An Interactive Pedestrian Environment Simulator for Cognitive Monitoring and Evaluation

Jason Orlosky, Markus Weber, Yecheng Gu, Daniel Sonntag, Sergey Sosnovsky
German Research Center for Artificial Intelligence (DFKI)
Stuhlsatzenhausweg 3, D-66123 Saarbruecken, Germany
{ jason.orlosky, markus.weber, yecheng.gu, daniel.sonntag, sergey.sosnovsky} @dfki.de

ABSTRACT

Recent advances in virtual and augmented reality have led to the development of a number of simulations for different applications. In particular, simulations for monitoring, evaluation, training, and education have started to emerge for the consumer market due to the availability and affordability of immersive display technology. In this work, we introduce a virtual reality environment that provides an immersive traffic simulation designed to observe behavior and monitor relevant skills and abilities of pedestrians who may be at risk, such as elderly persons with cognitive impairments. The system provides basic reactive functionality, such as display of navigation instructions and notifications of dangerous obstacles during navigation tasks. Methods for interaction using hand and arm gestures are also implemented to allow users explore the environment in a more natural manner.

Author Keywords

Simulation; cognitive monitoring; interaction; evaluation; virtual reality.

ACM Classification Keywords

H.5.1 Multimedia Information Systems: *Artificial, augmented, and virtual realities*; H.5.2 User Interfaces: *Training, help, and documentation*; I.6.6 Simulation Output Analysis

INTRODUCTION

In recent years, intelligent technologies have been proposed as a tool for assisting with education and training for a wide range of fields [3, 10]. A number of these assistive technologies are designed for tracking and remedying of cognitive disabilities, which often manifest in different ways. For example, a dementia patient may exhibit wandering behavior, the exact nature of which is difficult to determine since his or her exact movements and field of view may not be recorded. To address these challenges, we introduce a prototype simulator that can provide better mechanisms with which to study various types of activities and cognitive states in the domain of pedestrian safety. In order to improve monitoring, analysis, and training for such individuals, we combine a number of interactive technologies such as the Oculus Rift, LEAP Motion, and the Myo by Thalmic Labs. These tools enable hand tracking and sensing of arm motion and muscle activity that can be used for interactions such as button presses and walking within the simulated environment. The virtual environment is created with Unity 3D, and the simulation itself is viewed through the Rift to provide an immersive experience.

Additionally, two types of hand and gesture devices are used for both input and monitoring when users are engaged in virtual tasks. These devices allow users to directly interact with the simulated environment in a more natural way, and can also be used for online monitoring and post-analysis of the user's activities and movements relative to his or her field of view. Overall, the system functions as an adaptive test bed for studying impairments, and at the same time facilitates better interaction. This work describes the prototype system and demonstration in detail. Users can complete tasks in the virtual environment in different conditions, use hand tracking and integrated muscle sensors for movement, and interact with the safety notifications generated by the system. Tasks include navigating through the virtual environment in different conditions and reaching a specified location. Users will have to deal with both traffic and obstacles to complete each task successfully.

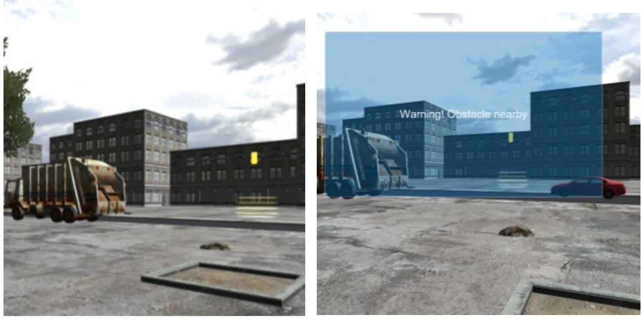

Figure 1. Simulated environment showing obstacles and waypoint (left) with reactive warning system alerting the user to a tripping hazard (right).

IUI'15 Companion, Mar 29 - Apr 01, 2015, Atlanta, GA, USA
ACM 978-1-4503-3308-5/15/03.
http://dx.doi.org/10.1145/2732158.2732175

RELATED WORK

The recent development of 3D tools and technologies for gestures, interaction, and control has led to new possibilities for 3D interactive applications and immersive environments [12]. One such example is hand-based interaction, where a user can directly control 3D objects with his or her hands in both virtual and augmented reality [7]. Haptic devices can also be used to interact and provide feedback for the user, which have been employed in simulations and for control of robotic systems [1]. Head worn displays have proven to be instrumental for implementing immersive simulations of situations with present danger or limited visibility, which provides further motivation for us to use virtual content to help evaluate perception of a dynamic environment [9]. Even more recently, interacting in a virtual immersive display has proven to be useful in treating amblyopia, as shown in [2].

These types of interactions and simulations have also had more specific applications in education and medicine. In 2010, a simulator was developed that showed benefits for training the elderly with respect to pedestrian safety. Gu et al. have more recently proposed the combination of semi-immersive virtual reality (VR) with intelligent tutoring approaches in order to support children learning pedestrian safety [5]. Research shows that children, the elderly, and the intoxicated are the most endangered categories of pedestrians; therefore, technology to evaluate and support these groups is in high demand [13].

SIMULATION HARDWARE AND SOFTWARE

In this work, we seek to improve on prior simulation technology to provide a more interactive and immersive environment. Moreover, we can extract more information about user actions than in more simplistic setups, which can be utilized for more thorough analysis and study of an individual's behavior.

Display and environment

As an immersive display, we use the Oculus Rift DK2 head worn display (HWD), which provides stereoscopic images to the user, conducts six-degree-of-freedom head tracking, and gives access to video and position data streams so that we can monitor the user's field of view and head orientation. Less immersive setups, for example those based on single or multiple monitor displays, may fail to provide a real sense of danger and/or spatial understanding during simulation. CAVE systems, which project images onto walls that surround the user, are also immersive, but are not as affordable and maybe difficult to obtain for a majority of potentially interested stakeholders such as educational institutions and healthcare facilities.

The simulation itself is implemented in Unity 3D, which enables generation of high resolution virtual worlds and an improved sense of presence. The simulation is set in a manually constructed urban environment, as shown on the left of Figure 1. The environment contains many stationary objects, such as trees, benches, rocks, roads, curbs, and other obstacles (e.g. parked cars and road curves), that might be of concern for an elderly patient or an individual with motor impairments. Additionally, various dynamic and interactive objects such as cars, traffic lights, waypoints, and buttons are implemented so that the user can complete tasks while dealing with realistic environmental obstacles.

Figure 2. Oculus Rift with attached LEAP Motion for hand interactions such as button presses (left) and the arm-worn Myo to facilitate walking via natural arm movements (right).

As a part of the simulation, reactive alarms are installed on potentially hazardous objects. For example, when a user fails to look at an object near his or her feet, an alarm appears in the display to alert him or her of the object, as shown on the right of Figure 1. The alarms are activated only when a user fails to notice an object. For example, an individual may not look down at a curb or rock in his or her path. By checking the distance to the object and current camera frustum, we know whether or not that object has entered the user's central field of vision. If the person has come too close to the object without it entering his or her field of view for some time, the alarm is triggered in the display. Though this alarm system is mainly designed to provide feedback on a user's cognitive ability and learning, it can also be used to study reactions to augmented elements in the real world. The simulated alarm also resembles a notification that someone might see in an optical see-through HWD used outside, so we can also gather feedback about virtual text notifications that might be presented in a mixed reality situation.

More importantly, by recording the user's head movements, we can measure reaction time, whether or not certain objects are noticed, and general movement tendencies. The change in response over time (i.e. learning) to certain alarms or notifications can also be used to monitor the patient's mental state, or gather data about any cognitive impairments that may be present. Based on the age, conditions, and ability of the user, tasks can be modified to require a higher level of cognition, such as adding a blind curve, an obstacle, haze, or darkness as shown on the right of Figure 3. Users must then think ahead in order to predict whether the situation poses a greater danger than when all traffic can be seen clearly from a distance. Consequently, the ability of the user to predict or think at a higher level can be analyzed since we have access to field of view data within the virtual environment.

Figure 3. Interactive traffic light (left), and simulated haze to increase task difficulty (right).

Physical interactions

We also needed a way for the user to physically interact with the environment. First, in order to move through the simulated world, he or she must have a method to engage walking. Many other implementations utilize keyboards or controllers for movement [5], but we sought to have something more natural, especially since our target users are individuals who may have a cognitive impairment. To accomplish this, we use the Myo, an arm-worn band by Thalmic Labs that contains sensors to measure movement and muscle activity. Since the device is arm-worn, it can be used in a non-invasive manner, and can take advantage of the user's natural arm movements as a form of input. In particular, we utilize the natural swinging of the arm to facilitate movement. Since a person's arms move naturally during walking, this becomes a much more familiar way of interacting than by pressing keys on a keyboard. Faster arm swinging corresponds to faster movement within the environment. Although walking machines for virtual reality such as the Omni Treadmill are available, they are still expensive and not portable. For the purposes of research and studying patients at a healthcare or research facility, the Myo provides us with a small, inexpensive input device.

Second, users must be able to physically interact with virtual elements in the simulated world, for example, by pressing one of the signal crossing buttons located on the traffic lights shown on the left of Figure 3, or signaling a taxi. For this purpose, we use the LEAP Motion, which is mounted to the HWD and provides near-field hand tracking and is ideal for on-demand interactions in our environment, specifically, button presses for road crossings. The role of these procedural tasks is especially important for evaluating user behavior and cognition. The user also sees his or her hands inside the simulation; thus, a more realistic physical interaction is achieved. Although haptic feedback via Myo has not been yet implemented, we plan to incorporate a tactile vibration upon a button press as a part of future work.

The purpose of using these tools is two-fold. While we need to provide the user with natural interaction, we also want to record as much data as possible about movements or judgments that may signify an impairment or general lack of cognition at a certain time. Examples include forgetting to press a button altogether, excessive search behavior, or severely delayed reactions to an important stimulus.

Discussion and application areas

Arm, hand, and finger gesture recognition are of particular interest for future natural user input scenarios with implicit gestures and, as a byproduct, for reducing cognitive load (compared to explicit gesture, which a user has to learn). Implicit gestures need to be grounded so that the human computation part (recognition and understanding of human gestures) can be achieved. Data collected by this kind of simulator has the potential to produce better recognition accuracy and hence more realistic VR scenarios and user behavior capture and interpretation. Collected samples of usage data, such as gestures or actions used to complete tasks in the simulator, can be labeled post-task, and categorized via classification algorithms. Our next steps include the creation of a 6D motion gesture database for implicit gestures and the application of new spatio-temporal event classification algorithms [8].

Reliable implicit gestures could also be used for gesture-based disambiguation of user intents. In addition to the already implemented alarms, providing reminders about consecutive stages in individual activities such as "press the traffic light button" to compensate for a lack of situational awareness can be useful for more severe dementia patients, and for caregivers who are evaluating the patient. In addition to evaluating patients, the virtual reality environment may develop into a safe, cost-effective, and engaging approach for future immersive training environments of dementia patients where training implicit gestures could help improve the performance in daily life (e.g., pressing buttons on home appliances to remember their functionality). This training could be greatly beneficial for increasing a user's retention of situations and the implicit usage of gestures for controlling electric and mechanical machines [6]. In addition to traffic scenarios, the reminders or alarms could also be evaluated for scenarios that may be difficult to evaluate in the real world due to privacy issues (e.g. in a bathroom).

CONCLUSION

In this work, we introduce a simulator designed to improve monitoring interaction for and analysis of cognitive abilities in a virtual environment. We construct a replica of a suburban environment, and provide a number of navigation tasks within the environment. This environment is viewed through the Oculus Rift, and includes integration of the Myo and LEAP Motion sensors into the framework. This allows for more natural input, and enables collection of more detailed data and feedback for individuals such as dementia patients. We hope this environment will promote more detailed study of cognitive abilities and can be used in other contexts, such as education and training.

ACKNOWLEDGEMENTS

This research has been supported by the Federal Ministry of Education and Research in the Kognit project, http://www.dfki.de/Kognit.

REFERENCES

1. Bar-Cohen, Y. (2003). Haptic devices for virtual reality, telepresence, and human-assistive robotics. *Biologically Inspired Intelligent Robots, 122, 73*.
2. Blaha, J., & Gupta, M. (2014, March). Diplopia: A virtual reality game designed to help amblyopics. In *Virtual Reality (VR), 2014 iEEE* (pp. 163-164). IEEE.
3. Chaudhri, V. K., Lane, H. C., Gunning, D. & Roschelle, J. (2013). Intelligent Learning Technologies Part 2: Applications of Artificial Intelligence to Contemporary and Emerging Educational Challenges. *AI Magazine 34(4)* 10-12.
4. Dommes, A., & Cavallo, V. (2010). The beneficial effects of a simulator-based training on the elderly pedestrian safety. In *Proceedings of 12th International Conference on Mobility and Transport for Elderly and Disabled Persons*.
5. Gu, Y., & Sosnovsky, S. (2014, February). Recognition of student intentions in a virtual reality training environment. In *Proceedings of the companion publication of the 19th international conference on Intelligent User Interfaces* (pp. 69-72). ACM.
6. Gupta, S. K., Anand, D. K., Brough, J., Schwartz, M., & Kavetsky, R. (2008). Training in Virtual Environments. A Safe, cost-effective, and engaging approach to training. University of Maryland.
7. He, Z., & Yang, X. (2014, November). Hand-based interaction for object manipulation with augmented reality glasses. In *Proceedings of the 13th ACM SIGGRAPH International Conference on Virtual-Reality Continuum and its Applications in Industry* (pp. 227-230). ACM.
8. Jeni, L. A., Lőrincz, A., Szabó, Z., Cohn, J. F., & Kanade, T. (2014). Spatio-temporal Event Classification using Time-series Kernel based Structured Sparsity. In *Computer Vision–ECCV 2014* (pp. 135-150). Springer International Publishing.
9. Orlosky, J., Toyama, T., Sonntag, D., Sarkany, A., & Lorincz, A. (2014, March). On-body multi-input indoor localization for dynamic emergency scenarios: fusion of magnetic tracking and optical character recognition with mixed-reality display. In *Pervasive Computing and Communications Workshops (PERCOM Workshops), 2014 IEEE International Conference on* (pp. 320-325). IEEE.
10. Petersen, N., & Stricker, D. (2012, November). Learning task structure from video examples for workflow tracking and authoring. In *Mixed and Augmented Reality (ISMAR), 2012 IEEE International Symposium on* (pp. 237-246). IEEE.
11. Percer, J. (2009). Child pedestrian safety education: Applying learning and developmental theories to develop safe street-crossing behaviors (No. HS-811 190).
12. Plemmons, D., & Holz, D. (2014, July). Creating next-gen 3D interactive apps with motion control and Unity3D. In *ACM SIGGRAPH 2014 Studio* (p. 24). ACM.
13. Whelan, M., Oxley, J., Charlton, J., D'Elia, A., & Muir, C. (2008). Child Pedestrians: Factors Associated with Ability to Cross Roads Safely and Development of a Training Package.