captured, the system will filter it and determine its category (belongs to which face organ). The representative points of the new stroke are extracted and added to current point set. New 3D feature points are also calculated and added to the deformation model. The 3D feature points can be used as control points, and then a control-point based deformation mechanism is performed to construct the intermediate face model, and the user can evaluate his idea immediately. As more strokes are drawn, the final novel 3D face model will be obtained.

Given a deformation model, the mesh deformation can be seen a function approximation problem. Radial basis functions (RBF) are typically used to build up function approximations of this form.

Let $P = \{p_1, p_2, \cdots, p_M\}$ be the original vertices of the given mesh, and $C = \{cp_1, cp_2, \cdots, cp_N\}$ be the set of control points (3D feature points), and $Displacement = \{dis_1, dis_2, \cdots, dis_N\}$ be the displacement vector of all the control points, where p_i, cp_i, and dis_i are all 3D vectors.

Then a set of equations are generated according to the following equations:

$$f(x) = \sum_{i=1}^{N} w_i \Phi(\|x - x_i\|) + px + q \tag{2}$$

$$\sum_{i=1}^{N} w_i = 0 \tag{3}$$

$$\sum_{i=1}^{N} w_i cp_i^T = 0 \tag{4}$$

The approximating function $f(x)$ is represented as a sum of N radial basis functions ($\Phi(r) = e^{-\frac{r}{64}}$), each associated with a different center x_i, and weighted by an appropriate coefficient w_i. After the solution of the equations has been calculated, the displacement of each vertex in P can be easily obtained. As a result, the 3D face model is deformed.

6 Experiments and Analysis

We have implemented a sketch-based face modeling prototype according to the proposed approaches and performed a few experiments. Currently, we do not deform the eyebrows in the 3D model because they are very difficult to be located in an arbitrary face mesh and are not as important as other organs in face modeling. Furthermore, because the ears are hard to be drawn precisely in the front view, we do not deform the ears in these primary experiments, neither.

Fig. 7 shows a typical face model design process in which a user drew a facial sketch step by step. In each column of the figure, the top was the original sketch, and the middle was the deformed mesh model, and the bottom was the 3D model with lighting. As we can see from the figure, the generated face models are similar to the sketches.

We have also recorded the time cost of our approaches. The experiments have been performed on a Pentium 4 (3.0 GHz) PC with 2048Mb RAM. Given

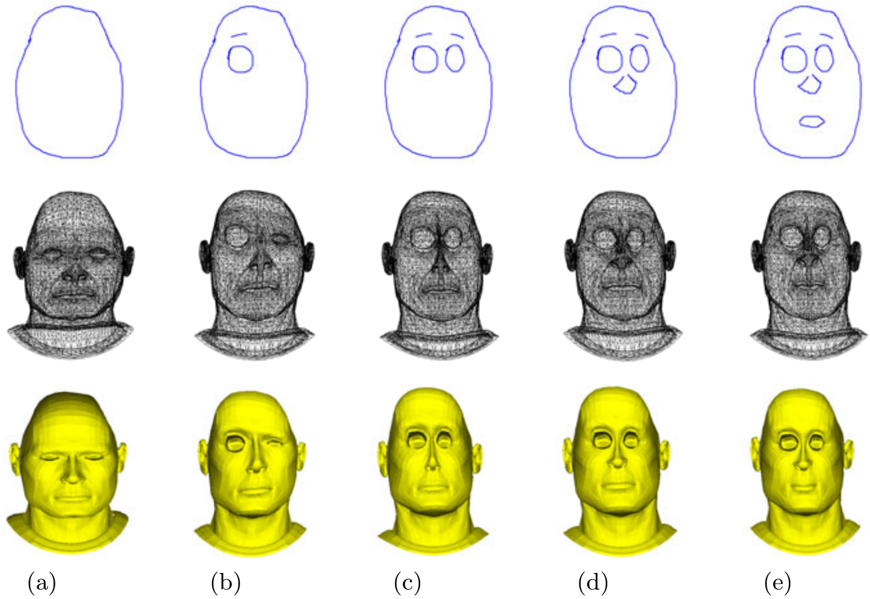


Fig. 7. A typical face model design process. (a)Outer face has been drawn. (b)Left eye has been drawn. (c)Right eye has been drawn. (d)Nose has been drawn. (e)Mouth has been drawn.

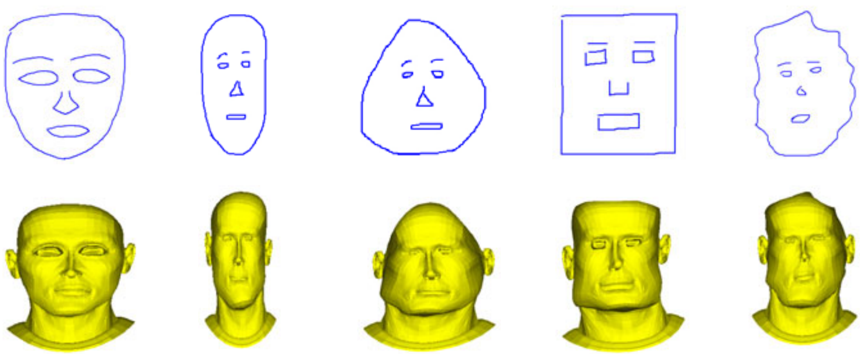
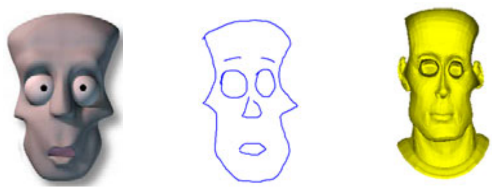


Fig. 8. A few novel samples and the result models

a template mesh model which has 10,000 facets, the mean time cost of mesh deformation is about 300 ms. Therefore, the proposed approaches fulfill the real-time requirement of creative design.

The primary objective of this study is to develop tools for designing face models for novel 3D characters by freehand sketch. Therefore, a few novel sketchy faces are collected, and the corresponding 3D face models are generated for evaluation. Fig. 8 demonstrates a few novel samples and the result models. As we can see in the figures, the generated models retain the shape originality in the sketchy input.

We have also gathered a few existed 3D character samples and imitated them by freehand sketch. Fig. 9 shows an example. Fig. 9(a) is the head model of a 3D character gathered from the Internet [16]; Fig. 9(b) is an imitated facial sketch drawn by a non-skillful student with ten minutes; Fig. 9(c) is the face model generated by our prototype system. By using the proposed sketch-based face modeling system, even a non-skillful user can create a novel face model quickly and evaluate it immediately.



(a) (b) (c)

Fig. 9. An example of face model imitation. (a) A model image is collected from the Internet. (b) A face sketch is drawn by a user. (c) A similar model is generated quickly.

7 Conclusion

This paper studies the utilization of freehand sketch in designing novel face models for 3D characters. The challenge is how to map a 2D sketchy face onto a 3D face model and retain the shape originalities. Accordingly, we have proposed a mapping mechanism on the basis of facial contours. Our proposed approaches have three characteristics. Firstly, we have defined and extracted a set of representative points from the sketchy input, which can retain both the global and the local shape originalities. Secondly, 3D face contours can be extracted automatically from a given 3D face model, and a set of 3D feature points are determined corresponding to the 2D representatives. Finally, the displacements between the 2D representatives and 3D feature points are calculated and used as the control parameters of model deformation. Compared to other 2D-based face modeling methods, our method is quite different and can be seen a reverse process, i.e., the 3D feature points are determined by the 2D representatives of a sketchy input. The experiments show that the proposed approaches can transfer the shape originalities into the constructed face model. Our approaches can create novel 3D face models quickly and fulfill the requirements of face model design. The approaches also can be used in modeling 3D objects (or characters) from which the main contours can be extracted. In the future, we plan to add side-view sketches to provide more depth information. Moreover, other drawing clues such as wrinkles should be considered to add more details to the sketchy face input.

Acknowledgments. This work was supported by the National Natural Science Foundation of China under Grant Nos. 69903006, 60373065, 60721002; the National High-Tech Research and Development Plan of China under Grant No. 2007AA01Z334; the Program for New Century Excellent Talents in University of China under Grant No. NCET-04-04605.

References

1. Parke, F.: Computer generated animation of faces. In: Proc. ACM National Conference, pp. 451–457. ACM, New York (1972)
2. Deng, Z., Noh, J.Y.: Computer facial animation: A survey. Data-Driven 3D Facial Animations, pp. 1–28. Springer, Heidelberg (2007)
3. Zha, H.B., Wang, P.: Realistic face modeling by registration of 3D mesh models and multi-view color images. In: Proc. of the 8th Int'l. Conf. Computer Aided Design and Computer Graphics, pp. 217–222. Welfare Printing Limited, Macao (2003)
4. Guenter, B., Grimm, C., Wood, D., Malvar, H., Pighin, F.: Making Faces. In: Proceedings of ACM SIGGRAPH 1998, pp. 55–66. ACM, New York (1998)
5. Liu, Z., Zhang, Z., Jacobs, C., Cohen, M.: Rapid Modeling of Animated Faces From Video. Technical Report, Microsoft Research (1999)
6. Maya: Autodesk Inc., http://www.autodesk.com/maya
7. Olsen, L., Samavati, F.F., Costa Sousa, M., Jorge, J.: Sketch-based modeling: A survey. Computer & Graphics 33(1), 85–103 (2009)
8. Blanz, V., Vetter, T.: Face recognition based on fitting a 3D morphable model. IEEE Transactions on Pattern Analysis and Machine Intelligence 25(9), 1063–1074 (2003)
9. Nataneli, G., Faloutsos, P.: Sketching Facial Expressions. In: ACM SIGGRAPH 2007 sketches, article No. 60 (2007)
10. Gunnarsson, O., Maddock, S.: A Statistically-Assisted Sketch-Based Interface for Creating Arbitrary 3-dimensional Faces. In: Proc. Theory and Practice of Computer Graphics 2007 (2007)
11. Igarashi, T., Matsuoka, S., Tanaka, H.: Teddy: a sketching interface for 3d freeform design. In: Proceedings of the SIGGRAPH 1999, pp. 409–416. ACM, New York (1999)
12. Chang, E., Jenkins, O.C.: Sketching articulation and pose for facial animation. In: Data-Driven 3D Facial Animation, pp. 132–144. Springer, Heidelberg (2007)
13. Sezgin, T.M., Davis, R.: Scale-space based feature point detection for digital ink. In: Making pen-based interaction intelligent and natural, AAAI fall symposium (2004)
14. Sun, Z., Liu, J.: Informal user interfaces for graphical computing. In: Tao, J., Tan, T., Picard, R.W. (eds.) ACII 2005. LNCS, vol. 3784, pp. 675–682. Springer, Heidelberg (2005)
15. Sklansky, J., Gonzalez, V.: Fast Polygonal Approximation of Digitized Curves. Pattern Recognition 12, 327–331 (1980)
16. http://www.kevman3d.com/images.asp?section=3dwip&id=3

A Framework for Virtual Hand Haptic Interaction

Xiaoxia Han[1] and Huagen Wan[2,*]

[1] College of Information Science and Electric Engineering,
Zhejiang University, Hangzhou 310027, China
[2] State Key Lab of CAD&CG, Zhejiang University, Hangzhou 310027, China
hgwan@cad.zju.edu.cn

Abstract. Virtual hand interactions play key roles in virtual environments. The recent addition of force feedback to virtual reality simulations has enhanced their realism, especially when dexterous manipulation of virtual objects is concerned. However, though there exist many works on virtual hand modeling regarding computer animation and interaction applications, much less attention is paid on haptic modeling of virtual hand. In this paper, we propose a framework for virtual hand haptic interaction. The modeling of a sophisticated virtual hand for haptic interactions is investigated with special attention paid on issues of visual realism, motion realism as well as force sensation. Specifically, our virtual hand model simulates natural anatomy in its appearance, motion and deformation, and reflects the feature of force feedback datagloves. Besides, virtual hand interactions based on the constructed virtual hand model are discussed. The proposed framework has been tested on a PC with a force feedback dataglove.

Keywords: Virtual Hand; Haptic Interaction; Dataglove; Collision Detection.

1 Introduction

Virtual reality (VR) technology is revolutionizing the way of interacting with computers. These interactions can be divided, in general, into either interactions with the system or interactions with virtual objects that consist of the virtual scene. For interactions with the system, much unlike the case of the 2D desktop dominated by the GUI/WIMP user interface style, multi-modal input and output (including stereo view, stereo sound, voice input, gesture recognition, 3D menu and haptic feedback, etc.) should be integrated. For interactions with 3D objects, the motion and behavior of virtual objects should be realistically simulated besides realistic visual feedback.

Most probably, interactions with objects in the virtual world involve the human hand, as in the case of performing operations in the real physical world. Dataglove provides the possibility of tracking the user's finger motions. It has been used as a main kind of VR input devices [1]. In this context, a virtual hand is used as an avatar

* Corresponding author.

Z. Pan et al. (Eds.): Transactions on Edutainment IV, LNCS 6250, pp. 229–240, 2010.

of the user's hand [2,3]. Moreover, force feedback dataglove provides the user with the sensations involved by force feedback during the manipulation of virtual objects with virtual hand [4]. Especially, haptic feedback is mandatory when the graphics is corrupted (simulating poor visibility) or when the manipulated object is partly or totally occluded, or when the environment is dark [5].

Indeed, a large number of applications have been foreseen for haptic interaction, for instances, virtual sculpture, virtual surgery, education, entertainment, and industry applications including virtual prototyping, training, and maintenance [6,7]. Therefore, the modeling of the virtual hand, including shape modeling, kinematics modeling, and haptic modeling, is fundamental and important for 3D interaction, and is required by a wide range of virtual reality applications. However, though there exist many works on virtual hand modeling regarding computer animation and interaction applications [8~12], much less attention is paid on haptic modeling of virtual hand [5,13,14].

In this paper, we investigate the modeling of a realistic virtual hand for haptic interactions. A framework for virtual hand haptic interaction is proposed in which the issues of visual realism, motion realism as well as feedback force accuracy are addressed. Specifically, our virtual hand model simulates natural anatomy in its appearance and motion, and reflects the feature of force feedback datagloves. We also discuss virtual grasping issues based on the constructed virtual hand model.

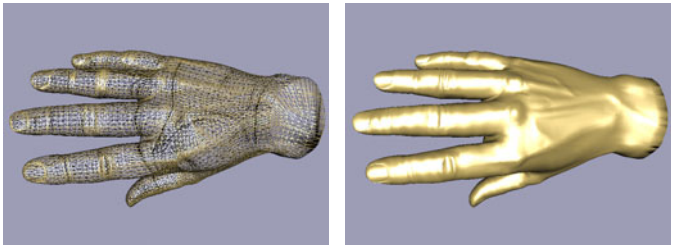
2 Virtual Hand Modeling

The human hand is a complex organ of a human being. It's not trivial to build the geometry model of a virtual hand, not to mention to set up its kinematics and haptic model. In this section, we'll discussion the construction of a virtual hand from various perspectives, including shape, motion, force feedback, and collision detection.

2.1 Shape Model

The complexity of the hand structure makes its shape modeling a complicated and tedious process. Many modeling techniques, from polygonal modeling, parametric surface modeling, to implicit modeling, have been proposed for modeling geometry of the human hand.

Polygon mesh is a fundamental representation of 3D objects. For virtual hand modeling, the built polygonal model should on the one hand be accurate enough to reflect the human hand shape, and on the other hand, it should not be too complicated to hinder the real time simulation of hand motion. Though it is often very tedious since large amount of user interaction is inevitable to construct the polygonal mesh of a human hand, the development of subdivision surfaces seems to help alleviate the heavy burden of lots of user input [15], and many mesh modeling tools exist. To facilitate rendering and motion control, we construct the geometry of the virtual hand with triangular mesh, based on the knowledge of the hand shape and its anatomic structure [16]. Besides, the balance between the amount of triangles and the visual appearance is also taken into account. Fig. 1 shows the virtual hand in wireframe as well as GPU pixel shading.



(a) wireframe (b) GPU pixel shading

Fig. 1. The virtual hand model

2.2 Motion Model

The human hand is a complex structure with extra articulation that enables us to grasp, hold, and operate a wide variety of objects. The kinematics model of the virtual hand is determined by the hand's skeleton structure. As illustrated in Fig. 2, each finger has three phalanges (proximal, middle, and distal); the thumb has two (proximal and distal). Correspondingly, each finger has three joints (distal interphalangeal joint (DIJ), proximal interphalangeal joint (PIJ) and metacarpophalangeal joint (MPJ)) with DIJ and PIJ each having one degree of freedom (DOF) and MPJ having two DOFs; the thumb also has three joints (thumb IJ, thumb MPJ, and trapeziometacarpal joint (TMJ)) with TMJ and thumb MPJ each having two DOFs and thumb IJ having one DOF. The palm has two DOFs. Therefore, the hand motion is highly constrained by the joints. More complicated, a joint is often constrained by other joints when in motion.

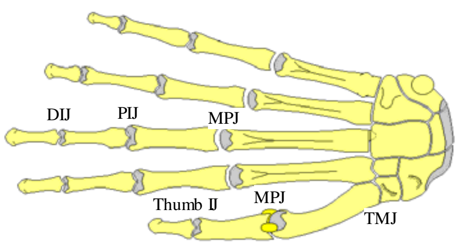


Fig. 2. Human hand anatomy (right hand)

The extra articulation of the hand makes it difficult to realistically simulate its motion and muscle deformation with simple kinematics models. Usually, a three-layer model which consists of the skeleton layer, the muscle layer and the skin layer is adopted to handle the virtual hand deformation [10]. However, we use a simpler two-layer model and employ skeletal subspace deformation (SSD) to handle virtual hand deformation [17, 18] due to strict computational time limit of haptic interaction.

As illustrated in Fig. 3, the kinematics model of the virtual hand consists of the skin layer and the skeleton layer. The skin layer is the triangular mesh used for displaying purpose. Its deformation is driven directly by the skeleton layer. The skeleton layer is built based on the anatomic structure of the human hand. Each finger is abstracted as a joint chain, where each joint has a local coordinate system, and includes such information as the joint position, orientation, rotation angle and a pointer to the next joint. The joint rotations are controlled by flex data (rotation angles) captured by a dataglove (e.g. the CyberGlove dataglove [1]).

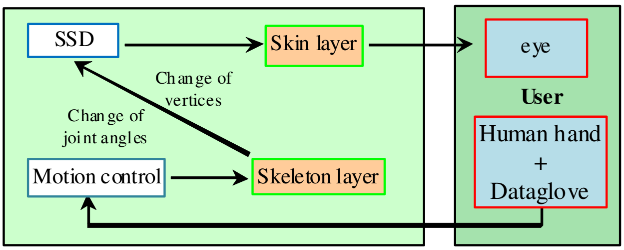


Fig. 3. Motion model of the virtual hand

The process of deforming the skin (i.e. the hand geometry) can be described as follows. Firstly, the user's finger rotations are captured by the dataglove, and the captured flex data directly control the rotations of the joints in the skeleton layer. Secondly, the rotations of the joints drive the deformation of the skin layer according to the SSD which is evaluated by the weighted blending of an affine transformation of each joint by Eq. 1.

$$v_j = \sum_{i=1}^{n} w_i M_i v_{j0} \tag{1}$$

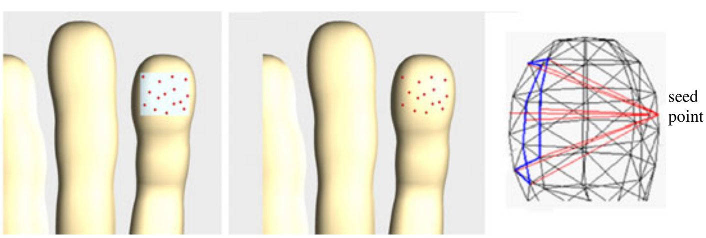
Where n is the number of joints, v_j is the j-th vertex in an arbitrary pose, v_{j0} is the j-th vertex in its rest pose, M_i is the affine transformation matrix defined by flex angles of joints and hand motion, and w_i is a joint weight that defines the contribution of the i-th joint's transformation to the j-th vertex. The weight w_i is assigned by the user to control deformation.

We use graphics processing unit (GPU) to accelerate the SSD computation. The vertices, normals, joint weights and joint indices are all stored in textures and transferred to GPU. The affine transformations of each joint are also transferred to GPU. Following is the pseudo code of the fragment program for SSD computation. Note that the joints which contribute to any vertex are limited to at most 4 to facilitate the GPU data storage.

```
// Pixel Shader - SSD Computation
uniform sampler2DRect vertex, normal, weight, jointIdx;
uniform mat4 world2joint[32], joint2world[32];
void main()
{
  vec2 coords = gl_TexCoord[0].xy;
  vec4 w = texture2DRect(weight, coords);
  ivec4 jIdx = ivec4(texture2DRect(jointIdx, coords));
  gl_FragData[0] = vec4(0.0);
  vec4 tmpnormal = vec4(0.0);
  for(int i = 0; i < 4; i++) {
    vec4 wlocal = world2joint[jIdx[i]]*
            texture2DRect(vertex, coords).rgba;
    gl_FragData[0] += joint2world[jIdx[i]]*wlocal*w[i];
    wlocal = world2joint[jIdx[i]]*
            texture2DRect(normal, coords).rgba;
    tmpnormal += joint2world[jIdx[i]]*wlocal*w[i];
  }
  gl_FragData[1] = vec4(normalize(tmpnormal.xyz), 0.0);
}
```

2.3 Haptic Model

Basically, the integration of haptic feedback within a virtual environment raises many problems at both hardware and software levels. During the past decade, much effort has been made to develop haptic rendering algorithms for various haptic devices. These methods can be classified into categories according to the avatars used: point-based methods, ray-based methods and object-based methods [7,19]. However, a current major limitation for the design of haptic interfaces is our poor knowledge concerning human haptic perception. Indeed, both psychological and physiological issues of haptic perception should be taken into account as far as haptic interaction is concerned.



(a) randomly distributed points (b) cylindridcally projected points (c) force proxy (side view)

Fig. 4. Haptic model

We think that the haptic model should fully respect the feature of area contacts between virtual fingertips and target objects in order to present realistic force via a force feedback dataglove. Our proposed haptic model is simple yet effective. As illustrated in Fig. 4, the force proxy is composed of a cluster of line segments which are generated by sampling the estimated contact area of a virtual fingertip. It is generated as follows. Firstly, a rectangle a bit narrower than the fingertip is created in front of the fingertip, and the center point of the rectangle is projected onto the back face of the fingertip to get a seed point. Secondly, a set of points is randomly generated inside the rectangle (Fig. 4(a)) and cylindrically projected onto the front face of the fingertip (Fig. 4(b)). Finally, the seed point is connected with each of the projected points to form a line segment cluster, and this line segment cluster (Fig. 4(c)) is called the force proxy of the virtual fingertip [14].

2.4 Collision Detection Model

Real-time collision detection is used to automatically identify whether there are interferences between the virtual hand and virtual objects. In general, collision detection requires a significant computational overhead, especially when involving deformable models. However, collision detection should be computationally efficient since real-time feedback is fundamental for haptic interactions.

To some extent, visual realism is of more interest, rather than accuracy, for virtual hand operations. For haptic interaction, collision detection between virtual hand and virtual objects is a more qualitative issue rather than a quantitative one. Realizing this, we build simplified structures for the palm and each joint of the virtual hand, and use these simplified geometries as the collision detection model. In order to prevent the penetration of virtual hand into virtual objects, a simplified structure is a bit larger than its corresponding geometry (Fig.5). The classic software toolkit, RAPID, is used as the collision detection engine between our collision detection model and virtual objects in the scene [20].

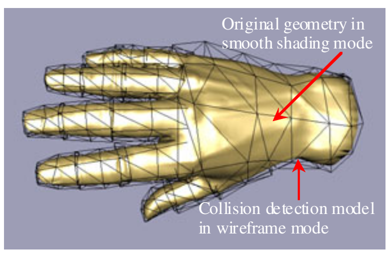


Fig. 5. Collision detection model

3 Virtual Hand Haptic Interaction Framework

We integrate the above models into a sophisticated virtual hand model to facilitate virtual hand haptic interaction. As illustrated in Fig.6 and Fig.3, the user's hand motion data are captured by 3D tracker and dataglove and used to drive the skeleton layer of the motion model of the virtual hand. While the skeleton transformations are directly transferred to the collision detection model and the haptic model, the shape model's deformation is driven by the skeleton transformations as described in the section 2.2. Note that the shape model's deformation is also constrained by whether there are contacts between the collision detection model and other virtual objects in the scene. The feedback force computation is performed between the haptic model and other virtual objects.

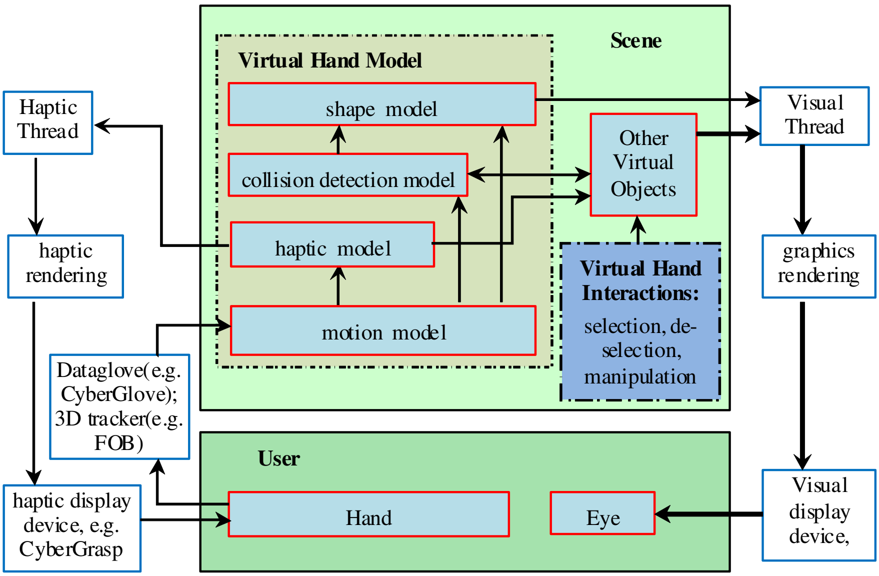


Fig. 6. Virtual hand haptic interaction framework

3.1 Feedback Force Computation

As the haptic model is composed of many line segments, we perform force computation first by detecting whether the line segments intersect with the target model. We perform voxelization on target models in the pre-processing stage to gain efficiency for intersection tests as it has proven to be a very significant way for accelerating such computations [14,21,22].

Given a line segment and a voxelized target model, following two steps are taken to check whether the line segment intersects with the target model. Firstly, all the voxels intersecting with the line segment are identified to reduce unnecessary computation since only the facets of the target model contained in these voxels may intersect with the line segment. We extend the method proposed by [23] to perform such intersection tests. Secondly, the contact point between the line segment and the target model is determined. As only triangles contained in the intersecting voxels need to be checked, we extend the algorithm presented in [24] to calculate the nearest intersections between the line segment and each triangle contained in the intersecting voxels. The above process continues till all line segments of the haptic model are dealt with, and the results (e.g. intersection points, indices of intersecting triangles) are recorded for force computation.

We use the Hooke's law to compute the feedback force to each fingertip based on the collision detection results [14]. Assume that a line segment L intersects a triangle M of the target model at the point P (Fig.7). Let P_s and P_e be the start point and end point of the line segment L (The equation of L is $L(t) = P_s + (P_e - P_s)t, (0 \le t \le 1)$, the intersection point P is represented by $L(t_0)$). Then the direction of the feedback force F_i generated by L is the same as the normal of M , and its magnitude is calculated by t

$$F_i = kx = kd\cos\alpha \tag{2}$$

where k is the stiffness of the target model, x is the penetration depth, and d is the penetration length along L:

$$d = \|P_e - P\| = (1 - t_0)\|P_e - P_s\| \tag{3}$$

It is worth noting that $\|P_e - P_s\|$ can be calculated in advance during the pre-processing stage.

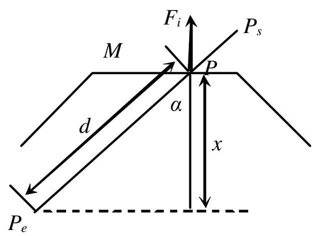


Fig. 7. Feedback force computation

Assume there are n line segments of the haptic proxy intersecting the target model, the feedback force can be calculated as an average of the feedback force generated by each line segment:

$$F_{proxy} = \sum_{i=1}^{n} F_i / n \tag{4}$$

3.2 Virtual Hand Interactions

Fundamental 3D interactions include selection, de-selection, and manipulation of virtual objects. By virtual hand interactions, the user can select and directly manipulate virtual objects with his/her hands. Typically, the virtual hand is used to visualize the current locus of user input. The position and orientation of the 3D tracker are mapped onto the virtual hand, while the flex angles of the dataglove are mapped to control the deformation of the virtual hand. To select an object, the user simply performs a grasp operation as in the real world to pick it up. The object is then attached to the virtual hand and can be easily translated and rotated within the virtual environment until the user releases it. Various grasping/releasing conditions can be used to trigger the grasping/releasing operation [25,26]. Many hand gestures allow users to interact with the virtual objects, for instances, five typical grasping patterns (palmar, lateral, pinch, spherical grasp and cylindrical grasp) can be set up for the virtual grasping of various virtual objects (Fig.8). For details on virtual grasping/releasing heuristics, please refer to [25].

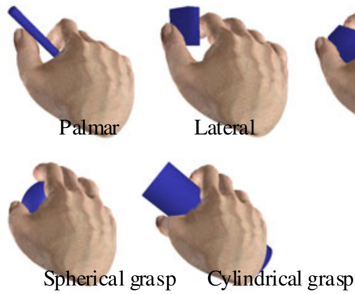


Fig. 8. Grasping patterns

4 Experiments

The proposed method was implemented with C++, and tested on a PC with Intel Core2 Quad 2.4GHz CPU with 2G RAM. The force feedback device used was the CAS-dataglove with PEDfinger which is a force feedback system developed by Institute of Automation Chinese Academy of Sciences to help users feel virtual 3D objects [27,28].

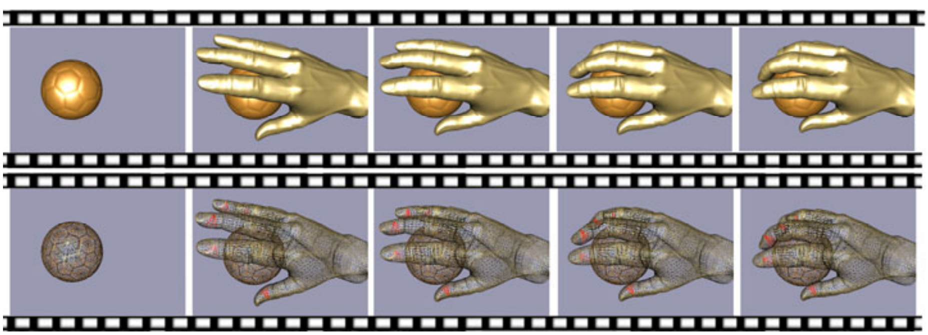


Fig. 9. Virtual grasping a soccer ball

We performed tests with the virtual hand haptic interaction framework. A soccer ball which consists of over 3,500 triangles was used as the target object. The virtual hand geometry shown in Fig. 1 consists of over 15,800 triangles. Fig.9 shows a sequence of snapshots of virtual grasping the soccer ball with feedback forces delivered onto the user's fingertips. Figure 10 shows the final frame of the sequence in which the haptic model is shown as red line clusters and is used for feedback force computation. In our tests, the user can feel continuous force feedback when he grasps the soccer ball.

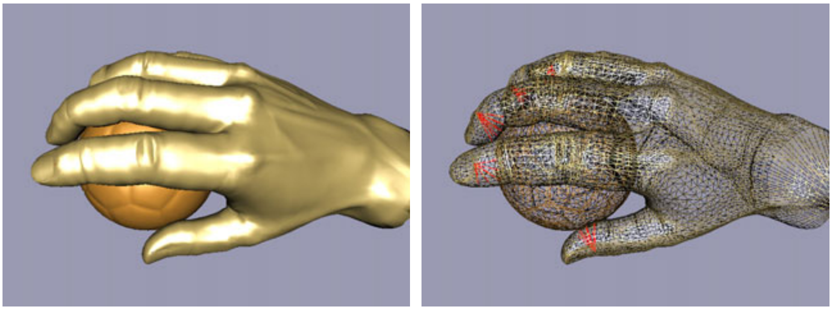


Fig. 10. Final frame of the sequence shown in Fig. 10

5 Conclusions

Haptic devices are used to extend a human being's sense of touch into a virtual world, in which the user can feel the geometry and other properties of virtual objects. Virtual hand haptic interactions play key roles in virtual environments, especially when dexterous manipulation of virtual objects is concerned. In this paper, a framework for virtual hand haptic interaction is proposed. The framework employs a sophisticated virtual hand for haptic interactions which consists of shape model, motion model, collision detection model and haptic model. These models are integrated together to serve for the purpose of virtual hand haptic interaction. The motion model which is one of the building blocks is directly driven by the user's hand motion, the collision detection model is used to facilitate interference check between the virtual hand and other objects, the haptic model is used for stable feedback force computation, and the shape model is used for visual presentation. While the collision detection model and the haptic model are directly driven by the skeleton layer of the motion model, the shape model deformation is driven by the motion model and constrained by the collision detection model. As natural anatomy of the human hand and the feature of force feedback datagloves are respected in the virtual hand model, the user can feel continuous force feedback when he interacts with virtual objects during our test on a PC with a force feedback dataglove.

There is yet a long way to go as to the haptic simulation of dexterous manipulation of virtual objects. However, recent advances in both haptic devices and software development are opening an array of opportunities for us. In the near future, we will try to haptically interact with deformable objects which needs more elaborate methods.

Acknowledgments. The authors would like to thank the 863 High Technology Plan of China (2006AA01Z130) and the National Natural Science Foundation of China (60673197) for financially supporting this research. Thanks would go to Professor Kui Yuan, Professor Qingxiu Du and Dr. Haibing Zhu for their help on the setup of the CAS dataglove.

References

1. Wireless Data Glove: The CyberGlove® II System (2008), http://www.immersion.com/3d/products/cyber_glove.php
2. Jayaram, U., Tirumali, H., Jayaram, S.: A tool/part/human interaction model for assembly in virtual environments. In: Proceedings of ASME DETC 2000, Baltimore, Maryland (2000)
3. Wan, H., Luo, Y., Gao, S., Peng, Q.: Realistic virtual hand modeling with applications for virtual grasping. In: Proceedings of ACM SIGGRAPH International Conference on VRCAI, Singapore (2004)
4. CyberGrasp™ Exoskeleton (2008), http://www.immersion.com/3d/products/cyber_grasp.php
5. Popescu, V., Burdea, G., Bouzit, M.: Virtual reality simulation modeling for a haptic glove. In: Proceedings of the Computer Animation 1999, Geneva, Switzerland (1999)
6. Stone, R.: Haptic feedback: a potted history, from telepresence to virtual reality. In: Brewster, S., Murray-Smith, R. (eds.) Haptic HCI 2000. LNCS, vol. 2058, pp. 1–7. Springer, Heidelberg (2001)
7. Laycock, S., Day, A.: A survey of haptic rendering techniques. Computer Graphics Forum 26(1), 50–65 (2007)
8. Hui, K., Wong, N.: Hands on a virtually elastic object. The Visual Computer 18(3), 150–163 (2002)
9. VirtualHand®, SDK (2008), http://www.immersion.com/3d/products/virtualhand_sdk.php
10. Moccozet, L., Huang, Z., Thalmann, M., Thalmann, D.: Virtual hand interactions with 3D world. In: Proceedings of Multimedia Modeling 1997, Singapore, pp. 307–322 (1997)
11. Lee, J., Yoon, S.: Realistic human hand deformation. Computer Animation and Virtual Worlds 17(3-4), 479–489 (2006)
12. Rhee, T., Neumann, U., Lewis, J.: Human hand modeling from surface anatomy. In: Proceedings of the 2006 symposium on Interactive 3D graphics and games, Redwood City, California (2006)
13. Maciel, A., Sarni, S., Buchwalder, O., Boulic, R., Thalmann, D.: Multi-finger haptic rendering of deformable objects. In: Proceedings of Eurographics Symposium on Virtual Environments (2004)
14. Wan, H., Han, X., Zhou, Z.: Realistic force rendering with applications for virtual assembly. Journal of Computational Information Systems 3, 303–308 (2007)
15. Sederberg, T., Zheng, J., Sewell, D., Sabin, M.: Non-uniform recursive subdivision surfaces. In: Proceedings of SIGGRAPH 1998, pp. 387–394 (1998)
16. Kapit, W., Elson, L.: The anatomy coloring book, 2nd edn. Harper Collins (1993)
17. Thalmann, M., Laperriere, R., Thalmann, D.: Joint–dependent local deformations for hand animation and object grasping. In: Proceedings of Graphics Interface 1988, pp. 26–33 (1988)
18. Rhee, T., Lewis, J., Neumann, U.: Real-time weighted pose-space deformation on the GPU. Computer Graphics Forum 25(3), 439–448 (2006); Proceedings of Eurographics 2006, Vienna, Austria (2006)
19. Srinivasan, M., Basdogan, C.: Haptics in virtual environments: taxonomy, research status, and challenges. Computers and Graphics 21(4), 393–404 (1997)
20. Gottschalk, S., Lin, M., Manocha, D.: Obb-Tree: A hierarchical structure for rapid interference detection. In: Proceedings of the ACM Siggraph 1996, New Orleans, LA (1996)

21. Kaufman, A.: Introduction to volume graphics. In: SIGGRAPH 1999 course notes, course, vol. 41, pp. 24–47 (1999)
22. Huang, J., Yagel, R., Filippov, V., Kurzion, Y.: An accurate method for voxelizing polygon meshes. In: Proceedings of ACM 1998 Symposium on Volume Visualization, Research, Triangle Park, NC (1998)
23. Andrew, W.: Fast ray-box intersection. In: Graphics gems. Academic Press Professional, Inc., San Diego (1990)
24. Mŏller, T., Trumbore, B.: Fast, minimum storage ray-triangle intersection. Journal of Graphics Tools 2(1), 21–28 (1997)
25. Wan, H., Luo, Y., Gao, S., Peng, Q.: Realistic virtual hand modeling with applications for virtual grasping. In: Proceedings of ACM SIGGRAPH International Conference on VRCAI, Singapore (2004)
26. Ullmann, T., Sauer, J.: Intuitive virtual grasping for non-haptic environment. In: Proceedings of PacificGraphics 2000, Hong Kong, pp. 373–377 (2000)
27. Fu, Y., Yuan, K., Zhu, H., Du, Q.: Motion Modeling and Software System Design for CAS-Glove Data Glove. Journal of System Simulation 16(4), 660–662 (2004)
28. Yuan, K., Zhu, H.: A new kind of force-reflecting device and its control. In: Proceedings of IEEE Conference on Intelligent Robots and System, Lausanne, Switzerland, pp. 2920–2924 (2002)

Phone, Email and Video Interactions with Characters in an Epidemiology Game: Towards Authenticity

Muriel Ney[1], Celso Gonçalves[1], Nicolas Blacheff[1],
Claudine Schwartz[2], and Jean-Luc Bosson[2]

[1] Laboratory of Informatics of Grenoble (LIG), CNRS and University Joseph Fourier
961 rue de la Houille Blanche, 38402 Grenoble Cedex, France
[2] Techniques for biomedical engineering and complexity management (TIMC),
University Joseph Fourier, Domaine de la Merci, 38710 La Tronche, France
{Celso.Goncalves,Muriel.Ney,Nicolas.balacheff,
Jean-Luc.Bosson}@imag.fr

Abstract. A key concern in game-based learning is the level of authenticity that the game requires in order to have an accurate match of what the learners can expect in the real world with what they need to learn. In this paper, we show how four challenges to the designer of authentic games have been addressed in a game for an undergraduate course in a medical school. We focus in particular on the system of interaction with different characters of the game, namely, the patients and a number of professionals. Students use their personal phone and email application, as well as various web sites. First, we analyze the authenticity of the game through four attributes, authenticity of the character, of the content of the feedback, of the mode and channel of communication and of the constraints. Second, the perceived authenticity (by students) is analyzed. The later is threefold and defined by an external authenticity (perceived likeness with a real life reference), an internal authenticity (perceived internal coherence of the proposed situations) and a didactical authenticity (perceived relevance with respect to learning goals).

Keywords: authenticity, immersion, simulation, role-play, epidemiology, communication.

1 Authentic and Immersive Games

Computer simulations afford the possibility of showing a phenomenon when it is impossible or impractical to confront learners with such a phenomenon in the physical world. However, a simulation alone may not be sufficient. An important component of an educational approach including a simulation is to allow an authentic learning experience, the virtual equivalent of a real-world experience. This is precisely what a learning game is about, simulating real-world possibilities to act, and consequently helping players to become more confident of his or her ability to recognize and handle similar situations, should it happen afterward in real life. However, authenticity does not mean a perfect reproduction of reality. What we mean by authentic are the main

Z. Pan et al. (Eds.): Transactions on Edutainment IV, LNCS 6250, pp. 241–255, 2010.

characteristics of the type of situations at stake, that is, those characteristics that require learners to mobilize the knowledge targeted (the learning goals) in order to be successful in the game.

One mean to create authenticity in learning games is immersion, making learners feel like a certain situation is real although they know it is not. According to Brown [1] the main indicator of immersion is the degree of involvement of the players. Game designers know that players may be engaged in different ways, mobilizing them differently, whether they are challenged on their ability to act rapidity, to find the best strategy, to communicate with different people, etc. There have been numerous definitions of immersion in games in previous studies (e.g. tactical, strategy or sensory immersion) [2], [3]. In this paper, we focus on immersion in an authentic situation, and more specifically, a situation that involves many moments of interactions with people. We call "interactional immersion" an immersion that relies mostly on interactions with other players or with characters of the game. Examples of such games that focus on the social aspects of the situation are Second Life or the Sims.

In this paper, we address the following four challenges posed to designers of authentic and immersive learning games:

- Find a compromise between the requirement of the real life reference and the learning context.
- Allow the learner to appropriate the meaning of the situation (what is the point in learning terms).
- Blur the line between the virtual and the real experiences.
- Create engaging interactions for players with realistic characters.

We show how these challenges have been taken up in a game for an undergraduate course in a medical school.

2 Design Methodology

The Laboratorium of Epidemiology (LoE) is collaboratively designed and then used by three parties: researchers, teachers and students [4], [5]. On one hand, medical students enrolled in a course of biostatistics and learn to develop critical thinking on statistics and epidemiology studies. On the other hand, researchers assess the conditions of learning with such a simulation. LoE is thus both an educational project integrated in a medical school, and a research project, a laboratorium, allowing repeated data collection campaigns that are not singular events in students and teachers lives. This integration is an attempt to reduce data collection bias and produce well-documented corpora.

LoE uses learning with live experiences of social interactions, learning by doing, and problem-based learning and as such follows a socio-constructivist approach. In LoE, students conduct their own epidemiology survey and present evidences in an article. By doing so, they acquire skills in critical reading of medical articles. Indeed, with LoE students learn about issues in applying methodologies in epidemiology surveys, in particular, when defining a main objective, designing a protocol, collecting data (with issues like sample size, data quality, ethical consideration, etc), using statistical tools and analyzing data, making a decision based on statistical results, and presenting results in a scientific form (article and talk) [5]. LoE also addresses patient-centered skills (history taking), communication skills (with different professionals), and teamwork skills [6].

LoE is based on a computer simulation and a role-play. Students play the role of public health physicians and experience an otherwise inaccessible professional situation: an epidemiological survey at the hospital. For the sake of authenticity, we used real data to simulate patients at different hospitals. This allowed us to tell students that everything was real. We used a national database containing about 10.000 patients suspected of being affected by one of the tromboembolic diseases. Then the mission given to students was first to evaluate the risk of these diseases at a hospital (look for signs and symptoms and give evidence based on statistics) and second to propose a decision tool (which patients are at great risk and should be tested first). The mission was sufficiently broad that students could further define their own objectives for their epidemiology survey.

To design an immersive simulation that offers students such a professional experience, we first performed a task tree analysis based on a series of workshops with an epidemiologist [4]. The epidemiologist was asked to focus on facts and actions rather than on a conceptual view of his practices. Combined with the learning goals of the course, this work resulted in a task tree presented in figure 1. It starts from the problem to be solved while tasks on the lower leaves correspond to the actions to be performed to solve that problem. Between the two, there are steps and sub-steps. On figure 1, seven main steps of the task tree are indicated. Only one of them is further detailed into sub-steps and actions.

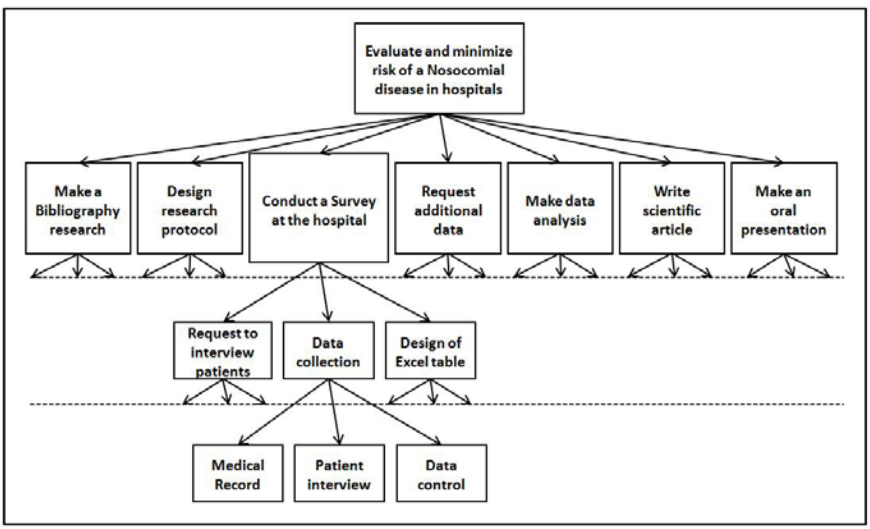


Fig. 1. Task tree analysis of the situation

The task tree approach allows visualizing the line of tasks that a learner with a given level of expertise may follow [7]. It was a basis to define a compromise between the requirement of the professional reference and the learning context. Finally, this work also helped unraveled the important interactions. We will now see how the different professionals and the patients were represented as characters of the game. The system of interactions between students and characters of the game is summarized in table 1. We indicated how the character is represented in the game (e.g. by a text giving it a name, gender and position) and whether the level of representation is low (few information) or high (many information including a picture, a voice, etc). Then the mode and channel of communication with characters are given. Some of them will be detailed below. Finally, the learning goals of the interactions are indicated.

The game computer environment is used by teams of three or four students. Teams may gather in groups at various times, and at least for the final simulated congress. Several groups may play the game in parallel. Therefore, the team is the main unit in the game. Although each student has a personal identifier for the web site, the environment will adapt to the team (never to a student individually). For instance, the patient "recognizes" the team when it comes back to a room, or the report on the article submitted to the congress is visible by the whole team. The only thing that teams of the same group share are regular meetings in class with a tutor. In our implementation, there are eight four-hour meetings and there are about 180 students distributed into six groups (with about 6 teams each).

As we will see in the following sections, patients are played by actors while professionals are played by tutors. In class, the tutor's role is to help students on the epidemiology methodologies. Tutors also play the different characters of table 1, but not necessarily for their own group of students. This way, they discover the characters' feedback content with their students and do not have to pretend they do not know it in advance. Another reason is that it allows attributing the characters to tutors according to their competences.

The environment is a pure Web platform that can be reached from any Internet-enabled computer. Server-side code is object-oriented PHP built upon open-source Symfony framework, and client-side code is rich HTML including Adobe© Flash and Ajax components. A second environment is used by researchers who can visualize computer trails of the users (not shown). A third environment is devoted to tutors, and mainly allow them to follow students' productions and to interact with them (table 1). This environment should not only facilitate the task they have to do, but also help them to "be" the character of the game that is supposed to interact with the students. In the following, we specify the immersive and authentic components of most interactions of table 1.

3 Description of the System of Interaction

We describe successively some of the email, phone, video and web interactions of table 1.

3.1 Interaction with Experts by Email

Students interact with two groups of experts (table 1). One is the Ethical Research Committee that, in real life, controls all the protocols before they can be applied on patients. Another is the Public Health Commission that commissions epidemiology surveys. We use similar communication mean for both, for the same reasons, and we describe in details only the first one.

Table 1. System of interactions between students and characters of the game

Character of the game	Representation level of the character	Mode of representation	Mode of communication	Channel of communication	Learning Goals
Head of hospital department	Low	Text	Oral and written	Telephone	Synthesize a protocol verbally
Expert of the ethical committee	Low	Web site	Written	e-Mail	Write and communicate a protocol
Patient	High	Video and text	Visual, oral and written	Online structured interview	Translate patient saying into medical data
Technician of the hospital information department	High	Video and Web site	Written	Web and e-mail	Make a data entry mask and fill it with data
Congress reviewer	Low	Web site	Written and oral	Web	Write and present statistical results to a scientific community
Public health commission	Low	Web site	Written	e-Mail	Propose a recommendation to a Public Health Organisation

Students have to send their protocol to the Ethical Research Committee for approval, before they can apply it at the hospital with patients. The learning goals of this interaction are to learn how to address to professionals in a formal way and to learn how to present a protocol in a written form (table 1). In real life, people address protocols to this committee by regular mail or, more and more often, by email. We choose the latter (figure 2). The benefit of this system is to give life to a character (member of this committee) that will receive students' productions. This way, we hope that students will address problems in terms of the underlying concepts and skills, rather than in terms of what they think that the teacher will expect them to do [8], which happens when the teacher is the receiver.

Students use their personal e-mail application. They send their protocol as a pdf file that should be introduced by a short message. This is compulsory since only approved protocols can be implemented at the hospital.

The immersive interface that we use for tutors is a standard Webmail interface. It is immersive in the sense that it is separated from their usual mailbox (they might probably answer in their name by mistake, otherwise) and it is organized with a dedicated signature and e-mail address. Furthermore, it is accessible to tutors only, from the game web platform.

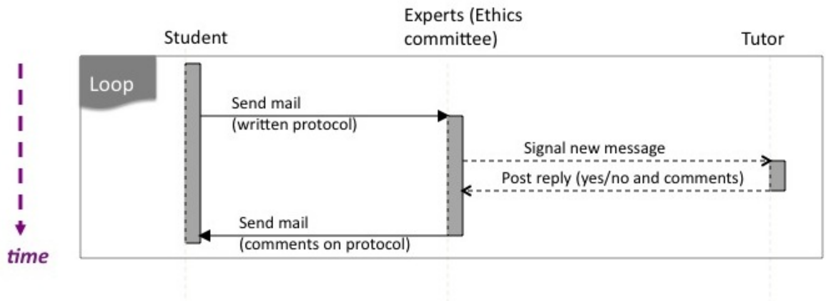


Fig. 2. Diagram of interaction between students and experts of the Ethics Research Committee

3.2 Interaction with Hospital by Phone

In order to be able to interview patients, students have to ask for the authorization of the head of each medical department of the hospital(s) they wish to visit (table 1). Students learn to be convincing, quick and thus prepared. It also put them in a rather intimidating situation since they are talking to a physician with a heading position. In real life, students do not have to meet him/her in person (this person is very busy and it is difficult to make an appointment) but rather they have to give a call.

We designed the following interaction (figure 3) where students have to use their personal phone. They make a call and, having listened to a brief message, they have to formulate their demand in the way they think useful to describe it to the head of the department. Later, a tutor listens to their message and writes an SMS. Students thus receive an answer on their personal phone by SMS that gives them an agreement or an argued refusal. An agreement allows the team to access to the patients' rooms. When a tutor sends a SMS with an agreement, it opens automatically the rooms to the given team of students. This phone call is thus a compulsory stage to continue the epidemiological survey. A team can make several calls, either to have access to several departments, or to repeat their demand after a refusal.

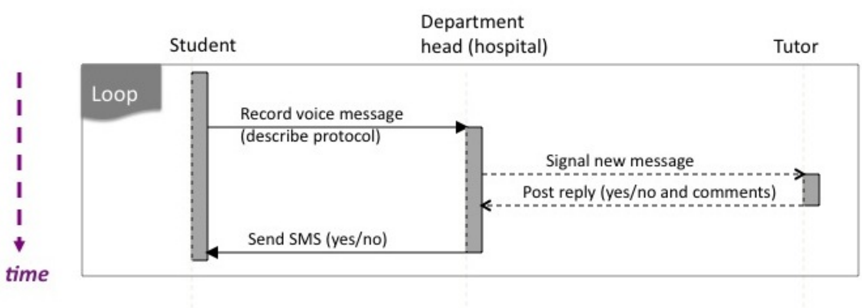


Fig. 3. Diagram of interaction between students and each head of the various hospital departments

In order to favor tutors' immersion, we designed a web interface for listening and answering to students' messages. We do not want tutors to answer in their name and for that reason, they do not use their personal phone. In the tutor web site, they can select the message and then they can take two actions, either accept or refuse the student's demand. Tutors may select an SMS answer among a list of possible answers. This list reminds the tutor of the validation criteria of student's messages, in the form of examples, e.g. a SMS may be "specify your objectives and constraints for patients and staff". We use IP convergence to integrate VOIP technology into the web application: Phone2Web via IMAP for tutors to check phone messages and Web2SMS for tutors to send SMS to students.

3.3 Interaction with Patients Based on Video

The interaction with patients is one that has been extensively simulated, obviously due to its central place in medical practices. Simulated patients are used to overcome the disadvantages of involving real patients that have limited availability and that provides low variability of cases [9]. Reviewing the literature, we built the following list of main strategies that have been used (table 2).

Table 2. Bibliographical search on patient interaction systems. We specify the representation model of the character (here the patient), the mean of communication with the patient, the type of activity the student can perform, and the learning goals.

Ref.	Character representation model	Mean of communication	Type of activity	Learning Goals
[10]	Text	None	Information research	Diagnostic
[11]	Hypertext Hypermedia	Website	Information research	Diagnostic
PULSE - U.S. Office of Naval Research	Avatar	Clicking	Information research Communication	Diagnostic Interview
[12], [13]	Video	Website	Information research Communication	Diagnostic Interview
iStan - Medical Education Technologies, Inc.	Model	Oral	Gestures Communication	Diagnostic Interview
[14]	Student	Oral	Gestures Communication	Diagnostic Interview
[6]	Actor	Oral	Gestures Communication	Diagnostic Interview

The patient can be physically present in the room (last three rows in table 2). First, the patient can be an actor. Simulated or standardized patients are individuals (actors) who are selected and trained to portray a patient, including accurate and consistent responses to questions [6]. Second, the patient can be a student. A student plays the role of a patient in front of peers, which is a cheaper alternative to simulated patients and sometimes leads to similar outcomes [14]. Third, the patient can be a model that talks and bleeds (like the commercial model iStan). The talking is done through the mouth of the model by a teacher. The teacher is behind the scene watching the student dealing with the model, thanks to a camera (the so-called Wizard of Oz method).

In all the other cases, the patient is represented on screen or paper. First, the patient can be described textually (first row in table 2). Many medical schools across the world including the one where our experiments are conducted have adopted Problem-Based Learning [6]. Patient cases are most often presented on paper. It has the drawback that the patient is only a source of information, and not a person with words, feelings, in a social and cultural context [10]. Second, the patient can be described through a multimedia online document. An authoring system that is currently in use in several healthcare disciplines worldwide is Web-SP [11]. The patient is represented by a number of textual information and images (his/her photo, X-rays, etc) accessible by checking boxes. Third, the patient can be a video-game character. The most accomplished example today is PULSE, which includes a 3D virtual hospital where students guide their avatar and interact with lifelike patients (doing more than interviews since physical examinations and interventions can be performed). Finally, the patient can be a person video recorded. Clips of either true patients or actors are used, but in most cases only for demonstration (physical examinations skills, doctor-patient relationships, etc), rather than for problem solving [6]. In a recent example, interviews in psychiatry are performed using a branching system of videos allowing students to personalize their interview [12], [13].

LoE includes video clips of actresses and actors portraying patients (up to 20 patients in each of the six hospitals). In our case, a benefit of creating virtual interviews was to script the responses according to learning goals. The scripted responses includes signs and symptoms, communicated verbally or not (e.g. breathing heavily can be a sign), related to the disease at stake or not (some information are irrelevant). The system allows the students to interact with the simulated patient by selecting one of five preformed questions (figure 4): present illness, present lifestyle, past history, family/social history, and medications. For each question, there is an appropriate video clip associated with it that provides the interviewer with the patient's response. The patient is speaking directly to the camera and thus to the students. This response lasts generally around one minute, and can only be seen once by the team of students. On-demand video has been implemented using Adobe© Flash technology.

When entering the room, a video is immediately starting: either the patient is absent (video of a nurse), not available (video of a physician doing a physical examination) or the patient is there (video of the patient either saying hello or telling that he/she saw the interviewer already). A probabilistic algorithm regulates the presence/absence of patients.

Students must visit at least fifteen patients before they can ask the Information Department to complete their database with hundreds of cases, which will allow doing statistics. With this system, students learn to listen to patients with great attention, to identify accurately symptoms and signs within the patient narrative, to be prepared (know what to look for in advance), to control data quality (e.g. by confronting notes from two interviewers of the same patient), and to manage time (patients are not always available). Moreover, since students have little opportunity to conduct interviews with real patients that early in their studies, this system could help them feel more comfortable when conducting their first patient interviews.

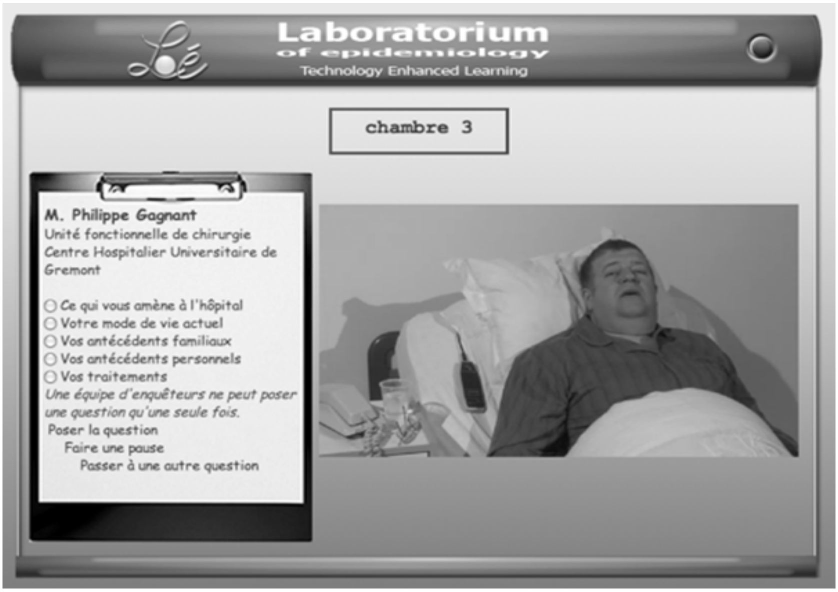


Fig. 4. Screenshot of one of the rooms with patient and questionnaire

3.4 Interaction with Institutions through Their Web Site

In order to be able to participate to the simulated medical congress at the end of the game, students need to submit their article. Students learn to write a scientific article, how it is structured, and how to choose and present their evidences (statistics may come in different forms). In real life, applicants to a congress interact with the scientific committee either by email or via the web site of the congress. We chose the latter.

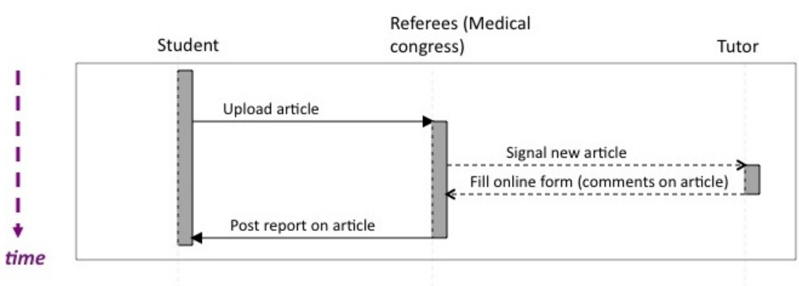


Fig. 5. Diagram of interaction between students and the scientific committee of the medical congress

We designed the following interaction (figure 5) where students have to upload their article on the congress web site. Later they will get a report on their article, visible on the web site. In most cases, they will be asked to revise their article and send it again. They also have access to the program of the congress. Articles are divided into two groups by the organizers of the congress and best articles are given a longer time of presentation.

The tutors who play the role of the referees connect to the congress website and fill a pre-structured form that will make a report on the article. They have a list of points they should pay attention to, to help them in this task. They play the role of referee for students they do not know.

4 Analyze Authenticity in a Game

Serious games are playing an important role for training people on real world situations. However, an entirely realistic simulation is neither practical nor desirable. On one hand, game designers add realism by adding more and more elements, and thus adding complexity to the situation. On the other hand, when students spend too much time to get familiar with too many details they can skip the main learning goals. In the context of learning, a key concern is thus the level of authenticity that the game requires in order to have an accurate match of what the learners can expect in the real world with what they need to learn. Failure to achieve the right level of authenticity runs the risk that the learners may adopt a different strategy in the virtual world than would be desired for learning.

Designed with the desire to make the right compromise between the professional requirements and the learning goals, LoE includes several characteristics to facilitate a move to the right degree of authenticity. Another very good example from which LoE was inspired is the WallCology project [15], which simulates a virtual ecosystem of insects within the unseen space of classroom walls. Interestingly, Moher and co-authors [15] pointed limits in this compromise, not only technical limits, but what we may call epistemic limits. For example, they decided to use imaginary creatures instead of authentic ones in order not to frighten young children and to avoid stereotyping the living conditions of some learners. The limit to authenticity was therefore defined by the learning context.

Authenticity is both a function of the game as well as of the perceiver, and we shall now look at each one separately in the context of LoE.

4.1 Authenticity of the Game

Considering the authenticity of our system of interaction, we identified four attributes (see table 3). First, the characters students are interacting with may be more or less authentic. This is decomposed into two issues. First, the authenticity of the character refers to its personification: what information is available and in what form (text, photos etc). Second, the content of the feedback the character sends to students may be more or less authentic, depending on how adapted to student's actions this feedback is. Another issue is the channel of communication (by email, phone, etc) and the mode (textual, verbal, visual). Finally, the constraints of the interaction may be considered. For instance, in the interaction with patients, it is important that a patient does not repeat at will his/her answers, which is a constraint that was reproduced in our system. This particular constraint is important to the learning context. Indeed, when a patient does not repeat oneself, students need to prepare their interview and also they need to look for ways to control their data, two learning goals of the course.

Table 3. Analysis of the degree of authenticity of three interactions in LoE

	Authenticity of character	Authenticity of feedback content	Authenticity of communication mode and channel	Authenticity of constraints
Head of hospital department	Low	High	High	High
Expert of the ethical committee	Low	High	High	High
Patient	High	High	Low	High

In Table 3, the authenticity of the character is said low when this character is hardly represented. It has only a name (head of hospital department) or not even a name but only a competence (e.g. for the group of experts of the Ethical Research Committee). Regarding the patient, it is only the mode and channel of communication that is low. Indeed, students do not really hold a dialogue with a patient as they would in real life. This table shows that several attributes may be tested when considering the impact of authenticity on learning outcomes.

4.2 Perceived Authenticity

Authenticity is also a function of the perceiver, in our case the student. We call it perceived authenticity. One aspect is students' judge likelihood that an event happening during the game could happen in the real world. This is also referred to as perceived realism [16]. This relates to the perceived fidelity of the game, that is, whether students feel that the simulated interactions mimic a reality.

To study perceived authenticity, we analyzed (1) the trails (verbal, written) of the interactions of seven four-student teams with different characters, and (2) the interviews of two students at three moments during the game. The seven teams (28 students) are second year medicine students at the Grenoble Medical School in a compulsory course of biostatistics paying the game LoE over a period of 12 weeks. We successively analyzed the phone, email and video interactions studying for each whether students behaved as they would with real people [5].

The first data set consisted of phone messages left by students on the hospital answering machine asking the head of the hospital department authorization to interview patients. We collected and analyzed 15 messages. A majority of the messages (10/15) addressed the person in charge of the hospital department in a formal way. This indicates a relation of hierarchy that is adequate in the context of this mission. This is to compare to the information available to students about the receiver of their message. In this case, all they know about the characters are name, gender and position of the receiver of their call. This proved to be enough. On the contrary, for most teams (9/15) patients are only sources of information rather than people. These characters (the patients) have no representation in the environment (students had no information about these characters at this stage of the game although they will later see them on videos).

The second data set consisted of e-mail messages. The students sent 12 messages to the experts of the Ethics Research Committee, introducing an attached research protocol. We analyzed whether students played the game, that is, addressed these messages to experts as they should, using the following six criteria: putting a message subject, giving a personalized name to the added file, greeting formally, presence of a message addressed to the receiver, closing formally, presence of a signature of the writer. Most messages (11/13) showed at least four of these criteria. Therefore, students used appropriate communication features to address to this character by email.

The third data set consisted of audio-taped intra-team interactions. Students' verbal interactions were recorded as they were using the web application to perform their patient interviews (i.e. view videos). We analyzed a four-hour session for one team. The verbal interactions were synchronized with the computer trails. Perceived authenticity is thus analyzed through students' verbal reactions to videos. Students had immediate reactions to what patients say, usually an interjection (e.g. “a tumble!”, “he’s deaf!”). After each video, they sometimes made inappropriate comments (e.g. “I hate these old people unable to answer a single question”). However, they might do the same after interviewing a real patient. We observed several indications of personification of the patients: Students showed feelings for the patient condition (“honestly, I am worried about him”, “Did you see his blood pressure? We’ll have to take care of him.”), students talked about a person rather than about a system (“the patient is not there, the nurse said that he is gone for some test”), students said hello to the patient after the hello-video. Moreover, students reacted to probabilistic events (presence/absence of patients): “look if the patient is back from the test” or “be careful, maybe the patient in room 4 will go away soon”. Therefore, we collected several indications that students perceived the situation as authentic in the sense that they played the game and behaved as in a real situation. The authenticity of the character, of the content of its feedback and of the constraints (see table 3) all contributed to their perceived authenticity.

Next, we looked for indications on perceived authenticity in three interviews with two students. After the phone interaction, these students mentioned that what makes the interaction more "interesting" and more "realistic" is: using their personal mobile phone, talking to an answering machine that limits the duration of the call (need to be prepared), leaving a phone message to a human being and not a machine. They said that they did not know exactly who was going to listen to their message and they had an image of a “person with a white blouse” (a physician). In a later interview, one of the students talked about the video interaction with patients. She pointed out that it was not necessary to interview patients, although it would be necessary in real world. Indeed, in her group it was not clear why students had to interview patients before they can ask for new data and this looked like an optional step. This pointed out a type of authenticity that we did not envision before, namely, an internal coherence of the mission proposed. Indeed inconsistencies within the mission can also disrupt students' engagement. Furthermore, the other student underlined that interviewing patients was useful for his training. In other words, he perceived this activity as relevant from the learning point of view. To summarize, a situation may be perceived as authentic from three points of view: it can be perceived as realistic, coherent and/or relevant. In our case, one student perceived the interviews with patients as realistic, relevant but not coherent. Another one perceived it as realistic, coherent but did not mention any thing about its relevance.

5 Conclusion

At the beginning of this paper, four challenges posed to the designer of authentic and immersive games were proposed. We are now going to show how these were tackled in the design of the game LoE and more specifically of the system of interaction with characters of the game.

First, to find a compromise between the requirement of the real life reference and learning context we combined a task tree analysis with learning goals. The task tree allows showing all the necessary actions to be taken towards the goal. Combined with the learning goal analysis, one can choose those actions that will be more important for learning and those that could be done in place of the students or skipped.

Second, to allow the learner's appropriation of the meaning of the situation, we introduced a form of monitoring on the learning process as part of the interaction system. This allowed them to revise their work and reflect on what has been learnt or missed. Furthermore, we included two moments of institutionalization. The latter is a term used in the Theory of Didactical Situations [17] to denote the process of giving a status to the activities done and the knowledge acquired by the students. First, after students had been working on their data analysis and when they started writing their paper, a teacher gave a two-hour interactive session on statistics. Second, at the very end and right after the simulated medical congress, the teacher debriefed the congress with the students and then decontextualizes the knowledge that was embedded in a specific problem during the game.

Third, to blur the line between the virtual and the real experiences, we allowed students to use their personal phone and email application. This design aimed at immersing students into the game. The ubiquity of mobiles phones is providing a fading of the line between the game space and real world experience [18]. An example is the game Majestic, one of the first so-called alternate reality games, in which players were receiving phone calls on their personal phones. In our case, phone calls and emails are both received and sent by learners, like no other game does, as far as we know.

Fourth, to create engaging players' interactions with realistic characters we addressed the authenticity of different attributes of the interaction: personification of the characters, content of the feedback from these characters, mode and channel of communication and constraints. Different degree of authenticity may be attained for these different attributes. Then we studied authenticity on the learner side, which is partly referred to as perceived realism [16]. In our case, we speak of the perceived authenticity of a mission proposed to students, which includes several situations of interaction.

We found indications of three types of perceived authenticity, that is, the extent to which each situation seems (1) similar to a real one (external authenticity of the mission), (2) coherent within the game (internal authenticity of the mission) and (3) relevant from a learning perspective (didactical authenticity of the mission). This is partly in line with research on film narratives [19] where internal realism and external realism are pointed out and correspond to (1) and (2). We suggest that the three types of authenticity should be taken into consideration when designing authentic learning games. However, we only caught a glimpse on how these three are related to authenticity attributes of the game with the game LoE, and further work is needed.

References

1. Brown, E., Cairns, P.: A grounded investigation of game immersion. Paper presented at the SIGCHI conference on Human factors in computing systems (CHI 2004), pp. 1297–1300. ACM Press, New York (2004)
2. Adams, E.: Postmodernism and the Three Types of Immersion. Gamasutra (2004) (Last retrieved online in November 2009), `http://designersnotebook.com/Columns/063_Postmodernism/063_postmodernism.htm`)
3. Björk, S., Holopaine, J.: Patterns In Game Design, p. 423. Charles River Media (2004)
4. Ney, M., Balacheff, N.: Learning aware environment: a Laboratorium of epidemiological studies. Paper Presented at the Adaptive Hypermedia Conference, Workshop on Technologies for Mobile and Wireless Adaptive E-learning Environments, Hanover, Germany (2008)
5. Gonçalves, C., Ney, M., Balacheff, N., Bosson, J.-L.: Student's Problem Appropriation in an Epidemiology Game. Paper presented at ECGBL 2009, European Conference on Game-Based Learning, Graz, Austria (2009)
6. Lane, J.L., Slavin, S., Ziv, A.: Simulation in Medical Education: A Review. Simulation & Gaming: An Interdisciplinary Journal 32, 297–314 (2001)
7. Wajeman, C., Girault, I., d'Ham, C., Ney, M., Sanchez, E.: Analysing experimental design tasks in scientific labwork. Paper presented at ESERA, the European Science Education Research Association, Barcelona, Spain (2005)
8. Millar, R., Leach, J., Osborne, J.: Improving science education: The contribution of research, pp. 27–47. Open University Press, Buckingham (2001)
9. Bokken, L., Rethans, J.J., Scherpbier, A.J., van der Vleuten, C.P.: Strengths and weaknesses of simulated and real patients in the teaching of skills to medical students: a review. Simulation in Healthcare 3(3), 161–169 (2008)
10. Kenny, N.P., Beagan, B.L.: The patient as text: a challenge for problem-based Learning. Medical Education 38, 1071–1079 (2004)
11. Zary, N., Johnson, G., Boberg, J., Fors, U.G.: Development, implementation and pilot evaluation of a Web-based Virtual Patient Case Simulation environment. Medical Education 21, 6–10 (2006)
12. Dagger, D., Rogers, C., Wade, V., Gaffney, C., Armstrong, K., Fitzmaurice, B., Gill, M., Walsh, E.: Adaptive Simulations for Communication Skills Training in Healthcare. Paper presented at the 2nd International Workshop on Personalisation for E-Health at the 11th International Conference on User Modelling, Corfu, Greece (2007)
13. Fitzmaurice, B., Armstrong, K., Carroll, V., Dagger, D., Gill, M.: Virtual interviews for students interacting online for psychiatry (VISIOn): a novel resource for learning clinical interview skills. Psychiatric Bulletin 31, 218–220 (2007)

14. Lane, C., Rollnick, S.: The use of simulated patients and role-play in communication skills training: A review of the literature. Patient Education and Counselling 67(1), 13–20 (2005)
15. Moher, T., Uphoff, B., Bhatt, D., Lopez-Silva, B., Malcolm, P.: WallCology: Designing interaction affordances for learner engagement in authentic science inquiry. Paper presented at the ACM CHI 2008 Conference on Human Factors in Computing Systems, pp. 163–172 (2008)
16. Shapiro, M.A., Pena-Herborn, J., Hancock, J.T.: Realism, imagination and narrative video games. In: Vorderer, P., Bryant, J. (eds.) Playing Video Games: Motives, Responses, and Consequences. Lawrence Erlbaum Associates, Mahwah (2006)
17. Brousseau, G.: Theory of Didactical Situations in Mathematics. Springer, Netherlands (1997)
18. de Freitas, S.: Learning in immersive worlds. Bristol. Joint information systems committee (2006) (Last retrieved online in November 2009), http://www.jisc.ac.uk/media/documents/programmes/elearninginnovation/gamingreport_v3.pdf
19. Busselle, R., Bilandzic, H.: Fictionality and Perceived Realism in Experiencing Stories: A Model of Narrative Comprehension and Engagement. Communication Theory 18, 255–280 (2008)

A Real-Time Interactive System for Facial Makeup of Peking Opera

FeiLong Cai[1,2] and JinHui Yu[1,⋆]

[1] State Key Lab of CAD/CG, Zhejiang University, Hangzhou 310027
[2] Faculty of Arts and Communication, Ningbo University, Ningbo 315211
{caifeilong,jhyu}@cad.zju.edu.cn
http://www.cad.zju.edu.cn/home/jhyu/

Abstract. In this paper we present a real-time interactive system for making facial makeup of Peking Opera. First, we analyze the process of drawing facial makeup and characteristics of the patterns used in it, and then construct a SVG pattern bank based on local features like eye, nose, mouth, etc. Next, we pick up some SVG patterns from the pattern bank and composed them to make a new facial makeup. We offer a vector-based free form deformation (FFD) tool to edit patterns and, based on editing, our system creates automatically texture maps for a template head model. Finally, the facial makeup is rendered on the 3D head model in real time. Our system offers flexibility in designing and synthesizing various 3D facial makeup. Potential applications of the system include decoration design, digital museum exhibition and education of Peking Opera.

Keywords: Peking Opera; Virtual makeup; FFD; Inheritance and protection.

1 Introduction

First we provide a brief historical background on the facial makeup of Peking Opera. The origin of facial makeup used in Peking Opera can be traced back to the Southern and Northern Dynasties Period more than 1,400 years ago in China, when leading actors used to wear masks. With the development of operatic arts, performers gradually took off their masks and painted colorful patterns on their faces instead, so audience could better see their facial expressions. Facial makeup enables audiences to grasp the personality of a character portrayed and the character's social status at a glance.

Nowadays the computer plays a vigorous role in promoting the protection of non-material cultural heritages as well as its inheritance. In 2006, the Peking Opera was classified as the first batch of China's non-material world heritage. Although facial makeup is an abstracted and exaggerated art form, it has some specific laws in colors and patterns used. We use these specific laws to mimic the

⋆ Corresponding author.

Z. Pan et al. (Eds.): Transactions on Edutainment IV, LNCS 6250, pp. 256–265, 2010.

creation process of facial makeup. Our goal is to build a computer system for facial makeup of Peking Opera, which can help the hobbyists of Peking Opera to study the process of makeup, and inherit and protect this folk art with Chinese characteristics.

In this paper, we propose an interactive facial makeup editing and rendering technique to produce facial makeup on an input 3D mesh head model in real-time. At first, we analyze the process of drawing facial makeup and characteristics of the patterns used in it, and then built a SVG pattern bank based on local features like eye, nose, mouth, and so on. Next, we pick up some SVG patterns from the pattern bank and composed them to make a new facial makeup. We offer a vector-based FFD tool to edit patterns which mismatch with each other. According the user adjusted results, our system created automatically texture maps for a template head model. Finally, the facial makeup is rendered on the 3D head model with a GPU and CG shader language. Our system offers flexibility in designing and synthesizing various 3D facial makeup, potential applications of the system include decoration design, digital museum exhibition and education of Peking Opera.

2 Related Work

Existing makeup systems can be classified into 2D bitmap-based and 3D model based approaches. Most simulators such as FaceFilter Studio [1], Virtual MakeOver [2], and Makeup Pilot [3], adapt a 2D bitmap-based environment as a lightweight solution with a simple interface. These systems have a simple 2D interface which supports various makeup tools, such as foundation, lip-liner, lip-stick, eye-shadow and eye-liner. However, they have low quality graphics and limited interactivity. VF Pro [4] is a 3D system that uses a template face model to provide more convincing visualization of a face, it is the most clearly related existing system to our work, we extend it to support vector-based texture map. Adams et al. [5] proposed interactive 3D painting system that uses a haptic interface.

In recent years, making use of computer technology, some researches had been carried out to protect non-material cultural heritages. Yu and Peng [6] proposed a framework to synthesize Chinese calligraphy using brush texture patches; Liu et al. [7] made a papercut-pattern as a composition of multiple sub-patterns that are either reflection symmetric or reflection asymmetric to synthesize new papercut image, then Li al.[8] made a 3d paper-cut modeling and animation system. Ye al. [9] used texture synthesis and image inpainting to virtual repair of Chinese bronze articles. Other researchers have attempted to render 3D objects with a variety of artistic effects, such as pen-and-ink illustrations[10], watercolor[11], impressionist[12], Chinese painting[13] and Islamic star patterns[14]. But as far as we know, no research concerned about facial makeup of Peking Opera has been made. We aim to mimic the process of facial makeup in Peking opera, while the user can modify freely patterns of facial makeup and the ability to display result in real time. We first created a 3D face model with diffuse map, specular map and normal map using a 3D modeling application. Then we built a patterns bank. Finally, all data were integrated into GPU to render a facial makeup with OPENGL and CG shader language.

3 System Framework

We designed an interactive facial makeup system for Peking Opera, the main of which is laid out as follows: building a bank of SVG-based makeup patterns in Section 3.1. Next, in Section 3.2, we describe the FFD algorithms for editing makeup patterns; Composing textures are provided in Section 3.3. Finally, we describe a physically based specular reflectance model for skin with facial makeup in Section 3.4. Fig.1 is a graphic overview of our system.

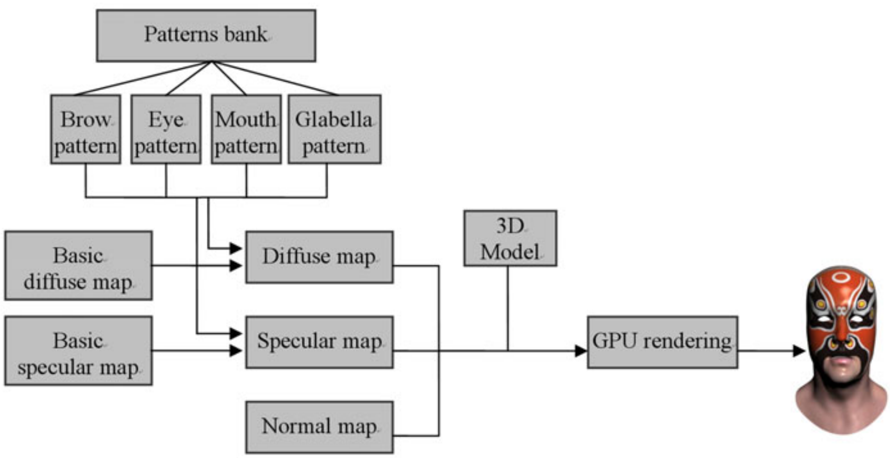


Fig. 1. Graphic overview of our system

3.1 Build a Bank of SVG-Based Makeup Patterns

Facial makeup refers to facial designs for Jing and Chou roles in Peking Opera. There are relatively few makeup patterns for a Chou role: the most common being a white nose for comic relief. The makeup pattern of Jing role are much more complicated and varied, such as the "whole-face", "three-tile face", "quartered face", "six-division face", "tiny-flowered face" and "lopsided face". Each pattern is rooted in his ability to reflect subtle and interesting changes in a human expression, and each pattern has its own symbolic meaning, and can reflect the identity, status, personality and appearance of the characters. Commonly, a facial makeup is divided into four basic parts: brow fossas, eye fossas, lipped fossas, glabella furrow, as shown in Fig.2 [15].

Fig. 2. Basic subobjects of a facial makeup

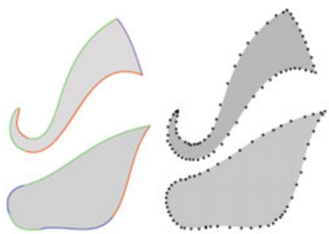


Fig. 3. Bzier based pattern and point based pattern

Fig. 4. Some patterns from the pattern bank

As an impressionistic and exaggerated art, types of facial makeup in Peking Opera is featured by painting brows, eyelids and jowls in various patterns such as bat, swallow wing and butterfly wing. Also, it is characterized by portraying facial expressions with exaggerated nasal fossas and lipped fossas. The age can be reflected by the height and shape of "Crow's-feet", temperament by the opening and closing of "Chordal furrow", and personality by different patterns of "Glabella furrow". Due to unchangeable rules in types of facial makeup, personalities of a character with certain facial makeup can be seen from the facial colors and figures. According to traditional facial makeup's categories of Peking Opera, we deconstruct a facial makeup into several pattern areas. And more, we notice that shape of these patterns is very smooth and curly, which can be modeled easily by Bzier curve. First, we manually draw these patterns with Bzier curves. Then, through De Casteljau Bzier subdivision arithmetic, each Bzier curve is subdivided to a group of vectorized points, as Fig.3 shows. When all

Bzier curve of a pattern are completed, it is saved into a pattern bank with SVG format. Here, we don't use bitmap image to store these pattern, because bitmap image will bring aliasing error when the user editing patterns. Fig.4 gives some patterns in the pattern bank.

3.2 Edit Makeup Patterns

As the user had chosen freely patterns with different style to compose a new facial makeup, there must be some places where patterns mismatch with each other, so some patterns need be adjusted in shape. For the convenience of the user interaction, we offer a technique of FFD for manipulating any shape in a free-form manner. Since all patterns are converted into groups of vector-based points, there is no problem of aliasing when they are being edited.

Pierre Bzier used FFD idea to manipulate large numbers of control points for Bzier surface patches, and the power of FFD as a modeling tool was more fully explored in [16]. In this paper we discuss mainly the 2D case of FFD. 2D FFD is a map from $R^2 \rightarrow R^2$. That is, it defines a new position for every point in a given region. Any lines or curves that lie in that region are thus altered.

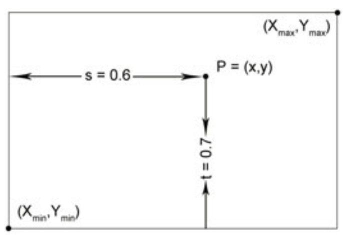


Fig. 5. The coordinates of a point in the deformation region range

We denote by (X_{min}, Y_{min}) and (X_{max}, Y_{max}) the corners of a deformation region, and by m and n the degrees of the FFD function. The deformation is defined in terms of a rational bivariate tensor product Bernstein polynomial which takes the form:

$$X(s,t) = \frac{\sum_{j=0}^{n}\sum_{i=0}^{m} w_{ij} B_i^m(s) B_j^n(t) P_{ij}}{\sum_{j=0}^{n}\sum_{i=0}^{m} w_{ij} B_i^m(s) B_j^n(t)} \tag{1}$$

where $B_i^m(s)$ and $B_t^n(s)$ are Bzier blending functions, and s and t are the local coordinates of a point with respect to the deformation region. The s and t coordinates of a point in the deformation region range between 0 and 1 (see Figure 5). Thus, for a point (x, y) within the rectangular region,

$$s = \frac{x - X_{\min}}{X_{\max} - X_{\min}}, t = \frac{y - Y_{\min}}{Y_{\max} - Y_{\min}} \tag{2}$$

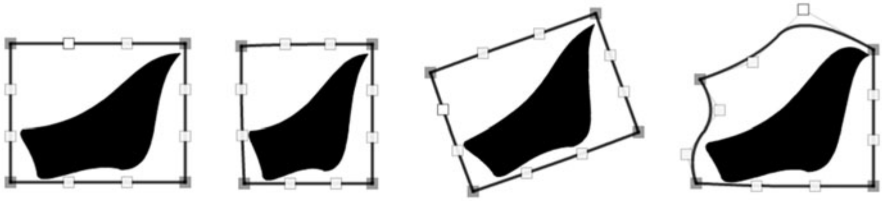


Fig. 6. Some results of using FFD transformation tool

To compute the position of a point which undergoes deformation, first compute its (s, t) coordinates using equation 2. If they are within the range [0, 1], then the point is repositioned using equation 1. Otherwise, it is not moved. Figure 6 show some results of using FFD transformation tool to a brow pattern: Fig.6(a) is a result of no any transformation, Fig.6(b) is a result of scaling in X-axis, Fig.6(c) is a result of rotation, Fig.6(d) is a result of free transformation.

Peking Opera uses different colors in facial makeup to exaggerate a performer's features. Commonly, each color has its own symbolic meaning. For example, red is the color of loyalty, integrity and courage. Black suggests a serious and taciturn disposition, including strength and roughness. White reveals a crafty and suspicious character, and gold and silver are sometimes used on the faces of immortals, demons and monsters. Different colors can also distinguish nobility from the common folk, goodness from evil or loyalty from treachery.

Meanwhile we found that a facial makeup has different highlight because of the difference of colors, especially gold and silver, and the difference of solvents for dissolution the color power. So we define an intensity attribute of highlight for each color. When the user complete all patterns adjustments and definitions for one facial makeup, he will get two texture images: a color image and a highlight gray image. Before rendering step, we need composing above two texture images onto the prepared textures.

3.3 Prepare Texture Maps

Because painting directly on the face, many features of human skin are remained. Now let us analyze the features of human skin. Typical types of skin features are wrinkles, pores, freckles, spots, moles, etc. Pores and wrinkles have 3D structures, freckles and spots have more or less 2D color variations, their optical properties strongly depend on their morphologies. They are clearly visible to our naked eyes, and we are very familiar with their appearances.

For mimicking these features of human skin, we prepare three basic texture maps: basic diffuse texture map Fig.7(a), basic specular texture map Fig.7(b) and normal texture map Fig.7(c). Now let explain the functions of these texture maps. Basic diffuse texture map is the color information of skin with freckles

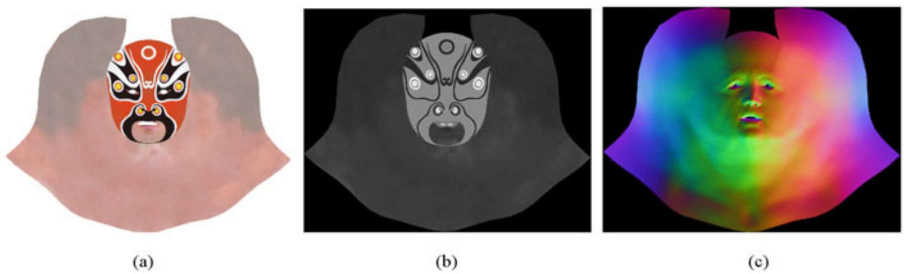


Fig. 7. Texture maps

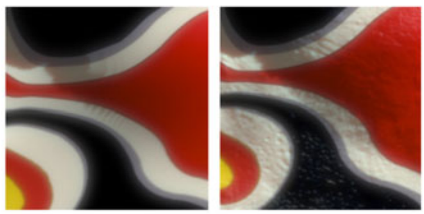


Fig. 8. A comparison of the rendering results with and without normal texture map

and spots, after adding facial makeup, diffuse texture map is completed. Basic specular texture is a gray image which is highlight distribution, after adding above mentioned the highlight attribute of color, specular texture map is composed. Normal texture map is the microfacet geometry information of wrinkles and pores, Fig.8 shows a comparison of the rendering results with and without normal texture map.

3.4 A Physically Based Specular Reflectance Model

Once all necessary data have been prepared, the rendering step is fully supported by graphics hardware and performed very fast with varying camera and lighting parameters.

Many physically based specular BRDFs from computer graphics literature can be used to improve the realism of skin rendering. In this paper, we adopt the Kelemen/Szirmay-Kalos specular BRDF model [17] based on the analysis of the photon collisions with the microfacets of the surface. This model is not only physically plausible, but provides other important features of real materials, including the off-specular peak and the mirroring limit case. Unlike most of other BRDF models, the proportion of the matte and specular components is not constant but varies with the viewing angle. The BRDF is

$$f_{r,TS} = \rho_s \frac{F(x, \overrightarrow{\omega_o}, \overrightarrow{\omega_i}) \cdot D(x, \overrightarrow{\omega_o}, \overrightarrow{\omega_i}, \sigma) \cdot G(x, \overrightarrow{\omega_o}, \overrightarrow{\omega_i})}{4(\overrightarrow{\omega_i} \cdot \overrightarrow{n})(\overrightarrow{\omega_o} \cdot \overrightarrow{n})} \quad (3)$$

where $\overrightarrow{\omega_i}$ and $\overrightarrow{\omega_o}$ are the incoming and outgoing lighting directions, F is the Fresnel reflectance, G is the geometry term, and D is the Beckmann microfacet distribution. The parameter σ controls the RMS slope of the microfacets, and is often referred to as the "roughness" parameter. We have found that values for σ in the range 0.2-0.4 would produce good results for skin, and we use σ=0.35 for all our results. ρ_s is used to control the oiliness of the skin surface..

A survey of human faces presented by Weyrich et al. [18] provides measured parameters for the Torrance/Sparrow specular BRDF model with the Beckmann microfacet distribution function. They assume such a model is valid for skin surface reflectance and measure roughness and intensity for ten regions of the face across 149 faces. The Torrance/Sparrow model is approximated closely by the Kelemen/Szirmay-Kalos model, and the measured parameters work well for either. Their data can be easily painted onto a face using a low-resolution two-channel map that specifies roughness and intensity for each facial region.

All physically based specular BRDF models contain a Fresnel reflectance term: This should be an unpolarized, dielectric Fresnel reflectance function with an F_0 parameter of 0.028. When computing a Fresnel term for a rough surface like skin, all θ terms should be measured from the half-angle vector H, and not from N.

The painting and oil in the outermost layer of skin are dielectric materials that reflect light without coloring it. In other words, the specular reflection of a white light from skin will be white, and the specular reflection of a colored light will be the same color as that light, regardless of the color of the underlying skin. Because of painting completely covered skins, almost all lights are reflected. So we assumed that no light transmits into the skin, subsurface scattering effect needn't be counted in our system.

4 Experimental Results

We implemented the BRDF rendering algorithm with OpenGL and CG Shading Language. The rendering examples in this paper were obtained on a Windows

Fig. 9. Rendering results with our system

PC with Intel Pentium processor E5200 with 2G memory; the graphic card is GeForce 8600 with 512M video memory. Consequently the rendering speed is quite fast, 25 to 30 frames per second for reasonably complicated meshes.

Fig.9 gives some rendering results of our system. Fig.9(a) (b) shows a comparison of the rendering results from different viewpoints; Fig.9(c) give a result in a red light; In Figure Fig.9(d) (e), giving results of two different color, blue and gold, which have different highlight attribute. In these examples, we show our system can let us observe the results from any viewpoint and any light condition in real time, and design freely all kinds of facial makeup.

5 Conclusions and Future Work

In this paper we presented a system to mimic the facial makeup of Peking Opera using an interface that integrates facial makeup editing with GPU rendering. The basic idea is first to analyze the characteristics of facial makeup and decompose facial makeup into some sub components that are structurally related to each other. After these patterns had been adjusted one by one, they were incorporated into the prepared texture maps. The step of the rendering are fully supported by graphics hardware and implemented by vertex and pixel shaders with multi-texturing. Using our system, the user can make a delicious facial makeup by picking up some pattern from the pattern bank and adjusting their shapes when necessary. The system gives instant feedback.

In the future research, we intend to make expressions and animations of the facial makeup. This may involve deformations of different patterns under varying face expressions.

Acknowledgments. The work described in this paper was supported by the grants from the National High Technology Research and Development Program of China (No. 2006AA01Z312), the National Natural Science Foundation of China (No. 60373007) and the Major Project of Science and Technology Development of Zhejiang Province (No. 2007C21043).

References

1. FaceFilter Studio 2, http://www.reallusion.com/facefilter
2. Virtual MakeOver, http://www.dailymakeover.com/makeover/virtualMakeover
3. Makeup Pilot, http://www.colorpilot.com/makeup.html
4. VirtualFashion Professional, http://www.virtual-fashion.com
5. Adams, B., Wicke, M., Dutre, P.: Interactive 3D Painting on Point-sampled Objects. In: Proceeding of Eurographics Symposium on Point Based Graphics, Grenoble, France, pp. 57–66 (2004)
6. Yu, J.H., Peng, Q.S.: Realistic synthesis of cao shu of Chinese calligraphy. Computers and Graphics 29, 145–153 (2005)
7. Liu, Y.X., James, H., Xu, Y.Q., et al.: Digital Paper cutting. In: Proceedings of ACM SIGGRAPH Sketches, Los Angeles, Article No.99 (2005)

8. Li, Y., Yu, J.H., Ma, K.L., et al.: 3D Paper-Cut Modeling and Animation. Computer Animation and Virtual World 18, 395–403 (2007)
9. Ye, X.F., Yu, J.H., Peng, Q.S.: Virtual repair of Chinese bronze articles. In: Proceeding of Computer Graphics International 2008, Istanbul, Turkey (2008)
10. Winkenbach, G., Salesin, D.: Rendeing parametric surfaces in pen and ink. In: Proceedings of ACM Conference Siggraph 1997, Los Angeles, pp. 91–100 (1997)
11. Curtis, C.J., Anderson, S.E., Seims, J.E., et al.: Computer-generated watercolor. In: Proceedings of ACM Conference SIGGRAPH 1997, Los Angeles, pp. 421–430 (1997)
12. Yu, J.H., Luo, G.M., Peng, Q.S.: Image based synthesis of Chinese landscape painting. Journal of Computer Science and Technology 18(1), 22–28 (2003)
13. Litwinowiz, P.: Processing images and video for an impressionist effect. In: Proceedings of ACM Conference Siggraph 1997, Los Angeles, pp. 407–414 (1997)
14. Kaplan, C.S., Salesin, D.H.: Islamic star patterns in absolute geometry. ACM Transactions on Graphics 23(2), 97–119 (2004)
15. Zhao, M.L.: Chinese Facial makeups in Peking Opera. Zhaohua Press, Beijing (2003)
16. Sederberg, T.W., Parry, S.R.: Free-form fundamentals of solid geometry. Computer Graph 20(4), 151–160 (1986)
17. Kelemen, C., Lszl, S.K.: A Microfacet Based Coupled Specular-Matte BRDF Model with Importance Sampling. In: Eurographics Short Presentations, pp. 25–34 (2001)
18. Weyrich, T., Matusik, W., Pfister, H., et al.: Analysis of Human Faces Using a Measurement-Based Skin Reflectance Model. ACM Transactions on Graphics 25(3), 1013–1024 (2006)

Design of Educational Game: A Literature Review

Ting Wei[1] and Yi Li[2,*]

[1] School of Journalism and Communication, Nanjing Xiaozhuang University, Jiangsu, Nanjing, 210017, China
[2] Educational Game Research Center of Nanjing Normal University, Jiangsu, Nanjing, 210097, China
weiting@yahoo.cn,yilisd@163.com

Abstract. This review focuses on educational game design, including the developmental stage of design thought, design theories, design approaches, etc. Specifically it highlights the balance and integration between educational character and playfulness in educational games. This review also suggests limited research areas in existing literature and illuminates the possible further research in educational game design.

Keywords: educational game, three generations, situated cognition, Edugaming.

1 Introduction

The research of educational games (computer and video games) can be traced back to the middle of 1950s, when attention was mainly paid to the design, development and application of commercial video games [1]. Till 1980s, some researchers began to concentrate on the educational potential of video games, studying how it motivated players' intrinsic motivation and trying to apply it into instruction. At the end of 1990s, with the development of computer technology and network, the research was not only restricted to video games, but extended to the field of electronic games and large-scale online games. In the late 1990s, the design, development and application of educational games were gradually concerned in the field of instruction and game industry.

We began our review with an exhaustive search of electronic databases. This search led us to several meta-analysis and reviews including those conducted by Kebritchi,M.et al. and Kowit Rapeepisarn et al. etc. on designing theories and practice in educational games. From this initial search, we began a more extensive review using basic search procedures and standard criteria for a comprehensive literature search. This process included library searches (both electronic and manual) in educational databases such as ERIC, PsycInfo, and Social Sciences Index using search terms including educational game design, computer game design, video game design, and so on.

The design is a key to the balance between educational potential and playfulness in educational games. This article comments the history of educational game design thought, design theories, design approaches, as well as design and development models

* Corresponding author.

Z. Pan et al. (Eds.): Transactions on Edutainment IV, LNCS 6250, pp. 266–276, 2010.

aiming to seek a balance between education and game. At the end of the article, the new trend in the field of educational game design is put forward.

2 Developmental Stages of Educational Game Design: Three Generations by Egenfeldt-Nielsen

Although the history of educational game research is relatively short, with the support of different philosophies of education and learning theories, the design pattern changed remarkably. Most of all, the impact of learning theories is the most profound. From Behaviorism to Cognitivism, then to Situated Learning Theory, it guides different educational game design. Egenfeldt-Nielsen [2] concluded the developmental stages of designing educational games according to the relationship between educational games and learning theories, and divided them into three generations, as shown in Figure 1:

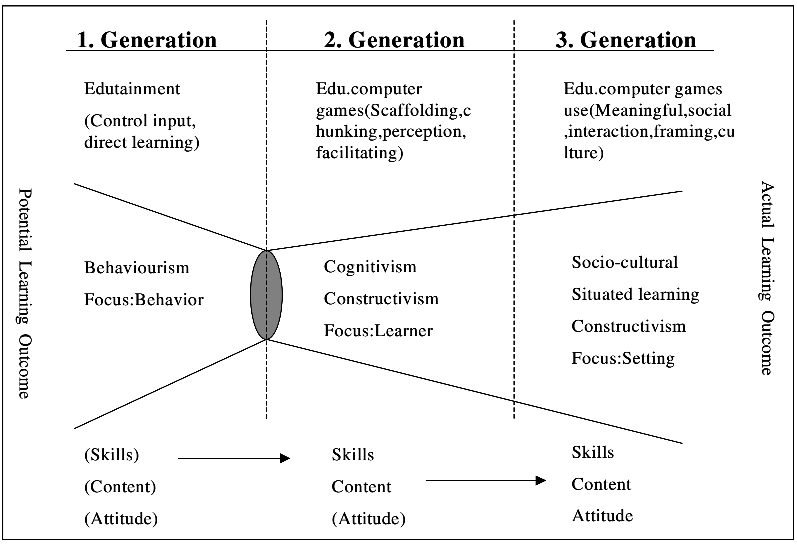


Fig. 1. The model shows the characteristics of three generations, and how they emphasize different learning theories [2]

The ideology of the first generation was mainly based on the early edutainment that assumes that learning occurs when learners unreflectively practice a skill enough times. Behaviorism was its basic theory, which focused on the behavior of learners, aiming at improving their action skills through games. The products of this period were mainly drill games, in which designers made the simple interaction of games as the reinforcement of learning.

The second generation was designed on the basis of Cognitivism which criticized the automatic relation presented in Behaviorism between stimuli and response. Learners became the center of attention. Designers began to pay attention to the designing of learning environment and the motivation of learning. Then there emerged a batch of educational games which integrated both the game and the learning relatively well, covering some courses including history, geography, and agriculture and so on. The design of educational games in second period was focused on learners, emphasizing that learners were the active knowledge constructors.

The third generation was based on Constructionism with a strong focus on learners, as well as setting what provided a platform for exploring new material, mostly individually, and some collaboratively. This is further stressed in situated learning and the social-cultural approach, where the learning process is regarded as mediated in a social context. Different learners have different understanding with the same knowledge under different circumstances; knowledge is part of activities, background and cultural products. The social, cultural and interactive activities play important roles. The third generation design plays a dominant role in current educational game design.

The three-generation division of Egenfeldt-Nielsen shows the designing ideology and change of the designing focus during different generations in educational game design. Egenfeldt-Nielsen emphasized that each generation is carried forward to the next, but de-emphasized. The learning mechanisms in Behaviourism are still partly at play in next-generations' titles and so are the ones from Cognitivism, but they are conceived in a broader overall frame [3]. They will be more or less adequate for explaining different aspects of learning.

3 Theories for Designing Educational Games

3.1 Situated Cognition

The theory of situated cognition suggests,"…that activity and perception are importantly and epistemologically prior, at a non-conceptual level, to conceptualization and that it is on them that more attention needs to be focused"[4].Knowledge is situated in its context, more specifically, knowledge is a product of its context, activity and culture within which it is developed and used.

In a symposium on learning theories for the analysis of educational games, Halverson, Shaffer, Squire, and Steinkuehlr[5]asserted that situated cognition provides a meaningful framework for the study of games, given that games have an ability to situate learning in an authentic context and engage players in a community of practice. In educational games designed well, educational content and context are smartly designed as a game-play environment for a learner to acquire various knowledge and skills in a particular subject domain. Gee [6] also considered when learners are engaged in games, they are able to form understandings based on activity and experience, understandings customizable to specific contexts of use, and from this basis they can eventually generalize their knowledge without losing the grounding of that knowledge in specific applications.

3.2 Flow

Flow is a term coined by M.Csikszentmihalyi. He has developed an empirically based model of enjoyment and an accompanying methodology called experience sampling method, which measures the qualities of enjoyable activities in various contexts of day-to-day life. Flow describes a state of complete absorption or engagement in an activity and refers to the optimal experience [7].

The flow elements are factors that contribute to the flow state and should be considered in educational game design. The elements of flow can be divided into three groups: flow antecedents, flow state (or flow experience) and flow consequences [8]. The state of flow includes "immersion, higher balance of challenge and skill, and control & satisfaction" [7]. It is such a merging of the player's total attention and the task at hand that all other sensory and cognitive distractions are invisible to the player. In these cases, the player's attention is totally drawn to the game environment and it is very difficult to distract him/her. The player is unaware of time passing, and may later remark on this.

According to Kiili[9], whenever people reflect on their flow experiences, they mention some, and often all of the following characteristics: concentration, time distortion, loss of self-consciousness, and sense of control. So in order to create flow in the educational game, Jones [10] outlines some methods. For example, reduce cognitive load on environmental operations and low-level cognitive tasks so that learners can concentrate on tasks; provide problems that are relevant to the learner and the content; tasks and information must flow smoothly from one to the other, and there can be no disjointed experience, such as stopping to figure out what a particular button does in the middle of task.

Generally, the aim of an educational game is to provide learners with challenges that match their skill level. Furthermore, challenges should be related to the main task so that flow experience is possible.

3.3 Experiential Learning Theory

Experiential learning builds upon the work of Piaget, Lewin, and Dewey Experiential learning theory defines learning as "the process whereby knowledge is created through the transformation of experience. Knowledge results from the combination of grasping and transforming experience"[11].During experiential learning, educators purposefully engage learners in direct experience and direct their focus on learning reflection to increase their knowledge, skills, and values. Experience occurs as a result of interaction between human beings and the environment in forms of thinking, seeing, feeling, handling, and doing. Experiential learning theory stresses the importance of direct experience and reflective observation.

Kebritchi,M.et al.[12] argued that computer games, which may be designed in the context of everyday life, can connect the players to every day life experience. Such concrete experience is the heart of the experiential learning approach in which knowledge is constructed, not transmitted, as a result of experiencing and interacting with the environment. Heidenreich [13] had an experiential learning course on the mathematics of games, and he asserted experiential learning is a pedagogy in which learners are put into an experience-rich environment and then construct their own learning as a community.

4 The Approach of Educational Game Design: Edugaming

How to achieve the balance of education and playfulness has always been a question that researchers try to solve in educational game design. On one hand, educational game designers get references from commercial games design. On the other hand, they devote themselves to developing new design methods of educational games, hoping that new methods can support learners' learning process in a better way. Fabricatore,Shang Junjie and David Ghozland etc. put forward "Edugaming" design approaches, and gave strategies from different perspectives.

4.1 The Content of "Edugaming"

Fabricatore[14] proposed "Edugaming" design approaches from the perspective of integration between learning task and game context. As suggested by the name, "Edugaming" is a combination of education and game. Fabricatore believed that education should be naturally integrated with games; the learning task needs to be taken into account as a factor of the game, and be apperceived by learners. At the same time, he also pointed out that, in order to design the game for instructional purpose, some related rules and outcomes of learning should be developed. Fabricatore found some educational games lacking the link between the game and the cognition tasks. Then, he advocated a new strategy for educational game design that is to create a simulated environment as well as a game experience, in which the knowledge could be integrated naturally. From the aspect of game, the learning task should be broken into steps in game context. As a factor of play, task can be perceived by learners. Fabricatore named this designing approach "Edugaming", which means there is no barrier between learning and playing naturally. Learning is human's nature, and the challenging play is just the process of exploring and learning.

Some researchers proposed some methods and strategies for designing "Edugaming", respectively from the aspects of the psychological demand of learners, challenging level of games and the learning objective, etc[15~19]. For example, during the design of educational games, providing learners chances of practicing skills should be taken into account; the game can be simple at the beginning in order to trigger learner's interests; the learner's demand can be considered when the game's pace and sustained time get settled; be sure that the game structure adapts to the learning objective; meanwhile, try to avert excessive objectives and obstacles that prevent learners from achieving goals; learners can control learning tools in order to meet their needs; constantly various types of challenge and rules of scoring, supply different levels of challenge; ensure the game could achieve upgrade through various skill; provide different feedbacks to encourage learners to pay attention to the process; encourage learners to review, evaluate and practice in games through discussing and assessing.

"Edugaming" offers a research way for the natural melting of education and game. However, it is in its initial stage. It still needs continuous exploration and development to put it into effect. Many researchers like Shang Junjie, Chong et al., and Kowit Rapeepisarn et al. studied Edugaming from different perspectives and put forward specific design methods.

4.2 Learners' Game Behavior and Edugaming Design

Shang Junjie et al. [20] started an empirical study about the features of learners' behavior during playing game in a case of the educational game named "VISOLE". The result showed that there occurred independent learning, collaborative learning and research learning in playing "VISOLE" game. However, there were also some deficiencies such as excessive trial and error, inefficient collaboration, superficial introspection, etc. According to learners' learning behavior features, he classified learning behavior into six types: creative learning, exploring learning, collaborative learning, trial and error learning, inquiring learning and entertaining learning.

According to the research, Shang, Junjie put forwarded some game design strategies in “VISOLE” (see Table 1).

Table 1. The genres of learners and game design strategies in VISOLE

The genres of learner’s game behavior	Design strategy
Creative learners	Provide the virtual environment for learners testing their “creative” ideas; set some bugs that can be controlled for learners exploring.
Exploring learners	“Force” the learner to change her/his behavior habits by some setting; let NPC question him/her for leading to explore.
Collaborative learners	Add effective interaction and communication.
Trial and error learners	Provide the learner supporting and helping automatically; NPC give some advice to enlighten the learner just like the teacher.
Inquiring learners	Set some units for discussing, collaborating to engage the learner to play the different roles.
Entertaining learners	Give the learner choices, challenges, as well as an engaging story.

Chong et al. [21] conducted a research about the impact of learning styles on the effectiveness of game-based learning. Kowit Rapeepisarn et al. [22] researched learning content, learning activity, learning style, learning behavior, game types and the relationships among each other. To design and develop an effective educational game, different game genres, learning activities and techniques, and learning styles are important issues for consideration.

4.3 Learners' Motivation and Edugaming Design

Motivation is a significant characteristic of educational games and that effective game design considers both intrinsic and extrinsic rewards for play. Intrinsic motivation pushes people to act freely on their own; and extrinsic motivation pulls people to act due to factors that are external to the activity itself, like reward or threat.

Bowman, Provenzo and Malone focused on the intrinsic motivation of learners and the game design for stimulating motivation. According to Bowman’s [23] research, learners’ response of vision and hearing in the games can not adequately explain the attraction of games. He supposed that the popularity of some computer games should be attributed to the flow experience which players processed in the games. Bowman thought the basic design of motivation is implemented by offering a certain mission to the player, the player’s choice and the player’s skill.

After studying the design of various video and computer games, Provenzo[24] pointed out that the reason why the educational game was so attractive was that most of games have goals, and few of them have a negative ending. At the same time the research made by Malone [25] showed that the motivation elements of a successful educational game include: challenge, control, fantasy, and curiosity. He pointed out that not all the games should contain these elements. However, each element would enhance the entertainment of educational games.

The research conducted by Bowman, Provenzo and Malone showed a common viewpoint---having a clear goal and mission, strengthening feedback, enhancing challenging, these elements could constantly inspire players and keep their motivation.

In a project about educational computer games, British Educational and Communications Technology Agency (BECTA) found the reason why learners are always involved in games was that these games could inspire the motive of learners [26].

Based on their research, BECTA classified relative content on the motivation in educational games (see Table 2).

Table 2. BECTA's findings on the motivation in educational games [26]

Questions	Findings About Motivation in Educational Games
What indicates motivation?	Independent work
	Self-directed problem posing
	Persistence
	Pleasure in learning
What generates motivation?	Active participation
	Intrinsic and prompt feedback
	Challenging but achievable goals
	A mix of uncertainty and open-endedness
What can motivation usefully support?	Collaborative interaction
	Peer scaffolding of learning
	Creative competition or co-operation
	Equal opportunities

Besides, David Ghozland[27] put forward the PNRC system for designing motivation in games. In this system, P means player state, that is the state of game variables of the player's avatar, for example his life, armor, and the quality of his equipment, etc. N are the needs at the moment when the challenge arises. These needs depend on the player's state and on his advancement into the game. R is the player's expectation of the reward, and it depends on the estimated difficulty and also on the player's past experience with the reward system. C is the player's expectation regarding the challenge. Most games build their motivation on the four PNRC functions. Because these functions are interdependent, so to succeed in an efficient management of motivation designers need to balance the four parameters.

5 Design and Development Models of Educational Games

A lot of research has been done in designing and developing educational games especially in how to keep the mergence and balance between education and playfulness in educational games. Designers are always seeking the right way to solve this problem. Some design models, development models and steps have been proposed.

Kristian Kiili[9] put forward a experiential gaming model combining education with play according to the theory of experiential learning, flow theory and game designing. The experiential gaming model consists of an ideation loop, an experience loop and a challenge bank. The challenges based on educational objectives form the heart of the model, which is to sustain the motivation and engagement of the player by pumping appropriate challenges to him or her. To overcome the challenges, a player generates solutions in the ideation loop. After the ideation phase the player tests solutions in the experience loop reflecting greater circulation and observes the outcomes of actions. Game should be usable and provide clear goals and appropriate feedback to the player in order to facilitate flow experience.

The experiential gaming model emphasized immediate feedback, clear goals and the suitable challenges for learners during playing. In the model the flow theory is used as a framework to facilitate positive user experience in order to maximize the impact of educational games. The experiential gaming model can be used to design and analyze educational games. However, the model works only as a link between educational theory and game design and does not provide the means to a whole game design project. In the future, the experiential gaming model need be tested and further developed.

Besides Kiili's experiential gaming model, Garris[28] and Jim Bizzocchi & Brad Paras[29] proposed their models according to the flow theory, but they had some differences. The IPOGM (Input-Process-Outcome Game Model) put forward by Garris stressed the psychological experience of learners when they are entirely involved in the game and ignore the things which have nothing to do with the games. By doing this, learners can excel by challenging themselves. Jim Bizzocchi & Brad Paras proposed the model which combined play, motivation and efficient learning, and the model emphasized the experience designing introduced by learners' inner motivation during playing.

The GOM (Game Object Model), presented by Amory and Seagram, integrates educational theory and game design and outlines a systematic approach needed to develop sound educational games. In the GOM, Amory et al. [30]consider an educational game to consist of a number of components each of which is described through abstract and concrete interfaces(represented by circles: abstract, black; concrete, white). Abstract interfaces refer to all pedagogical and theoretical constructs, and concrete interfaces refer to design elements. Therefore educational game designers make use of the abstract interfaces in the conceptualization phase of game design while game developers realize these pedagogical aspects of an educational game by including the concrete interfaces into the game software and gameplay. The GOM has successfully been used to design academic adventure games.

The development of a number of models to explore the relationships between educational theory and game design provides developers with a conceptual and practical framework that can support the development process. Future research should investigate these models as tools to both developing educational games, and to evaluating the educational value of traditional computer games.

6 Discussions and Conclusions

This review about the design of educational games, with focus on the balance between pedagogy and game, investigates the developmental stage of design thought, design approaches, structural elements, and design models. The following can be concluded from the literature review made in this paper:

- The three-generation model of Egenfeldt-Nielsen shows the characteristics of three generations, and how they emphasis different learning theories. The design ideology based on Constructionism leads to the current educational game design, which emphasizes meaningful, social interaction and so on. Three-generation division of educational game design mentioned by Egenfeldt-Nielsen has offered a historical view for research, helping the researchers to outline the development features and directions of educational games in different historical periods.
- “Edugaming” proposed by Fabricatore is a series of design approaches from the perspective of integration between learning task and the game context including challenge levels design, learning objectives design, and game feedbacks, etc. Shang, J.J. designed the educational game in the light of the genres of learner’s game behavior. Bowman, Provenzo and Malone found some elements, e.g. a clear goal and mission, strengthening feedback etc., could constantly motivate players. However, the games studied by Bowman and Malone etc. were relatively simpler compared with current computer games. Nowadays, the design of video and computer games has changed a lot. The way of narrating, the role playing, the multi-user, and the representation of three-dimensional space far more exceed those used in the games they researched. So, further research will be done on current computer games, especially on online games.
- Along with “Edugaming”, there are a lot of things worth exploring. For example, how to explore an educational game design suitable for a certain kind of learning content in the light of different learning types and learning outcomes, and how to effectively find a game's features such as goal and rule so as to integrate learning content with learning goal?
- The production of educational games is both complex and technically challenging. The use of models such as experiential game model, IPOGM, and GOM, etc. allows for the conceptualization and assessment of educational games based on contemporary educational ideas. More research should be made to investigate whether these models can be effectively used to design and develop educational games.

The design of educational games is the key to the balance between education and games. Due to its special significance and inner complexity, research can’t obtain instant effects. On the contrary, the design of educational game is an attractive field that deserves long-time research and thorough exploration.

In future research, communities and groups may bring additional dimensions to educational game design. In online educational game, social mutuality becomes an important factor to fascinate players. The social factors, such as competition and cooperation, identity presentation, the association among different groups and the distribution of knowledge, should be taken into account. As for the researching methods, quality analysis is mainly employed, and quantitative research and empirical research are far from enough.

References

1. Butler, T.: Games and simulations: Creative education alternatives. TechTrends 33(4), 20–24 (1988)
2. Egenfeldt-Nielsen, S.: Beyond Edutainment. Unpublished dissertation, University of Copenhagen, Copenhagen, Denmark (2005)
3. Egenfeldt-Nielsen, S.: Third generation education use of computer games. Journal of Educational Multimedia and Hypermedia 16(3), 264–281 (2007)
4. Brown, S.B., Collins, A., Duguid, P.: Situated Cognition and the Culture of Learning. Educational Researcher 18(1), 32–42 (1989)
5. Halverson, R., Shaffer, D., Squire, K., Steinkuehlr, C.: Theorizing games in education. Paper presented at the 7th international conference on Learning Sciences, Bloomington, IN (2006)
6. Gee, J.P.: Situated Language and Learning: A Critique of Traditional Schooling. Routledge, London (2004)
7. Csikszentmihalyi, M.: Flow: The Psychology of Optimal Experience. Harper Collins, New York (1990)
8. Chen, H., Wigand, R., Nilan, M.S.: Optimal Experience of Web Activities. Computers in Human Behavior 15, 585–608 (1999)
9. Kiili, K.: Digital Game-Based Learning: Towards an Experiential Gaming Model. Internet and Higher Education, 13–24 (2005)
10. Jones, M.G.: Creating Engagement in Computer-based Learning Environments (1998), http://itech1.coe.uga.edu/itforum/paper30/paper30.html (accessed March 2009)
11. Kolb, D.: Experiential Learning: Experience as the source of learning and development, p. 256. Prentice Hall, Englewood Cliffs (1984)
12. Kebritchi, M., Hirumi, A., Bai, H.: The Effects of Modern Mathematics Computer Games on Mathematics Achievement and Class Motivation. Computers & Education (2010)
13. Heidenreich, J.R.: The Mathematics of Games as Experiential Learning. Paper presented at the annual meeting of the Mathematical Association of America MathFest, TBA, Madison, Wisconsin (2008), http://www.allacademic.com/meta/p274088_index.html (accessed January 2010)
14. Fabricatore, C.: Learning and Video Games:an Unexploited Synergy (2000), http://www.learndev.org/dl/FabricatoreAECT2000.pdf (accessed May 2009)
15. Burguillo, J.C.: Using Game-Theory and Competition-based Learning to Stimulate Student Motivation and Performance. Computer & Education (2010)
16. Robertson, J., Howells, C.: Computer game design: Opportunities for successful learning. Computer & Education 50, 559–578 (2008)
17. Yue, T., Ling, Z.: On Research of Designing Digital Gaming for Education. Primary and Middle School Educational Technology 4, 27–29 (2007)

18. Yun, R.W., Pan, Z.G., Li, Y.: An Educational Virtual Environment for Studying Physics Concept in High Schools. In: Lau, R., Li, Q., Cheung, R., Liu, W. (eds.) ICWL 2005. LNCS, vol. 3583, pp. 326–331. Springer, Heidelberg (2005)
19. Yun, R.W., Wang, M., Li, Y.: Design and Implementation of Game-Based Learning Environment for Scientific Inquiry. In: Pan, Z., Zhang, X., El Rhalibi, A., Woo, W., Li, Y. (eds.) Edutainment 2008. LNCS, vol. 5093, pp. 60–69. Springer, Heidelberg (2008)
20. Junjie, S., Jong, M.S.Y., Lee, F.L., Lee, J.H.M.: Case Study on Learners' Behaviors in Game-based Learning and the Enlightenment for Educational Game Design. China Educational Technology 2, 65–71 (2008)
21. Chong, Y., Wong, M., Thomson Fredrik, E.: The Impact of Learning Styles on the Effectiveness of Digital Games in Education. In: Proceedings of the Symposium on Information Technology in Education, KDU College, Patailing Java, Malaysia (2005)
22. Rapeepisarn, K., Wong, K.W., et al.: The Relationship between Game Genres,Learning Techniques and Learning Styles in Educational Computer Games. In: Paper presented at the 3rd International Conference on E-learning and Games, Nanjing, China (2008), http://liama.ia.ac.cn/Edutainment2008/acceptance.html (accessed May 2009)
23. Bowman, R.F.: A Pac-Man Theory of Motivation: Tactile Implications for Classroom Instruction. Educational Technology 22(9), 14–17 (1982)
24. Provenzo, E.F.: Video Kids:Making Sense of Nintendo. Harvard University Press, Cambridge (1991)
25. Malone, T.W.: Toward a Theory of Intrinsically Motivating Instruction. Cognitive Science (4), 333–369 (1981a)
26. BECTA: Computer Games in Education project, http://www.BECTA.org.uk/research/research.cfm?section=1&id=2826 (accessed Feburary 2009)
27. Ghozland, D.: Designing for Motivation (2008), http://www.gamasutra.com/view/feature/1419/designing_for_motivation.php (accessed May 2009)
28. Garris, R., Ahlers, R., Driskell, J.E.: Games, Motivation, and Learning: A Research and Practice Model. Simulation & Gaming 33, 441–467 (2002)
29. Bizzocchi, J., Paras, B.: Game, Motivation, and Effective Learning: An Integrated Model for Educational Game Design (2005), http://ir.lib.sfu.ca/retrieve/1648/628be2dd11d8b6e8d3157783c9b7.doc (accessed Feburary 2009)
30. Amory, A., Seagram, R.: Educational Game Models: Conceptualization and Evaluation. South African Journal of Higher Education 17, 206–217 (2003)

Author Index