CORE
Metadata, citation and similar papers at core.ac.uk
Provided by Elsevier - Publisher Connector

ELSEVIER
Available online at www.sciencedirect.com
SciVerse ScienceDirect
Procedia Computer Science 12 (2012) 200 – 205
Procedia Computer Science

Complex Adaptive Systems, Publication 2
Cihan H. Dagli, Editor in Chief
Conference Organized by Missouri University of Science and Technology
2012- Washington D.C.

Utilizing Depth Based Sensors and Customizable Software Frameworks for Experiential Application

Brian Moriarty[a], Elizabeth Lennon[b], Frank DiCola[a], Kyle Buzby[a], Morisa Manzella[a], Emily Hromada[a]

[a]*Stevens Institute of Technology, 1 Castle Point on Hudson, Hoboken, NJ, USA 07086*
[b]*Information Sciences Institute, USC, Marina Del Rey, CA 90292*

Abstract

Using depth sensor cameras such as the Kinect and highly customizable software development frameworks in conjunction with artificial intelligence methodologies offer significant opportunities in a variety of applications, such as undergraduate science, technology, engineering, and math (STEM) education, professional or military training simulation, and individually-tailored cultural and media arts immersion. Designing a participatory educational experience where users are able to actively interface and experience given subject matter in a practical experiential manner can enhance the user's ability to learn and retain presented information. Such natural gesture user interfaces have potential for broad application in disciplines ranging from systems engineering education to process simulation. This paper will discuss progress on the development of testing environments for interactive educational methods in conjunction with artificial intelligent systems that have the ability to adjust the educational user experience based on individual user identification. This will be achieved through depth sensor skeletal tracking, allowing experience adaptation based on the nature and effectiveness of the interactive educational experience.

Keywords: Kinect; Natural Gesture Interfaces; STEM education; Process Simulator; Depth Sensors; Skeletal Tracking, Artificial Intelligence

1. Introduction

1.1 Immersive Visual Environments (IVEs)

Immersive interactive technologies, which have their basis in environmental simulators and Cave automatic virtual environments (CAVEs), provide users with a vector for situational experience via a software-based environment [1]. Since their inception, extensive research has been done to extend the functionality of CAVE Environments to other purposes, such as education, entertainment, and simulation [2]. Advances in depth sensor cameras and natural gesture human interfaces provide a more lifelike alternative to outmoded object-based control systems for exploring virtual environments [3]. Developments in motion sensing technology and high-resolution image capture have accelerated research in new interface methods for experiential applications. Immersive Visual Environments (IVEs), such as the iDome [4], Virtual Room [5], and Advanced Visualization & Interaction Environment are environments that allow kinesthetic and embodied experiences for education. Each of these representation methods utilize imaging technologies in order to enhance interactive experience: the iDome provides

1877-0509 © 2012 Published by Elsevier B.V. Selection and/or peer-review under responsibility of Missouri University of Science and Technology.
Open access under CC BY-NC-ND license. doi:10.1016/j.procs.2012.09.054

a means to stimulate the full field of human view, the Virtual Room provides participants with a full immersive 360 stereoscopic environment, and the Advanced Visualization & Interaction Environment provided an early methodology for constructing interactive virtual spaces.

1.2 Depth Sensor Cameras

In addition to advancements in 3D virtual environments, depth sensors have been on the market since the year 2000 when the first time of flight (TOF) camera was released under the name ZCam by 3DV Systems. The very first ZCam was released with the intention of generating a live key of a scene and providing mattes to the foreground or background with the thought of future abilities to provide gesture recognition, among other features [6]. Since then, the company that started the movement in depth sensing cameras has been bought by Microsoft accelerating the development of advanced depth sensing cameras [7].

The field is currently being led by Microsoft with its red, blue, green, and depth (RGB-D) sensor camera, the Kinect. However, competitors have come out with their own version of this sensor, such as Panasonic with the D-Imager, a device that serves the same purpose, but gathers the data in a much different way. The technology employed in the Kinect utilizes a single infrared (IR) projector to disperse an array of IR beams that are then gathered by the Kinect system's receiver. The D-Imager does not include a RGB sensor, but can provide greater detail in the depth information with its specialized charge-coupled device (CCD) sensor. The sensor receives individual beams of IR emitted from a wide array of light-emitting diodes (LEDs). This technology allows for the Panasonic devices to have the ability to specialize for specific environments. The "High Illumination type" model of the D-Imager allows for the device to be in ambient lighting 5 times greater than that of the regular or "High Precision type" model with the addition of six LEDs to the array [8]. Some work has been done to utilize multiple Kinect sensors but calibration is necessary and extra image processing is required due to image blurring [9]. However, expanding upon the CCD and LED array technology of the D-IMager a full virtual environment could be modeled using the sensors through a single CCD sensor receiving data from multiple arrays of LEDs strategically placed.

The current software development kits (SDKs) released by Microsoft and Omek Interactive have been successful with general gesture recognition and control [10, 11]. Although the use of gestures for human computer interaction can be convenient, they can be difficult due to the complex forms the body, specifically the hands, can take [13]. The future of these depth sensing cameras will address these issues as resolution of the sensors is increased and the entire environment can be analyzed.

Making its debut in May 2012, the Leap Motion may be the beginning of the new era of depth sensing cameras. Although the device has a minimal range, it allows for near field gesture controls at an even greater accuracy. Claiming to be over "100 times more accurate than any motion sensing/natural user interface on earth" with "the ability to track movements down to the 1/100th of a millimeter" the Leap Motion could address the issues that Chaudhary references due to the complexity of human form [12, 13].

1.3 Explorable Visual Environment (EVE)

Through the development of Explorable Visual Environments (EVE), we have extended IVE operational concepts further by merging several technologies, including the Kinect depth sensor camera technology, into a single interface, providing an environment that is immersive, interactive, and informational in nature. In utilizing Kinect gesture recognition, interactive augmented reality informational output, as well as the ability to walk through environments in a manner approximating actual experience, the participant is provided with a greater contextual understanding of an environment than is otherwise possible. In effect, the immersive architecture of EVE is a tool for providing tangible knowledge and experience of a given subject area. As described in previous publications, the continuing goal of this work is to encourage the user to immerse in and investigate the virtual space, a virtual space that may be tailored to teaching specific fields of study [14]. The EVE system extends the concept of an IVE by encouraging users to interact with the visual environment before them. Whereas the content for most IVEs consists of linear narrative, the subject matter displayed by EVE allows users to pick and choose parts that interest them, and spend as much time as they want exploring their choices. By providing an environment that is fully interactive and navigable, education is based heavily on a user's curiosity and eagerness to learn, rather than their ability to pay attention for extended periods of time. The non-linear style of learning brought about by an interactive environment allows users to skip material they already know, and select topics relevant to their individual needs and interests. Using depth-sensing cameras for gesture recognition, the natural interface utilized by EVE, ensures an easy learning curve for all age brackets. Another added benefit of this system of control is the feeling of unity the learner experiences with the subject they are studying, one of the primary goals of EVE [15]. Explorable educational interfaces have the potential to accelerate learning through enhanced student engagement in turn improving the overall quality of the learning experience."

1.4 EVE in undergraduate STEM Applications

To remain globally competitive and address the growing challenges of the twenty first century across sectors such as healthcare, energy and defense, it is critically important that the United States increase its percentage of STEM professionals [16, 17, 18]. However, of all students starting a STEM program as an undergraduate, approximately 15% complete his or her degree program, and 30% to 35% of those starting in STEM field switch to alternate degree programs outside of STEM disciplines. Though the factors contributing to STEM degree completion rates vary, they clearly indicate one way of increasing the number of STEM professionals in the near-term is to target the retention of students in undergraduate STEM programs [17]. A critical aspect of achievement and completion of STEM-related curricula at any level (elementary, secondary, or post-secondary) is the student's degree of engagement with course content [19].

Many, if not all, of the STEM undergraduate courses taught involve oral lectures and/or Microsoft Powerpoint presentations by the professor to the students. Many students find this mode of learning boring, in turn impacting their understanding and retention of the presented materials. Lack of interest and difficulty in learning the course content are factors contributing to low rates of STEM degree completion. Though several online, educational game-based systems, e.g. Mathletics (http://www.mathletics.com) or Tutpup (http://www.tutpup.com/) have emerged as self-described, successful lesson aids, they may be fee-based and are still constrained by an individual computer display and conventional mouse and keyboard interface devices. Furthermore, many of these systems are geared for elementary and secondary school levels. EVE is unique in that it is initially focused on undergraduate STEM courses and enables interactions in three dimensions with relevant content via natural gestures. To address this factor, This paper details progress on the ongoing development of the EVE system in which users can interact with visual and or audio content using a natural-gesture based interface. In this work, EVE's flexible software framework and natural gesture library (described further in the Methods section) is being adapted for undergraduate STEM applications. More specifically, the EVE system is undergoing development to be used as a testbed platform in order to investigate the level of engagement of STEM undergraduate students with core STEM concepts using EVE as a new interactive educational technology system. Furthermore, EVE integrates commercially available, low-cost hardware components and open source software packages in order to realize an affordable, interactive, STEM education delivery system.

Previous applications of the EVE system focused on a single user's interaction with near-tactile, giga-pixel images to increase the historical or cultural context of presented images for the user [14, 20]. In an analogous fashion, the current effort focuses on building an explorable virtual environment in which STEM undergraduates can interactively engage with essential STEM-related course content. The natural-gesture based interface coupled with the representation of concepts or ideas in multiple dimensions adds greater context and insights into the presentations of the STEM- related content. Such context and ease of conceptual exploration has yet to be achieved in a conventional undergraduate course setting, even considering the increased usage of online materials and computer modeling systems. Since the open SDK kit for Kinect was released in November 2011 and corresponding interactive, post-secondary STEM educational systems, like EVE, remain in their infancy, a means to compare conventional, i.e. lecture style, versus emerging interactive and intuitive modes of education is also necessary.

2. Methods

2.1 EVE System Development

In previous implementations of EVE, we extended the functionality of Ideum Systems' Open Exhibits Software Development Kit (SDK), which includes a small gesture library and the tools for relatively simple extensibility. A primary complication of our work was that the included SDK was originally purposed for touch-screen applications such as commercial tablets and touch screens. We had to modify this library and adapt it to be used with an Xbox Kinect so that one could use gesture inputs rather than touch inputs. Coding was done in ActionScript3, using Adobe Flash CS5. The Xbox Kinect detects bodies and their movement in a 3D space using visuals and infrared. The Kinect had a specific depth of field and field of view - a region in front of the Kinect where it most accurately detects gestures with a minimum and maximum focal distance, an issue that has been recently addressed through the Kinect near mode. A human body within this region is recognized by the Kinect as an Open Sound Control (OSC) Contour Input, or "blob." Each of these blobs is recognized as an appendage and monitors movements for appropriate input. Appendage coordinates are then tracked and pass through a Flash OpenSound Control (FLOSC) Communications Bridge, allowing for two-way interactive communication, and then translated into a visual Flash Display Output where the image will change depending on the gesture. EVE's gesture library is intuitive and easy to use to investigate a visual space or displayed visual content. Panning is a one-hand gesture and requires the user to stick out his/her hand and move it in the direction s/he wants the image to move. Zooming in and out is acquired by swiping two hands outwards and inwards. The movements are natural, intuitive and simple, allowing for easy image exploration. In addition, hovering gestures will result in access to augmented reality (AR) information, where facts and information about highlighted details are displayed on the visual interface. Using both hands and lifting up in front of the screen accesses icons that denote AR components. Pointing and hovering over these icons reveal a window with additional information regarding a specific detail in the image. The participant can swipe the window away and swiping both hands downwards removes the AR icons and leaves a clean visual. The release of the most recent version of the Microsoft Kinect SDK affords further opportunities for development of gesture-based navigable environments as near mode and finer grain joint recognition have been built into the development tools, allowing for more natural navigation, providing for the ability to create simulations akin to real experience[14].

2.2. Testing Details

In order to prove that making an interactive educational environment will be worthwhile, a means of measuring the learning outcomes of the EVE system must be established. Tests are being formed in order to evaluate the mode of content delivery via EVE with regards to a participant's knowledge and motivation of a subject, before, during, directly after, and following at least a two week period of time.

A group of at least fifty undergraduate STEM students enrolled at Stevens Institute of Technology in a summer research program will be selected for the study. The group of students will be given a preliminary test on a subject which all of the participants will have roughly the same base knowledge. This test will test both knowledge and the individual motivation to learn about the given subject material; the study will also include a brief survey about the individual's typical study habits and how they use and implement conventional educational materials commonly presented to them.Participants will then be split into two even groups, one which will be taught lessons by the conventional method, the control, and one that will be taught using the EVE delivery mode. During the process, which will take place over about two months, students will take follow-up quizzes after each lesson to evaluate content learned, mood, and motivation after the lesson [21]. For the EVE method, recordings of the lessons will be captured and looked at for purposes of validating mood, and engagement during the lesson. After the learning process, a post-test with similar material to the preliminary test will be given [22]. Two weeks after the post-test is given, an altered version of the pre and post-tests focused on retention will be administered to the groups to evaluate the students' lesson retention.Throughout the testing and experiential process, each student's progress will be documented and broken down by their major, GPA, gender, and other individual characteristics, such as hobbies and involvement on campus, as well as the time for completion of tests and quizzes, behavioral observations during lessons, and tests and quizzes. Participants will also be graded on their performances in 'real-world' scenarios, which will be done through interacting with the EVE method. All of the collected data will be analyzed to compare the student outcomes among the control group and those in the group assigned to use the EVE system.

3. Discussion of Upcoming EVE User Data Collection and Testing

3.1 Participant Demographics

All of the participants in this study are undergraduate students entering sophomore, junior, or senior year pursuing a STEM program at Stevens Institute of Technology in Hoboken, NJ. Each of the participants has a GPA between 3.2 and 4.0 based on a 4.0 scale. The participants are also part of an academic scholars, on-campus, summer research program. The study participants are pursuing STEM programs, ranging from biology and physics to chemical, civil, environmental, and mechanical engineering among others. Though this sample of students is limited to those with high GPA's who are already motivated enough to pursue summer research projects, restricting the initial study to these undergraduate students constrains the factors related to baseline student performance. This permits the initial study to focus on the comparison of results between the conventional and EVE modes of information delivery. Studies to start in the upcoming academic year will incorporate freshman through junior students with a wider distribution of grade point averages (GPAs) in order to explore EVE's capabilities in STEM undergraduate applications from a broader user base. Since all of those in the study are presently in STEM programs, the nature of the content delivered using both the conventional delivery and EVE delivery modes were based on a typical core engineering course, such as second year calculus.

3.2 Definition of User and EVE System Performance Metrics

In order to compare a conventional method of course content delivery with the equivalent course content delivered using the EVE system, metrics related to session time, user engagement, and lesson-content retention have been defined. The metric for lesson time is defined simply as the total amount of time it takes the user to go through all of the presented content based on their own indication that they feel ready to answer follow-up questions on the concepts presented. Participant, i.e. user, engagement is measured according to the following three defined metrics: mood state, i.e., smile intensity as defined in [23], lines of notes recorded, and specifically for the EVE system, the number of interactive gestures made normalized against the total number of interactive gestures possible for the presented lesson plan. In order to measure the actual effectiveness of the mode of content delivery, the participants also complete a pre-test and post-test in the vein of a typical student assessment. The pre-test and post-test have identical lesson-content related questions. This provides a differential metric for the user's amount of content learned.

3.3 Anticipated Outcomes

Earlier implementations of the EVE system were designed to realize immersive 3D environments [9, 22]. Though much hype has surrounded the advent of interactive gesture technologies in classroom settings quantitative studies on the performance of such systems for undergraduate STEM applications has not yet been reported. We expect that the ongoing development of the system detailed in this work, in combination with the results of controlled undergraduate STEM user testing to occur this summer will support findings that increased student-engagement leads to improved student outcomes. In addition we anticipate that the level of user-engagement use the EVE system for lesson content delivery will be greater than that of the conventional lecture-based mode of delivery. Following the initial summer testing scenario, additional studies to begin in the fall will further assess these variables based on a broader sample of participants with a greater distribution of GPAs.

4. Conclusions and Future Work

We have made progress on the extension of the EVE system to investigate the broader effectiveness of interactive gestures via depth based sensors, like Kinect, in an undergraduate STEM application. Applying EVE's capabilities for increasing context and engagement of STEM undergraduates with specific core course materials may improve the overall learning experience of these students and increase the likelihood of STEM program retention. In addition to the EVE system realization for the STEM undergraduate application reported here, the complete findings of the corollary effectiveness study described above will be fully analyzed by August. In addition to these near-term efforts, we are working towards the design of adaptive learning modules for EVE that utilizes a user's mood data as a feedback variable to adapt the lesson in response to the user's level of engagement and/or enjoyment. Though many studies encourage the use of alternate modes of interaction and engagement of content through technology in classroom settings, the adoption of such systems and practices remains sparse, as professors and institutions are resistant to make changes to the conventional tried and true modes of content delivery. However given the stark statistics regarding the STEM professional pipeline in the US, it is time to complete rigorous studies that really evaluate the effectiveness of interactive, adaptive technology systems that may ultimately be capable of addressing the undergraduate STEM education challenges facing the US. Intended as a long term study, the future work on a mood-adaptive EVE system, would help address questions regarding long-term effectiveness of new technology systems in undergraduate STEM education. Furthermore, studies involving a multi-user EVE system are recommended to see how the learning tool behaves when it has to adapt to multiple users which forms the basis of another project entitled EDEN, which stands for (EDucational ENvironments).

The concept of EDEN entails an educational environment where participants can effectively learn through interacting with an immersive interface. Using EVE as a general framework, EDEN becomes an adaptable system that can be applied to any learning environment or target audience. Whether students prefer to learn with others in a classroom setting, with others online, or by themselves, EDEN can adapt to the scenario. For group settings, a network of EDEN systems working in tandem would allow for the collection of data from various sources, leading to greater adaptability based on the usage and performance data of the group. Due to the flexible nature of the project, EDEN applications can be created for multiple age groups, from K-12 students to the elderly, as its functions are not solely tied to the classroom. Special needs children are another audience that EDEN apps can reach out to, to further develop the existing research [24] that has been done to educate those with learning disabilities.

EDEN will be deployed in Stevens Institute of Technology's entry in the U.S. Department of Energy Solar Decathlon 2013, where the home learns about its inhabitants while they themselves learn about their dwelling as well. In addition to EDEN's usage in a home setting, we plan to execute this system in a "classroom of the future" setting where students can be actively involved in class material through an interactive surface. We believe that motivating students to remain engaged with their subject of study will lead to a positive learning experience, as well as extended use of EDEN systems in classrooms worldwide.

References

[1] Fernando, A.T., Murray, N., Tan, K., Wimalaratne, P. Software Architecture for a Constraint-based Virtual Environment, Proc. VRST '99. ACM pp. 147-9, 1999.
[2] Kenderdine S. 2003. This is not a peep show! The Virtual Room at Melbourne Museum (VROOM). ICHIM 2003.
[3] Fernando, T., Marcelino, L., Tan, K., Wimalaratne, P. Interactive Assembly Modelling within a CAVE Environment, Eurographics-Portuguese Chapter, pp. 43-49, 2000.
[4] Bourke, P.D. Proceedings of the Computer Games & Allied Technology 09 (CGAT09), iDome: Immersive gaming with the Unity game engine, pp. 265-272, 2009.
[5] Childers, L., Disz, T.L., Hereld, M., Hudson, R., Judson, I., Olson, R. et. al., Parallelldatorcentrum Kungl Tekniska Hogskolan Seventh Annual Conference, Activespaces on the grid: the construction of advanced visualization and interaction environments, (1999).
[6] Gvili, R., Kaplan, A., Ofek, E., Yahav, G. Depth keying, Proc. SPIE,vol. 5006, pp. 554–563, 2003.
[7] Elron Group Company 3DV Signs Definitive Asset Sale Agreement, (2009, March 29), [Online]. Available: http://www.elron.com/uploaded/29%203%2009%20PR%203DV%20asset%20sale%20FINAL.pdf
[8] Panasonic D-Imager: 3D sensing technology Product, [Online]. Available: http://www2.panasonic.biz/es/densetsu/device/3DImageSensor/en/product.html
[9] A. Maimone and H. Fuchs, Reducing Interference Between Multiple Structure Light Sensors Using Motion, vr, pp.51-54, 2012 IEEE Virtual Reality, 2012
[10] Microsoft Kinect: Develop for Kinect, [Online]. Available: http://www.microsoft.com/en-us/kinectforwindows/develop/new.aspx
[11] Omek Interactive. (2011, Oct.) Omek Beckon Development Suite. [Online]. Available: http://www.omekinteractive.com/products.html
[12] Leap Motion: About, [Online]. Available http://live.leapmotion.com/about.html
[13] Chaudhary, A., Rajeha, J.L., Das, K., Rajeha, S. Intelligent Approaches to interact with Machines using Hand Gesture Recognition in a Natural Way: A Survey. International Journal of Computer Science and Engineering Survey, vol. 2, no. 1, pp. 122–133, 2011.
[14] Moriarty, B., Amaba, C., Moy, Z. Building Immersive Image-Based Environments: Explorable Visual Environments, International Conference on Emerging Wireless and Information Technology 2011.
[15] Star, S.L., Griesemer, J.R., Institutional ecology 'translations' and boundary objects: Amateurs and professionals in Berkeley's Museum of Vertebrate Zoology, Soc. Stud. Sci. 19, (1989) 1907-39.
[16] Augustine, N.R., Barrett, C., Cassell, G., Chu, S., Gates, R., Grasmick, N. et al., National Academy of Sciences, National Academy of Engineering, and Institute of Medicine, Rising Above the Gathering Storm: Energizing and Employing America for a Brighter Economic Future, (2007) 4.
[17] Business Higher Education Forum, Increasing the Number of STEM Graduates: Insights from the U.S. STEM Education & Modeling Project, 2010.
[18] Association of American Universities, "National defense education and innovation initiative: Meeting America's Economic and Security Challenges in the 21st Century," 2006.
[19] Vilhjalmsson, J. Chiu, H. Lee, & L. Qu Enhancing ITS instruction with integrated assessments of learner mood, motivation and gender, Proc. of the 12th International Conference on Artificial Intelligence in Education (AIED), 2005. Amsterdam: IOS Press.
[20] Moriarty, B., Amaba, C., Moy, Z. Utilizing Explorable Visual Environments for Experiential Applications. Conference on Systems Engineering Research, 2012.
[21] Beal, C. R., Shaw, E., Chui, Y., Lee, H., Vilhjalmsson, H., & Qu, L. Enhancing ITS instruction with integrated assessments of learner mood, motivation and gender. 12th International Conference on Artificial Intelligence in Education. 2005.
[22] Ferreira, S., and Faezipour, M. An Analysis of Processes, Risks, and Best Practices for Use in Developing Systems Engineering Process Simulators, Conference on Systems Engineering Research, 2012.
[23] Hoque, M., Hernandez, J., and Picard, R.W. Mood Meter: Large-Scale and Long-Term Smile Monitoring System, Accepted Siggraph Conference 2012
[24] Stromer, R., Kimball, J. W., Kinney, E. M., & Taylor, B. A. Activity Schedules, Computer Technology, and Teaching Children With Autism Spectrum Disorders. Focus On Autism & Other Developmental Disabilities, 21(1), 14-24. 2006.