Bio-reckoning: Perceptual User Interface Design for Military Training

Tami Griffith[1], Deanna Rumble[2], Pankaj Mahajan[2], and Cali M. Fidopiastis[2]

[1] U.S. Army Research Laboratory, Human Research and Engineering Directorate, Simulation and Training Technology Center, Orlando, FL USA
Tami.Griffith@us.army.mil
[2] University of Alabama at Birmingham, School of Health Professions, Birmingham, AL, USA
{ddrumble,pmahajan,cfidopia}@uab.edu

Abstract. Simulation based training is one way to attain operational realism for training complex military tasks in a safe, task relevant manner. For successful transfer of knowledge, skills, and abilities to the dynamically changing military environment, the human-computer interface should minimally support learning during the training process and provide congruent action plans that facilitate understanding of the overall training goal. While there are emerging controller technologies, simulators still rely on such input devices as mouse and keyboard. These devices potentially cause information and training bottlenecks as they limit naturalistic interactivity within the more advanced serious gaming platforms. Given the shortcomings of current interface design, we suggest a human-computer interface framework that includes perceptual user interface components and an open source serious game testbed. We discuss a multimodal framework called bio-reckoning that integrates brain-computer interface techniques, eye tracking, and facial recognition within EDGE, the U.S. Army's newest serious game based training tool.

Keywords: simulation based training, perceptual user interfaces, brain-computer interfaces, serious games, military training, augmented cognition.

1 Introduction

Human Computer Interface (HCI) techniques do not enjoy the same timely advances as do computer components and related hardware. This lag is apparent when reviewing interface design for military training simulations, especially those following a serious game platform [1]. Smith's review of the use of games in military training makes clear that throughout history the game play or simulation supported cognitive function and related action needed for battlefield success. However, regardless of the technological advances (e.g., high fidelity terrain maps and realistic avatars) in today's serious game training paradigms current HCIs execute our 'plans to perform' an action in computer space by means of intermediary physical manipulations such as pressing keys or directing a joystick. Transferring actions through these traditional

D.D. Schmorrow and C.M. Fidopiastis (Eds.): AC/HCII 2013, LNAI 8027, pp. 31–40, 2013.

input devices places an intermediary between the human operator and the training simulation that can detract from training goals and objectives, and more importantly fail to support transfer of training to the field environment [2].

For example, computer-aided training that relies on joystick manipulations may hinder necessary cognitive processes (e.g., focused attention) through the bottleneck of translating intended action through an unnatural modality. Further, mapping computer interactions through these peripherals requires time to learn and to operate pre-training, while demanding time to translate a user's physical action into a limited predefined set of object behaviors during training. Given the artificiality of these input devices, physically operating interfaces for multiple objects may demand more cognitive resources and may lead to overload simply through motor control processing and motoric interference.

Also neglected in the traditional user-interface paradigm is the affective training that is necessary to provide the correct amount of emotion regulation required to react appropriately under stress [3]. The gaming industry leads the software and hardware development for many serious games for military use [4] [5]. Matching interactivity and emotional regulation design elements to the training environment is not important. The gaming industry's goal is entertainment, and not transfer of training to the theatre of war. Poor transfer of training from simulation to field is not only costly from a financial perspective, but also can lead to loss of life. Offering a more naturalistic interaction within military training simulations is an overlooked necessity.

How to proceed in creating appropriate user interfaces for supporting transfer of training is a non-trivial task. Defining types of actions necessary for task training and mapping them to the serious game action codes is a primary concern. Once actions are chosen, how to instantiate these codes in the serious game environment by choosing or developing serious game controllers is key. With new gaming technologies such as the Microsoft Kinect, a motion based controller, one strategy is to gather all state-of-the art controllers and user test for ease of use and improved performance. Problems quickly arise in that these controllers work optimally with a particular gaming console. Even controllers with their own software development kits pose interfacing issues that may require knowledge from a highly trained technician to integrate.

Another issue when choosing off-the-shelf gaming environments for training is that the metaphor used for interface design may not match the one needed for training. Controller technology and subsequent action code responses are gaming specific and serve the purpose of gaming goal, of which high entertainment value is one. These elements are also chosen as part of the gamming narrative or story that provides the nature of the interaction as a gaming element. How challenging these action pairings to gaming objectives are to learn depends on the overall goal of the game. For example, discovering how to launch a weapon may be a gaming objective for an action game. Action codes within the game environment support this aspect of exploration. In contrast, guessing how to change mission critical entities within military simulations for training is never appropriate. More importantly, the inability to access the source code of proprietary serious game platforms does not allow changes to controller parameters and their associated action maps further limiting the number, variety, and type of controller.

Low Cost Game Interfaces. Nintendo opened the game industry to non-gamers by creating an interface, Wii Mote, that made playing games more natural and engaging [6]. Researchers [7] explored the use of the Wii Mote and Numchuk for navigation, object manipulation, and object selection within a First Person Shooter (FPS) type gaming environment. The user navigated using the Wii Mote; the combination of the Wii Mote and the Numchuk performed the action of selecting and manipulating objects. Experienced FPS gamers reported the navigation strategy as frustrating; yet, many found the manipulation tasks more pleasant using the Wii tools as compared to a mouse. In this example, user satisfaction could be due to the optimized interactivity when using the new controllers. This research demonstrates the ease of use of these controllers for manual tasks, but falls short when describing the entire user experience.

Microsoft™ responded to the market success of the Nintendo Wii with the groundbreaking Kinect depth sensor camera [8]. The Kinect uses an infrared laser projector combined with an image sensor, which captures video data in 3D [9]. It is capable of simultaneously tracking six people with two active players at a time, allowing facial feature extraction or the ability to "recognize" players and the ability to track 20 joints per player [10]. The system includes a directional microphone to support voice control. In June of 2011, Microsoft™ released a non-commercial Software Development Kit (SDK) for use with Windows [11]. The Kinect interfaces with a standard PC via USB connector. This system has arguably changed the face of the interface world by bringing the player into the game more accurately than ever, however there are still unsolved problems that with the Kinect. Though the microphone is useful for administrative functions, it is still not reliable enough to replace the keyboard (or in this case, joystick). Gestures used to do interactions or administrative functions can be awkward or might not be readily recognized [12].

Voice as a Controller. Voice recognition for use in games is still at an early state of research. The goal of using voice in a FPS game is to assist in interacting with objects. For example, if the user approaches a vehicle, an action menu appears asking if the user wants to enter the vehicle. To activate the menu the player would use a specific word. Drawbacks to using voice as an interface are that background noise and casual conversation may unintentionally activate a task. It may be necessary to confirm direction to avoid false positive responses. The proposed strategy for this research would be to use a Small-Vocabulary/Many User system, with only a small set of words in use at specific times. Mohanram [13] showed that speech recognition as a game interface was not yet ready for public adoption with only 40% of the users considering speech as a better input strategy than voice. The greatest issue was misrecognition of the voice cues.

Apple has since released Siri as an alternative means of inputting data into its smart phone. Siri “understands” conversational context and is surprisingly accurate in converting spoken word into text [14]. Siri and PC applications (e.g., Dragon Dictation) demonstrate that voice recognition as a natural and intuitive interface tool is beginning to ‘come of age’. This technology clearly brings new functionality that was not readily available in the past.

Controller Testbed. A more efficacious serious game design strategy would be to start with a framework that provides a more systematic manner of testing controllers, interfaces, and content for their effects on military training and training transfer. The serious game platform should allow for full access to technology and action codes, as well as allow for the integration of multiple action controllers. This work uses Enhanced Dynamic Geo-Social Environment (EDGE), a military relevant gaming environment, to provide a testbed from which to assess gaming and training elements. Within EDGE, the Bio-Reckoning Interface (BRI) integrates multiple psychophysiological and body (e.g., facial features and limb movement) measures and uses these measures as naturalistic input control to the EDGE platform. For the purpose of this paper, we discuss the development of a BRI interface that uses brain computer-interface techniques, eye tracking, and face recognition to provide action codes to EDGE.

2 Perceptual User Interfaces for Military Training

Perceptual User Interface (PUI) design takes into account naturalistic human nonverbal and verbal human responses as part of the device or sensor input to the human-computer system [15]. This relationship between the trainee and simulator is symbiotic, like that in an intelligent automated system. This idea extends the concepts of Augmented Cognition, where the system uses the trainee's psychophysiological data to determine learner biophysical states that impede the learning process during training [16]. However, unlike the Augmented Cognition closed-loop system, the multimodal interaction capabilities of a PUI based system would allow for user control over the type of interaction that accounts for individual differences in how a person processes complex cues and related action in operational environments.

PUIs integrate concepts from perceptive, multimodal, and multimedia user interface designs. According to [15]: 1) perceptive interfaces are aware of the learner's body, face, and hands; 2) multimodal interfaces use several learner perceptual modalities such as speech and eye tracking as system input; and 3) multimedia systems include the use of text, graphics, animation, voice, and touch to best deliver training content. For these interface styles to be effective they must reciprocally monitor user behavior, model the goals and objectives of the training, flexibly change to learner preferences, positively support learning acquisition, support multi-tasking, and motivate the learner to interact meaningfully with to-be-learned material. Turk [17] contended that an ideal user interface should seamlessly transfer the intent of the user to the system, and the system response should appropriately support the user experience (e.g., reduce extraneous cognitive load).

2.1 PUI Military Examples

QuickSet is one of the first examples of a military based PUI that was a wireless, handheld capability that could control distributed interactive simulations based on

Modular Semi-automated Forces representing training at 29 Palms, California [18]. QuickSet used multiple input sensors such as speech, gesture, and direct object manipulation to support platoon leaders and company commanders with decision making involving multiple distributed assets (e.g., vehicles or personnel). The use of data fusion algorithms such maximum likelihood estimators coupled with artificial neural networks assisted to disambiguate the sensor inputs and increase reliability in noisy military exercises [19]. The system was also extensible to support 3-D terrain visualization.

Improved sensor technology and data fusion algorithms allowed for the miniaturization of sensors such that wireless, unobtrusive biosensors fit into wearable systems that convey real-time data acquisition [20]. The Virtual Locomotion Controller is an example of a wearable multimodal capability that used solid-state gyros and accelerometers, ultrasonic range sensors, and force sensitive footpads to provide naturalistic motion (e.g., crouching and running) within Military Operations on Urbanized Terrain (MOUT) simulated environments [21]. These advances in sensor and algorithm development also allow for redundant HCIs that present a combination of system features (e.g., face recognition, eye tracking, and graphics) to the operator, as well as user selection modes where the operator chooses the type of feature based on task relevance.

2.2 Bio-Reckoning Interface Components

The BRI is an example of a state-of-the-art PUI that provides data fusion across a multimodal sensor suite. These data streams are synchronized and applied to the serious game either as emulators for controllers in the case of proprietary games or as actual controllers as in training simulations. Their output can also provide information on the efficacy of the training system (i.e., eye tracking). Within this BRI prototype, we explored the use of BCI techniques, eye tracking, and face recognition.

Brain-Computer Interface Techniques. Brain-Computer Interfaces (BCIs) afford the possibility of removing the interface-as-middleman in both gaming and virtual reality contexts [22]. A typical BCI system consists of three processing modules: 1) a brain activity-monitoring device (i.e., electroencephalography-EEG) that records brain activity, 2) a signal-processing module that identifies specific brain patterns or features related to a person's intention to initiate action, 3) and a translator that converts these brain features into meaningful control commands [23]. Electrophysiological sources of control (ESC) are the mental activities and their associated EEG measures that become the control mechanism that perform actions within a given application. ESC are currently elicited in an active (user conscious control without external stimulation), a reactive (external stimuli elicits user brain response), or a passive (brain activity associate with a cognitive state drives system change) manner. The proposed BRI system combines an easy-to-apply wireless EEG sensor headset made by Advanced Brain Monitoring.

Eye Tracking. Blinks, direction of gaze, and fixations are all candidate eye movements that can act as naturalistic input to a serious game [24]. Additionally, gaze patterns provide information on how naturalistic a task appears to the operator [25]. These gaze patterns can also address training design elements through the evaluation of fixation patterns during task performance. For this work, we used an EyeTech TM3 eye tracker that monitors head movement as well as the gaze pattern of both eyes.

Face Recognition. Humans have a biologically mediated expertise in identifying faces versus objects [26]. Researchers in computer vision have sought to replicate this process in computer software since the 1960's. Facial recognition software measures various generalizable features common to all human faces (e.g., spacing between the eyes). While facial recognition systems are becoming more accurate at their primary task of identifying faces, techniques for transferring these facial features and their related meaning (e.g. smile) to an avatar is not readily available. Open source solutions exist; however, we chose a more robust commercial product, faceAPI created by Seeing Machines.

3 Enhanced Dynamic Geo-Social Environment (EDGE)

Experiential learning is one of the benefits of using a serious gaming platform for training military tasks. However, there are challenges to serious game use in training. Besides adoption, there needs to be a clear training benefit to using this training paradigm. Additionally, PUI features should serve to augment training or otherwise not be a part of the training system. Proprietary games for training do not allow code access to develop and test appropriate training content integrated with candidate PUIs. A solution that 1) provides access to source code; 2) allows for community input assisting with extensibility and updates; and 3) leverages coding expertise from a global network [27] affords the opportunity to test different types of PUI features within an operationally relevant training environment.

Fig. 1. Screen shots from EDGE showing accurate physics, terrain, and visual representations of military relevant operational environments

EDGE is a government owned architecture designed using AMSAA approved standards (e.g., *OneSAF*) to provide highly accurate virtual simulations of military operational environments utilizing state-of-the-art Multiplayer Online Gaming

(MOG) technologies [28]. Access to EDGE requires Federal Government sponsorship for use. This feature allows the community to upgrade the software to meet the challenge and pace of changing technology. Figure 1 depicts the fidelity of the computer-generated models, along with accurate physics, and military relevant operational environments.

The first level developed within EDGE was a tutorial level that requires the user to walk and run in each direction, complete a high and low crawl through small openings, walk a balance beam, drive a vehicle and shoot a weapon. Throughout the level, smart menus appear to interact with objects, such as entering a vehicle. If the user is at a keyboard, the “F” key activates these menus. This level ensures that users are familiar with controls prior to using a training level. However, for this research, the level ensures that the user can complete tasks within a FPS game environment. Additional levels allow free exploration of a small village and an urban environment.

4 Future Work Additional Controllers Integrated with the BRI

The next phase of BRI development will include the Playstation Kinect sensor, which will extract body positioning information and collect voice data. In addition, a Nintendo Wii will have a dual purpose of controlling a tactical weapon to allow the user to engage an enemy and a steering wheel to allow the user to have a sense of driving a vehicle. This combination will allow the user to move forward or backward using the BCI, jump, kneel, turn using the Kinect, shoot and drive using the Wii, and interact with the user prompts using voice. While the implementation of these controllers is initially to improve interaction within the EDGE platform, the intention is that the interface is not platform specific. Future research includes a phased approach described below.

4.1 Phase I: Establish Prototype and Measure User Experience

Implement a Prototype Interface Set to Engage with a First-Person-Shooter Environment. A difficult challenge is ensuring that all integrated controllers function in a complimentary manner. For example, if a player intends to jump across a hole in the training simulation, the Microsoft™ Kinect should detect that the player is jumping *and* the BCI must move the avatar forward concurrently. Otherwise, the player will not effectively complete the jump across the opening. Additionally, a single Graphical User Interface (GUI) should setup pairings for the controllers and their associated actions, as well as monitor pairing functionality throughout the experience.

Demonstrate the Prototype and Establish Measures of User Engagement. The first study will compare traditional keyboard and mouse input to the experience of using the BRI controllers. Measures of user experience will include: 1) time to successfully complete the tutorial level, 2) usability of the combined interface during interaction in a simulated small town, and, 3) psychophysiological measures of engagement and distraction (e.g., brain activity and skin conductance). Finally, participants will respond to the NASA TLX and a user questionnaire based on five criteria for user acceptance of the interface (i.e., intuitive, readily available, augments existing user capabilities, accessible through an open toolkit, and fun).

4.2 Phase II: Comparison between Virtual and Live Experiences

Create a Scenario in the Virtual Environment That Replicates a Live Military Training Scenario. A MOUTexercise translatable into a live experience may provide the testing environment for this phase of research. A training exercise of this type would take place in a small simulated or mock-up town. The live exercise would use laser training weapons rather than live-fire weapons to reduce the risk of injury. The target location for the scenario is at the Maneuver Center of Excellence at Fort Benning, Georgia.

Compare Brain Activity and Skin Conductance Readings. To determine the how the level of realism experienced in a FPS game with an immersive interface suite compares to a live training experience at the same level of physical risk, trainees would experience both environments and perform similar tasks. Brain activity and skin conductance provide comparison measures of engagement and distraction or stress in each environment.

Compare Performance in the Live Environment after Practice in the Virtual Environment. Simulations historically reduce the cost and risk associated with live training. Further, simulation based training may better prepare a trainee for live training and ultimately for operational engagements. This study will compare the performance of trainees at the Maneuver Center of Excellence live training environment with and without preparatory virtual training. The expectation is that Trainees with virtual training preparation will perform better in a MOUT operation (building clearing, hostage rescue, etc) than those moving directly into a live training environment. This fits well with the described research because the level of realism established while in the virtual environment may be a critical factor in live training preparedness. To further establish that realism is a factor, three study groups will be established; one with no virtual training, one with keyboard and mouse at a desktop and one using the prototype BRI.

5 Conclusion

The ultimate goal of this research is to show that a combination of off-the-shelf emerging controller technologies integrated within a simulation-based trainer can improve the interaction between the human and computer. This improved interface can increase the user's sense of presence, immersion and flow, which may lead to improved human performance [29] and potentially to training realism. By creating a

PUI testbed using a military relevant simulation based training environment, this could benefit the gaming world as well as support military training applications.

It is clear that there is no one-size-fits-all interface in existence today. The premise of this research is that a combination of interface tools may begin to close the gap between the user and the immersive environment. Various modalities are mixed and matched and can be adjusted to support specific training needs; however, this type of experimentation should occur in a valid testing environment. In most cases, a traditional interface is sufficient; however, when total immersion is the goal for training or even for entertainment, a combination of interfaces may provide a better user experience.

References

1. Smith, R.: The long history of gaming in military training. Simulat Gaming 41(1), 6–19 (2010)
2. Grant, S.T., Barnett, J.S.: Evaluation of wearable simulation interface for military training. Hum. Fact. (2012), doi:10.1177/0018720812466892
3. Rahman, M., Balakrishanan, G., Bergin, T.: Designing human–machine interfaces for naturalistic perceptions,decisions and actions occurring in emergency situations. Theoretical Issues in Ergonomics Science 13(3), 358–379 (2012)
4. Smith, R.: Game Impact Theory: The Five Forces That Are Driving the Adoption of Game Technologies within Multiple Established Industries. Games and Society Yearbook, 1–32 (2006)
5. Smith, R.: The Disruptive Potential of Game Technologies. Research Technology Management 50(2), 57–64 (2007)
6. Klochek, C., MacKenzie, I.S.: Performance measures of game controllers in a three-dimensional environment. In: Proceedings of Graphics Interface 2006, pp. 73–79. CIPS, Toronto (2006)
7. Fischer, L., Oliveira, G., Osmari, D., Nedel, L.: Finding Hidden Objects in Large 3D Environments: the Supermarket Problem. In: Proceedings of 2011 XIII Symposium on Virtual Reality, pp. 79–88. IEEE Press, Brazil (2011)
8. Thorpe, A., Ma, M., Oikonomou, A.: History and Alternative Game Input Methods. In: The proceedings of the 2011 16th International Conference on Computer Games (CGAMES), Derby, UK, pp. 76–93 (2011)
9. Xia, L., Chen, C.C., Aggerwal, J.K.: Human Detection using depth information by Kinect. In: The proceedings of IEEE Computer Society Conference on Computer Vision and Pattern Recognition, pp. 15–22. IEEE (2011)
10. Joystiq, `http://www.joystiq.com/2010/06/19/kinect-how-it-works-from-the-company-behind-the-tech`
11. Microsoft, `http://www.microsoft.com/en-us/kinectforwindows/`
12. Totilo, S.: `http://kotaku.com/5680501/review-kinect?skyline=true&s=i`
13. Mohanram, N.K.: A Speech-based Quiz Game. Final Dissertation Report (2003)
14. Sadun, E., Sande, S.: Talking to Siri: Learning the Language of Apple's Intelligent Assistant. Que Publishing, USA (2012)
15. Turk, M., Robertson, G.: Perceptual user interfaces. Communications of the ACM 43(3), 33–34 (2000)

16. Nicholson, D.M., Fidopiastis, C.M., Davis, L.D., Schmorrow, D.D., Stanney, K.M.: An adaptive instructional architecture for training and education. In: Schmorrow, D.D., Reeves, L.M. (eds.) Augmented Cognition, HCII 2007. LNCS (LNAI), vol. 4565, pp. 380–384. Springer, Heidelberg (2007)
17. Turk, M.: Perceptual User Interfaces. In: NSF Workshop (2006)
18. Cohen, P.R., Johnston, M., McGee, D., Oviatt, S., Pittman, J., Smith, I., Chen, L., Clow, J.: QuickSet: Multimodal Interaction for Simulation Set-up and Control. In: The proceedings of the Fifth Conference on Applied Natural Language Processing, Washington, DC, USA (March 1997)
19. Pittman, J., Smith, I., Cohen, P.R., Oviatt, S.L., Yang, T.C.: QuickSet: A multimodal interface for militarysimulation. In: The Proceedings of the Sixth Conference on Computer Generated Forces and Behavioral Representation, pp. 217–224. Univ. of Central Florida, Orlando (1996)
20. Fidopiastis, C.M., Wiederhold, M.: Mindscape Retuning and Brain Reorganization with Hybrid Universes: The Future of Virtual Rehabilitation. In: Schmorrow, D., Cohn, J., Nicholson, D. (eds.) The PSI Handbook of Virtual Environments for Training & Education: Developments for the Military and Beyond, vol. 3, pp. 427–434. Praeger Security International, Westport (2008)
21. Lane, S.H., Marshall, H., Roberts, T.: Control interface for driving interactive characters in immersive virtual environments Technical Report, US Army Research (2006)
22. Lalor, E.C., Kelly, S.P., Finucane, C., Burke, R., Smith, R., Reilly, R.B., McDarby, G.: Steady-state VEP-based brain-computer interface control in an immersive 3D gaming environment. Eurasip J. on Appl. Sign Process. 19, 3156–3164 (2005)
23. Wolpaw, J.R., Birbaumer, N., McFarland, D.J., Pfurtschellere, G., Vaughan, T.M.: Brain-computer interfaces for communication and control. Clin. Neurophysiol. 113(6), 767–791 (2002)
24. Smith, J.D., Graham, T.C.N.: Use of eye movements for video game control. In: The proceedings of the ACE 2006, Hollywood, California, USA, June 14-16 (2006)
25. Hayhoe, M., Ballard, D.: Eye movements in natural behavior. Trends Cogn. Sci. 9(4), 188–194 (2005)
26. Tanaka, J., Gautghier, I.: Expertise in object and face recognition. The Psychology of Learning and Motivation 36, 83–125 (1997)
27. Darken, R., McDowell, P., Murphy, C.: Open Source Game Engines: Disruptive Technologies in Training and Education. In: Proceedings of theInterservice/Industry Training, Simulation and Education Conference (I/ITSEC). National Defense Industrial Association, Orlando (2005)
28. Dwyer, T., Griffith, T., Maxwell, D.: Rapid Simulation Development Using a Game Engine-Enhanced Dynamic Geo-Social Environment. In: The proceedings of Interservice/Industry Training, Simulation & Education Conference (I/ITSEC). NTSA, Orlando (2011)
29. Eisenberger, R., Jones, J.R., Stinglhamber, F., Shanock, L., Randall, A.T.: Flow Experiences at work; for high need achievers alone? J. Organ Behav. 26, 755–775 (2005)